

HYPER-TEXT-LINGUISTIK

TAH: EIN TEXTLINGUISTISCHES ANALYSEMODELL FÜR HYPERTEXTE

Theoretisch und praktisch exemplifiziert am Problemfeld der typisierten Links
von Hypertexten im World Wide Web

Inauguraldissertation
zur Erlangung des Doktorgrades der Philosophie
(Fach Germanistische Linguistik)
an der Ludwig-Maximilians-Universität München

vorgelegt von Oliver Huber aus München

für claire, für alles

*Linien aus Licht,
aufgereiht im Nichtraum des Verstandes,
Ballungen und Anordnungen von Daten.
Wie die Lichter einer Stadt, die sich langsam entfernen ...*

William Gibson *Neuromancer*

Danksagung

Die primäre Danksagung geht an den Betreuer meiner Arbeit, Herrn PD Dr. Wolfgang Schindler, für dessen kompetente Beratung aber auch äußerst liebenswürdige Unterstützung. Er stand in jeder Phase der Entstehung dieser Dissertation als Ansprechpartner in fachlichen, methodischen aber auch terminlichen Nöten zur Verfügung. Herzlicher Dank für konstruktive Kritik an der texttheoretischen Ausrichtung dieser Arbeit gebührt auch meinem Zweitkorrektor, Herrn Prof. Elmar Seebold. Späte Dankesworte gehen an meinen ersten Lehrer der Linguistik, Herrn Dr. Hubert Gindele. Ohne dessen begeisternde Lehre wäre meine Leidenschaft für die Sprachwissenschaft wohl nie in dieser Form geweckt worden.

Für viele linguistische und informationswissenschaftliche Fachgespräche und praktische Hilfen zur Thematik der Dissertation möchte ich mich ausdrücklich bei Frank Sanders bedanken. Für das Layout gab Gunter Mehler zahlreiche wichtige Hilfestellungen, für die Durchführung der Umfrage Kathi Schuler. Beim Korrekturlesen der Arbeit standen Bianca und Claire hilfreich zur Seite. Für die von ihnen erbrachten zeitlichen und nervlichen Opfer nochmals ein großes *Dankeschön*. Für sämtliche Fehler und Versäumnisse bleiben natürlich alleine ich selbst, ein gewisser Mangel an Schlaf und MS Word verantwortlich.

Intellektuelle Arbeit braucht auch emotionale Grundlagen. Für das Ertragen sämtlicher Höhen und Tiefen, die das Verfassen eines so umfangreichen Werkes mit sich bringt, danke ich besonders Claire sowie meiner Familie.

Abstract

This doctoral thesis suggests a text-linguistic approach dealing with linked electronic texts, so called HYPERTEXTS. After explaining the concept and philosophy of the hypertext-idea and elaborating on a pragmatic definition of WWW-hypertexts (Part A of the thesis), a linguistic framework, the *Textlinguistisches Analysemodell für Hypertexte (TAH)*, is proposed (Part B). TAH provides a pragmatic-functional, top-down model to analyze CONTEXT, FUNCTION and STRUCTURE of complete hypertexts and the hypertext-components NODE and LINK. Main aim of TAH is to strictly separate the different aspects of each hypertext- and linguistic level – without isolating them.

The theoretical approach of TAH is used for a case study (Part C). Therefore a corpus of WWW-hypertexts, mainly informative and instructive ones, is examined with TAH-methods. Aim of this study is on the one hand to provide some sort of evidence for the plausibility of TAH and is on the other hand linked with a specific question of research: It deals with possibilities provided by linguistics to enhance hypertext-links with logically, semantically and/or rhetorically typed information. The conclusion of the study is to distinguish exactly between semantics of a deepstructure-knowledge-presentation, semantic and rhetoric relations of the surface, i.e. the hypertext the user can see, and the layout or traversal behavior respectively of the user interface, i.e. the browser-software. Part C mainly deals with problems of typed links on the surface of hypertexts.

Keywords: text, hypertext, text linguistics, HTML, WWW, XML, typed links

Kompaktes Inhaltsverzeichnis

| | | |
|----------------|---|------------|
| Prolog: | Eine Bestandsaufnahme..... | 1 |
| 1 | Ausgangslage | 1 |
| 2 | Angaben zum Corpus | 6 |
| 3 | Angaben zur Literaturlage..... | 10 |
| Teil A: | Hypertext als Untersuchungsgegenstand..... | 13 |
| 4 | Hypertext: Ein historischer Abriß..... | 14 |
| 5 | Text, E-Text, Hypertext und Hypertext-Netz..... | 20 |
| 6 | Komponenten eines Hypertextes | 24 |
| 7 | Zusammenfassung: Hypertexte im WWW..... | 45 |
| Teil B: | Ein Textlinguistisches Analysemodell für Hypertexte (TAH) .. | 46 |
| 8 | Textdefinitionen..... | 47 |
| 9 | Textualitätskriterien | 54 |
| 10 | Dimensionen einer textlinguistischen Analyse..... | 74 |
| 11 | Paratextuelle Elemente | 82 |
| 12 | Makrostruktur und Makrostrukturanalyse..... | 86 |
| 13 | Referenzstruktur und referentielle Bewegung | 92 |
| 14 | Zusammenfassung: Das Analysemodell TAH..... | 100 |
| Teil C: | Fallbeispiel – Typisierte Links | 108 |
| 15 | Technische Voraussetzungen: Links im WWW..... | 110 |
| 16 | Vokabularien typisierter Links..... | 138 |
| 17 | Analyse des Corpus' mit Hilfe von TAH..... | 176 |
| Epilog: | Einsatzmöglichkeiten des TAH..... | 230 |
| Anhang: | Literatur und Verzeichnisse..... | 233 |
| A | Glossar | 234 |
| B | Fragebogen der Pilotstudie | 238 |
| C | Corpusverzeichnis | 239 |
| D | Literaturverzeichnis..... | 240 |

Komplettes Inhaltsverzeichnis

| | |
|---|-----------|
| Prolog: Eine Bestandsaufnahme..... | 1 |
| 1 Ausgangslage | 1 |
| 1.1 Textlinguistische Motivation | 1 |
| 1.2 Problemstellung..... | 2 |
| 1.3 Vorgehensweise | 3 |
| 1.4 Einschränkungen..... | 4 |
| 1.5 Zielgruppe | 5 |
| 2 Angaben zum Corpus | 6 |
| 2.1 WWW als technische Plattform..... | 6 |
| 2.2 Informationstexte als primäre Textsorte | 6 |
| 2.3 Die Texte des Corpus' | 8 |
| 3 Angaben zur Literaturlage..... | 10 |
| 3.1 Literatur zu Hypertext | 10 |
| 3.2 Literatur zur Textlinguistik | 12 |
| 3.3 Literatur zu Textlinguistik und Hypertext | 12 |
| Teil A: Hypertext als Untersuchungsgegenstand..... | 13 |
| 4 Hypertext: Ein historischer Abriss..... | 14 |
| 4.1 Die Pioniere: Bush, Nelson und Engelbart | 14 |
| 4.2 Der Durchbruch: Berners-Lee und das WWW | 17 |
| 5 Text, E-Text, Hypertext und Hypertext-Netz..... | 20 |
| 5.1 Abgrenzung Text vs. Teilttext, Hypertext vs. Knoten | 20 |
| 5.2 Abgrenzung E-Text, Hypertext und Hypertext-Netz..... | 22 |
| 6 Komponenten eines Hypertextes | 24 |
| 6.1 Allgemein | 24 |
| 6.2 Komponente I: Knoten..... | 24 |
| 6.2.1 Struktur..... | 25 |
| 6.2.2 Inhalt und Größe | 26 |
| 6.2.3 Darstellungsform..... | 27 |
| 6.2.4 Knoten-Typen..... | 27 |
| 6.3 Komponente II: Links | 28 |
| 6.3.1 Allgemeine Eigenschaften | 29 |
| 6.3.2 Formale Unterscheidungsmöglichkeiten | 30 |
| 6.3.3 Inhaltliche Unterscheidungsmöglichkeiten..... | 32 |
| 6.3.4 Unterscheidung nach dem Travesierungsverhalten | 33 |

| | | |
|----------------|---|-----------|
| 6.3.5 | Kritik der Link-Typologie..... | 34 |
| 6.4 | Navigations- und Orientierungskomponenten | 35 |
| 6.4.1 | Textsequenzierung, Navigation und Browsing..... | 35 |
| 6.4.2 | Orientierungsprobleme und kognitive Mehrbelastung | 36 |
| 6.4.3 | Traditionelle Orientierungshilfen..... | 38 |
| 6.4.4 | Verzeichnisse | 39 |
| 6.4.5 | Metainformationen | 39 |
| 6.4.6 | Typographische Auszeichnungen | 40 |
| 6.4.7 | Hypertext-spezifische Orientierungs- und Navigationshilfen | 40 |
| 6.4.8 | Browser-Software | 40 |
| 6.4.9 | Backtracking, Historie, Bookmarks | 42 |
| 6.4.10 | Graphische Übersichten | 43 |
| 6.4.11 | Anfrage- bzw. Suchmechanismen: Information Retrieval..... | 43 |
| 6.4.12 | Guided Tours..... | 44 |
| 7 | Zusammenfassung: Hypertexte im WWW..... | 45 |
| Teil B: | Ein Textlinguistisches Analysemodell für Hypertexte (TAH) .. | 46 |
| 8 | Textdefinitionen..... | 47 |
| 8.1 | <i>Text</i> als vorwissenschaftliche Größe..... | 47 |
| 8.2 | Text: Produkt und Produktion..... | 47 |
| 8.3 | Die Entwicklung des Terminus' <i>Text</i> | 48 |
| 8.4 | Exkurs: Einheiten zwischen Text und Satz..... | 52 |
| 9 | Textualitätskriterien | 54 |
| 9.1 | Kohäsion..... | 55 |
| 9.2 | Kohärenz | 58 |
| 9.2.1 | Kohärenz im Hypertext..... | 60 |
| 9.2.2 | Exkurs: Kohärenzerzeugung durch Schemata | 60 |
| 9.3 | Intentionalität | 65 |
| 9.4 | Akzeptabilität | 66 |
| 9.5 | Informativität..... | 68 |
| 9.6 | Situationalität | 69 |
| 9.7 | Intertextualität | 70 |
| 9.8 | Zusammenfassung..... | 72 |
| 10 | Dimensionen einer textlinguistischen Analyse..... | 74 |
| 10.1 | Allgemein | 74 |
| 10.2 | Kontextuelle Bedingungen..... | 74 |
| 10.2.1 | Kommunikationsform | 75 |

| | | |
|----------------|---|------------|
| 10.2.2 | Handlungsbereich..... | 76 |
| 10.3 | Textfunktion..... | 77 |
| 10.4 | Textstruktur..... | 78 |
| 10.5 | Textsorte..... | 79 |
| 11 | Paratextuelle Elemente..... | 82 |
| 11.1 | Traditionelle paratextuelle Elemente von Texten..... | 82 |
| 11.2 | Paratextuelle Elemente von Hypertexten..... | 83 |
| 12 | Makrostruktur und Makrostrukturanalyse..... | 86 |
| 12.1 | Allgemeines: Textthema und Makrostruktur..... | 86 |
| 12.2 | Exemplarische Makrostrukturanalyse eines Hypertextes..... | 88 |
| 13 | Referenzstruktur und referentielle Bewegung..... | 92 |
| 13.1 | Referenzbeziehungen in Texten..... | 92 |
| 13.2 | Modell der referentiellen Bewegung..... | 93 |
| 13.3 | Referentielle Bewegung in Hypertexten..... | 94 |
| 13.4 | Die Rolle der Quaestio..... | 98 |
| 14 | Zusammenfassung: Das Analysemodell TAH..... | 100 |
| 14.1 | Ausrichtung und Vorteile des Analysemodells..... | 100 |
| 14.2 | Einschränkungen..... | 101 |
| 14.3 | Überblicksdarstellung von TAH..... | 103 |
| 14.4 | Anmerkungen zu den Untersuchungsebenen..... | 105 |
| 14.4.1 | Anmerkungen zur Ebene <i>Hypertext</i> | 105 |
| 14.4.2 | Anmerkungen zu den Ebenen <i>Knoten</i> , <i>Absatz</i> und <i>Satz</i> | 105 |
| 14.4.3 | Anmerkungen zur Ebene <i>Link</i> | 106 |
| Teil C: | Fallbeispiel – Typisierte Links..... | 108 |
| 15 | Technische Voraussetzungen: Links im WWW..... | 110 |
| 15.1 | Datenhaltung, Zwischenschicht und Oberflächenstruktur..... | 110 |
| 15.1.1 | Von HTML zum Semantic Web..... | 110 |
| 15.1.2 | Die Ebenen einer Link-Typisierung..... | 111 |
| 15.1.3 | Die Oberflächenstruktur als Problembereich der Untersuchung .. | 115 |
| 15.2 | Linking in HTML und XML..... | 117 |
| 15.2.1 | HTML als SGML-Anwendung..... | 118 |
| 15.2.2 | HTML als XML-Anwendung..... | 124 |
| 15.2.3 | Das Hyperlink-Modell von XML: Xlink..... | 129 |
| 15.2.4 | Implementierung typisierter Links im Corpus’..... | 130 |
| 15.2.5 | Mögliche zukünftige Implementierungen..... | 132 |

| | | |
|----------------|--|------------|
| 16 | Vokabularien typisierter Links..... | 138 |
| 16.1 | Problemstellung..... | 138 |
| 16.2 | Linguistische Ansätze | 139 |
| 16.2.1 | Relation Element Theory..... | 139 |
| 16.2.2 | Rhetorical Structure Theory..... | 145 |
| 16.3 | Hypertext-spezifische Ansätze..... | 154 |
| 16.3.1 | Allgemeine Anmerkungen und kleine Pilotstudie | 154 |
| 16.3.2 | Typisierte Links in älteren Hypertext-Systemen..... | 157 |
| 16.3.3 | Typisierte Links in HTML | 161 |
| 16.4 | Zusammenfassung: Geeignete Vokabularien | 175 |
| 17 | Analyse des Corpus' mit Hilfe von TAH..... | 176 |
| 17.1 | Vorbemerkung zur Analyse | 176 |
| 17.2 | Allgemeine, exemplarische Analyse des Corpus' | 177 |
| 17.2.1 | Kontextuelle Aspekte | 177 |
| 17.2.2 | Kommunikativ-funktionale Aspekte..... | 179 |
| 17.2.3 | Konventionelle Aspekte | 182 |
| 17.2.4 | Strukturelle und paratextuelle Aspekte | 182 |
| 17.2.5 | Intertextuelle Aspekte | 190 |
| 17.3 | Analyse der Typisierten Links | 194 |
| 17.3.1 | Analyse der proprietären Codierung von SELF..... | 194 |
| 17.3.2 | Analyse der proprietären Codierung von SELF8..... | 203 |
| 17.3.3 | Analyse der proprietären Codierung von GRAMMIS..... | 206 |
| 17.3.4 | Analyse der proprietären Codierung von SPORT1 | 211 |
| 17.3.5 | Versuch einer RST-Codierung des Corpus' | 216 |
| 17.3.6 | Versuch einer Codierung des Corpus' mit Hilfe des Dublin Core..... | 218 |
| 17.3.7 | Versuch einer Codierung des Corpus' mit Hilfe des <i>rel</i> -Attributs..... | 219 |
| 17.4 | Zusammenfassung: Auswertung der Ergebnisse | 220 |
| 17.4.1 | Aussagen zur Methode..... | 220 |
| 17.4.2 | Aussagen zu inhaltlichen Fragestellungen bzgl. der Ebene <i>Link</i> | 222 |
| Epilog: | Einsatzmöglichkeiten des TAH..... | 230 |
| Anhang: | Literatur und Verzeichnisse..... | 233 |
| A | Glossar | 234 |
| B | Fragebogen der Pilotstudie | 238 |
| C | Corpusverzeichnis | 239 |
| D | Literaturverzeichnis..... | 240 |

Abbildungsverzeichnis

| | | |
|----|--|-----|
| 1 | Knotenstruktur | 25 |
| 2 | Intra-, inter- und extra-hypertextuelle Links | 31 |
| 3 | Link-Klassifizierung nach Lokalität vs. Globalität | 31 |
| 4 | Oberfläche einer Browser-Software | 41 |
| 5 | Link-Struktur eines einfachen Hypertextes | 89 |
| 6 | MODELL DER REFERENTIELLEN BEWEGUNG | 94 |
| 7 | Ebenen der Typisierung | 112 |
| 8 | Typisierung durch ‚Mißbrauch‘ des title-Attributs | 131 |
| 9 | Link-Typisierung durch Piktogramme | 132 |
| 10 | TOPIC MAPS – TOPICS UND OCCURRENCES | 135 |
| 11 | TOPIC MAPS – TOPIC ASSOCIATIONS | 137 |
| 12 | Einfach Textanalyse mit RST | 151 |
| 13 | SELF [tcab.htm] Abschnitt <i>SGML-gerechten Dokumenttyp angeben</i> | 169 |
| 14 | Windows-Ordnerstruktur als Metapher für die Knoten-Hierarchie in SELF8 | 184 |
| 15 | Pull-Down-Menü <i>Quickbar</i> als hypertext-interne Navigationshilfe von SELF8 | 186 |
| 16 | Pull-Down-Menü <i>Übersicht</i> als hypertext-interne Navigationshilfe von GRAMMIS | 188 |
| 17 | Navigationsleisten und Suchfunktion in W3C | 190 |
| 18 | Untypisierte Links in W3C [/Math/] Abschnitt <i>Highlights</i> | 191 |
| 19 | Extra-hypertextueller Link in SELF [tgba.htm] Abschnitt <i>WWW-Server für MS</i> | 191 |
| 20 | SELF8 [cgiperl/intro/webserverpc.htm] Abschnitt <i>Geeignete Produkte</i> | 192 |
| 21 | Platzhalter-Knoten in Sport1 [dynamisch erzeugter Knoten] | 193 |
| 22 | SELF [tdbc.htm] Abschnitt <i>Schnellformatierung im Text</i> | 195 |
| 23 | SELF [tdca.htm] Abschnitt <i>Textschatten (text-shadow)</i> | 201 |
| 24 | SELF8 [editorial/hinweise.htm] Abschnitt <i>Neu in Version 8.0</i> | 203 |
| 25 | SELF8 [css\formate\direkt.htm] Abschnitt <i>Formatdefinitionen</i> | 205 |
| 26 | SELF8 [css\formate\anzeige\styleattr.htm] Abschnitt <i>Die Seite ...</i> | 206 |
| 27 | GRAMMIS [grammis_sys.ansicht] Abschnitt <i>Konnektoren (Auswahl Konnektoren)</i> | 207 |
| 28 | Piktogramme in SPORT1 [dynamisch erzeugter Knoten] | 211 |
| 29 | Mouse-Over-Text in Sport1 [dynamisch erzeugter Knoten] | 213 |
| 30 | RST-typisierter Knoten SELF [tdca.htm] Abschnitt <i>Textschatten (text-shadow)</i> | 217 |
| 31 | Style-Sheets der Link-Typenklassen | 228 |

Tabellenverzeichnis

| | | |
|----|---|-----|
| 1 | Hypertexte des Corpus' | 9 |
| 2 | Makrostrukturanalyse von Knoten <i>Buchvorstellung</i> | 90 |
| 3 | Makrostrukturanalyse von Knoten <i>Seebold</i> | 90 |
| 4 | Referentielle Bewegung in einem Hypertext | 95 |
| 5 | Das TEXTLINGUISTISCHE ANALYSEMODELL FÜR HYPERTEXTE (TAH) – Ebene <i>Hypertext</i> | 103 |
| 6 | Das TEXTLINGUISTISCHE ANALYSEMODELL FÜR HYPERTEXTE (TAH) – Ebene <i>Knoten</i> | 103 |
| 7 | Das TEXTLINGUISTISCHE ANALYSEMODELL FÜR HYPERTEXTE (TAH) – Ebene <i>Link</i> | 104 |
| 8 | RET-Familie 1: ELEMENTS OF INTENSIONAL FORCE..... | 140 |
| 9 | RET- Familie 2: DIMENSIONAL ELEMENTS | 141 |
| 10 | RET-Familie 3: ELEMENTS OF AGREEMENT | 141 |
| 11 | RET-Familie 4: PROPOSITIONAL ELEMENTS | 141 |
| 12 | RET-Familie 5: ELEMENTS OF PART-WHOLE INCLUSION | 141 |
| 13 | Komplexe RET-Familie 1: CONTRAST..... | 142 |
| 14 | Komplexe RET-Familie 2: SIMILARS..... | 142 |
| 15 | Komplexe RET-Familie 3: CLASS INCLUSION | 142 |
| 16 | Komplexe RET-Familie 3: CASE RELATION | 142 |
| 17 | Komplexe RET-Familie 5: PART-WHOLE..... | 142 |
| 18 | Presentational Relations der RST..... | 147 |
| 19 | Subject Matter Relations der RST..... | 148 |
| 20 | Multinuclear Relations der RST..... | 148 |
| 21 | Relation <i>Evidence</i> als Beispiel einer PRESENTATIONAL RELATION der RST | 149 |
| 22 | Relation <i>Condition</i> als Beispiel einer SUBJECT MATTER RELATION der RST | 149 |
| 23 | Relation <i>Contrast</i> als Beispiel einer MULTINUCLEAR RELATION der RST..... | 149 |
| 24 | Ergebnis der Pilotstudie | 156 |
| 25 | Heuristische Typensammlung der Pilotstudie | 156 |
| 26 | Link-Typen nach Duncan..... | 158 |
| 27 | Link-Typen des AUTHOR'S ARGUMENTATION ASSISTANT (AAA)..... | 160 |
| 28 | Relationen der <i>rel</i> - und <i>ref</i> -Attribute..... | 163 |
| 29 | Relationen des DUBLIN CORE..... | 165 |
| 30 | Piktogramme zur Link-Typisierung in SELF..... | 168 |
| 31 | Zusätzliche Piktogramme zur Link-Typisierung in SELF8 | 170 |
| 32 | QBullets – Navigationssymbole..... | 172 |
| 33 | QBullets – Multimediale Symbole..... | 172 |
| 34 | QBullets – Semantische Symbole | 172 |
| 35 | QBullets – Netzwerkspezifische Symbole | 172 |
| 36 | QBullets – Vermischte Symbole | 172 |
| 37 | Funktion der Piktogramme in SPORT1 | 213 |
| 38 | Vorschlag einer Link-Typologie | 227 |
| 39 | Mediale Modifizierer | 228 |
| 40 | Modifizierer des Browser-Verhaltens | 228 |
| 41 | Glossar | 237 |

Abkürzungen

Für diese Arbeit geprägte Abkürzung:

- TAH: TEXTLINGUISTISCHES ANALYSEMODELL FÜR HYPERTEXTE

Wichtige allgemeine Abkürzungen (Erläuterungen im Glossar):

- CSS: CASCADING STYLE-SHEETS
- DC: DUBLIN CORE
- HTML: HYPERTEXT MARKUP LANGUAGE
- RDF: RESOURCE DESCRIPTION FRAMEWORK
- RET: RELATION ELEMENT THEORY
- RST: RHETORICAL STRUCTURE THEORY
- TM: TOPIC MAP
- XML: EXTENSIBLE MARKUP LANGUAGE

Die Abkürzungen der Corpus-Hypertexte sind in Kapitel 2.3 aufgelistet.

Typographische Konventionen

- | | |
|---|---|
| - Fachtermini in Kapitälchen | KOHÄRENZ ist die ... |
| - Objektsprache kursiv | der Eigenname <i>Peter</i> |
| - Objektsprachliche Termini in Kapitälchen | der Terminus KOHÄRENZ |
| - Texte des Corpus' in Großbuchstaben | SELF |
| - Zitate in doppelten Anführungszeichen | „Hiermit ist gemeint [...]“ |
| - Indirektes in einfachen Anführungszeichen | der sogenannte ‚Erfolg‘ |
| - Hyper-Links in Zitaten unterstrichen | „mehr dazu <u>hier</u> “ |
| - Link-Ziele in Zitaten kursiv | „ <i>Etymologie</i> : Theorie der“ |
| - Hervorhebungen im Fließtext unterstrichen | <u>die</u> Definition von Text |
| - Knotengrenzen in Zitaten als Längsstrich | „... <u>hier</u> ... <i>Dazu ist</i> ...“ |

Zitierkonventionen

- Zitate aus gedruckter Sekundärliteratur werden wie gewöhnlich durch ein Kürzel, bestehend aus Autorennamen und Jahr, sowie der Seitenzahl wiedergegeben. Beispiel: Brinker 2001 S. 112
- Zitate aus elektronischer Sekundärliteratur werden durch das Kürzel des gesamten Hypertextes, den Dateinamen des Knotens sowie die Abschnittsüberschrift der fraglichen Textstelle wiedergegeben. Beispiel: Münz 1997b [tcai.htm] Abschnitt *Verweistexte*
- Hypertext-Knoten des Corpus' werden durch das Corpuskürzel, den Dateinamen des Knotens sowie die Abschnittsüberschrift der fraglichen Textstelle zitiert. Beispiel: GRAMMIS [grammis_sys.ansicht] Abschnitt *Konnektoren*

Rechtschreibkonventionen

- Die Arbeit hält sich an die alten Rechtschreibregeln.
- Sämtliche Komposita, die eine informationswissenschaftliche Komponente enthalten, sind mit Bindestrich geschrieben (z.B.: *Ziel-Knoten*, *WWW-Hypertext*).

Prolog: Eine Bestandsaufnahme

1 Ausgangslage

Die technische Entwicklung der Informationstechnologien schreitet mit großer Geschwindigkeit voran. Während Begriffe wie INTERNET, HTML und WWW vor fünf Jahren höchstens einigen Interessierten bekannt waren, wird heute damit geworben, daß vom Enkel bis zur Oma die gesamte Familie ‚drin‘ ist – im weltweiten Datennetz. Während gesellschaftliche und wissenschaftliche Auswirkungen dieser Änderungen noch nicht mal im Ansatz aufgearbeitet worden sind, hat auf technischer Seite bereits die zweite ‚Internet-Revolution‘ begonnen: die der Umwandlung des heutigen darstellungsorientierten WWW in ein Netzwerk von Bedeutungen, in ein SEMANTIC WEB. Eines bleibt jedoch unverändert: Das WWW wird nach wie vor aus verknüpften elektronischen Texten bestehen – aus sogenannten HYPERTEXTEN.

Den sich immer schneller entwickelnden Technologien muß fachwissenschaftliche Grundlagenforschung entgegen bzw. zur Seite gestellt werden. Die vorliegende Arbeit ist ein technisch orientierter Beitrag zur textlinguistischen Grundlagenforschung. Ihr Ziel ist zu zeigen, daß mit aktuellen textlinguistischen Möglichkeiten durchaus fruchtbare Untersuchungen von Hypertexten möglich sind.

1.1 Textlinguistische Motivation

Eine Frage drängt sich freilich gleich zu Beginn auf: Ist die Textlinguistik denn überhaupt für Hypertexte zuständig? Hierauf sei mit einem Bonmot aus der textlinguistischen Hypertext-Forschung geantwortet:

„Wenn sich nicht herausstellen sollte, daß HT [= Hypertext] im selben Umfang Texte sind wie Walfische Fische, dann gilt: Die Linguistik beschäftigt sich unter anderem mit Texten. HT sind Texte. Also beschäftigt sich die Linguistik mit HT.“

(Freisler 1994 S. 5)

Ob nun wirklich dort ‚Text drin ist, wo Text drauf steht‘, wird erst die genauere Betrachtung des Untersuchungsgegenstandes zeigen. Nichts scheint jedoch auf den ersten Blick näher zu liegen, als daß sich die für Texte zuständige linguistische Teildisziplin der TEXTLINGUSTIK dieser elektronischen Gebilde annimmt. Problematisch ist hierbei nur, daß die Textlinguistik noch immer um die Definition der Größe TEXT ringt. Zugegebenermaßen liegen hier die

Verhältnisse deutlich unklarer als etwa bei der linguistischen Größe SATZ. Allerdings arbeiten auch Syntaktiker nicht mit einer allgemein anerkannten Festlegung letzterer Größe, sondern mit unterschiedlichen, der jeweiligen Fragestellung angepaßten, theoretisch determinierten Satzdefinitionen – und dies mit beachtlichen Erfolgen. In der vorliegenden Arbeit wird ein KOMMUNIKATIV-FUNKTIONALER TEXTBEGRIFF vorgestellt, der sich für praktische Untersuchungen von Hypertexten eignet. Auf (durchaus existierende) theoretische Unzulänglichkeiten bzw. Gegenentwürfe wird hierbei stets hingewiesen.

Bis jetzt wurden hauptsächlich Untersuchungen zu einzelnen Aspekten der TEXTUALITÄT von Hypertexten (meist kontrastiv zu herkömmlichen Texten) durchgeführt. Im Vordergrund standen primär Probleme von KOHÄSION und KOHÄRENZ. Die vorliegende Arbeit möchte hingegen einen textlinguistischen Orientierungsrahmen für sich anschließende, breiter angelegte praktische textwissenschaftliche (siehe unten) Untersuchungen von Hypertexten liefern. Hypertexte werden hierfür in mehrere Ebenen unterteilt und auf verschiedene Aspekte hin untersucht. Ein solcher Beschreibungsapparat wird in dieser Arbeit als TAH – als TEXTLINGUISTISCHES ANALYSEMODELL FÜR HYPERTEXTE – eingeführt und exemplarisch auf dessen Anwendbarkeit hin geprüft.

1.2 Problemstellung

Als Untersuchungsgegenstand werden, wie gesagt, keine gedruckten Texte oder Diskurse der gesprochenen Sprache verwendet, sondern Hypertexte. Die Besonderheit dieser Texte ist neben ihrer elektronischen Realisierung eine tendenziell ‚nicht-lineare‘ Anordnung der Information in Teiltexen, sogenannte KNOTEN, die miteinander über Verweise, sogenannte LINKS, verbunden sind. Bei der Lektüre entscheidet der Leser in der Regel relativ frei über die Reihenfolge der zu rezipierenden Knoten. *Relativ* deshalb, weil der Autor mit dem Setzen von Links natürlich die Zahl der NAVIGATIONS-MÖGLICHKEITEN¹ während der Lektüre begrenzt: Je mehr Links innerhalb eines Knotens existieren, um so mehr mögliche LESEPFADEN kann der Leser bei der Lektüre einschlagen.

Eine allgemein akzeptierte Definition des Phänomens HYPERTEXT steht noch aus, was wiederum nur verständlich ist, da Hypertexte ja selbst (jedenfalls zum

¹ Dieser Vorgang der Textsequenzierung innerhalb eines Hypertextes durch den Leser wird oft mit *Navigieren in einem Hypertext* paraphrasiert.

größten Teil) aus der noch nicht exakt festgelegten linguistischen Größe TEXT bestehen. Ebenso wie bei der Diskussion des Textbegriffes werden für den Terminus HYPERTEXT unterschiedliche Ansätze der Sekundärliteratur kritisch vorgestellt. An dieser Stelle sollte jedoch klar sein, daß nicht auf alle offenen Fragen der HYPERTEXT-THEORIE eingegangen werden kann. Es wird sich jedoch zeigen, daß für die Struktur von Hypertexten eine relativ gut verwendbare und technisch saubere Arbeitsdefinition zu finden ist.¹ Auch die für Teil C dieser Arbeit relevanteste Komponente des Hypertextes, der LINK, ist zwar intuitiv gut faßbar, theoretisch jedoch nicht ganz einfach festzulegen. Aber auch hier läßt sich eine Definition erarbeiten, die ausreichend für die vorliegende Untersuchung ist.

1.3 Vorgehensweise

Der Untersuchungsgegenstand HYPERTEXT wird im Hauptteil A ausführlich eingeführt, das textlinguistische Instrumentarium in Teil B erarbeitet. Als Kernstück der Arbeit wird am Ende des Hauptteils B ein TEXTLINGUISTISCHES ANALYSEMODELL FÜR HYPERTEXTE (TAH) vorgeschlagen, das im folgenden Fallbeispiel des Hauptteils C zur Untersuchung der folgenden Problematik exemplarisch verwendet wird: Jeder Link stellt den Leser eines Hypertextes vor die Entscheidung, ob er im aktuellen Knoten weiterlesen oder aber dem Link folgend zu dessen ZIEL-KNOTEN² wechseln will. Dies kann unter Umständen zu einer kognitiven Mehrbelastung des Lesers führen und ihm die Orientierung in komplexeren Hypertexten erschweren.³ Eine Verbesserung der Situation wäre eventuell dadurch zu erreichen, daß man dem Leser, bevor dieser einen Link tatsächlich verfolgt, signalisiert, um was für eine Art von Knoten es sich beim Ziel des Links handelt; also etwa, ob dort ein *Beispiel* für den beschriebenen Sachverhalt des Ausgangs-Knotens zu finden ist, eine *Gegenposition* zu vorher Gesagtem o.ä. Auch Aussagen zum Medium des Ziel-Knotens sind denkbar, etwa ob es sich um einen textuellen Knoten oder ein Film- bzw. Tondokument handelt. Der Fachausdruck für solche Links mit Signalfunktion ist TYPISIERTER LINK. Typisierte Links werden im Rahmen des Fallbeispiels mit Hilfe von TAH untersucht: Dort wird gefragt, welche Verknüpfungsmöglichkeiten

¹ Dies freilich auf Grundlage des Textbegriffes, der in dieser Arbeit verwendet wird und auf dessen theoretische Unvollständigkeiten bereits hingewiesen wurde.

² Die Corpus-Hypertexte haben keine Links, die über mehr als einen Zielknoten verfügen. Mit neuen Technologien ist dies jedoch denk- und machbar, mehr dazu auch in Kapitel 15.2.3.

³ Siehe hierzu Kapitel 6.4.2.

rhetorischer, semantischer und/oder logischer Natur prinzipiell zwischen Textteilen existieren und ob, bzw. wie, diese zu klassifizieren sind. Herangezogen werden hierfür linguistische und informationswissenschaftliche Typologien sowie die proprietären Typisierungsansätze der Corpus-Hypertexte.

1.4 Einschränkungen

Vorliegende Untersuchung ist, wie gesagt, eine genuin textlinguistische. Dies scheint auf den ersten Blick insofern problematisch, als daß sich seit längerem eine tendenziell interdisziplinäre Textwissenschaft¹ auszubilden versucht, die neben der Linguistik auch Disziplinen wie *Psychologie*, *Neurobiologie*, *Soziologie* und weitere Teilbereiche integriert. Diese Herangehensweise erhält durch das sich in (bestimmten) sprachwissenschaftlichen Kreisen immer mehr durchsetzende Selbstverständnis der Linguistik als Teildisziplin einer allgemeinen Kognitionswissenschaft² vermehrt Rückendeckung. Auch der Untersuchungsgegenstand HYPERTEXT ist natürlich für die unterschiedlichsten Disziplinen relevant, wie die Veröffentlichungsflut aus BWL, Soziologie, Philosophie, Literaturtheorie u.ä. zu diesem Thema zeigt (siehe Kapitel 3).

Meiner Ansicht nach hängt ein interdisziplinäres Projekt bei allen möglichen Synergieeffekten jedoch stark vom Input der Einzeldisziplinen ab. Die Textlinguistik als relativ neue sprachwissenschaftliche Teildisziplin kann im Moment noch nicht die Ergebnisse liefern, die linguistische Grundlage für eine fächerübergreifende Untersuchung von kognitiver TEXTVERARBEITUNG oder gar HYPER-TEXT-VERARBEITUNG sein könnte. Auf die Integration des deskriptiven Analysemodells in ein umfassendes kognitives Modell wurde daher bewußt verzichtet, TAH soll als rein textlinguistischer Beitrag zur Beschreibung hyper-textueller Funktionen und Strukturen gesehen werden.³ Insofern versucht die vorliegende Untersuchung auch, einen (bescheidenen) Teil der noch ausstehenden ‚Hausaufgaben‘ aus dem Bereich der Textlinguistik zu erledigen. Die Ergebnisse könnten dann ihrerseits Input für empirische Untersuchungen und angewandte Forschung sein, etwa für die Implementierung einer semantisch angereicherten Hypertext-Oberfläche (siehe Kapitel 15.1.3) unter Benutzung

¹ Vgl. hierzu etwa Kapitel 1 von van Dijk 1980a, Strohner 1990 sowie Friederici 1999.

² Vgl. hierzu etwa Rickheit u. Strohner 1993, Sucharowski 1996 sowie Pörings u. Schmitz 1999.

³ Auf eine mögliche, voraussichtlich unproblematische Einbettung des deskriptiven TAH in ein größeres kognitives Erklärungsmodell der Hypertext-Verarbeitung wird im Kapitel *Ausblick* dieser Arbeit hingewiesen.

neuer Texttechnologien wie XLINK, XML-TOPIC-MAP und weiterer XML-DERIVATE.

1.5 Zielgruppe

Die Arbeit wendet sich explizit an ein linguistisches Fachpublikum mit gewissen interdisziplinären Neigungen sowie einem besonderen Interesse an moderner Texttechnologie. Allgemeine linguistische Fachbegriffe werden daher ohne gesonderte Vorstellung verwendet. Textlinguistische Termini werden hingegen ausführlich eingeführt und kritisch besprochen. Für das Verständnis ist eine gewisse Vertrautheit mit dem Internet sowie die Kenntnis informationswissenschaftlicher Grundbegriffe hilfreich – aber nicht unabdingbar. Die für die Arbeit relevanten technischen Konzepte, Implementierungen und Spezifikationen werden von Grund auf eingeführt. Hierbei wird, soweit dies die Komplexität der Materie ermöglicht, ‚behutsam‘ vorgegangen – ohne jedoch übermäßig stark zu vereinfachen. Der technisch interessierte Leser wird hauptsächlich in den Fußnoten auf Einschränkungen, Problembereiche und weiterführende Literatur aus der Informatik verwiesen. Die wichtigsten technischen Begriffe sind darüber hinaus in einem Glossar aufgeführt.

Die Arbeit dürfte durchaus auch für Informatiker interessant sein, die gewillt sind, sich mit sprachwissenschaftlichen Fragestellungen auseinanderzusetzen. In terminologischen Nöten sei auf die allgemeinen linguistischen Nachschlagewerke aus Kapitel 3.2 verwiesen.

Leser, die mit den Konzepten von Hypertext und den Implementierungen von Markup-Sprachen bereits vertraut sind, können die ausführlichen informationswissenschaftlichen Einführungen des Kapitels 15 getrost ‚querlesen‘.

2 Angaben zum Corpus

Nachdem die allgemeine Ausrichtung der Untersuchung im letzten Kapitel erläutert wurde, sollen nun die Präferenzen genannt werden, die für die Auswahl der Hypertexte des Corpus' von Bedeutung waren.

2.1 WWW als technische Plattform

Bei der Auswahl des Corpus' müssen gewisse Festlegungen gemacht werden, um die Vergleichbarkeit der Corpus-Hypertexte zu gewährleisten: Zuerst sollte man sich auf eine mediale Realisierung, eine technische Plattform der zu untersuchenden Hypertexte einigen. Zahlreiche Softwaresysteme stehen nämlich zur Implementierung eines Hypertextes zur Verfügung (vgl. Schnupp 1992 S. 97 f.). Als derzeit mit Abstand bedeutendste elektronische Realisierung darf das WORLD WIDE WEB (WWW) gelten.¹ Alle anderen Hypertext-Systeme führen heute mehr oder weniger ein Nischendasein.² Aus diesem Grund wird sich die Arbeit ausschließlich mit im WWW präsenten und in der Sprache HTML³ geschriebenen Hypertexten beschäftigen. Nur bei relevanten Besonderheiten anderer Systeme wird auf diese Implementierungen kurz eingegangen.

2.2 Informationstexte als primäre Textsorte

Die zweite Einschränkung betrifft die TEXTSORTEN⁴, die es zu untersuchen gilt. Im WWW sind verschiedene Textsorten zu finden, HYPertext selbst hingegen ist nicht als eine Textsorte zu interpretieren (siehe Kapitel 10.5)! Zweifelsohne werfen die verschiedenen, hypertextuell realisierten Textsorten jeweils spezifische textlinguistische Fragestellungen auf. Zu denken ist hier auch an die unter dem Begriff HYPERFICTION bekanntgewordene, elektronisch realisierte Belletristik, die versucht, im neuen Medium einen ästhetischen Mehrwert zu erlangen.

¹ Die Entwicklung des Word Wide Web und des Internet wird in Kapitel 4.2 zusammenfassend beschrieben.

² Drei Gründe lassen sich nennen, wieso Hypertext-Realisierungen überwiegend im WWW stattfinden: Die Hypertexte können mit der im Hinblick auf das Betriebssystem unabhängigen Hypertext-Beschreibungssprache HTML auf einer nahezu beliebigen Rechnerarchitektur implementiert werden. Die erstellten Hypertexte werden in das weltweite Computernetz INTERNET eingespielt, das noch immer mit einem exponentiellen Wachstum der Benutzerzahlen aufwarten kann. Die zur Ansicht von WWW-Hypertexten notwendige Browser-Software wird von den meisten Firmen aus Marketinggründen verschenkt (vgl. auch: Münz 1997a [htxt212.htm]).

³ Eine kurze Einführung in HTML gibt das Kapitel 15.2.

⁴ Sämtliche textlinguistischen Begriffe werden ausführlich in Teil B eingeführt.

Für die vorliegende Untersuchung werden jedoch hauptsächlich Informations- und Instruktionstexte herangezogen, die sich mit Informationstechnologien oder linguistischen Fragestellungen beschäftigen. Diese Einschränkung hat mehrere Ursachen methodischer, praktischer und persönlicher Natur: Zum einen ist die Spezialisierung auf diese Textsorte durch den beruflichen Werdegang des Autors zu erklären, der sich über mehrere Jahre hinweg mit dem Verfassen von Anleitungen und Lernmedien für Software sowie mit Programmierung beschäftigt hat. Schon aus diesem Grund lag es nahe, vorhandenes Expertenwissen für eine spezielle Textsorte in eine textlinguistische Untersuchung einfließen zu lassen. Andererseits wurden gerade Informations- und Instruktionstexte von Anfang an massiv als Hypertexte realisiert und machen auch heute (noch) die Mehrheit der Hypertexte im WWW aus.¹ Auch lautet eine Hypothese des Autors, daß solche Textsorten tendenziell geeigneter für eine semanto-pragmatische Typisierung der Links sind, da sie vermutlich vermehrt über standardisierte inhaltliche Verknüpfungen zwischen Textteilen (bzw. Knoten) verfügen. Ein Beispiel aus dem Hypertext SELF des Corpus' kann dies veranschaulichen:²

„Spaltenabstand (column-gap)
Anzeigebeispiel: So sieht's aus

Sie können bestimmen, wie groß der Abstand zwischen den einzelnen Spalten sein soll. Bei mehr als zwei Spalten gilt diese Angabe für alle Spaltenabstände.

Beispiel (Style-Sheet-Definition für HTML-Tag im Text):
<div style="columns:3; column-gap:0.5cm">Inhalt</div>

Erläuterung:

Mit column-gap: können Sie den Spaltenabstand bestimmen. Erlaubt ist eine numerische Angabe.“

(SELF [tdci.htm] Abschnitt: *Spaltenabstand (column-gap)*)

Die obige Erklärung gliedert sich in mehrere Teile:

- Eine Einleitung, die anreißt, um was es überhaupt geht (*Sie können bestimmen, wie ...*)
- Ein Beispiel, welches einen exemplarischen Quellcode zeigt, sowie ein Anzeigebeispiel, welches die Auswirkungen des eben genannten Quellcodes

¹ In Kapitel 4.2 wird über die Motivationen zur Gründung des WWW zu lesen sein, daß diese primär darin bestand, das „problem of information access“ zu lösen (vgl. Berners-Lee 1989).

² Die Unterstreichungen symbolisieren Links (siehe hierzu auch den Abschnitt *Typographische Konventionen* dieser Arbeit).

auf dem Bildschirm illustriert. Das Anzeigebeispiel ist in einen separaten Knoten ausgelagert und kann per Klick auf den LINK *So sieht's aus* aufgerufen werden.

- Eine abschließende Erläuterung mit weiterführenden Informationen.

Schon nach einem kurzen Blick auf lediglich ein Beispiel kann vermutet werden, daß Teile von Informationstexten aus der Technischen Dokumentation wiederum aus bestimmten Teiltextrn bestehen, die sich einer Typisierung nicht von vornherein verschließen dürften. Im obigen Fall könnte man etwa die ad hoc gebildeten LINK-Etiketten *Visualisierung* (im obigen Beispiel: *So sieht's aus*) oder *Zusatzinfo* (im obigen Beispiel: *numerische Angabe*) einführen. Ob diese ersten Vermutungen Bestand haben, wird sich jedoch erst nach umfangreicheren theoretischen und praktischen Untersuchungen feststellen lassen.

2.3 Die Texte des Corpus'

Wichtigster Hypertext des Corpus' ist zweifelsohne der Titel *SelfHTML. Die Energie des Verstehens: HTML-Dateien selbst erstellen* von Stefan Münz, der unter dem Kürzel SELF verwendet wird. Besonders geeignet ist dieser Text auch aufgrund seiner hohen Link-Dichte im Fließtext. SELF wurde im Jahr 1998 unter der Adresse <http://www.teamone.de/selfhtml/> online gestellt. Kurz vor Abgabe der vorliegenden Untersuchung ist jedoch eine Neuauflage unter der Adresse <http://selfhtml.teamone.de/> ins WWW gestellt worden. Die alte Adresse ist nun nicht mehr gültig. Durch die Neuauflage bot sich die Chance, die beiden Versionen vergleichend zu untersuchen. Dies ist besonders interessant, da sich die in SELF benutzten typisierten LINKS zum Teil geändert haben. Aus diesen Gründen ist sowohl die alte als auch die neue Auflage in das Corpus aufgenommen. Die Neuauflage wird im Rahmen der vorliegenden Arbeit mit SELF8 abgekürzt, da diese mit der Versionsnummer 8.0 versehen wurde. Die alte Version befindet sich seit kurzem wieder auf dem neuen Server des SelfHTML-Teams, und zwar im Archiv unter <http://selfaktuell.teamone.de/archiv/doku/7.0/>. So bleibt die Überprüfbarkeit der Analyse der alten Version gewährleistet.

Ein weiterer Text kommt aus der wissenschaftlichen Fachkommunikation: Das Grammatik-Informationssystem GRAMMIS, welches unter diesem Kürzel verwendet wird. GRAMMIS ist insofern interessant, als daß dieses Projekt in zweifacher Weise TYPISIERTE LINKS verwendet: Zum einen sind die LINKS durch

farbige Markierungen bestimmten Bereichen zugeordnet. Zum anderen wird diese Zuordnung zum Teil durch MOUSE-OVER-TEXTE¹ unterstützt.

Ein Hypertext, der thematisch aus dem Rahmen fällt, ist das Sportportal der Sport 1 GmbH, das unter dem Kürzel SPORT1 verwendet wird. Dieser Hypertext lohnt aufgrund seiner zahlreichen PIKTOGRAMME², die für die unterschiedlichen LINK-Typen verwendet werden, einer genaueren Betrachtung. Außerdem stellt SPORT1 eine im WWW im häufiger anzutreffende Sorte von Hypertexten dar, die durch eine geschickte Verquickung von Informations- und Werbetexten gekennzeichnet ist. Hier können zum Teil sehr interessante Unterschiede bzgl. Funktion und Struktur einzelner Knoten im Vergleich zu den restlichen, rein informativen Hypertexten des Corpus' herausgearbeitet werden.

Ebenfalls als Primärtext dient die Informationsseite des W3C. Der besondere Reiz dieses Hypertextes resultiert nicht nur aus der Tatsache, daß das W3C die Standards im Internet festlegt und verwaltet – und damit als die maßgebende Informationsseite zur Internet-Technologie gesehen werden muß –, sondern auch aus der Herausforderung für die W3C-Verantwortlichen, im Rahmen der sich schnell ändernden Technologien, unterschiedliche Versionen von technischen Spezifikationen – und damit von Hypertext-Knoten – zu verwalten und, im Idealfall, diese zeitlichen Relationen dem Leser auch zu visualisieren.

Folgende Tabelle stellt die Corpus-Hypertexte nochmals übersichtlich zusammen:

| Kürzel | Name | Adresse |
|---------|---|---|
| GRAMMIS | Grammis. Das grammatische Informationssystem des IDS | http://hypermedia.ids-mannheim.de/grammis/ |
| SELF | SelfHTML. Die Energie des Verstehens: HTML-Dateien selbst erstellen (Version 7) | http://selfaktuell.teamone.de/archiv/ |
| SELF8 | SelfHTML. Die Energie des Verstehens: HTML-Dateien selbst erstellen (Version 8) | http://selfhtml.teamone.de/ |
| SPORT1 | Sport 1 (Sportportal) | http://www.sport1.de/ |
| W3C | W3C. World Wide Web Consortium | http://www.w3c.org/ |

Tabelle 1: Hypertexte des Corpus'

¹ Darunter versteht man Texte, die auf dem Bildschirm erscheinen, sobald der Mauszeiger auf einen Link bewegt wird.

² PIKTOGRAMME sind kleine, ikonenhafte Bilder, die zur Visualisierung von Link-Typen benutzt werden können. Sie werden in Kapitel 15.2.4 vorgestellt.

3 Angaben zur Literaturlage

Im folgenden wird ein kurzer Überblick über die Literaturlage in den Bereichen HYPERTEXT und TEXTLINGUISTIK gegeben.

3.1 Literatur zu Hypertext

Die Arbeiten zum Thema HYPERTEXT sind zahlenmäßig kaum noch zu bewältigen, daher beschränkt sich die Vorstellung auf einige ‚Meilensteine‘ aus unterschiedlichen Disziplinen.

Historische Primärtexte, durch deren Lektüre sich die Entwicklung des Hypertext-Gedankens nacherleben läßt, sind Bush 1945, Engelbart 1963, Engelbart u. English 1968, Nelson 1974, ACM Hypertext 1987 und 1989 sowie Berners-Lee 1989. Auf die Bedeutung dieser Texte wird in Kapitel 4 genauer eingegangen.

Einen Überblick über die Forschung bis 1990 gibt die kommentierte Bibliographie in Atkinson u. Knee 1990. Die historische Entwicklung beschreiben Berk u. Devlin 1991. Einen geschichtlichen Abriß, der das WWW einschließt, gibt Lennon (Lennon 1997).

Grob kann folgende Gliederung der vorhandenen Literatur vorgeschlagen werden:

(1) Die meisten Untersuchungen zum Thema stammen aus dem Gebiet der Informatik. Als deutschsprachiges Standardwerk gilt Kuhlen 1991. Kuhlen beschäftigt sich neben Fragestellungen der Informationswissenschaften auch am Rande mit textlinguistischen Problemen wie der KOHÄRENZ von Teiltextrn. Schnupp 1992 befaßt sich dagegen besonders ausführlich mit der technischen Implementierung von Hypertexten. In weiteren Titeln der Informatik geht es hauptsächlich um die Verbindung der Hypertext-Idee mit Konzepten der KÜNSTLICHEN INTELLIGENZ, der WISSENSREPRÄSENTATION sowie des INFORMATION RETRIEVAL.¹

(2) Ferner existieren zahlreiche Arbeiten aus der Psychologie, die kognitions-wissenschaftliche und pädagogische Fragen an die neue Technologie stellen. Hierbei soll hauptsächlich ein möglicher Mehrwert hypertextueller Lernmedien kritisch untersucht werden. Ansel Suter dokumentiert den Versuch, Kenntnisse der englischen Sprachwissenschaft mittels eines Hypertextes zu vermitteln (Ansel Suter 1995). Heike Gerdes faßt den Forschungsstand psychologisch-

¹ z.B.: Sarre1991, Brown 1991, Cordes u. Streit 1992, Neubert 1994, Agosti 1996, Haake 1997 und Riggert 1998

pädagogischer Disziplinen zusammen, zitiert die Ergebnisse der wichtigsten Fallstudien und kommt aufgrund eigener Experimente zu einer kritischen Einschätzung des sogenannten KOGNITIVEN MEHRWERTS von Hypertexten (Gerdes 1997). Mit allgemeinen, erziehungswissenschaftlich relevanten Fragen des Informationsproblems in elektronischen Datennetzen aller Art beschäftigt sich Astleitner (Astleitner 1997).¹ Eine weitere psychologische Untersuchung ist Unz 2000, die mit empirischen Methoden Hypertexte und gedruckte Texte kontrastiv bzgl. Informationssuche und Navigation untersucht. Die aktuellste Studie zum Thema dürfte Bromme 2002 sein, die allerdings vor Redaktionsschluß nicht mehr gesichtet werden konnte.

(3) Auch die Literatur- und Texttheorie hat sich des Themas mit Begierde angenommen. Offen ist hierbei vor allem das veränderte Verhältnis *Autor-Leser*, da jeder Benutzer eines Hypertextes einen eigenen Lesepfad durch das Werk wählen kann. HYPERTEXT wird hier als neue Kulturtechnik verstanden und als solche analysiert. Bolter untersucht die Auswirkungen des Computers als „writing space“ und geht dabei besonders auf die nicht-lineare Repräsentationsform von Wissen in Hypermedien ein (Bolter 1991). Landow untersucht den Bezug zwischen hypertextuellen Schreibarten und der Kritischen Theorie (Landow 1992 u. 1994), während Gaggi 1997 eine Dezentralisierung des Subjekts durch die – bzw. innerhalb der – neuen Medien sieht und gleichzeitig Nelsons' Vision eines „alle Texte umfassenden Hypertextes“ neu belebt.² Eine poststrukturalistische Herangehensweise an den ‚Hypertext-Raum‘ verfolgt Wenz 2001a. Wenz 2001 hingegen unternimmt eine erste Pilotuntersuchung zur unterschiedlichen Akzeptanz von gedrucktem Text, textuellem Hypertext sowie multimedialen Hypertext.

(4) Technische Artikel zu sogenannten MARKUP-SPRACHEN wie HTML und XML, mit denen Hypertexte implementiert werden können, finden sich in großer Anzahl im Internet. Die maßgeblichen Spezifikationen stehen hierbei auf den Seiten des Internet Konsortiums W3C (<http://www.w3c.org>). Dort sind auch die programmatischen Artikel zum sogenannten SEMANTIC WEB – eine Art semantische Tiefenstruktur für das heutige World Wide Web – zu finden.

(5) So gut wie alle Abhandlungen über Hypertexte gehen in unterschiedlicher

¹ Weitere psychologisch bzw. erziehungswissenschaftlich motivierte Arbeiten sind Ambron u. Hooper 1988, Marchionini u. Shneiderman 1988, Jonassen 1990, Nix u. Spiro 1990, Retterer 1991, Fickert 1992, Hammwöhner 1993, McKnight 1993, Tergan 1993 u. 1995, Hasebrook 1994 u. 1995, Meyerhoff 1994, Glowalla 1995, Rada 1995 sowie Rouet 1996.

² Weitere Titel aus der Literaturtheorie sind: Doland 1988, Bolter 1989, Moulthrop 1991, Andersen 1992, Idensen 1993 und Ramm 1994.

Intensität auch auf das Problem der typisierten Links ein. Die These ist meistens, daß typisierte Links einen Mehrwert brächten, ein passendes allgemeines Vokabularium zur Typisierung aber noch nicht existiert.

3.2 Literatur zur Textlinguistik

Als allgemeine linguistische Nachschlagewerke dienen Bußmann 1990 und Glück 2000. Eine erste Annäherung sowie einen Einblick in die Geschichte der Textlinguistik bringt der Sammelband *Textlinguistik* (Dressler 1978). Grundlegende textlinguistische Gesamtdarstellung für diese Arbeit ist Brinker 2001. Neben weiteren spezifischen Darstellungen zur TEXTUALITÄT (de Beaugrande u. Dressler 1981), REFERENZSTRUKTUREN (Vater 1994) und THEMENENTFALTUNG (van Dijk 1980 a u. b) sind am Rande auch Spezialuntersuchungen interessant, die das Vorkommen von Texten in technischen Distributionen untersuchen (Ott 1996 und Labarta Postigo 1997).

3.3 Literatur zu Textlinguistik und Hypertext

Die textlinguistische Beschäftigung mit Hypertexten ist vergleichsweise jung. Die Zahl der einschlägigen Literatur ist noch überschaubar, steigt seit den letzten Jahren aber sprunghaft an. Die Beiträge sind oftmals texttheoretischer Natur. Sie arbeiten allgemeine Unterschiede und Ähnlichkeiten von traditionellen Texten und Hypertexten heraus und lassen sich unter die oben genannten Arbeiten zur Literaturtheorie subsumieren. Andere Untersuchungen greifen die erarbeiteten Unterschiede auf und machen sie zur Grundlage für Konversionen zwischen Texten und Hypertexten. Besonders wird hierbei auf Aspekte der Kohärenzplanung eingegangen (Campel 1995, Lobin 1999 und Storrer 1999).

Mit den Besonderheiten von Sprache im Internet sowie den Eigenschaften schriftlicher Texte in multimedialen Kontexten beschäftigt sich Schmitz (Schmitz 1995, 1996 u. 1997). Angelika Storrer hat eine rege Publikationstätigkeit zur hypertext-spezifischen Textlinguistik entwickelt (Storrer 1999, 2000, 2001a, 2001b und 2001c; Ferner befindet sich ihre Habilitationsschrift zum selben Thema derzeit in Druck.). Interessante Sammelbände sind außerdem Weingarten 1997a, Runkehl u.a. 1998, Holly u. Biere 1998 und besonders Lobin 1999a. Darüber hinaus befaßt sich Nummer 63 der Zeitschrift *Osnabrücker Beiträge zur Sprachtheorie (Obst)* mit der Thematik *Hypermedien und Wissenskonstruktion* und versammelt hierfür einige interessante Aufsätze – z.B. Rapp u. Zumbach 2001.

Teil A: Hypertext als Untersuchungsgegenstand

Teil A führt ausführlich in den Untersuchungsgegenstand HYPERTEXT ein.

Hierbei werden zuerst historische Entwicklungslinien und mediale Realisierungen gezeigt. Besonderes Augenmerk gilt hierbei der Geschichte des WWW, als dem für heutige Hypertexte wichtigsten Mediums (Kapitel 4).

Anschließend erfolgt in Kapitel 5 eine differenzierte Betrachtung und Klassifizierung ‚hypertext-verdächtiger‘ Phänomene im WWW. Bei einer kontrastiven Betrachtung von traditionellen Texten und Hypertexten wird sich herausstellen, daß nicht jeder elektronische Text im WWW unbedingt ein Hypertext ist – auch wenn häufig vom WWW als dem *größten Hypertext der Welt* gesprochen wird.

Im letzten großen Abschnitt von Teil A, Kapitel 6, werden dann die einzelnen Komponenten von Hypertexten vorgestellt. Fokus liegt hierbei auf der Definition aktueller WWW-Hypertexte, die sich zum Teil deutlich von Vertretern anderer Hypertext-Systeme unterscheiden. Wichtig ist in diesem Zusammenhang auch die Beschreibung der zur Betrachtung von WWW-Hypertexten nötigen Software, dem sogenannten BROWSER.

Abschließend folgt nochmals eine kurze Zusammenfassung der wichtigsten Designprinzipien, Eigenschaften und medialen Bedingungen heutiger Hypertexte im WWW, bevor zur Einführung der für Hypertexte relevanten textlinguistischen Begriffe und Methoden in Teil B übergeleitet wird.

4 Hypertext: Ein historischer Abriss

Dieses Kapitel zeichnet die Geschichte des Hypertextes von den ersten theoretischen Konstrukten bis hin zum Durchbruch im Rahmen des World Wide Web und des Internet nach.

4.1 Die Pioniere: Bush, Nelson und Engelbart

Nach allgemeiner Meinung begründete Vannevar Bush mit seinem 1945 veröffentlichten Artikel *As we may think* (Bush 1945)¹ seine Position als „Vater des Hypertext-Konzepts“ (Gerdes 1997 S. 1 [Auszeichnung im Original]). Bush entwirft hier ein futuristisches System namens *Memex*, mit dem Information in der Form gespeichert werden kann, wie es heute auch für Hypertexte üblich ist:

„Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and, to coin one at random, "memex" will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory. [...]

Thereafter, at any time, when one of these items is in view, the other can be instantly recalled merely by tapping a button below the corresponding code space. Moreover, when numerous items have been thus joined together to form a trail, they can be reviewed in turn, rapidly or slowly, by deflecting a lever like that used for turning the pages of a book. It is exactly as though the physical items had been gathered together from widely separated sources and bound together to form a new book. It is more than this, for any item can be joined into numerous trails.“

(Bush 1945 S. 107f.)

Die Idee der Memex ist also, daß atomare Informationseinheiten separat gespeichert werden. Da diese zumeist für mehrere Kontexte relevant sind, können sie stets in neuen Kombinationen zusammengefügt werden. Den dadurch entstehenden Lesepfad bezeichnet Bush als „new book“.

Der Begriff *Hypertext* als solcher ist auf eine Prägung von Theodor Nelson aus dem Jahr 1965 zurückzuführen.² In dem Artikel *Professor Nelson Talk Analyzes P.R.I.D.E.* der Zeitung *Vassar College Miscellany News* vom 3. Februar 1965, wird der Vorschlag für ein neues Informationssystem beschrieben:

¹ Die Bedeutung dieses Grundsatzartikels arbeiteten Nyce u. Kahn 1991 heraus.

² vgl. Kuhlen 1991 S. 38, Nielsen 1990 S. 33 u. Hasebrook 1995 S. 169; Nelson selbst bestätigt, daß die Begriffe *Hypertext* und *Hypermedia* von ihm stammen (vgl. hierzu auch den offiziellen *Newsletter* von Nelson im Internet unter der Adresse <http://www.picosof.com/993>[momentan offline!]);

„[...] he [Nelson] reminded the audience of the problems of organizing material into a coherent piece of writing. As a new organization method Mr. Nelson has invented the PRIDE (Personalized Retrieval Indexing and Documentary Evolution) system.“

(Wedeles 1965)

Nelson wies auf mögliche Probleme hin, die sich bei der Linearisierung von Informationsmaterial zu einem einzigen „coherent piece of writing“ ergeben. Sein System *P.R.I.D.E* versucht, diese Probleme zu lösen:

„ In this system passages of material would be translated into machine language and filed in the machine in any sequence. With the proper instructions the machine would print out any sequence the writer wished to try, freeing him from the necessity of keeping the ideas in his head.“

(Wedeles 1965)

Nelson glaubte, diese Form der Informationsrepräsentation entspräche eher der menschlichen Form des Denkens als die der rein linearen Variante. In diesem Zusammenhang prägte er das Wort *hyper-text*:

„Mr. Nelson pointed out that we often do not think in linear sequences but rather in ‘swirls’ and in footnotes. He introduced the concept of the hyper-text, which would be a more flexible, more generalized, non-linear presentation of material on a particular subject.“

(Wedeles 1965)

Der Schöpfer des Wortes *hyper-text* nahm also an, menschliches Denken verlaufe nicht in linearen Sequenzen, sondern sei vielmehr mit „swirls“ oder „footnotes“ zu vergleichen. *Hypertext* ist für Nelson eine nicht-lineare, verallgemeinerte und flexiblere Weise, Wissen über ein bestimmtes Thema zu repräsentieren.

Als zukünftiges Einsatzgebiet von Hypertexten gibt Nelson folgendes an:

„For example, it is possible that basic texts on a subject could be interindexed, so that the necessity and difficulty of tracing footnotes and rare sources would be eliminated.“

(Wedeles 1965)

Zusammenfassend kann also festgehalten werden, daß HYPERTEXT im Sinne des Wortschöpfers eine nicht-lineare Form der Wissensrepräsentation zu einem bestimmten Thema ist. Aufgrund dieser nicht-linearen Form (und mittels geeigneter Implementierung) kann je nach Bedarf eine bestimmte Sequenz des gespeicherten Informationsmaterials abgerufen werden. Eine wichtige Beobachtung ist, daß die Linearisierung der Information erst im Moment des

Lesens vollzogen wird: Aus einer bestimmten Menge von Informationseinheiten wird also eine Teilmenge (die eventuell auch identisch mit der ursprünglichen Menge sein kann) mittels Verweisen verbunden. Dadurch entwickelt der Leser den Text zum Teil selbst. Wortwahl, Aufteilung der Gesamtinformation in kleinere Einheiten, sowie das Setzen der Verweise bleibt Aufgabe des Autors. Das Aneinanderfügen dieser Informationspakete erfolgt dann durch einen, im Wortsinn kreativen Akt des Lesens. Verschiedenes Kombinieren der Informationseinheiten führt somit für den Leser zu unterschiedlichen Texten. Diese Leseptide bleiben aber freilich immer eine Teilmenge des vom Autor zusammengestellten Gesamtfundus'. Insofern wird auch beim Hypertext das Konzept *Autor* nicht aufgegeben!

Auch die heutige (de facto vorhandene) Beschränkung von Hypertexten auf das elektronische Medium wird bereits in dieser ersten Definition vorweggenommen.¹

Auffällig ist ferner, daß sich diese erste Definition nicht zwangsläufig auf eine textuelle Form von Information beruft, sondern vielmehr von „material“ aller Art spricht. Hier kommt eine frühe multimediale Auffassung von HYPERTEXT zum Ausdruck:

„Programs can include words, pictures, printing, or strings of words.“

(Wedeles 1965)

Dies nimmt in gewisser Hinsicht den sich heute vollziehenden Wandel von HYPERTEXT zu HYPERMEDIA voraus.

Nelson politisierte später mit der Arbeit *Dream Machines: new freedoms through computer screens – A minority Report* (Nelson 1974) die Hypertext-Idee:

„Nelson vertrat (und vertritt) die Auffassung, daß alle Informationsquellen in einer demokratischen Informationsgesellschaft frei zugänglich sein sollten. Er meinte, daß ein freier Zugang dann am besten gesichert sei, wenn alle Dokumente in einer Art gigantischem Hypertext miteinander verknüpft seien.“

(Hasebrook 1995 S. 169)

¹ Dies bezeugt ebenfalls eine Stelle aus Nelson 1967. Hier definiert der Autor HYPERTEXT als Sammelbegriff für „[...] any text, which cannot be printed [...] on a conventional page“ (Nelson 1967, S. 195). Dieser Sammelbegriff schließt aber Printmedien, welche über ‚hypertextuelle‘ Strukturen verfügen (etwa Enzyklopädien) explizit nicht mit ein.

Die erste elektronische Realisierung des Hypertext-Konzeptes gelang Douglas C. Engelbart, deren Grundlagen er in seinen Publikationen beschreibt (Engelbart 1963, Engelbart u. English 1968). Allerdings wurden diese Implementierungen fast ausschließlich von Spezialisten benutzt.

4.2 Der Durchbruch: Berners-Lee und das WWW

Der Hypertext-Idee zum Durchbruch verholfen hat erst Tim Berners-Lee mit der Erfindung des WORLD WIDE WEB (WWW). Berners-Lee begann 1980 seine Arbeit am wissenschaftlichen Zentrum CERN in Genf. Hier schrieb er für den Eigenbedarf sein erstes Programm zur Informationsspeicherung unter Verwendung von assoziativen Verweisen. Er nannte die Software *Enquire*, unter Bezugnahme auf einen "verstaubten Ratgeber aus viktorianischer Zeit" mit dem Titel "Enquire Within upon Everything". Dieses Buch bildete so etwas wie "das Tor zu einer Welt der Information, in der von der Fleckenentfernung bis zur Geldanlage alle möglichen Fragen beantwortet wurden." (Berners-Lee 1999a S. 11) Die Software selbst wurde niemals veröffentlicht, bildet aber so etwas wie die konzeptionelle Grundlage für die Entwicklung des World Wide Web.

1989 schlug er ein globales Hypertext-Projekt vor (Berners-Lee 1989). Das Wissen der Leute sollte mittels eines Netzes aus Hypertext-Dokumenten kombiniert werden. Aufgrund der Systemunabhängigkeit entschied sich Berners-Lee für die INTERNETPROTOKOLLE TCP/IP (siehe unten) als technische Grundlage für den Austausch von Hypertext-Dokumenten. Auf dieser TCP/IP-Basis entwickelte er 1990 das HYPERTEXT TRANSFER PROTOCOL (HTTP). Auch schrieb er den ersten WWW-Server¹, *httpd*, und den ersten BROWSER², *WorldWideWeb*, den er später in *Nexus* umbenannte.³ Die Webtechnologie machte er am CERN und im Sommer 1991 im Internet verfügbar.⁴ Das WWW verhalf dem Internet sowie der Hypertext-Idee auf breiter Front zum Durchbruch, da es intuitiv und ohne

¹ Der Begriff SERVER ist im Glossar erläutert.

² Ein BROWSER ist ein Programm zum Lesen von Hypertexten im WWW. En détail wird diese Software in Kapitel 6.4.8 vorgestellt.

³ Nähere Details stehen im Internet unter <http://www.w3.org/People/Berners-Lee/WorldWideWeb.html>

⁴ Interessant ist, daß es sich bei dem ersten Programm sowohl um einen Browser als auch um einen Editor handelte. Mit Hilfe des Editors sollte jedermann ermöglicht werden, zu bestehenden Dokumenten Bemerkungen hinzuzufügen bzw. von bestehenden Dokumenten aus selbständig zusätzliche Links einzubauen. Hierbei handelt es sich um eine Kernforderung von Berners-Lee, die jedoch in dieser Form nicht erfüllt wurde. Die späteren Entwicklungen konzentrierten sich ausschließlich auf die Navigationsfunktionen. Mögliche Gründe für die Vernachlässigung der Editierfunktion werden bei Berners-Lee 1999a S. 111f. genannt.

Expertenwissen zu benutzen war. In der Presse wird heute oftmals das Internet mit dem WWW-Hypertext gleichgesetzt. Dies ist jedoch nicht korrekt, was ein kurzer Blick auf die Geschichte des Internet zeigen soll:

"The Web is an abstract (imaginary) space of information. On the Net, you find computers — on the Web, you find document, sounds, videos,... information. On the Net, the connections are cables between computers; on the Web, connections are hypertext links. The Web exists because of programs which communicate between computers on the Net. The Web could not be without the Net. The Web made the Net useful because people are really interested in information (not to mention knowledge and wisdom!) and don't really want to have know about computers and cables [...]"

(Berners-Lee FAQ Abschnitt *General Questions*, 1998)

Das Internet selbst ist deutlich älter als das WWW.¹ Seine Ursprünge hat es im militärischen Bereich: Im amerikanischen Verteidigungsministerium dachte man in den 60'er Jahren verstärkt über Möglichkeiten nach, wie militärische Daten selbst vor atomaren Angriffen zu sichern seien. Möglich war dies nur durch ein Datennetz, bei dem die Daten auf mehreren, weit voneinander entfernten Rechnern lagerten. Wurden Daten modifiziert, sollten diese vom veränderten Rechner an alle angeschlossenen Rechner weitergegeben werden. Ein wichtiger Punkt hierbei war, daß jeder Rechner über mehrere Wege mit allen anderen kommunizieren konnte. Im Notfall konnten so Ausfälle von einzelnen Maschinen oder Verbindungskabeln durch Umleitungen überbrückt werden. Die ersten Rechner wurden 1969 im sogenannten ARPA-NET (ADVANCED RESEARCH PROJECTS AGENCY)² zusammengeschlossen.

In einem zweiten Schritt erhielten Forschungseinrichtungen der USA Zugang zum ARPA-NET. Für die Wissenschaftler stand nun weniger das Synchronisieren von Daten als vielmehr die schnelle Kommunikation mit Kollegen sowie der unkomplizierte Datenaustausch im Vordergrund. Vor dem Hintergrund der heterogenen EDV-Architekturen der Institute entstand die

¹ Die Geschichte des Internet wird hier nur kurz angerissen. Für weitere Informationen gibt es sehr gute Ressourcen in ausreichender Anzahl im WWW. Ein Ausgangspunkt hierfür ist die Kategorie *Internet/History* bei *Yahoo!* [http://dir.yahoo.com/Computers_and_Internet/Internet/History/]. Einen schnellen Überblick in Tabellenform gewährt die Seite *see.think. The History of the Net.* [<http://members.magnet.at/dmayr/history.htm>]. Ausführlich über den Entwicklungsstand im Jahre 1994, also kurz vor dem weltweiten Durchbruch des WWW, informiert die Seite *Entering the Web: A Guide to Cyberspace* von Kevin Hughes [<http://sunsite.sut.ac.jp/coll/guide.61.html/guide.toc.html>]. Besonders zu erwähnen sind auch die umfangreichen Archive des W3C [<http://www.w3.org/History/>], die die verschiedenen Entwicklungsphasen des Internet und besonders die des WWW nachvollziehbar machen.

² Dieses Netz existiert noch heute und hat unter <http://www.arpa.mil> eine Schnittstelle zum WWW.

Notwendigkeit der Entwicklung eines gemeinsamen ÜBERTRAGUNGSPROTOKOLLS für Daten: diese Rolle übernahm das TRANSMISSION CONTROL PROTOCOL (TCP) in Zusammenarbeit mit dem INTERNET PROTOCOL (IP). Die Kombination TCP/IP ist noch heute das ‚Rückrat‘ des gesamten Datenverkehrs im Internet. Erst durch dieses Protokoll wurden die getrennten heterogenen Universitäts- und Institutsnetze zu einem ‚Netz der Netze‘.

TCP/IP arbeitet wie folgt: Wenn heutzutage mittels eines Browsers eine HTML-Datei¹ im WWW aufgerufen wird, wird diese Datei bei der Übertragung via Internet in kleine Pakete zerstückelt. An jedes Paket ‚heftet‘ das IP einen Adreßaufkleber, auf dem der Empfänger des Datenpakets steht², also der Computer, auf dem der Browser läuft, mit welchem die fragliche Datei angefordert wurde. Als zusätzliche Information gibt IP an, um das wievielte Paket innerhalb der jeweiligen Sendung es sich handelt. Den korrekten Eingang der Datenpakete beim Empfänger erledigt TCP. Sind alle Pakete eingetroffen, beginnt der Browser mit der Darstellung der Datei.

TCP/IP ist wie gesagt ein universelles Protokoll und wäre prinzipiell zur weltweiten Datenübertragung ausreichend. Die Kommunikation wurde jedoch bald in dem Maß durch die Komplexität der Bedienung erschwert, als daß immer mehr Menschen Daten austauschen wollten, die selbst keine Experten der Informationstechnologie waren. Aus diesem Grund wurden ‚höhere‘ PROTOKOLLE entwickelt, die sich TCP/IP bedienen aber wesentlich einfacher zu benutzen waren. Beispiele hierfür sind FTP, TELNET und EMAIL. Dem Internet zum eigentlichen Durchbruch verholfen hat jedoch, wie oben bereits angesprochen, das HYPERTEXT TRANSFER PROTOCOL (HTTP) von Berners-Lee.³ HTTP ist das INTERNET-PROTOKOLL zur Übermittlung von Hypertext-Dateien, die mittels der Sprache HTML kodiert sind. Dieses Zusammenspiel von HTTP und HTML wird als WWW bezeichnet – eine Anwendung also, die sich der technischen Infrastruktur des eigentlichen Internet bedient. Das WWW hat sich – neben Email – zur populärsten Anwendung des Internet entwickelt.

¹ In das Konzept der HYPERTEXT MARKUP LANGUAGE (HTML) wird in Kapitel 15.2 eingeführt.

² Dazu gibt es ein Adressierungsschema – die sogenannten IP-Adressen. Eine typische IP-Adresse sieht in Dezimalschreibweise beispielsweise etwa so aus: 127.174.211.5 – vier höchstens dreistellige Zahlen, getrennt durch Punkte. Im Rahmen dieser Arbeit ist kein weiteres Verständnis der IP-Adressierung bzw. der Namensauflösung mittels sogenannter DOMAIN NAME SERVER notwendig. Für den geneigten Leser sei jedoch auf die *TCP/IP Resources List* [http://www.private.org.il/tcpip_rl.html] bzw. die *DNS Resources Directory* [<http://www.dns.net/dnsrd/>] verwiesen.

³ Ein ausgezeichnete Einstieg in HTTP ist der Aufsatz von Tim Furche [<http://www.pms.informatik.uni-muenchen.de/lehre/seminar/html-metamorphosen/00ss/ausarbeitungen/HTTP/>]. Sämtliche relevanten Spezifikationen des W3C werden hier verständlich erläutert. Die Spezifikationen selbst finden sich unter <http://www.w3.org/Protocols/>.

Um den Ausführungen der Untersuchung folgen zu können, ist eine tiefere Einsicht in das Zusammenspiel der oben eingeführten geschichteten Protokolle nicht notwendig. Vielmehr sollte in diesem Kapitel nur kurz der technische Hintergrund angedeutet werden, der zum Austausch von HTML-Seiten notwendig ist. Philosophie, Struktur und Anwendung von HTML selbst sind jedoch zum großen Teil Gegenstand dieser Arbeit und werden in Kapitel 15 besprochen.

5 Text, E-Text, Hypertext und Hypertext-Netz¹

Im letzten Kapitel wurden die historischen Voraussetzungen geschildert, die letztendlich zu den WWW-Hypertexten unserer Tage geführt haben. Im vorliegenden Kapitel sowie dem nächsten geht es darum, diese heutigen Hypertexte genauer zu bestimmen. Zunächst soll eine extensionale Festlegung der elektronischen Gebilde im WWW geschehen, bevor im nächsten Kapitel dann eine mehr intensionale Definition der WWW-Hypertexte versucht wird.

5.1 Abgrenzung Text vs. Teiltex, Hypertext vs. Knoten

Bei gedruckten belletristischen Werken oder Fachbüchern fällt es noch verhältnismäßig leicht, die Dimension eines Textes aufgrund klarer Begrenzungen festzustellen: Der Roman *1984* von George Orwell ist als TEXT aufzufassen.² Er verfügt über eine einheitliche TEXTFUNKTION sowie durchgängige KOHÄRENZ.³ Schwieriger hingegen ist die Situation beim Betrachten einer Tageszeitung. Ähnlich wie bei einem gedruckten Roman sendet auch die Beschaffenheit einer Zeitung als physische Entität klare Begrenzungssignale aus. Allerdings bereitet es Schwierigkeiten, eine einheitliche Textfunktion einer Tageszeitung zu erkennen. Zwar kann man in der Regel bei seriösen Zeitungen von einer Überzahl INFORMATIVER Texte ausgehen, die jedoch flankiert werden von Texten mit APPELLFUNKTION (Kommentar, Werbung) und KONTAKTFUNKTION (Öffentliche Glückwünsche,

¹ In diesem und dem nächsten Kapitel müssen bereits einige textlinguistische Termini wie KOHÄSION oder KOHÄRENZ verwendet werden, die explizit erst im Teil B eingeführt werden. Vorab zum besseren Verständnis sei erwähnt: Unter KOHÄSION versteht man den durch grammatische Mittel hergestellten Textzusammenhang. Der Begriff der KOHÄRENZ hingegen bezeichnet den semantischen Zusammenhalt des Textes.

² Der eigentliche Text wird flankiert von PARATEXTUELLEN Elementen, auf die in Kapitel 11 näher eingegangen wird.

³ In Kapitel 9 wird sich zeigen, daß KOHÄRENZ das wichtigste Kriterium für die Bestimmung der TEXTUALITÄT sind.

Beileidsbekundungen). Zwischen den einzelnen textuellen Phänomenen bestehen in der Regel relativ selten Kohärenzbeziehungen: Die Großzahl der Nachrichten steht isoliert im Gesamtkontext der Zeitung. Lediglich die entscheidenden Themen der Ausgabe werden durch mehrere Nachrichten und Kommentare abgedeckt, die miteinander durch Referenzen, wie etwa *siehe auch* etc., verbunden sind.

Die Teiltexthe einer Zeitung verfügen mit einer eigenständigen Textfunktion über das wichtigste TEXTUALITÄTSKRITERIUM, ich fasse sie daher als vollwertige, eigenständige Texte auf, die lediglich zum Teil lose – via Referenzen – in den Gesamtkontext einer medialen Realisierung namens *Zeitung* eingebettet sind.

Wie verhält es sich nun mit Hypertexten? Sind die Knoten hier, vergleichbar mit den Teiltexthen einer Zeitung, ebenfalls als eigenständige Texte zu sehen? Eventuell ist es hilfreich, zur Beantwortung dieser Frage zuerst das Medium zu betrachten, in dem die Hypertexte des Corpus' realisiert sind: das WWW. Es ist davon auszugehen, daß so gut wie alle Hypertexte (und auch die einzelnen Knoten von Hypertexten) durch hinreichend geduldige Navigation von einem beliebigen Startpunkt im WWW aus untereinander erreichbar sind.¹ Handelt es sich daher beim WWW um einen einzigen Text?

Ein Argument für diese Vermutung ist zwar die weltweite Vernetzung einzelner WWW-Hypertexte mittels Links, die als explizite KOHÄSIVE MITTEL gelten dürfen. Allerdings mußte festgestellt werden, daß eine gemeinsame Kohärenz, eine gemeinsame KOMMUNIKATIVE FUNKTION sowie ein einheitliches THEMA für das WWW als Ganzes nicht festzustellen sind. Da aber strukturelle Aspekte wie Kohäsion nie alleinstehend relevant sind, sondern lediglich das Fundament für funktional-kommunikative Aspekte des Textes darstellen, sind Phänomene wie Kohärenz oder Textfunktion als wichtigere Kriterien für die Texthaftigkeit eines Gebildes anzusehen. Hinsichtlich dieser funktional-kommunikativen Kriterien ist das WWW, wie erwähnt, in seiner Gesamtheit aber kein singulärer Text.

Verläßt man nun die Makro-Perspektive des WWW und zieht die Mikro-Perspektive eines einzelnen Knotens – etwa des Textes SELF – heran, so läßt sich wie folgt argumentieren: Jedem Knoten von SELF kann durchaus ein

¹ Dies gilt vor allem, wenn man die Ergebnisseiten von SUCHMASCHINEN im WWW ebenfalls als – dynamisch generierte – Hypertexte betrachtet. Die Behauptung gilt jedoch nicht für sogenannte HYPERTEXT-INSELN, die bewußt keine Anbindung an solche Knoten haben, die öffentlich zugänglich sind. Technisch möglich ist das durch die Arbeitsweise der Suchmaschinen: Diese navigieren Link für Link durch das WWW und bauen einen Index der besuchten Seiten auf. Existiert www-weit kein einziger LINK zur Hypertext-Insel, ist diese auch nicht zu indexieren. Zu denken ist hier etwa an Bewerbungsunterlagen, deren Adresse nur der jeweilige potentielle Arbeitgeber vom Bewerber erhält. Solche Hypertexte setzen jedoch lediglich auf der Technik des WWW auf, gehören aber nicht zum eigentlichen WWW-HYPERTEXT-NETZ (siehe unten).

eigenes Textthema sowie einen Textfunktion zugestanden werden. Diese sind jedoch nur Teilthemen und Teiltextfunktionen von SELF selbst. Dies unterscheidet einen Hypertext wie SELF von Tageszeitungen, macht ihn jedoch vergleichbar mit gedruckten technischen Dokumentationen.

5.2 Abgrenzung E-Text, Hypertext und Hypertext-Netz

Im letzten Abschnitt wurde festgestellt, daß das WWW nach kommunikativ-funktionalen Kriterien beurteilt keinen singulären Text darstellt. Die einzelnen Hypertexte im WWW können, nach erster Betrachtung, wohl als Texte aufgefaßt werden, die sich aus Teiltextrn – den Knoten – zusammensetzen. Eine naheliegende Frage ist nun, ob denn sämtliche im WWW auftretenden Texte auch automatisch Hypertexte sind.

Zur Beantwortung dieser Frage muß eine Grobgliederung der im WWW vertretenen Phänomene erstellt werden. Laut Storrer 1999 läßt sich das WWW als HYPERTEXT-NETZ interpretieren, in dem sowohl Hypertexte als auch sogenannte E-TEXTE¹ zu finden sind:

„Ein Hypertextnetz verknüpft mehrere Hypertexte und E-Texte sowie Paratexte zu Hypertexten und E-Texten durch Links. Das WWW kann in diesem Sinne als weltumspannendes Hypertextnetz angesehen werden, das [aus] einer wachsenden Anzahl von Teilnetzen besteht, deren Aufbau und Inhalte sich in steter Veränderung befinden und in ihrer Gesamtheit von niemandem mehr überblickt werden.“²

(Storrer 1999 S. 38)

Die beiden möglichen Komponenten von Hypertext-Netzen, Hypertexte und E-Texte, unterscheidet die Autorin wie folgt:

„Ein Hypertext ist ein von einem Hypertextsystem verwalteter nicht-linear organisierter Text mit einer erkennbaren Textfunktion und einer thematischen Gesamtvorstellung, die als übergeordnete Einordnungsinstanz fungiert. [...] Ein Hypertext kann für sich allein stehen und z.B. auf CD-ROM publiziert sein; typischerweise wird er jedoch als Teilnetz in ein größeres Hypertextnetz eingebunden. Hypertexte sind typischerweise keine abgeschlossenen Texte; sie haben ‚offene Enden‘ [...]

Als E-Texte bezeichne ich Texte, die als linear organisierte Texte in ein Hypertextnetz eingebunden sind. [...] E-Texte müssen auf nachvollziehbare Art und Weise in das übergreifende Hypertextnetz eingebunden werden, ansonsten ergeben sich im Hinblick auf die

¹ Der Begriff E-TEXT wurde laut Storrer zuerst von Zimmer in dessen *Digitaler Bibliothek* geprägt (Online im WWW unter <http://www.zeit.de/digbib/index1.html>).

² Der Begriff PARATEXT wird von Storrer nicht näher erläutert. Es kann aber wohl davon ausgegangen werden, daß hiermit in etwa die paratextuellen Phänomene gemeint sind, welche auch in Kapitel 11 dieser Arbeit beschrieben werden.

Kohärenzbildung jedoch keine wesentlichen Unterschiede gegenüber dem linear organisierten Printtext.“

(Storrer 1999 S. 38 f.)

Wer textuelle Gebilde für das WWW verfaßt, erstellt damit also nicht automatisch Hypertexte sondern eventuell traditionelle, lineare, lediglich elektronisch realisierte Texte – E-Texte eben. Vor- und Nachteile solcher E-Texte gaben und geben Stoff für zahlreiche Untersuchungen¹, die vorliegende Arbeit beschäftigt sich jedoch ausschließlich mit Hypertexten – der obigen Definition folgend. Die Komponenten solcher Hypertexte wurden bereits oben kurz angesprochen und werden nun im folgenden Kapitel 6 ausführlich eingeführt.

¹ Als Vorteile sind der schnelle Zugriff, die Möglichkeit der Volltextsuche, das kostengünstige und verlagsunabhängige Publizieren sowie die schnelle Wiederverwendbarkeit der Texte zu werten. Außerdem muß sich der Leser bei der Lektüre von E-Texten nicht auf hypertext-spezifische Eigenschaften einlassen. Als Nachteil dürfte sich das unkomfortable Lesen längerer Texte am Bildschirm herausstellen. Storrer 2001b geht genauer auf diese Punkte ein.

6 Komponenten eines Hypertextes¹

Nachdem zuletzt die Extension des Untersuchungsgegenstandes konkretisiert wurde, geht es in diesem Kapitel nun um die einzelnen Komponenten – also um die intensionale Bedeutung – von Hypertexten.

6.1 Allgemein

Manche Komponenten eines Hypertextes sind mit den Bestandteilen traditioneller Texte zu vergleichen oder gar mit diesen identisch. So stellen die sogenannten KNOTEN den eigentlichen ‚Textkörper‘ zur Verfügung. Ferner haben viele Hypertexte Orientierungsmittel, die auch in normalen Texten vorkommen. Hierbei ist etwa an Inhaltsverzeichnisse oder Indices zu denken. Andere Komponenten hingegen – elektronisch direkt ausführbare Verweise, sogenannte LINKS, und bestimmte NAVIGATIONS-KOMPONENTEN – finden sich nicht in Printmedien und können daher als hypertext-spezifisch angesehen werden.

6.2 Komponente I: Knoten

Die Information eines Hypertextes ist in viele disjunkte Einheiten aufgeteilt. Diese Knoten als „grundlegende, atomare Informationseinheiten“² (vgl. Gerdes 1997 S. 12f. u. Kuhlen 1991 S. 79f.) sollen im folgenden hinsichtlich Größe, Inhalt, Kohärenz³, Funktionalität sowie weiterer Eigenschaften näher charakterisiert werden.

¹ Dieses Kapitel ist eine überarbeitete und erweiterte Version des Abschnitts 2.2 meiner unveröffentlichten Magisterarbeit. Die Arbeit ist über das Internet als PDF-Datei zu beziehen [<http://www.huberoliver.de/Magister.pdf>].

² „Knoten sind [...] nicht immer atomar.“ (Schnupp 1992 S. 58) In manchen Hypertext-Systemen ist es möglich, Knoten in einen anderen Knoten einzubauen, und/oder die Information ein und desselben Knotens in einer zusammengefaßten sowie einer ausführlichen Version, die vom Leser bei Bedarf mittels eines Software-Mechanismus⁷ aktiviert werden kann, darzustellen. In diesem Fall spricht man von STRETCHTEXTEN (vgl. Schnupp 1992 S. 40; Ansel Suter S. 19 f.). Mit XLINK, dem LINKING-Mechanismus von XML, sind solche Stretchtexte ebenfalls möglich (siehe Kapitel 15.2.3). Eine Unterscheidung bzgl. dieses TRAVESIERUNGSVERHALTENS unternimmt Kapitel 6.3.4.

³ Die Definitionen textlinguistischer Termini finden Sie in den Kapiteln von Teil B.

6.2.1 Struktur

Der Knoten *Buchvorstellung* der untenstehenden Abbildung läßt sich in drei Komponenten einteilen: KNOTEN-NAME, KNOTEN-INHALT und LINKS.¹

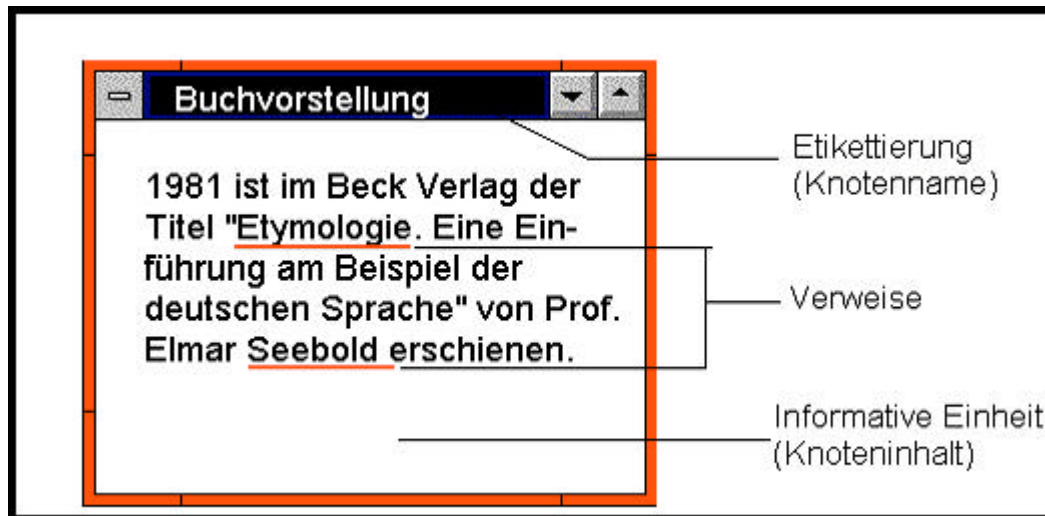


Abbildung 1: Knotenstruktur

In der Regel verfügt jeder Knoten über eine eindeutige Etikettierung. Durch diesen Knoten-Namen ist der Knoten im Netzwerk identifizierbar und kann indiziert werden².

Informative Einheiten stellen die Inhalte des Knotens. Die Einheiten können dabei prinzipiell multimedial sein. Im obigen Beispiel ist der Knoten lediglich mit Fließtext und Links (siehe unten) gefüllt.

Durch solche direkt durch Mausklick ausführbare Links kann der Leser vom aktuellen Knoten zu einem anderen navigieren: In obiger Abbildung stellen die markierten Wörter *Etymologie* und *Seebold* sogenannte LINK-ANKER – bzw. LINK-TEXTE bei rein textueller Realisierung – dar, die auf Knoten (oder Teile von Knoten) verweisen, welche in irgendeiner Form ‚mehr Information‘ zu den besagten Begriffen bereithalten.

¹ Kuhlen definiert neben den genannten drei Komponenten noch zwei weitere: einen BEGRIFFSORIENTIERTEN REFERENZTEIL (Schlüsselwörter) sowie einen ZUSAMMENFASSENDEN REFERENZTEIL (Kurzinhalt). Allerdings sind diese Komponenten in erster Linie für das INFORMATION RETRIEVAL in größeren Hypertexten und somit für informationswissenschaftliche Disziplinen interessant. (vgl. Kuhlen 1991 S. 79 f. u. 97).

² Die Etikettierung wird bei Horn „labeling principle“ genannt: Der Knoteninhalte soll sich für den Leser möglichst einfach durch den Knotennamen erschließen lassen (Horn 1989 S. 86 f.).

6.2.2 Inhalt und Größe

Lassen sich die formalen Aspekte eines Knotens noch relativ eindeutig festlegen, so sind Fragen nach dessen Inhalt und Größe schwerer zu beantworten.

Der Inhalt eines Knotens läßt sich laut Kuhlen „weder in intensionaler noch in extensionaler Sicht [...] exakt definieren“ (Kuhlen 1991 S. 79). Nach Bolter sind es „topics“, die mittels entsprechender „connections“ miteinander verbunden sind. Inhalt dieser „topics“ können entweder „paragraphs, sentences, individual words, or indeed digitized graphics“ sein (Bolter 1991 S. 24).

Hierbei stellt sich freilich die Frage nach der richtigen Fragmentierung des Textes durch den Autor, also die Frage nach dem *Wieviel an Füllung* pro Knoten. Von manchen Autoren werden hierfür sehr konkrete Lösungen vorgeschlagen; man solle den einzelnen Knoten „short“ gestalten, nicht länger als etwa „hundred lines“ (Shneiderman u.a. 1991 S. 147). Andere greifen auf Ergebnisse der Erforschung der Kapazität des Kurzzeitgedächtnisses zurück, die besagen, der Mensch könne maximal sieben kognitive Einheiten (CHUNKS) auf einmal kurzfristig im Gedächtnis speichern. Horn folgert daraus, daß ein Knoten nicht mehr als sieben Sätze enthalten solle (sogenanntes „chunking principle“; Horn 1989 S. 86f.; vgl. auch Gerdes 1997 S. 13 sowie Kuhlen 1991 S. 87).

Gerdes bezeichnet die „Festlegung des Knoten-Inhalts“ (und somit auch der Knoten-Größe) als „ungelöstes Problem der Hypertext-Technologie“. Nach ihr bestehen die Knoten – abhängig vom spezifischen Hypertext-System – „nur aus einer einzigen Idee, bzw. Proposition, oder auch aus umfangreichen Textdateien“ (Gerdes 1997 S. 12). Das „relevance principle“ von Horn besagt ähnliches, nämlich, daß ein Knoten nur die zentrale Aussage einer einzigen Idee beinhalten darf. Weitere, periphere Aussagen zur selben Idee sollten, durch Links verbunden, in anderen Knoten abgelegt werden.

Wichtig ist an dieser Stelle festzuhalten, daß ein Knoten sowohl eine einzige Proposition als auch eine kohärente Kette mehrerer Propositionen enthalten kann. Der einzelne Knoten hat also potentiell die Möglichkeit, einen Text mit mehreren Kapiteln zu enthalten. In Kapitel 5 wurde festgestellt, daß solche traditionellen, elektronischen Texte, die in einem einzigen Knoten abgelegt sind, als E-TEXTE bezeichnet werden und nicht in den Bereich einer textlinguistischen Untersuchung fallen, die speziell auf Hypertexte ausgerichtet ist.

6.2.3 Darstellungsform

In der Forschung gibt es eine Kontroverse, ob das SCROLLEN (Hoch- oder Herunterfahren) von Text verboten sein sollte oder nicht.¹

Die Befürworter dieses SCROLLING tendieren dazu, einem Knoten keinen bestimmten, bildschirmbedingten Umfang aufzuzwingen. Sie verstehen Hypertexte als elektronische Bücher. Die Knoten werden dabei mit Textabschnitten bzw. Kapiteln von variabler, den Gegebenheiten angepaßter Länge, verglichen.²

Die Gegner des Scrolling hingegen sehen Hypertexte als Karteikartensysteme. Grundelement ist dabei die Karteikarte mit unveränderbarer Größe: Der Knoteninhalte sollte daher der Bildschirmgröße angepaßt werden.³

Ein Blick auf Hypertexte im WWW läßt erkennen, daß sich weitgehend Realisierungen durchgesetzt haben, die einem Knoten keine Längenbeschränkung auferlegen. Allerdings gibt es auch hier eine Gegenbewegung: mittels sogenannter FRAMES⁴ wird häufig eine ‚scroll-freie Zone‘ innerhalb eines Knotens angestrebt.

6.2.4 Knoten-Typen

Knoten können mindestens nach folgenden Kriterien unterschieden werden:

- (1) Inhalt – Text vs. Multimedia
- (2) Funktion – Inhaltlicher Knoten vs. Strukturknoten vs. Navigationsknoten
- (3) Komplexität – atomar vs. zusammengesetzt

ad (1): Für diese Untersuchung sind primär textuelle Knoten relevant. Dies bedeutet jedoch nicht, daß in den Hypertexten des Corpus' keine Grafiken o.ä. vorkommen. In vorliegender Arbeit steht aber nicht das exakte Herausarbeiten der Struktur von in Hypertexten auftretenden nichtsprachlichen semiotischen

¹ Den aktuellen Stand der Diskussion um diese Designalternativen – besonders aus lernpsychologischer Sichtweise – beschreibt Gerdes 1997 S. 15 f.

² Meiner Meinung nach ist aber eher eine Analogie von SCROLLING zum *Lesen einer Papyrusrolle* anzusetzen, da Kapitel in Büchern oft über mehrere Seiten gehen. Aufgrund des dadurch nötigen Umblätterns ist die Buchmetapher für zu scrollende Knoten nur bedingt aussagefähig.

³ Der Umfang des Knotens müßte allerdings relational definiert werden. Aufgrund sehr unterschiedlichen Bildschirmgrößen und Auflösungen könnte die Karteikartenmetapher ansonsten nicht aufrechterhalten werden.

⁴ Frames unterteilen die Bildschirmanzeige in mehrere Bereiche. Jeder dieser Bereiche kann eigene Inhalte haben. Oftmals sind einige der Frame-Bereiche an die Bildschirmgröße angepaßt, wodurch ein Scrolling nicht mehr nötig ist.

Systemen im Vordergrund. Vielmehr geht es um die funktionale Gesamtwirkung von Knoten im Rahmen der jeweiligen Lesepfade.

ad (2): NAVIGATIONSKNOTEN beinhalten in erster Linie Hilfsfunktionen zur Orientierung innerhalb des Hypertextes sowie Angaben zur Handhabung des Systems. STRUKTURKNOTEN hingegen bieten Übersichtsfunktionen an, die sich auf inhaltliche Knoten beziehen. Hierunter fallen zum Beispiel Knoten, die Inhaltsverzeichnisse oder Indices enthalten. Strukturknoten und Navigationsknoten werden öfters zusammengefaßt, so sind zum Beispiel bei Meyerhoff alle Knoten, die „keine inhaltlichen Knoten sind“, ORGANISIERENDE KNOTEN (Meyerhoff 1994 S. 40). Ferner ist wichtig festzuhalten, daß ein einziger Knoten sowohl Inhalts- als auch Struktur- bzw. Navigationsinformationen enthalten kann. Dies ist auch bei Kapiteln normaler Texte der Fall, die neben dem Textkörper PARATEXTUELLE Struktur- (z.B. dem einzelnen Kapitel vorangestellte Kapitelüberschriften) und ‚Navigationsinformation‘ (etwa Seitenzahlen) enthalten können.

ad (3): Bei einigen Hypertext-Systemen besteht die Möglichkeit, sogenannte META-KNOTEN zu konstruieren, die ihrerseits wiederum Knoten enthalten. So könnten die drei Knoten unseres Beispiel-Hypertextes aus Abbildung 5 etwa in einen Meta-Knoten mit der Bezeichnung *Beispiel für einen Hypertext* integriert werden. Per Klick auf einen entsprechenden Link könnte der integrierte Knoten dann innerhalb des Meta-Knotens als STRETCHTEXT angezeigt werden (vgl. hierzu auch Kapitel 6.3.4). In HTML wiederum ist es möglich, mittels der bereits angesprochenen Frames mehrere Knoten in separaten Bereichen auf einer Bildschirmseite anzuzeigen.

6.3 Komponente II: Links

Kurzdefinition: Im WWW wird der Begriff HYPER-LINK als Synonym für VERKNÜPFUNG, LINK, HYPERTEXT-LINK oder auch HYPERTEXT-VERKNÜPFUNG gebraucht. Im Rahmen dieser Arbeit wird zumeist aufgrund der Kürze die Form LINK verwendet.

Im Rahmen vieler Hypertext-Modelle, etwa HYTIME¹ oder auch XLINK (siehe Kapitel 15.2.3), versteht man unter LINKING das Verknüpfen von Knoten (oder Teilen von Knoten) auf ‚irgendeine‘ Art und Weise. Diese Verknüpfungen können als ausführbare Links realisiert sein, müssen dies aber nicht – der Fokus liegt auf dem Aspekt der Verknüpfung, nicht auf dem der Ausführung!

Links im heutigen WWW sind hingegen ausführbare Verweise und stellen das hypertext-spezifische Element dar. Wobei diese Aussage sogleich etwas relativiert werden muß: Das eigentlich Neue ist die schnelle, direkte Realisierung der Links aufgrund der Möglichkeiten, die das elektronische Medium bietet. Prinzipiell sind nämlich auch Printtexte durchzogen von einem Geflecht mehrerer Verweisebenen: Zu erwähnen ist hier zuallererst das DEIKTISCHE VERWEISSYSTEM. Aber auch mit Inhaltsverzeichnissen o.ä. (siehe Kapitel 6.4.3) wird auf eine bestimmte Seite des fortlaufenden Textes verwiesen. Diese ‚Links‘ üben eine hierarchisierende bzw. organisatorische Funktion aus und helfen, den Text als Ganzen zu strukturieren. Besonders in Enzyklopädien und Fachbüchern existieren zusätzlich assoziative Querverweise, die einen Bezug zu anderen Artikeln des Werks herstellen.

Alle drei soeben angesprochenen Verweisklassen existieren – in unterschiedlicher Quantität – auch in Hypertexten. Während die zuletzt genannten assoziativen Links dort sehr häufig anzutreffen und fast als stilbildend anzusehen sind, ist das deiktische Verweissystem hingegen aufgrund der modularen Informationsrepräsentation bestimmten Einschränkungen unterworfen.² Hierarchisierende Links spielen auch in Hypertexten eine nicht zu unterschätzende Rolle, da die meisten großen Hypertexte eben auch über Verzeichnisse unterschiedlichster Art verfügen.

6.3.1 Allgemeine Eigenschaften

Links stellen, wie gesagt, die zweite Hauptkomponente von Hypertexten dar. Hierbei wird in der Regel ein Wort, ein Satz, ein Teiltext oder auch ein graphisches Element im Text des aktiven Knotens in irgendeiner Form graphisch abgesetzt und somit als Ausgangspunkt des Links definiert, welcher

¹ Mehr Information zum Standard HYTIME bietet das WWW [<http://www.hytime.org/>]. XLink erbt einige Eigenschaften dieses Hypertext-Modells, darunter die multidirektionalen Verweise. Auf die weitere Philosophie von HyTime wird im Rahmen dieser Arbeit jedoch nicht eingegangen.

² Die größte Einschränkung liegt darin, daß deiktische Mittel in der Regel nicht ohne weiteres knotenübergreifende Verweisstrukturen erstellen können. Ähnlichen Einschränkungen unterliegen auch die einzelnen Artikel einer Enzyklopädie.

auf ein bestimmtes LINK-ZIEL verweist. Das Ziel kann nun seinerseits ein Wort, ein Satz, ein Teiltex oder auch ein graphisches bzw. multimediales Element im selben oder einem anderen Hypertext-Knoten sein. Für alle Links gilt, daß sie sowohl Ausgangs- als auch Zielpunkt haben müssen. Die Ausgangspunkte werden in der Literatur häufig als REFERENCES, LINK-POINTS, HOT SPOTS, LINK-INDIKATOREN, LINK-ICONS, BUTTONS oder ANKER bezeichnet, während man für die Zielpunkte Bezeichnungen wie DESTINATION POINTS, LINK REGIONS, REFERENCE POINTS, ZIEL-RESSOURCE oder LINK-ZIEL verwendet (vgl. Conklin 1987 S. 34, Kuhlen 1991 S. 108 sowie Gerdes 1997 S. 18). In dieser Arbeit werden hauptsächlich die Begriffe LINK-ANKER bzw. LINK-TEXT für den Ausgangspunkt und LINK-ZIEL für den Zielpunkt von Links verwendet.

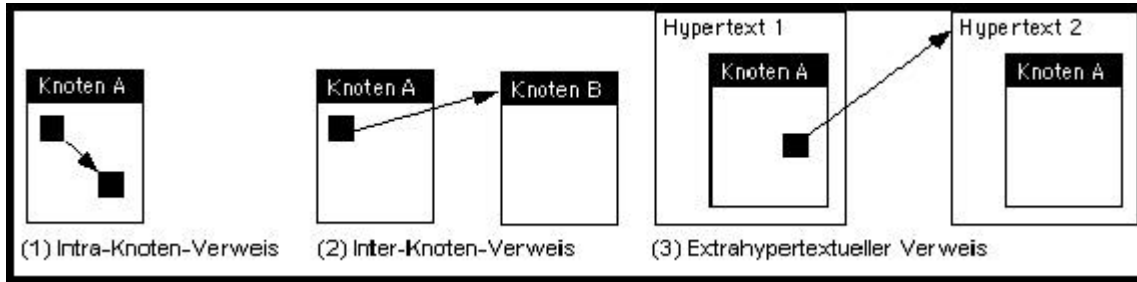
Links werden bei Kuhlen als „informationelle Funktionen erster Ordnung“ gesehen (Kuhlen 1991 S. 98). Dies geschieht in Abgrenzung zu den Navigationshilfen (Kuhlen 1991 S. 124), welche nach seiner Einteilung „Informationelle Funktionen zweiter Ordnung“ darstellen.

Eine Klassifizierung der Links kann inhaltlich-funktional (siehe Kapitel 6.3.3) und/oder formal (siehe hierzu das nächste Kapitel) erfolgen. Auch eine Klassifizierung nach dem TRAVESIERUNGSVERHALTEN wäre denkbar (vgl. Kapitel 6.3.4).

6.3.2 Formale Unterscheidungsmöglichkeiten

Ein erstes Unterscheidungskriterium ist, ob einem Link nur entlang einer Richtung gefolgt werden kann (= UNIDIREKTIONAL) oder ob dieser BIDIREKTIONAL ist, also ein Link-Anker gleichzeitig potentiell Link-Ziel (innerhalb desselben Links!) sein kann. Bei unidirektionalen Verweisen ist die Reihenfolge *Ausgangsknoten – Zielknoten* streng vorgegeben.¹ Der Zielpunkt eines Ankers kann entweder im aktuellen Knoten, in einem anderen Knoten oder in einem anderen Hypertext liegen. Die folgende Grafik soll diese Verhältnisse visualisieren:

¹ In HTML lassen sich nur unidirektionale Links konstruieren. Allerdings ist über rückwärtsgerichtete Navigationshilfen des Browsers eine Rückkehr zum Ausgangspunkt möglich. Sämtliche Browser für das WWW verfügen beispielsweise über eine sogenannte Backtrack-Funktion, die Rückwärtsnavigation ermöglicht (siehe Kapitel 6.4.8 und 6.4.9; vgl. auch Utting u. Yankelovich 1989 S. 59, Kuhlen 1991 S. 104 sowie Gerdes 1997 S. 18).

Abbildung 2: Intra-, inter- und extra-hypertextuelle Links¹

Links können des weiteren bezüglich GLOBALITÄT bzw. LOKALITÄT unterschieden werden. GLOBAL bedeutet, daß ein gesamter Knoten als Ausgangs- bzw. Zielpunkt fungiert. Als LOKALER Ausgangs- bzw. Zielpunkt dient ein Bereich (Wort, Satz, Abbildung) innerhalb eines Knotens. Auf diese Art können vier verschiedene Verweisarten klassifiziert werden:

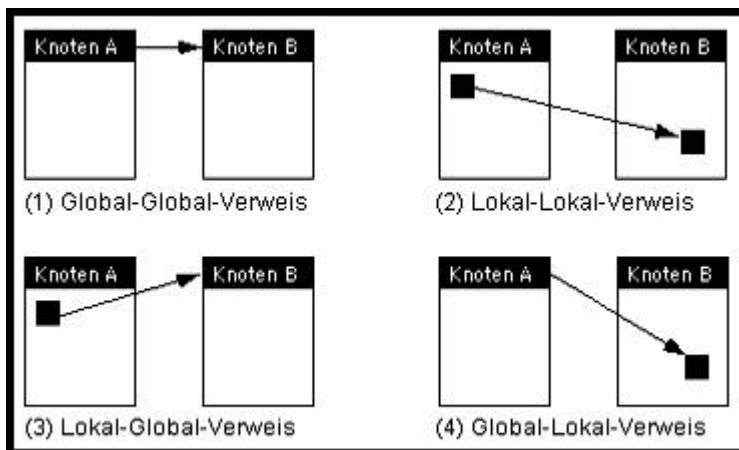


Abbildung 3: Link-Klassifizierung nach Lokalität vs. Globalität

Die INTER-KNOTEN-LINKS und EXTRA-HYPertextUELLEN LINKS (Abbildung 2, Nummer 2 und 3) können prinzipiell jede der in Abbildung 3 beschriebenen Varianten annehmen. Die Intra-Knoten-Links (Abbildung 2, Nummer 1) hingegen schließen die Variante global-global (Abbildung 3, Nummer 1) aus, da ansonsten der globale Knoten nur auf sich selbst verweisen könnte. Im WWW

¹ Die von Kuhlen eingeführten und von Gerdes und anderen übernommenen Begriffe INTRA- und INTER-HYPertextUELL wurden in dieser Arbeit durch INTRA-KNOTEN bzw. EXTRA-KNOTEN ersetzt. Sie treffen meiner Meinung nach den Sachverhalt besser, da es sich bei den seitens der Autoren beschriebenen Relationen um Verhältnisse innerhalb eines Knotens bzw. zwischen zwei Knoten handelt. Es geht also nicht um gesamt-hypertextuelle, sondern vielmehr um knoten-interne Verweisrelationen. Einen Begriff zu verwenden, der sich auf den Hypertext als Gesamtstruktur bezieht, ist nur bei Beziehung (3) sinnvoll, die als EXTRA-HYPertextUELL eine Relation zu einem anderen Hypertext beschreibt (siehe Kuhlen 1991 S. 107 f. und Gerdes 1997 S. 20).

werden jedoch gesamte Knoten äußerst selten als Link-Anker benutzt. Die meisten Links sind dort Lokal-Lokal bzw. Lokal-Global.

6.3.3 Inhaltliche Unterscheidungsmöglichkeiten

Conklin 1987 schlägt die Unterscheidung in REFERENTIELLE und ORGANISATORISCHE LINKS vor. Diese von Kuhlen modifizierte Einteilung wurde von einem Großteil der Sekundärliteratur übernommen.¹

Bei referentiellen Links wird der Grund für das Setzen eines Verweises nicht explizit gemacht. Sie sind semantisch nicht näher spezifiziert. So entstehen assoziative Relationen zwischen Knoten, „[...] ohne daß der Autor dies in jedem Fall genau begründen könnte“ (Kuhlen 1991 S. 114). Organisatorische Links definieren dagegen explizite semantische Beziehungen. Sie drücken die „primär logische Struktur eines Hypertextes“ aus (Ansel Suter 1995 S. 16). Kuhlen nimmt in diese Klasse auch Links auf, die argumentative Strukturen aufbauen. In der neueren Literatur wurde diese erweiterte Definition größtenteils übernommen.

Referentielle (= Assoziative) Links

In Anlehnung an die informationswissenschaftliche Relationentheorie definiert Kuhlen die referentiellen Links als „Relationen, die nicht Hierarchie- oder Äquivalenzrelationen zugeordnet werden können“ (Kuhlen 1991 S. 104).

Durch diese Klasse von Verweisen entsteht ein assoziatives Relationsgefüge zwischen den einzelnen Knoten. Gerade die vernetzte Struktur, welche durch diese Klasse nicht-hierarchischer Verknüpfungen entsteht, ist angeblich charakteristisch für das Medium Hypertext.

Auch wenn die Motivation des Hypertext-Autors, den Link an einer bestimmten Stelle zu definieren, nicht explizit sichtbar gemacht wird, so wird dieser zumeist doch nicht willkürlich gesetzt:

„Anlaß der Verknüpfung ist in der Regel ein in der Einheit aus irgendwelchen Gründen attraktiv erscheinendes Element, z.B. ein Teil einer Graphik, eine Bezeichnung, eine Aussage, ein Literaturhinweis.“

(Kuhlen 1991 S. 114)

¹ Siehe Conklin 1987 S. 33 f. (ferner: Kuhlen 1991 S. 113 f., Astleitner 1997 und Gerdes 1997 S. 21. Eine andere Terminologie findet sich hingegen bei Schnupp 1992 S. 134.)

Das Vorkommen eines solchen Elements wird dann zum Anlaß genommen um

„a) entweder auf andere informationelle Einheiten zu verweisen, in denen mehr über dieses der Bezeichnung entsprechendes [sic!] Konzept gesagt wird (ohne daß explizit ausgedrückt wird, was dieses mehr ausmacht),

b) oder Verknüpfungen zu anderen Bezeichnungen in anderen informationellen Einheiten (oder zu anderen Passagen in derselben Einheit) her[zu]stellen, durch die spezifizierende oder generalisierende Aussagen zu dem Ausgangskonzept gewonnen werden können.“

(Kuhlen 1991 S. 114)¹

Organisatorische (= Typisierte) Links

Organisatorische Links definieren, wie gesagt, explizit gemachte semantische bzw. argumentative Beziehungen. Sie strukturieren Hypertexte, indem sie die hierarchische, klassifikatorische und/oder konzeptuelle Natur der Verknüpfungen zwischen einzelnen Knoten explizit visualisieren. Die Notwendigkeit hierarchisch-strukturierender Links als Ergänzung zu den assoziativen ist nahezu unumstritten, da sie

„gewohnten Denkstrukturen oder Lesestrategien, zuerst das Allgemeine und dann das Besondere [sic!] zu rezipieren, entgegenkommen. Hierarchisierung ist auch in Hypertexten eine nicht aufzugebende Ordnungshilfe.“

(Kuhlen 1991 S. 106)

Das Explizieren dieser Verweisklasse ist naturgemäß sehr vom verwendeten Hypertext-System bzw. der verwendeten Beschreibungssprache abhängig. Bei HTML-Hypertexten ist es lediglich über Umwege möglich, frei definierte typisierte Links² zu setzen (siehe Kapitel 16.3.3).

6.3.4 Unterscheidung nach dem Travesierungsverhalten

Links können auch bzgl. ihres TRAVESIERUNGSVERHALTENS (= TRAVERSAL) kategorisiert werden. TRAVESIERUNG meint hierbei die Art, wie der Ziel-Knoten (besser: die ZIEL-RESSOURCE³) eines Links dargestellt wird. Mit HTML können

¹ In der Klassifikation von Kuhlen werden die referentiellen Links noch weiter unterteilt. Unterschieden werden ANNOTATIVE, SYNTAGMATISCHE, EXTRA-REFERENTIELLE sowie ASSOZIATIVE LINKS.

² Dies sind Link-Typen, die vom jeweiligen Autor frei definiert werden können, wie etwa der Typ *Beispiel*. Der HTML-Container *rel* stellt ein kontrolliertes Vokabular von Link-Typen zur Verfügung, so etwa *glossary*, *index* oder *start*.

³ Im XML-Sprachgebrauch wird RESSOURCE für sämtliche adressierbaren Einheiten verwendet – also auch für Knoten oder Teile von Knoten. XML wird in Kapitel 15.2 vorgestellt.

für Links zwei Verhalten definiert werden: Der Ziel-Knoten ersetzt die aktuelle Anzeige (REPLACE) oder wird in einem neuen Fenster der Browsersoftware dargestellt (NEW). Mittels XML lassen sich zusätzlich sogenannte STRETCHTEXTE realisieren: Ein Klick auf einen Link mit dem Travesierungsverhalten EMBED, zeigt den Ziel-Knoten innerhalb des Ausgangsknotens an. Der Text des Ausgangsknotens, der unterhalb des Links steht, wird hierbei nach unten verschoben, um dem Text des Ziel-Knotens Platz zu machen.

6.3.5 Kritik der Link-Typologie

Wie erwähnt, basiert die derzeitige Klassifizierung der Links auf der Einteilung von Conklin 1987, welche wiederum von Kuhlen 1991 modifiziert wurde und seitdem als Standardeinteilung nach inhaltlich-funktionalen Kriterien gelten darf.

Als Kritik läßt sich anführen: Kuhlen bestreitet keineswegs die funktionale Ähnlichkeit der referentiellen Links im Vergleich zu den typisierten Varianten.

„Diese [referentiellen] Verknüpfungen haben in der Regel durchaus semantische oder argumentative Funktion und leisten ihren Beitrag zu dem, was wir Hypertextkohärenz genannt haben.“

(Kuhlen 1991 S. 105)

Neben den rein assoziativen Verknüpfungen zählt Kuhlen zu der Klasse der referentiellen Links auch Metaverknüpfungen – etwa Links in Inhaltsverzeichnissen. Auf diese Weise sollen „textspezifische Übersichts- bzw. Metainformationsmittel“ zur Verfügung gestellt werden (vgl. Kuhlen 1991 S. 113f.). Hierbei handelt es sich aber um konventionell festgelegte (typisierte) Strukturinformationen, die inhaltlich-funktional sehr schwer von den eigentlichen typisierten Verweisen zu unterscheiden sind. Kuhlens Hauptunterscheidungskriterium ist daher auch formaler Art: die Expliztheit der typisierten Links im Gegensatz zu den nicht explizit gemachten referentiellen Verweisen. Nach dieser Logik verschwinden aber die Unterschiede zwischen assoziativen und typisierten Links, da jeder assoziative Link typisiert werden kann, falls ein geeignetes kontrolliertes Vokabular vorliegt. Genau solche Vokabularien werden im Laufe dieser Arbeit untersucht.

6.4 Navigations- und Orientierungskomponenten

6.4.1 Textsequenzierung, Navigation und Browsing

Wie im Laufe dieser Arbeit schon mehrere Male angesprochen, unterscheidet sich das Lesen eines Hypertextes von dem eines traditionellen, hauptsächlich linearen Textes:

„Gedruckte Texte werden normalerweise von vorne nach hinten gelesen, es gibt immer nur ein Vor oder Zurück. Man kann (meistens) der Autorin dahingehend vertrauen, daß sie die einzelnen Kapitel in einer verständlichen Reihenfolge angeordnet hat, daß neue Abschnitte auf zuvor Gesagtem aufbauen, kurz, daß man durch den Text geführt wird. In Hypertexten bleibt die Aufgabe der Sequenzierung der Leserin überlassen.“

(Gerdes 1997 S. 28f.)

Dieser Aussage ist jedoch nur bedingt zuzustimmen. Richtig ist, daß vor allem die thematische Struktur eines traditionellen Textes in der Regel eine lineare ist. Der Leser ist aber nicht unbedingt an diese vom Autor vorgegebene Struktur gebunden. So kann er etwa ‚durch den Text blättern‘ (ein Vorgang der durchaus mit dem weiter unten beschriebenen BROWSING vergleichbar ist) oder mittels Verzeichnissen, die in gewisser Hinsicht ein traditionelles INFORMATION RETRIEVAL ermöglichen, direkt an einer bestimmten Stelle des Textes mit der Lektüre beginnen.

Auch brechen Querverweise die Linearität der Textstruktur. Die vom Autor vorgegebene lineare Ursprungssequenz mutiert damit zu einem bloßen Angebot an den Leser und kann von diesem durch eine neu geschaffene Lektüre-reihenfolge ersetzt werden. Ähnliches geschieht auch in Hypertexten – mit dem Unterschied, daß keine vom Autor vorgegebene globale, lineare Ursprungssequenz existiert. Innerhalb der einzelnen Knoten jedoch ist vom Autor sehr wohl eine lineare Struktur vorgegeben. Diese kann mittels der Hypertext-Verweise – ähnlich wie mit Querverweisen in linearen Texten – verlassen werden.

In Hypertexten kann dies allerdings wesentlich schneller (= per Mausklick) als in Printtexten erfolgen, in denen erst zur passenden Seite geblättert werden muß. Die Versuchung, die lineare Knoten-Struktur zu durchbrechen, ist daher als höher zu bewerten.

Dieser Vorgang der Textsequenzierung innerhalb eines Hypertextes durch den Leser wird allgemein als NAVIGIEREN bzw. als BROWSING beschrieben. Gerdes

sieht in dieser Navigation den „dynamischen Aspekt“ des Hypertextes, während sie Knoten und Links als „statische Aspekte“ beschreibt (Gerdes 1997 S. 28). *To browse* bedeutet *weiden, grasen*. Übertragen bedeutet es in *to browse around* soviel wie *sich umsehen*, bzw. ist in *to browse through a book* gleichbedeutend mit *in einem Buch schmökern*. Marchionini und Shneiderman 1988 unterscheiden „browsing knowledge“ (= explorative Strategie zur Informationssuche) von „finding facts“ (= systematische Suchstrategie), das hauptsächlich in traditionellen Datenbanken aber auch beim INFORMATION RETRIEVAL innerhalb eines Hypertext-Systems eingesetzt wird.¹

6.4.2 Orientierungsprobleme und kognitive Mehrbelastung

Conklin deutet als erster die möglichen Probleme beim Lesen eines Hypertextes an: Orientierungsprobleme aufgrund einer mangelhaften Struktur („getting lost in hyperspace“) sowie eine allgemeine kognitive Mehr- bzw. Überbelastung des Lesers („cognitive overhead“, siehe Conklin 1987 S. 38).

Im Rahmen dieser Arbeit kann der Forschungsstand zur KOGNITIVEN PLAUSIBILITÄT von Hypertexten nur sehr verkürzt wiedergegeben werden:²

Angeblich soll die Informationsaufnahme aus Hypertexten effektiver sein als die aus linearen Texten. Autoren, die diese Position vertreten, argumentieren damit, daß sich die in einem hypertextuellen Netz präsentierte Information ‚natürlicher‘ in das kognitive Netz des Lesers übertragen ließe. Der Umweg der Entlinearisierung, die beim Umsetzen linear strukturierter Information nötig ist, entfielen bei Hypertexten. Gerdes 1997 weist jedoch nach, daß diese angebliche kognitive Plausibilität von Hypertext zumindest als fragwürdig gelten muß. Hierbei ist auf mögliche Probleme der Kompatibilität kognitiver und hypertextueller Strukturen hinzuweisen. Ferner versuchen hypertextuelle Netze in der Regel nicht – wie etwa semantische Netze –, einen Knoten lediglich mit einem singulären Konzept zu besetzen. Ob solche Hypertexte nun in der Lage

¹ Die Literatur definiert Termini wie BROWSING, SUCHEN, INFORMATION RETRIEVAL oder NAVIGATION häufig sehr unterschiedlich. Jonassen 1989 etwa differenziert nicht zwischen BROWSING und NAVIGIEREN, Nielsen 1990 hingegen macht eine Unterscheidung zwischen SUCHEN und NAVIGIEREN während Gloor u. Streitz 1990 SUCHEN mit INFORMATION RETRIEVAL gleichsetzt.

Die verschiedenen Arten von BROWSING – ASSOZIATIVES BROWSING, UNGERICHTETES BROWSING bzw. BROWSING MIT MITNAHME-EFFEKT – haben für das Thema der vorliegenden Arbeit eher untergeordnete Bedeutung. Hier sei auf Astleitner 1997, Gerdes 1997 sowie Kuhlen 1991 verwiesen. Gleiches gilt für kognitive bzw. lernpsychologische Vor- und Nachteile aufgrund unterschiedlicher Formen der Informationsrepräsentation und -verarbeitung.

² Für weitere Informationen siehe Kuhlen u.a. 1989, Utting u. Yankelovich 1989, Kuhlen 1991 ausführlicher Hammwöhner 1993, Gerdes 1997 sowie Rapp u. Zumbach 2001.

sind, die kognitiven Strukturen des Menschen nachzubilden, darf tatsächlich eher als fraglich gelten. Sehr kritisch zur kognitiven Plausibilität von Hypertexten äußern sich auch Rapp u. Zumbach 2001.

Einem möglichen kognitiven Mehrwert von Hypertexten steht jedoch in jedem Fall eine kognitive Mehrbelastung aufgrund der Hypertext-Links gegenüber: Man kann nämlich jeden Link als Aufforderung an den Leser verstehen, sich Gedanken über den Ziel-Knoten des Links zu machen, denn:

"Links präsentieren Antworten auf nicht gestellte Fragen. Und indem Antworten angeboten werden, werden diese Stellen als fraglich eingefärbt. Der Hypertext richtet sich also nicht nur die Textvorlage zurecht, sondern modelliert zugleich auch den Leser."

(Wingert 1993 S. 35)

Je mehr Verweismöglichkeiten – und damit Fragestellungen – dem Leser eines Knotens angeboten werden, um so mehr muß er sich mit diesem Angebot auseinandersetzen – was von der eigentlichen Informationsverarbeitung ablenken kann.

Der Leser eines elektronischen Hypertextes sieht sich aber auch grundsätzlich einer kognitiven Mehrbelastung gegenüber gestellt. Hierzu lassen sich vier Ebenen unterscheiden, die kognitiv verarbeitet werden müssen:

Ebene des Hypertext-Systems: Anders als beim Buch stellt sich für viele Leser im Umgang mit dem Computer (und damit auch mit der Benutzerschnittstelle eines Hypertext-Systems) das anfängliche Problem, mit den jeweiligen Funktionen vertraut zu werden und diese nutzbringend für die eigenen Zwecke einzusetzen.

Strukturebene der Hypertext-Basis: Der Leser muß – wie oben bereits ausgeführt – über die Relevanz der ihm zur Verfügung stehenden Knoten und Verweise entscheiden, da die Sequenzierung des Textes ja zusätzlich zum eigentlichen Lesen geleistet werden muß. Je mehr Verweise angeboten werden, desto größer wird vermutlich die kognitive Belastung beim Verarbeiten der Strukturinformationen sein, da sich bei jeder Navigationsentscheidung die Frage nach der Relevanz eines Verweises für den jeweiligen Kontext stellt.

Ebene der typisierten Links: Sind Verweise typisiert, muß vom Leser das dahinterstehende System begriffen werden. Das vom Autor zur Link-

Typisierung eingeführte sekundäre Zeichensystem muß vom Leser gelernt und ‚verinnerlicht‘ werden. Typisierte Verweise können also Probleme lösen, die auf der Strukturebene der Hypertext-Basis entstehen. Um aber tatsächliche Teil der Lösung zu sein ohne dabei neue Probleme aufzuwerfen, muß die Typisierung so schlüssig und einfach zugänglich wie nur eben möglich sein. Vorliegende Arbeit versucht, genau hierfür einen Beitrag zu leisten.

Inhaltsebene der Hypertext-Basis: Erst auf dieser Ebene setzt sich der Leser mit dem eigentlichen informativen Gehalt des jeweiligen Knotens auseinander.

Um diese Probleme zu verringern, stehen in Hypertext-Systemen diverse Orientierungs- und Navigationsinstrumente zur Verfügung. Hierbei kann man einerseits sogenannte VORWÄRTSGEWANDTE von RÜCKWÄRTSGEWANDTEN Instrumenten unterscheiden, und andererseits untersuchen, ob das jeweilige Hilfsmittel hypertext-spezifisch ist oder in ähnlicher Form auch in traditionellen Medien existiert.

6.4.3 Traditionelle Orientierungshilfen¹

Konventionell sind bei Büchern einige Orientierungshilfen üblich, ohne die sich die Gefahr eines Orientierungsverlustes erhöhen würde. Die folgenden Kapitel bezeichnen Orientierungshilfen wie Verzeichnisse, Metainformationen und Typographische Auszeichnungen, die Konventionen im Bereich der Buchproduktion sind und vom Leser in der Regel erwartet werden.²

Die eben aufgezählten Komponenten stellen so etwas wie die ‚mediale Struktur‘ umfangreicherer Printtexte dar, die auf den eigentlichen Textkörper aufgesetzt ist.³ Gewisse Komponenten sind für bestimmte Textsorten typisch (etwa Literaturverzeichnisse in wissenschaftlichen Arbeiten). Ein Fehlen dieser medialen Strukturelemente führt ebenso zu Problemen bei der Textverarbeitung, wie Verstöße gegen die von der Textsorte geforderten Besonderheiten im

¹ Unter textlinguistischen Aspekten werden diese paratextuellen Elemente in Kapitel 11 besprochen.

² Diese traditionellen Orientierungshilfen sind für Hypertexte ebenfalls zu empfehlen (Gerdes 1997 S. 36) und dort in unterschiedlicher Qualität und Quantität – sowie in teils dem elektronischen Medium angepaßter Form – auch anzutreffen.

Auffällig ist, daß die elektronische Variante traditioneller Orientierungshilfen oft Teil der Hypertext-Basis ist, während die hypertext-spezifischen Hilfen oft Bestandteile der Browser-Software bzw. deren Navigationskomponente darstellen.

³ Siehe auch Kapitel 11.

Aufbau des eigentlichen Textkörpers.¹ Oft werden diese traditionellen Orientierungshilfen auch von Hypertext-Autoren verwendet.

6.4.4 Verzeichnisse

Zur Extension von *Verzeichnis* zählen Inhaltsverzeichnisse, Indices, Abkürzungsverzeichnisse, Glossare sowie eventuell Literaturverzeichnisse.

Inhaltsverzeichnisse strukturieren längere Texte (z.B. Bücher) und sollen Überblick und erste Orientierung gewährleisten. Ferner geben sie dem Leser ebenso wie Indices die Möglichkeit, über Seitenangaben direkt auf bestimmte Textpassagen zuzugreifen. Kuhlen sieht diese Orientierungshilfen aufgrund der direkten Einstiegsmöglichkeiten als „nicht-lineare“ Mittel an (Kuhlen 1991 S. 136). Mittels der vergleichbaren organisatorischen Links kann in Hypertexten per Mausklick direkt zur jeweiligen Textpassage gewechselt werden.

Fachbücher verfügen in der Regel zusätzlich über Glossare, Literatur- und Abkürzungsverzeichnisse. Bei Unsicherheit über vom Autor verwendete Terminologie, Sekundärliteratur oder Abkürzung muß der Leser den linearen Lesefluß unterbrechen und das entsprechende Verzeichnis benutzen.

6.4.5 Metainformationen

Bücher im herkömmlichen Sinne stehen Leser und Autor materiell zur Verfügung.² Bereits der Umfang eines Buches gibt dem Leser gewisse Informationen über den vorliegenden Text. So wird er wohl an die Lektüre einer 1000-Seiten-Dokumentation einer Software anders herangehen als an die einer dünnen Schnell-Start-Broschüre für dasselbe Produkt. Hypertexte liefern diese Information in beschriebener Weise nicht.³ Auch die Textsorte liefert indirekte Anweisungen an den Leser, wie er mit dem Text umzugehen hat: Ein Handbuch zur Geschichte wird anders verwendet als eine historische Enzyklopädie.

Eine nicht zu unterschätzende Metainformation ist die aktuelle Lektürepröposition innerhalb des Buches, die etwa über typographische Auszeichnungen leicht zu erschließen ist. Ein Hypertext-Autor (insbesondere der eines nicht-hierarchischen Hypertextes) kann diese Information deutlich schwerer geben.

Oft gibt auch das Vorwort wichtige Lesehilfen, wenn es beispielsweise vorab

¹ Dies wäre der Fall, wenn etwa die wesentliche Information eines Textes der Sorte *Nachricht* im letzten Satz ‚versteckt‘ ist.

² Allgemeines zur Thematik *Übergang vom Materiellen ins Virtuelle* findet sich bei Negroponte 1995.

³ Informationen wie *Der vorliegende Hypertext hat x Knoten* liefert hier zwar Abhilfe, jedoch kann dadurch das ‚materielle Erfahren‘ nicht ersetzt werden.

die Gliederung der einzelnen Kapitel eines Sprachlehrbuchs in Abschnitte wie *Überblick*, *Thema* und *Übungen* erläutert.

6.4.6 Typographische Auszeichnungen

Die Schrift- bzw. Raumgestaltung eines Druckerzeugnisses trägt wesentlich zur Orientierung bei: Seitenzahl, Kopf- und Fußzeile verraten dem Leser dessen aktuelle Position im Buch. Überschriften fassen den Inhalt des folgenden Abschnittes zusammen bzw. geben dessen Thema an. Randnotizen enthalten Schlüsselwörter des nebenstehenden Absatzes und erleichtern so das ‚Überfliegen‘ des Textes. Auch wird manchmal mit Farbauszeichnung oder Fettdruck auf Schlüsselbegriffe hingewiesen. Objektsprachliche Ausdrücke werden oft kursiv bzw. unterstrichen gesetzt.

6.4.7 Hypertext-spezifische Orientierungs- und Navigationshilfen

Nach der Vorstellung der althergebrachten Lese- und Orientierungshilfen handeln die folgenden Abschnitte nun von den neuen Orientierungs- und Navigationshilfen, die für Hypertexte – und besonders für solche aus dem WWW – typisch sind.

6.4.8 Browser-Software

Erst die BROWSER-SOFTWARE ermöglicht die Darstellung eines elektronisch realisierten HTML-Hypertextes auf dem Monitor. Ein Großteil der hypertext-spezifischen Orientierungs- und Navigationshilfen sind im Programmumfang dieser Software enthalten, andere können vom Autor in die Hypertext-Basis integriert werden.

Ein BROWSER ist die Benutzerschnittstelle eines im WWW realisierten Hypertextes.¹ Er zeigt den Inhalt der einzelnen Knoten an und ermöglicht dem Leser im Hypertext zu navigieren. Folgende Abbildung auf der nächsten Seite zeigt die Oberfläche einer aktuellen Browser-Software.

¹ Nicht zu verwechseln ist die hier behandelte Browser-Software, kurz BROWSER, mit graphischen Übersichts-Knoten, welche in der Literatur auch des öfteren als BROWSER bezeichnet werden (vgl. Kapitel 6.4.10 sowie Gerdes 1997 S. 37).

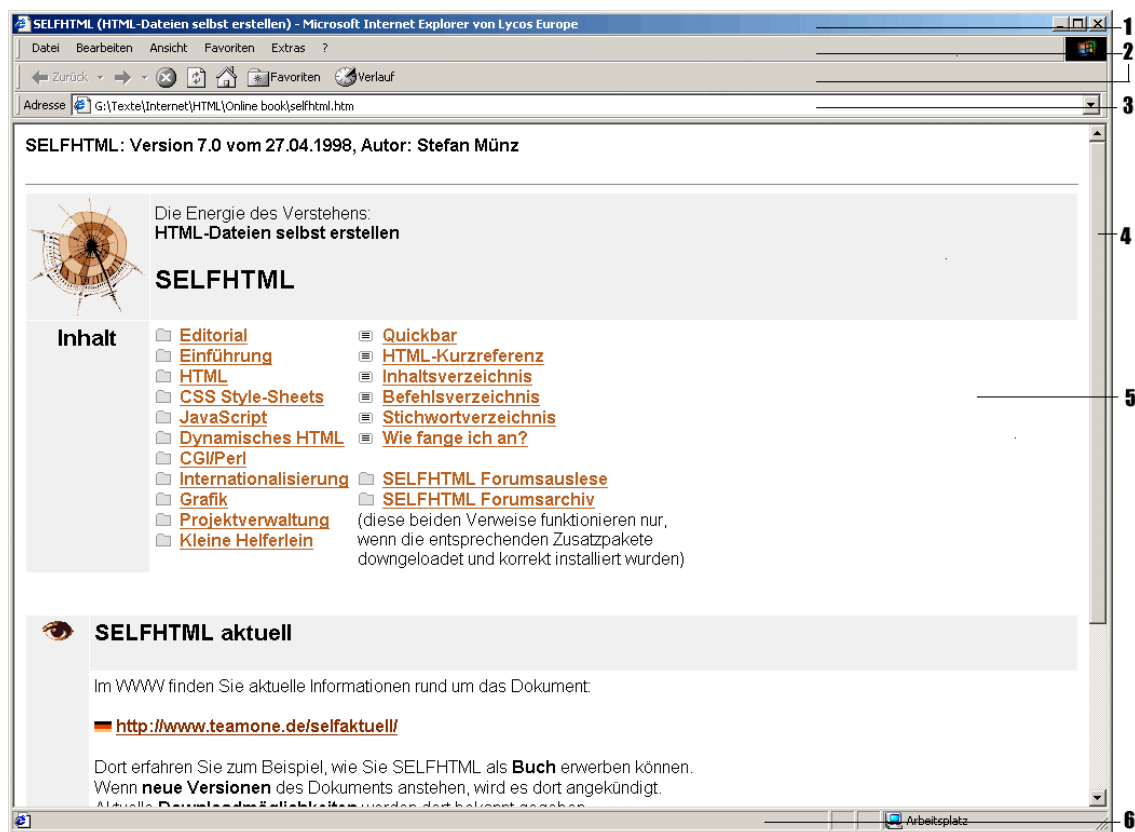


Abbildung 4: Oberfläche einer Browser-Software

Die verschiedenen Komponenten der Browser-Software sind in der Abbildung mit eins bis sechs beziffert:

ad 1: TITELLEISTE: Die Titelleiste zeigt den Namen des aktuellen Knotens an (*SelfHTML: HTML Dateien selbst erstellen*). In der Regel ist dieser Name mit der Hauptüberschrift des Knotens identisch.¹

ad 2: MENÜLEISTE und SYMBOLLEISTE: Die Menüleiste aktiviert per Mausklick sogenannte PULL-DOWN-MENÜS, die Befehle enthalten. Mit diesen kann der Leser z.B. bestimmte Knoten für späteren Gebrauch abspeichern (Menü *Datei*), Textpassagen in andere Anwendungen kopieren (Menü *Bearbeiten*) oder allgemeine Parameter zur Darstellung der Knoten – Schriftgröße, Anzeige des HTML-Quellcodes – festlegen (Menü *Ansicht*). Ferner sind über das Menü ? diverse Hilfestellungen zum Umgang mit der Browser-Software erhältlich. Die Symbolleiste enthält die eigentlichen Navigationselemente: Sie ermöglicht dem Leser das Vorwärts- und Rückwärtsbewegen im Hypertext, das direkte Ansteuern eines vorher definierten Start- bzw. Suchknotens sowie das Festlegen und Aktivieren von FAVORITEN, die als elektronische Variante eines Le-

¹ Zur Problematik, eindeutige und treffende Namen für Knoten zu finden siehe Horn 1989 S. 60 f.

sezeichens zu verstehen sind.¹

ad 3: ADRESSLEISTE: Hier wird die Adresse des Knotens im WWW bzw. auf dem lokalen Speichermedium angezeigt.

ad 4: SCROLLBAR: Die Scrollbar ermöglicht die Bewegung innerhalb von Knoten, die mehr Platz als eine Bildschirmseite zur Darstellung beanspruchen.

ad 5: DARSTELLUNGSFENSTER: Hier wird der Knoten-Inhalt dargestellt.

ad 6: STATUSZEILE: In der Statuszeile wird Auskunft über aktuelle Operationen der Software gegeben (z.B. Laden von Knoten). Außerdem werden nähere Informationen zu Links (z.B. das Link-Ziel) angezeigt, sobald der Leser den Mauszeiger auf den Link-Anker positioniert hat.

6.4.9 Backtracking, Historie, Bookmarks

Bei sämtlichen in diesem Kapitel vorgestellten Hilfen handelt es sich um rückwärtsgerichtete Navigationsmittel. Sie strukturieren und organisieren den vom Leser bereits zurückgelegten Lesepfad. Die Kapitel 6.4.10 bis 6.4.12 stellen hingegen vorwärtsgerichtete Orientierungshilfen vor, mit denen die Gesamtstruktur veranschaulicht und somit auch Auskunft über mögliche zukünftige Lesepfade gegeben wird. Das sogenannte BACKTRACKING ermöglicht dem Leser ein schrittweises ‚Zurückgehen‘ im Hypertext, indem zu sämtlichen bereits besuchten Knoten zurück navigiert werden kann. Die Browser-Software ermöglicht diese Funktion über das Symbol *Zurück*.² Vom aktuellen Knoten ausgehend kann also das „schrittweise Zurückverfolgen eines eingeschlagenen Suchwegs“ (Kuhlen 1991 S. 156) beginnen:

„Tauchen Orientierungsprobleme auf, kann die Leserin ihren Weg zurückverfolgen, bis sie auf vertraute Knoten stößt oder wieder am Anfang ist.“

(Gerdes 1997 S. 43)

Ferner verfügt der Leser über die Möglichkeit, per Klick auf das Symbol *Verlauf* eine (nach Adressen und Knotennamen sortierte) alphabetische Liste sämtlicher zuvor aktivierter Knoten einzusehen und den gewünschten nochmals zur Ansicht zu bringen. In der Literatur wird diese Funktion DIALOG-GESCHICHTE bzw. HISTORIE genannt (Gerdes 1997 S. 44 sowie Kuhlen 1991 S. 156).

¹ Symbolleiste und Menüleiste wurden zusammengefaßt, da sämtliche als Symbole dargestellte Navigationsfunktionen ebenfalls über die Menüs *Explorer* und *Favoriten* zugänglich sind.

² Ebenso kann mit *Vorwärts* zum jeweils nächsten Knoten einer bereits durchlaufenen Lesereihenfolge navigiert werden.

Mit Hilfe des Symbols *Favoriten* bzw. des gleichnamigen Menüs werden elektronische Lesezeichen, sogenannte BOOKMARKS, gesetzt, die jederzeit wieder aktiviert werden können. Bereits angewählte Verweise werden von der Software per Standardeinstellung graphisch abgesetzt dargestellt. Auf diese Weise wird insbesondere bei Übersichtsknoten (etwa Inhaltsverzeichnissen) der Lesefortschritt visualisiert.

6.4.10 Graphische Übersichten

Die hier vorgestellten Navigationshilfen werden oftmals mit dem Namen BROWSER bezeichnet. Sie sind keine Komponenten der oben beschriebenen BROWSER-SOFTWARE! Sie können (mit größerem Aufwand) mittels sogenannter CLICKABLE IMAGE MAPS auf einem Knoten des Hypertextes zugänglich gemacht werden. Allerdings wird dies nach Kenntnisstand des Autors eher selten realisiert. Folglich sollen diese graphischen Übersichten auch nur der Vollständigkeit halber kurz angesprochen werden.¹

Ein BROWSER ist laut Gerdes ein spezieller Knoten der

„[...] ähnlich einer Landkarte eine graphische Übersicht über die Vernetzungsstruktur des gesamten Hypertextes oder auch nur eines Teilausschnittes bietet.“

(Gerdes 1997 S. 37)

6.4.11 Anfrage- bzw. Suchmechanismen: Information Retrieval

Viele komplexe Hypertexte im WWW werden mittels eines Mangementsystems auf der Grundlage von Datenbanken verwaltet. Dadurch ist es möglich, mit Hilfe von Suchfunktionen Anfragen an das System zu richten. Je nach Ausprägung und Komplexität von Datenbank und Suchkomponente kann dabei nach Knoten-Titeln gefragt oder aber auch eine Volltextsuche durchgeführt werden (siehe Kuhlen S. 18, 25f. u. 212f.). Eine genaue Schilderung der hier vorliegenden Sachverhalte ist aber eher Aufgabe der informationswissenschaftlichen Disziplinen.²

¹ Ausführlicher informiert Kuhlen 1991 S. 142 f.

² Siehe hierzu Kuhlen 1991, Lucarella u. Zanzi 1993 sowie Agosti 1996. Zur Abgrenzung von Information Retrieval in Hypertexten bzw. dem in Expertensystemen siehe Woodhead 1991.

6.4.12 Guided Tours

Gibt der Autor die Lesereihenfolge eines Hypertextes in Teilen oder auch komplett vor, so bezeichnet man diesen vordefinierten Pfad in der Literatur als GUIDED TOUR.¹ Die prinzipielle Navigationsfreiheit im Hypertext wird hierbei eingeschränkt. Ist der ganze Text als Guided Tour realisiert, läßt sich im strengen Sinn nicht mehr von einem Hypertext sprechen, da er in diesem Fall vom Autor rein linear angelegt ist.

¹ vgl. Conklin 1987, Marshall u. Irish 1989, Kuhlen 1991, Gerdes 1997

7 Zusammenfassung: Hypertexte im WWW

Im vorliegenden Teil A wurde der Untersuchungsgegenstand HYPERTEXT in historischer, intensionaler und extensionaler Hinsicht festgelegt. Die wichtigsten Beobachtungen seien hier nochmals zusammengefaßt:

Allgemein formal besteht ein Hypertext aus Knoten und per Mausklick direkt ausführbaren Links. Knoten sind die elementaren Informationseinheiten, die von Inhalt und Größe her nicht festgelegt sind: Sie können textuell oder auch multimedial ausgerichtet sein, einzelne Propositionen oder auch ganze Bücher enthalten. Links sind gekennzeichnete ‚Sprungstellen‘ innerhalb von Knoten, die auf andere Knoten verweisen. Die Arten der Verknüpfungen sind hierbei sehr verschieden, sie können nach formalen, inhaltlichen oder typisierungsspezifischen Kriterien unterschieden werden. Der Leser legt durch seine Navigation (= Klicken auf Links) einen Lese pfad an. Dieser Lese pfad bewegt sich in den eng gesteckten Grenzen der vom Autor angelegten Knoten- und Link-Struktur. Von einer poststrukturalistischen Marginalisierung des Autors kann beim Hypertext also in der Form nicht gesprochen werden: Die Struktur von Hypertexten ist nicht-linear ausgerichtet, die Knoten sowie die tatsächlichen Lese pfade sind hingegen linear.¹

Zum Durchbruch verholfen hat der Hypertext-Idee das WWW. Sämtliche Texte des Corpus' sind aus dem WWW entnommen. Dort sind einzelne Hypertexte, lineare E-Texte und kleinere Hypertext-Netze, die alle in der Sprache HTML verfaßt sind, zum größten Hypertext-Netz der Welt verknüpft. Zum Lesen solcher Hypertexte benötigt man eine Browser-Software, deren zahlreiche eingebauten Navigationshilfen die traditionellen Orientierungshilfen ergänzen, welche zumeist zusätzlich als hypertext-interne Navigation umgesetzt sind.

Vereinfacht ließe sich eine Definition des Untersuchungsgegenstands auf folgenden Satz verdichten: Hypertexte sind im elektronischen Medium realisierte, tendenziell nicht-lineare und potentiell multimedial ausgerichtete Texte. Das Sichten des Untersuchungsgegenstandes hat daher keine Evidenz dafür gebracht, daß primär textuelle Hypertexte nicht mit textlinguistischen Methoden zu untersuchen seien. Im Teil B werden daher nun Größen und Methoden der Textlinguistik vorgestellt, mit deren Hilfe ein Analysemodell für Hypertexte entwickelt werden soll.

¹ Das bedeutet, sie verfügen genauso wie gedruckte Texte über eine lineare Abfolge, die aber von zahlreichen nicht-linearen Mitteln der Kohäsion und Kohärenz durchzogen ist (siehe Kapitel 9).

Teil B: Ein Textlinguistisches Analysemodell für Hypertexte (TAH)

Teil A hat gezeigt, daß aufgrund Extension und Intension von HYPERTEXT eine textlinguistische Herangehensweise sinnvoll sein kann. Daher wird im vorliegenden Teil B ein TEXTLINGUISTISCHES ANALYSEMODELL FÜR HYPERTEXTES entwickelt, welches in Kapitel 14 zusammenfassend unter dem Kürzel TAH eingeführt wird. TAH ist ein in die Ebenen *Hypertext*, *Knoten* und *Link* unterteiltes Modell und folgt auf jeder der Ebenen einer primär pragmatisch ausgerichteten Top-Down-Analyse.

Die Entwicklung des Modells geschieht wie folgt: Ausgehend vom vorwissenschaftlichen Textverständnis wird in Kapitel 8 kurz die geschichtliche Entwicklung des linguistischen Textbegriffes geschildert: Die ältere strukturbezogene sprachsystematische Textdefinition wird hierbei für diese Arbeit verworfen, da Texte mehr als lediglich das Bemühen um satzübergreifende Strukturen sind. Verwendet wird hingegen ein kommunikativ-funktionaler Textbegriff, der prinzipiell auch offen für multimedial ausgerichtete Hypertexte ist.

In Kapitel 9 wird näher auf die einzelnen Kriterien der TEXTUALITÄT von textuellen Gebilden eingegangen. Die Kriterien werden hierbei hinsichtlich ihrer Relevanz für Texte und Hypertexte besprochen. Dies führt letztendlich zur Feststellung, daß in einem textzentrierten Analysemodell wie dem TAH lediglich textzentrierte Textualitätskriterien wie KOHÄSION und KOHÄRENZ besprochen werden können. Für die leserzentrierten Textualitätskriterien wären empirische Untersuchungen mit einer größeren Anzahl von Testpersonen nötig.

Nach dem Festlegen der allgemeinen Ausrichtung des Textbegriffes und Auswahl der entscheidenden Textualitätskriterien werden in Kapitel 10 die Dimensionen einer Textanalyse eingeführt. Hierbei handelt es sich hauptsächlich um TEXTFUNKTION, TEXTSTRUKTUR und TEXTSORTE. Auf PARATEXTUELLE ‚Randbereiche‘ von Texten wird gesondert in Kapitel 11 eingegangen.

Die Kapitel 12 und 13 führen zwei eher formale Bottom-Up-Methoden ein, mit deren Hilfe sich Unklarheiten bei – für unsere Untersuchung in Teil C besonders interessanten – thematisch-strukturellen Aspekten von Texten beseitigen lassen. Bei diesen Instrumenten handelt es sich um das MODELL DER REFERENTIELLEN BEWEGUNG sowie die MAKROSTRUKTURANALYSE.

Die Einführung der textlinguistischen Größen und Methode erfolgt nicht isoliert, sondern ist von Anfang an direkt auf den Untersuchungsgegenstand bezogen. Stets wird also überprüft, ob textlinguistische Aussagen nicht nur auf Texte sondern auch auf Hypertexte anzuwenden sind.

8 Textdefinitionen

In diesem Kapitel wird die geschichtliche Entwicklung des linguistischen Textbegriffes geschildert und darauf aufbauend eine für diese Arbeit maßgebende kommunikativ-funktionale Definition von TEXT festgelegt.

8.1 Text als vorwissenschaftliche Größe

Brinker 2001 und Vater 1994 gehen auf die Verwendungsweisen des Begriffes TEXT in der Alltagssprache ein. Neben Gebrauchsmustern wie *Text zu einer Abbildung*, *über einen Text predigen* oder der *Text eines Liedes* läßt sich eine Kernbedeutung im Sinne von TEXT als „(schriftlich) fixierte sprachliche Einheit, die in der Regel mehr als einen Satz umfaßt“ herauslesen. (Brinker 2001 S. 12)

Diese vorwissenschaftliche Definition hält jedoch nach Vater 1994 einer genaueren Überprüfung nicht stand: Vater kritisiert hierbei die ausschließliche Fixierung auf das schriftliche Medium sowie die Einschränkung auf Äußerungen, die mehr als einen Satz umfassen. Als Argumente gegen die letztere Einschränkung führt Vater Sprichwörter als Ein-Satz-Texte ins Feld. Der Umfang eines sprachlichen Gebildes scheint also keine Rolle für dessen TEXTHAFTIGKEIT zu spielen (Wawrzyniak 1980 S. 7, Vater 1994 S. 16).

8.2 Text: Produkt und Produktion

Ein weiterer Einwand seitens Vater an der im oberen Abschnitt eingeführten vorwissenschaftlichen Textdefinition ist der folgende:

„Text als ‚fixierte sprachliche Einheit‘ bedeutet eine einseitige Festlegung auf einen Aspekt von Texten, nämlich Text als Ergebnis einer sprachlichen Kommunikationshandlung; doch spielt der Kommunikationsvorgang selbst (die ‚Textkonstruktion‘) eine mindestens ebenso große Rolle;“

(Vater 1994 S. 15)

Bei der Untersuchung von Hypertexten stellt sich die Frage nach der Textkonstruktion in verstärktem Maß, kann man doch dem Leser eines Hypertextes eine gewisse ‚Ko-Autorenschaft‘ zugestehen: Der Leser bewegt sich durch die vom Autor mit Knoten und Links geschaffene Hypertext-Struktur.

Durch Klicken auf die Links konstruiert er den konkreten Text als Lesepfad zum Teil erst beim Lesen selbst. Die Struktur des Hypertextes ist also insofern Ergebnis der sprachlichen Kommunikationshandlung des Autors, während ein

Lesepfad sozusagen in ‚Teamwork‘ durch Strukturentscheidungen des Autors und Navigationsentscheidungen des Lesers entsteht.

Beide Phänomene sind aber Produkte und nicht Produktion sprachlicher Handlungen: Der gesamte Hypertext sowie auch einzelne Lesepfade gehören der TEXTINTERNEN Dimension an. Über kognitive Vorgänge beim eigentlichen Kommunikationsvorgang, also über Konstruktions- bzw. Rezeptionsaspekte lassen sich bei einer TEXTZENTRIERTEN Analyse zwar Vermutungen anstellen – aber keine gesicherten Aussagen! Hierfür bräuchte es empirische AUTOR- bzw. LESERZENTRIERTE Untersuchungen, die vom Wissen des Autors und des Lesers ausgehend Rückschlüsse ziehen auf hierbei stattfindende kognitive Vorgänge (etwa dem Erschließen von im Text nicht Gesagtem durch mentale Schemata o.ä.; siehe hierzu auch Kapitel 9.2.2). Die KOGNITIVE LINGUISTIK vertritt demnach einen konstruktivistischen Textbegriff, der weit über den in dieser Arbeit vertretenen textinternen, kommunikativ-funktionalen hinausgeht:

„Die jeweilige sprachliche Datenstruktur (also der Text) ist somit nur in einem kommunikationsabstrakten Sinn real; empirisch real ist nur das Rezeptionskonstrukt eines bestimmten Rezipienten. Die Aktivität des Rezipienten bei der Textverarbeitung wird experimentell u.a. mit Hilfe von Reproduktionsexperimenten untersucht. Die Gegenüberstellung von Originaltext (also dem im Experiment vorgelegten Text) und Resultattext (also der Repräsentation, die von der Vpn [Versuchsperson] geliefert wird) soll dabei kognitive Operationen transparent machen. Die spezifischen Veränderungen im reproduzierten Resultattext (wie Hinzufügungen, Paraphrasen) werden als Spuren der kognitiven Konstruktivität des Rezipienten bewertet. [...] Dabei geht es in den prozeduralen Textmodellen nicht nur um die aktuellen Rezeptionsprozesse in bestimmten Situationen, sondern vielmehr um die allgemeinen Prinzipien, die diesen Prozessen zugrunde liegen. Ziel der Forschung ist die Erklärung der textuellen Kompetenz, die für das Verstehen bzw. Hervorbringen von Texten verantwortlich ist.“

(Schwarz 1992 S. 159 f.)

Das Analysemodell, das in diesem Teil der Arbeit entwickelt wird, ist ein tendenziell statisches und textzentriertes, untersucht also deskriptiv das Produkt sprachlicher Handlung.

8.3 Die Entwicklung des Terminus' *Text*

Neben so heterogenen Disziplinen wie Psychologie, Geschichtswissenschaft, Literaturwissenschaft, Rechtswissenschaft, Theologie und Anthropologie, die sich alle auf ihre Weise mit der Interpretation von Texten beschäftigen, existiert eine genuin linguistische Teildisziplin, die sich um die Entwicklung eines Textbegriffes sowie um den Aufbau eines Untersuchungsinventars zur Interpretation von Texten bemüht: Die TEXTLINGUISTIK stellt eine sprachwissenschaftliche Disziplin dar,

„[...] die sich mit der Analyse satzübergreifender sprachlicher Regularitäten beschäftigt und das Ziel hat, die konstitutiven Merkmale der sprachlichen Einheit Text zu bestimmen und damit eine Texttheorie zu begründen.“

(Bußmann 1990 S. 779)

Ausgehend von strukturalistischen Ansätzen und von einer Textgrammatik im Sinne einer erweiterten Satzgrammatik, versteht die moderne Textlinguistik den Begriff TEXT zumeist nicht mehr nur als „Folge von Sätzen und damit als Einheit des Sprachsystems“ sondern unter Einbeziehung pragmatischer Aspekte als „kommunikative Einheit“ (Bußmann 1990 S. 779).¹

Als Problembereiche der Textlinguistik werden genannt:

- (a) Allgemeine Aspekte der strukturellen und funktionalen Textkonstitution, d.h. der Textualität;
- (b) Klassifikation von Texten im Rahmen einer Texttypologie;
- (c) Probleme einer textlinguistischen Integration von Stilistik und Rhetorik;
- (d) Interdisziplinär orientierte Forschungen zur Textverarbeitung und Verständlichkeit.

(vgl. Bußmann 1990 S. 779)

Vorliegende Arbeit befaßt sich hauptsächlich mit dem Problembereich (a). Die Bereiche (b), (c) und (d) werden lediglich am Rande behandelt.

Namensgebender Untersuchungsgegenstand der Textlinguistik ist das Phänomen TEXT. Vater nennt folgende Dimensionen, in denen Texte variieren können (vgl. Vater 1994 S. 16f.):

mündlich – schriftlich²
 einsätzig – mehrsätzig
 monologisch – dialogisch
 lingual – außerlingual – gemischt

Die letzte Unterscheidungsdimension ist besonders interessant für hypermediale Gebilde. Sie weicht jedoch von der für diese Arbeit maßgebenden Definition von Text (weiter unten, siehe auch Brinker 2001 S. 17) in der Weise ab, als daß sie nicht eine „Folge sprachlicher Zeichen“ ist. Die Frage ist nun, wie eine Textanalyse mit außerlingualen bzw. gemischten Hypertext-Knoten umzugehen hat. Eine kommunikativ-funktionale Analyse dürfte von der prinzipiellen

¹ Zur Geschichte der Textlinguistik siehe Sowinski 1983 S. 19 f. und Brinker 1994 S. 12 f.

² Siehe Helbig 1975 S. 48, Mackeldey 1987 S. 11 sowie Vater 1994 S. 16.

Ausrichtung her keine Probleme mit der Beschreibung nichtsprachlicher Erscheinungen haben – sofern eine (TEXT)FUNKTION (siehe Kapitel 10) ersichtlich ist. Auf thematisch- und grammatisch-struktureller Ebene versagen dann aber die linguistischen Methoden, wenn es an die Beschreibung außerlingualer ‚Texte‘ geht. Aus diesem Grund wird das kommende Hypertext-Analysemodell ausschließlich für linguale Bestandteile konstruiert sein, durch seine kommunikativ-funktionale Ausrichtung aber tendenziell offen für Analysen außerlingualer Elemente bleiben. Die technischen Methoden müßten dann allerdings die Disziplinen liefern, welche sich mit der Analyse solcher nichtsprachlicher ‚Texte‘ beschäftigen.

Brinker 2001 faßt die linguistisch motivierten Versuche, eine Definition von TEXT zu erlangen, zusammen und muß feststellen, daß „es verschiedene Textdefinitionen“ gibt, man aber keinen „allgemein gültigen Textbegriff“ bestimmen kann (Brinker 2001 S. 12). Er nimmt eine Unterscheidung zwischen dem Textbegriff einer sprachsystematisch ausgerichteten sowie dem einer kommunikationsorientierten Textlinguistik vor.

Die erstgenannte erweitert lediglich die bisher angenommene Hierarchie sprachlicher Einheiten (PHONEM, MORPHEM/WORT, PHRASE, SATZ) um die Einheit TEXT. Dies drückt die Auffassung aus,

„[...] daß nicht nur die Wort- und Satzbildung, sondern auch die Textbildung (die Textkonstitution) durch das Regelsystem der Sprache gesteuert wird und auf allgemeinen, sprachsystematisch zu erklärenden Gesetzmäßigkeiten gründet.

[...]

Text wird definiert als eine kohärente Folge von Sätzen. Das bedeutet aber, daß der Satz nach wie vor als Markstein in der Hierarchie sprachlicher Einheiten angesehen wird; er gilt als die Struktureinheit des Textes. Die wichtigste Konsequenz dieser Konzeption ist, daß der für die Textlinguistik zentrale Begriff der Textkohärenz rein grammatisch aufgefaßt wird.“

(Brinker 2001 S. 14)

Die zweite Richtung der Textlinguistik ist die kommunikationsorientierte.¹ Sie wendet sich gegen die sprachsystematisch ausgerichtete Forschung, der sie eine idealisierte Untersuchung von Texten vorwirft, indem diese ausschließlich als isolierte, statische Objekte behandelt werden. Es müsse aber vielmehr berücksichtigt werden,

¹ Vertreter sind: Isenberg 1970 u. 1976, Schmidt 1973, Sandig 1973 sowie Kallmeyer u.a. 1974.

„[...] daß Texte immer eingebettet sind in eine Kommunikationssituation, daß sie immer in einem konkreten Kommunikationsprozeß stehen, in dem Sprecher und Hörer bzw. Autor und Leser mit ihren sozialen und situativen Voraussetzungen und Beziehungen die wichtigsten Faktoren darstellen.“

(Brinker 2001 S. 15)

Diese zweite Richtung beruft sich auf Pragmatik und Sprechakttheorie und sieht den Text nicht mehr ausschließlich als kohäsive Satzfolge, sondern vielmehr als mehr oder weniger komplexe Sprechhandlung. Ziel einer kommunikationsorientierten Textanalyse ist das Herausfiltern der kommunikativen TEXTFUNKTION eines speziellen Textes.

Abschließend versucht Brinker einen Textbegriff zu finden, der beide Richtungen zu integrieren vermag. Er kommt dabei zu folgender Definition:

„Der Terminus *Text* bezeichnet eine begrenzte Folge von sprachlichen Zeichen, die in sich kohärent ist und die als Ganzes eine erkennbare kommunikative Funktion signalisiert.“

(Brinker 2001 S. 17)

Diese Definition soll für die sich anschließende Analyse als grundsätzliche Arbeitsdefinition gelten¹, da sie folgende Stärken aufweist: TEXT wird „als begrenzte Folge von Zeichen“ gesehen. Nach diesem Kriterium beurteilt, kann es sich beim WWW in seiner Gesamtheit schon formal nicht um einen Text nach obiger Definition handeln. Dem WWW werden nämlich in jeder Sekunde neue Hypertext-Knoten hinzugefügt, so daß es praktisch unmöglich ist, die Grenzen dieser Folge von sprachlichen Zeichen aufzuzeigen.² Bereits dieses formale Kriterium deutet darauf hin, daß als Texte im herkömmlichen Sinne wohl eher Hypertexte denn Hypertext-Netze zu verstehen sind (siehe auch Kapitel 5). Auch werden sich eher Kohärenzstrukturen zwischen Knoten eines Hypertextes finden lassen, wohingegen Übergänge zwischen zwei Hypertexten innerhalb eines Hypertext-Netzes tendenziell eher weniger kohärent sein dürften. Daher spricht auch das zweite Kriterium der obigen Textdefinition für die

¹ Wichtig ist abermals, die Definition von Brinker 2001 nicht als die allgemeingültige Textdefinition zu begreifen. In zwei Sätzen kann wohl kaum ein so komplexes Phänomen wie TEXT beschrieben werden. Die Definition soll lediglich die allgemeine Ausrichtung für die Arbeit darstellen, in dem sie ein grobes Raster für die Entscheidung über die Texthaftigkeit des Untersuchungsgegenstandes zur Verfügung stellt.

² Die einzige Möglichkeit wäre, sämtliche im WWW existierenden Dokumente zu einem Zeitpunkt *X* zu speichern und diese Datensammlung als synchrone Momentaufnahme zu betrachten. In der Praxis ist dies jedoch nicht zu realisieren, da sämtliche ins WWW integrierte Computer zum selben Zeitpunkt von einer Instanz abgefragt werden müßten.

Texthaftigkeit von Hypertexten und gegen die von Hypertext-Netzen. Das letzte und meiner Meinung nach entscheidende Kriterium¹ der Definition von Brinker ist die KOMMUNIKATIVE FUNKTION, die die „Folge von sprachlichen Zeichen als Ganzes“ signalisiert. Dies trifft, wie sich zeigen wird, für Knoten eines Hypertextes sowie (mit Einschränkungen) für Hypertexte zu. Das WWW als momentan größtes Hypertext-Netz verfügt über keine einheitliche Textfunktion.

Insgesamt handelt es sich bei Brinker um eine sehr allgemeine Definition. Dies bringt sowohl Vorteile als auch Nachteile mit sich: Als positiv darf gelten, daß mit einer weiten Definition ein Großteil der doch sehr unterschiedlichen Repräsentanten von TEXT untersucht werden kann. Dies geht so weit, daß durch eine rein kommunikativ-funktionale Herangehensweise sozusagen Schnittstellen für die Analyse von ‚Texten‘ angeboten werden, die auf völlig andere Weise medial realisiert sind. Die grundlegende Ausrichtung der Analyse zu definieren, war Absicht dieses Kapitels. So wichtig diese prinzipielle Feststellung aber auch ist: Gewonnen ist durch sie für die praktische Arbeit der Textanalyse freilich noch relativ wenig. Daher müssen in den folgenden Kapiteln Einzelaspekte der textuellen Dimensionen herausgegriffen werden, auf ihre Brauchbarkeit zur Untersuchung von Hypertexten hin untersucht und zuletzt zu einem schlüssigen Gesamtanalyse- und Interpretationsmodell für die einzelnen Ebenen von Hypertexten zusammengesetzt werden.

8.4 Exkurs: Einheiten zwischen Text und Satz

Bevor dies erfolgt, soll der Blick noch kurz auf weitere mögliche Einheiten zwischen SATZ und TEXT gerichtet werden.

Umfangreichere Texte, z.B. Bücher, sind oft in Kapitel (vgl. Stevick 1970) unterteilt. Kapitel stellen Teiltex te dar, die zumeist über eine eigene Überschrift verfügen und eine inhaltliche Einheit bilden:

„Die Einteilung eines Prosawerks in K[apitel] dient der Verdeutlichung von Gliederung und Proportion, die K[apitel]-Überschriften der Erweckung und Steuerung der Lesererwartung.“

(Schweikle 1990 S. 232)

¹ Das PRIMAT DES FUNKTIONALEN tritt noch deutlicher bei der Kurzdefinition von Bußmann 1990 zu Tage: Text wird hier gesehen als „sprachliche Äußerungsform einer kommunikativen Handlung“ die ihrerseits zuallererst bestimmt ist von „pragmatischen, *textexternen* Kriterien einer kommunikativen Intention, die situationsspezifisch ist und auf eine entsprechende Hörererwartung trifft“. Erst an zweiter Stelle spricht sie von „sprachlichen, *textinternen* Merkmalen“ eines Textes. (Bußmann 1990 S. 776)

Weitere, dem Text untergeordnete Ebenen, auf denen eine textlinguistische Analyse betrieben werden kann, sind *Abschnitt* bzw. *Absatz* und *Satz*. Der Begriff *Abschnitt* ist schwieriger zu definieren als der oben beschriebene, in der Regel übergeordnete Terminus *Kapitel* oder der untergeordnete Begriff *Absatz*. Ebenso wie die beiden verwandten Begriffe kann *Absatz* in etwa paraphrasiert werden mit *inhaltlich zusammengehöriger Teil eines Textes*. Formal-typographisch ist *Abschnitt* jedoch, im Gegensatz zu *Kapitel* und *Absatz*, weniger festgelegt.

Welcher der oben vorgestellten Begriffe am besten auf hypertextuelle Knoten anzuwenden ist, ist nicht klar. Ein Knoten kann eine einzige Proposition, einen Satz oder auch – als E-Text – einen vollständigen Text samt aller untergeordneten Einheiten (etwa Kapitel, Abschnitte etc.) enthalten.

9 Textualitätskriterien

Im letzten Kapitel wurde das allgemeine, kommunikativ-funktionale Textverständnis dieser Arbeit erläutert. Bevor die Dimensionen eines Textes bzw. einer Textanalyse – also TEXTFUNKTION, TEXTSTRUKTUR, TEXTSORTE sowie PARATEXTUELLE Elemente – eingeführt werden können, muß näher definiert werden, was überhaupt die konstituierenden Elemente eines Textes ausmacht – was also die TEXTUALITÄTSKRITERIEN eines textuellen Gebildes sind und welche Dimensionen des Textes sie betreffen. Dies geschieht im vorliegenden Kapitel.

De Beaugrande und Dressler interpretieren das Phänomen TEXT als "KOMMUNIKATIVE OKKURRENZ [...], die sieben Kriterien der Textualität erfüllt" (de Beaugrande/Dressler 1981 S. 3 [Auszeichnung im Original]). Unter TEXTUALITÄT versteht man „die Gesamtheit der Eigenschaften, die einen Text zum Text machen“ (Vater 1992 S. 31).¹ Ist eines dieser TEXTUALITÄTSKRITERIEN nicht erfüllt, so kann man de Beaugrande und Dressler zufolge nicht mehr von einem Text sprechen, denn:

„Wenn irgendeines dieser Kriterien als nicht erfüllt betrachtet wird, so gilt der Text nicht als kommunikativ. Daher werden nicht-kommunikative Texte als Nicht-Texte behandelt [...]“

(de Beaugrande/Dressler 1981 S. 3)

Die Kriterien im einzelnen sind: KOHÄSION, KOHÄRENZ, INTENTIONALITÄT, AKZEPTABILITÄT, INFORMATIVITÄT, SITUATIONALITÄT UND INTERTEXTUALITÄT.

Vater 1992 stellt meiner Meinung nach zu Recht die Frage, ob stets alle Textualitätskriterien erfüllt sein müssen, um einen Text zu konstituieren und ob bei einem, in welcher Form auch immer, nicht-kommunikativen Text von einem NICHT-TEXT gesprochen werden kann. Im weiteren Verlauf der Arbeit wird sich ergeben, daß die von de Beaugrande und Dressler vertretenen Thesen in dieser Schärfe nicht zu halten sind. Lediglich das Kriterium der KOHÄRENZ, d.h. die innere, kognitive „SINNKONTINUITÄT“ (de Beaugrande/Dressler 1981, S. 88 [Auszeichnung im Original]) eines Textes, muß unbedingt erfüllt sein, um sicherzustellen, daß der Leser ein sprachliches Gebilde als Text begreift. Kohärenz stellt damit das entscheidende Kriterium dar (vgl. auch Vater 1992

¹ Vater weist darauf hin, daß sich TEXTUALITÄT auf den statisch-strukturellen Zustand eines Textes bezieht, während der Begriff VERTEXTUNG (Wawrzyniak 1980) den Vorgang der Textproduktion in den Vordergrund stellt.

S. 65).¹ Im folgenden werden die sieben Textualitätskriterien kritisch besprochen und auf deren Verwendbarkeit für Hypertexte hin untersucht. In einer kurzen Zusammenfassung wird nochmals aufgeführt, welche Kriterien in das zu entwickelnde Hypertext-Analysemodell aufgenommen werden.

9.1 Kohäsion

Das Prinzip der KOHÄSION beschreibt den Textzusammenhang, der durch formale Mittel der Grammatik hergestellt wird. In der Regel ist Kohäsion die Grundlage der Kohärenz. Als kohäsives Mittel gilt hauptsächlich die TEXTPHORIK (= intratextuelles Verweissystem auf Grundlage von Wiederaufnahmestrukturen einzelner Textelemente). Aber auch Textverdichtungsstrukturen (Ellipsen, Proformen) und morpho-syntaktische Mittel, mit denen Beziehungen zwischen einzelnen Elementen eines Textes ausgedrückt werden können (z.B. Kongruenzbeziehungen), unterstützen die Kohäsion eines Textes.

Im folgenden werden nun die kohäsiven Mittel ausführlicher vorgestellt: Die Textphorik, als wichtigstes Mittel der Kohäsion, beruht semantisch auf kompletter oder teilweiser KOREFERENZ². Der Begriff subsumiert sowohl syntagmatische Substitutionen von Lexemen durch Proformen (Textverweise) als auch rein semantische (Teil)Wiederaufnahmestrukturen (etwa ISOTOPIE).

Folgende Satzsequenz ist ein beliebtes Beispiel zur Veranschaulichung kohäsiver Mittel: Koreferenz entsteht in diesem Fall durch eine Referenzkette zwischen Anapher, hier ein Pronomen, und seinem Antezedenten.

Maximilianus hat angerufen. Er kommt um 15 Uhr.

¹ Ob dieser kohärente Text dann in jeder Situation für jeden Leser ‚Sinn macht‘, ist wiederum eine Frage, die de Beaugrande und Dressler mit anderen Textualitätskriterien beantworten wollen. Ich schließe mich jedoch Vater an, der meint, daß ein kohärenter Text auch dann noch ein Text sei, wenn er für bestimmte Leser aufgrund mangelnden Vorwissens in bestimmten Situationen ‚keinen Sinn macht‘. Wäre dies nicht der Fall, so würde eine textzentrierte Analyse, wie sie in dieser Arbeit durchgeführt wird, von vornherein sinnlos sein.

² Mit KOREFERENZ von Ausdrücken wird die EXTENSIONALE IDENTITÄT von Referenten hergestellt (z.B. *Noam Chomsky* und *Der Verfasser von ‚Syntactic Structures‘*), nicht etwa eine INTENSIONALE IDENTITÄT, wie sie u.a. durch (teilweise) Synonymie geleistet wird (z.B. *der Sprachwissenschaftler* und *der Linguist*). Bei Koreferenz handelt es sich also um den Aufbau einer Identität der designierten Referenten – nicht um das Gleichsetzen von Bedeutungen. Während Synonymie im mentalen Lexikon der Sprachbenutzer fest verankert ist, wird Koreferenz erst aus dem Kontext bzw. mit Hilfe des Weltwissens konstruiert.

ISOTOPIE ist als semantische Rekurrenz bzw. Teilrekurrenz zu verstehen: Eine Grundannahme des Isotopiekonzepts ist, daß die Wortbedeutung mittels Komponentenanalyse in kleinste semantische Merkmale zerlegt werden kann. Einzelne semantische Merkmale können in „Wörter[n] desselben Bedeutungs- bzw. Erfahrungsbereiches“ gehäuft auftreten (z. B. *Titel, Autor, Veröffentlichung*; siehe auch Bußmann 1990 S. 357). Auf diese Weise verbinden sich Wortbedeutungen zu semantischen Komplexen, welche als ISOTOPIEEBENEN bezeichnet werden. Diese Ebenen stellen das semantische Gerüst eines Textes dar und wirken somit textkonstituierend:

„In der Anzahl der I[sotopie]-Ebenen spiegelt sich die thematische Komplexität eines Textes.“

(Bußmann 1990 S. 357)

Der Übergang von einer Isotopieebene zur nächsten kann mit einem Wechsel vom Textthema zu einem Teilthema einhergehen.

Kohäsion in Hypertext

Betrachtet man das Textualitätskriterium KOHÄSION in Bezug auf Hypertext, so läßt sich folgendes feststellen: Die Verwendung kohäsiver Mittel bewirkt, wie gesagt, die Herstellung von Verbindungen zwischen einzelnen Texteinheiten. Zugleich verursacht das Benutzen von Referenzketten, die ja im Grunde nichts anderes als Querverweise auf bestimmte Stellen im Text sind, daß die Linearität eines Textes teilweise aufgehoben wird.¹ Kuhlen argumentiert in diesem Zusammenhang, daß die in normalen Texten häufig implizite Verkettung in Hypertexten "als explizit gemachte Verkettung" angesehen werden könnte (Kuhlen 1991 S. 34).

Durch die Aufteilung in Knoten werden in Hypertexten jedoch bestimmte

¹ Der Text selbst sowie die Rezeption bleiben freilich weiterhin linear, d.h. – in unserem Kulturkreis – von oben nach unten sowie von links nach rechts. Gemeint ist mit *nicht-linear* vielmehr, daß bei der kognitiven Verarbeitung des linearen Inputs Rückgriffe auf die bereits rezipierte TEXTBASIS erfolgen, um etwa Proformen mit dem zugehörigen Antezedenz zu verbinden. Tendenziell nicht-lineare Mittel im gedruckten Text selbst, wie z.B. Fußnoten, Querverweise etc. sind an dieser Stelle nicht gemeint.

Desweiteren sollte klar sein, daß diese textinternen Rückgriffe auf bereits Gelesenes flankiert werden von INTENDIERTEN sowie ELABORATIVEN INFERENZEN. Ersteres meint den Rückgriff auf im Text implizit mitgemeintes Wissen, daß explizit nicht erwähnt wurde (*Hans geht ins Bistro. Der Kellner bringt die Karte. Unausgesprochen bleibt etwa, daß Hans die Tür öffnet, einen freien Platz sucht und sich setzt*). Letzteres, der Rückgriff auf elaborative Inferenzen, meint hingegen die ständige Interaktion mit dem WELTWISSEN des Lesers, das möglicherweise in sogenannten SCHEMATA organisiert ist. Diese Schemata kann man sich als mentale Repräsentation von Standardsituationen des Lebens vorstellen. Die Schematheorie wird in Kapitel 9.2.2 nochmals kurz besprochen.

kohäsive Mittel (z.B. Proformen) über Knotengrenzen hinweg zum größten Teil unwirksam, "da ihr Gebrauch eine eindeutig definierte Lesefolge voraussetzt" (Hammwöhner 1993 S. 25). Kuhlen leitet daraus die Forderung nach KOHÄSIVER GESCHLOSSENHEIT der einzelnen Knoten ab. Damit ist gemeint, daß innerhalb eines Knotens keine impliziten kohäsiven Bezüge auf andere Knoten existieren sollten. Durch diese Geschlossenheit wird gewährleistet, daß ein Knoten von beliebigen anderen referenziert werden und somit Bestandteil verschiedener Leseptide sein kann.

Sowohl bei Kuhlen 1991 als auch bei Hammwöhner 1993 läßt sich beobachten, daß die Autoren stets über knotenübergreifende Kohäsion sprechen, auch wenn sie dies nicht explizit deutlich machen. Nur so lassen sich etwa folgende Sätze erklären:

„Traditionelle textuelle kohäsive Strukturen spielen in Hypertexten kaum eine Rolle; vielmehr werden sie entweder ganz aufgelöst oder durch die Verknüpfungstechnik explizit gemacht.“

(Kuhlen 1991 S. 37)

Dem ist nur bedingt zuzustimmen: Auf Knotenebene spielen sämtliche oben angesprochene Kohäsionsphänomene dieselbe Rolle, wie sie dies in traditionellen Texten ebenfalls tun. Aber auch in knotenübergreifenden Strukturen kann Kohäsion durchaus mit von der Partie sein: Wenn auch anaphorische bzw. kataphorische Wiederaufnahmestrukturen in der Regel dort keinen Sinn machen dürften, so sind doch explizite Wiederaufnahmen durch Repetitionsstrukturen¹ denkbar. Ebenso möglich sind implizite Wiederaufnahmen des Referenten durch Ausdrücke, die in einer gewissen Beziehung zu diesem stehen. Hier wäre etwa an Teil-von-Relationen zu denken: Wird zum Beispiel im ersten Knoten *Bahnhof* als Referent eingeführt, so kann in einem zweiten Knoten ohne Sinnverlust – sowie bei beliebiger Rezeptionsreihenfolge der beiden Knoten – der Referent *Bahnpolizei* eingeführt werden. Mentale SCHEMATA dürften hier ebenfalls eine bedeutende Rolle spielen (siehe auch Kapitel 9.2.2).

Ein weiterer Einwand gegen die strikte Forderung nach kohäsiver Geschlossenheit von Knoten, wie sie von Hammwöhner und Kuhlen vertreten wird, liefert das zwischen Kohäsion und Kohärenz (vgl. hierzu den nächsten

¹ Wobei über Knotengrenzen hinweg wohl nur totale Rekurrenz, etwa *Der Programmierer dieser Seite... | Der Programmierer* funktionieren dürfte, während partielle Rekurrenz wahrscheinlich problematisch sein dürfte.

Abschnitt) gelegene Phänomen der oben bereits angesprochenen Isotopie: In einem Hypertext, der Werk und Leben von Noam Chomsky zum Thema hat, kann der Beginn eines Knotens durchaus so aussehen: *Der Autor der Syntactic Structures...*

Chomsky kann hier – das entsprechende Wissen des Lesers vorausgesetzt – als Autor eindeutig zugeordnet werden, obwohl der Knoten nicht kohäsiv geschlossen zu sein scheint, weil er mit einer bestimmten NP eingeleitet wird. Da jedoch in einem ‚Chomsky-Hypertext‘ ein informierter Leser weiß, daß *Der Autor der Syntactic Structures* und *Noam Chomsky* ein und dieselbe Person sind, kann der Referent eindeutig identifiziert werden.

9.2 Kohärenz

Während einige Autoren nicht zwischen der oben besprochenen Kohäsion und der sogenannten KOHÄRENZ unterscheiden¹, bezeichnen de Beaugrande und Dressler mit KOHÄRENZ ausschließlich den semantisch-kognitiven, inhaltlichen Zusammenhalt eines Textes.

Hierbei spielen Kausalitäts-, Referenz- und Zeitbeziehungen eine Rolle. Obwohl die Kohäsion, wie gesagt, das Gerüst für die Kohärenz darstellt, ist zu erkennen, daß der Leser zumeist versucht, auch dort inhaltlich-kognitive Bezüge herzustellen, wo diese nicht durch kohäsive Mittel markiert sind. Nicht nur die Interpretation eines Textes gilt laut Aussage der Autoren als kognitiver Prozeß, sondern die gesamte, dem Text zugrundeliegende Konstellation, die sogenannte TEXTWELT, wird als kognitive Größe gesehen. Bei dieser Textwelt handelt es sich um die Summe aller KONZEPTE² eines Textes sowie deren Relationen zueinander. Diese Textwelt ist eine virtuelle Einheit,

„[...] die mit der gewöhnlich angenommenen ‚realen Welt‘, d.h. mit der von einer Gesellschaft oder sozialen Gruppe als gültig angesehenen Auffassung der menschlichen Lage, nicht unbedingt übereinstimmen muß [...]“

(de Beaugrande u. Dressler 1981 S. 88)

Ein kohärenter Text zeichnet sich durch SINNKONTINUITÄT seiner Textwelt aus. Oft kann jedoch aufgrund der kohäsiven Mittel keine Sinnkontinuität als Grundlage für die Schaffung von Kohärenz hergestellt werden, beim Verstehen

¹ In einer breiteren Verwendungsweise wird KOHÄRENZ definiert als „textbildender Zusammenhang von Sätzen, der alle Arten satzübergreifender grammatischer [...] und semantischer Beziehungen umfaßt“ (Bußmann 1990 S. 389; vgl. hierzu auch Vater 1992 S. 41f.).

² Die genaue Definition dieser Konzepte wird weiter unten erläutert.

der Textwelt muß der Leser daher zumeist zusätzlich auf bereits vorhandenes Wissen zurückgreifen. De Beaugrande und Dressler merken hierzu an,

„[...] daß die Textwelt *mehr* als den Sinn der Ausdrücke des Oberflächentexts enthält, denn kognitive Prozesse steuern ein gewisses Maß an ALLTAGSWISSEN bei, das von Erwartungen und Erfahrungen der Kommunikationsteilnehmer bezüglich der Organisation von Ereignissen und Situationen abgeleitet wird.“

(de Beaugrande u. Dressler 1981 S. 89)

Was von de Beaugrande und Dressler als „Alltagswissen“ bezeichnet wird, läßt sich in verschiedene Teilphänomene zerlegen: Nur wenn Autor und Leser über tendenziell gleiches SPRACH-, WELT- und (eventuell) FACHWISSEN verfügen, wird der Leser den Text tendenziell so ‚verstehen‘, wie dieser vom Autor ‚gemeint‘ wurde. Auf diese unterschiedlichen Formen von WISSEN soll nun ein genauerer Blick geworfen werden.

Oft sichern, wie gesagt, nicht nur kohäsive Mittel wie Anaphern das Verständnis eines Textes, sondern erst die Verwendung dieser Mittel im entsprechenden Kontext und im Bezug auf das erwartbare Wissen des Lesers:

Lange Zeit bemühte sich die Münchner Mannschaft auswärts. Dann gewannen die 60'er.

Weiß der Leser nicht um die Koreferenz von *Münchner Mannschaft* und *60'er*, kann er diesen Text nicht verstehen.

Lassen sich Koreferenzen nicht durch sprachliches Wissen, also mit Hilfe syntaktischer Mittel und semantischer Relationen aufschlüsseln (wie z.B. *Heinz sammelt Kakteen, die Gewächse hatten es ihm angetan.*) greift der Leser auf Weltwissen (*Er langte auf den Herd und zog blitzartig die Hand zurück*) bzw. Fachwissen (*Der Dozent sprach über Hyperonymie. Die semantische Relation dient dazu ...*) zurück. Oftmals können tendenziell mehrdeutige Sätze bzw. Satzfolgen (*Sie brachte die Katze zum Tierarzt. Whisky mußte eingeschlüpfert werden.*) lediglich über Inferenzen, also Schlußfolgerungen im Hinblick auf Weltwissen aufgelöst werden: *Whisky* wird hier eine Wiederaufnahme von *Katze* und keine von *Sie* sein. Desweiteren ist es oft auch möglich, Zusammenhänge zu erkennen, die lediglich implizit vorhanden sind:

Ostersonntag, nachmittags. Heinzi sucht noch immer sein Nest.

Die beiden Sätze verfügen über keine Wiederaufnahmestrukturen, ihr inhaltlicher Zusammenhang wird lediglich vom Leser durch Weltwissen konstruiert. Hierbei spielt wahrscheinlich die Isotopieebene – bzw. das Schema

– *Ostern* eine entscheidende Rolle, da hier die Konzepte *Ostersonntag* und *Nest* in Beziehung gesetzt werden.

Der Grenzbereich zwischen Kohäsion und Kohärenz sowie deren Inferenzen sind äußerst komplex, exemplarisch soll hierfür lediglich ein knappes Beispiel angeführt werden: Beim Decodieren impliziter Koreferenz darf eine wiederaufnehmende NP-Struktur scheinbar über keine Demonstrativa verfügen, während bestimmte Artikel keine Probleme bereiten:

HTML ist eine Markup-Sprache. Die Befehle / * Diese Befehle sind leicht zu erlernen.

Wird die Koreferenz jedoch mit einem Abstraktum realisiert, sind Demonstrativa oftmals möglich:

Er hat sich HTML beigebracht. Diese Leistung wurde finanziell belohnt.

9.2.1 Kohärenz im Hypertext

Wurde Kohärenz bis jetzt als allgemeines Textualitätskriterium eingeführt, so sollen nun kurz hypertext-spezifische Aspekte der Kohärenz besprochen werden: Bei der Untersuchung von Kohärenz in Hypertexten muß zumindest zwischen einer lokalen und einer globalen Kohärenz unterschieden werden: Auf Knotenebene ist die lokale Kohärenz des Teiltextes zu analysieren. Hier gelten ähnliche Regeln wie bei traditionellen Texten. Auf der Ebene des gesamten Hypertextes wird sich die Frage nach einer globalen Kohärenz stellen. Möglicherweise kann hier auf schemabasierte Ansätze bzw. Isotopieebenen zurückgegriffen werden: Die Vermutung ist, daß der Leser mit einer gewissen Erwartungshaltung an die Lektüre des Hypertextes herangeht und daß es dadurch zur Aktivierung von schemaorganisiertem Wissen kommt. In Abhängigkeit der aktivierten Schemata sind eventuell bestimmte Aktanten bereits im Lektüre-Fokus des Lesers, so daß diese auch zu Knotenbeginn (oder nach Navigation mittels eines Links auf einen anderen Knoten) zugeordnet werden können – auch wenn sie nicht explizit eingeführt bzw. durch mangelnde kohäsive Mittel nicht direkt erschlossen werden können.

9.2.2 Exkurs: Kohärenzzerzeugung durch Schemata

Kohärenz wird sowohl text- als auch lesergeleitet aufgebaut. Zwar interessiert in dieser Arbeit primär die textgeleitete Ausprägung, ein gewisser Einblick in den lesergeleiteten Aspekt des Kohärenzaufbaus und der mentalen Textverarbeitung

als solche ist aber durchaus sinnvoll, um so das später vorgestellte Analysemodell besser in den Gesamtkontext der menschlichen Textverarbeitung einordnen zu können.

Die kognitive Verarbeitung eines Textes wird durch drei Faktoren bestimmt: durch den Text selbst, durch die individuelle Disposition des Rezipienten sowie durch situative Einflüsse. In manchen Textverarbeitungsmodellen treten dabei textuelle Faktoren hinter textexternen zurück (siehe auch Kapitel 8.3):

„A given text does not exist objectively or per se, but has a structure which is unique for each individual and depends on how the individual has understood and memorized the text.“

(Rickheit u. Bock 1983 S. 188)

Folgt man einem solchen Textverständnis, so ist der eigentliche Text

„[...] nur in einem kommunikationsabstrakten Sinn real; empirisch real ist nur das Rezeptionskonstrukt eines bestimmten Rezipienten.“

(Schwarz 1992 S. 159)

Bei der Textrezeption werden wahrscheinlich in ständigen, interaktiven Top-Down- und Bottom-Up-Prozessen die im Text unmittelbar vorliegenden Sinnzusammenhänge mit Einträgen im Gedächtnis abgeglichen, ergänzt oder modifiziert. Ohne ständiges Zurückgreifen auf bereits gespeicherte Erfahrungen könnte der Leser in kürzester Zeit einen völlig neuen Text wohl kaum verstehen. Rein iterative Modelle, die Textverstehen als zyklischen Prozeß sehen, bei dem Propositionen in einem ‚Arbeitsspeicher‘ des Gehirns repräsentiert sind und mit im Text folgenden Propositionen verbunden werden¹, gelten heute als widerlegt.

In der kognitiven Linguistik, sowie in Arbeiten zur KÜNSTLICHEN INTELLIGENZ beschäftigt man sich mit der Frage, wie es möglich ist, diese sehr großen Mengen an gespeicherter Information in Sekundenbruchteilen bei der Textrezeption zu reaktivieren. Als Erklärung dienen sogenannte SCHEMABASIERTE Ansätze. Hier wird davon ausgegangen,

„daß Prozesse des Textverstehens wesentlich darauf basieren, Vorwissen, das in der Form von Schemata existiert, auf die im aktuell zu bearbeitenden Text geschilderten Inhalte abzubilden.“

¹ Nach und nach würde somit eine Kohärenzstruktur im Langzeitgedächtnis des Lesers aufgebaut. Ein solches Modell schlagen Kintsch u. van Dijk 1978 vor. Eine knappe und verständliche Einführung hierzu bietet Schwarz 1992. Kritik zu diesem Modell äußern de Beaugrande u. Dressler 1981. Kintsch und van Dijk sehen in der Koreferenzbeziehung zwischen aufeinander folgenden Propositionen die wichtigste Voraussetzung für den Aufbau einer kohärenten Vorstellung vom Text beim Leser. Hierzu wurde Kritik laut: „Daß Koreferenz weder ausreichend noch notwendig für Textkohärenz ist, ist inzwischen hinreichend belegt worden“ (Schwarz 1992 S. 157).

Dies bedeutet, daß Textverstehen im wesentlichen ein Rekonstruktionsprozeß [bereits gespeicherter Erfahrungen; Ergänzung von O. H.] ist [...]"

(Bußmann 1990 S. 665)

Um sich klarzumachen, was SCHEMATA eigentlich sind, sollte man zuerst einen Blick auf kleinere mentale Organisationsformen werfen, auf die sogenannten KONZEPTE. Konzepte haben die Funktion,

„[...] Wissen über die Welt zu speichern. Als Baustein unseres Kognitionssystems ermöglichen sie die ökonomische Speicherung und Verarbeitung subjektiver Erfahrungseinheiten durch die Einteilung der Informationen in Klassen nach bestimmten Merkmalen.“

(Schwarz 1992 S. 84)

Mit Hilfe dieser Konzepte speichert der Mensch das Wissen über die Welt, organisiert und kategorisiert dieses. Damit können Konzepte als elementare Bausteine der Kognition angesehen werden. Durch Konzepte wird das Erkennen von IDENTITÄT und ÄQUIVALENZ¹ von Objekten im Langzeitgedächtnis verankert. Erst wenn die Kommunikationspartner, also etwa der Autor und seine Leser, in etwa dieselben Konzepte und damit ein ähnliches mentales Modell der Welt haben, können Verstehensprozesse beim Lesen erfolgreich verlaufen.²

Eine Annahme der kognitiven Linguistik ist nun, daß diese Konzepte nicht voneinander getrennt im Langzeitgedächtnis gespeichert, sondern in CLUSTERN organisiert sind, welche menschliche Annahmen über komplexe standardisierte bzw. prototypische Objekte, Handlungen und Situationen repräsentieren. Für komplexe stereotype Objekt- und Situationsinformationen haben sich die Begriffe SCHEMATA, bzw. FRAMES durchgesetzt, bei standardisierten Handlungen spricht man eher von SCRIPTS. Schemata und Scripts können ihrerseits wieder untergeordnete Schemata bzw. Scripts enthalten. Das (in der Literatur gern verwendete) Script für einen Restaurantbesuch könnte sich aus folgenden Teilscripts zusammensetzen:

¹ IDENTITÄT (auch: TYPE-KONZEPT) meint hier, daß ein und dasselbe Objekt der außersprachlichen Wirklichkeit zu verschiedenen Zeitpunkten an verschiedenen Orten sein und dennoch als ein einziges Objekt eindeutig identifiziert werden kann (z.B. Der Linguistikdozent wird im Seminar von den Studenten genauso eindeutig identifiziert wie bei einer zufälligen Begegnung im Hochgebirge.). ÄQUIVALENZ (auch: TOKEN-KONZEPT) bedeutet die eindeutige Zuordnung zweier Objekte zu ein und derselben Klasse (z.B. der VW Golf auf der Straße und ein anderer Golf im Parkhaus können eindeutig der Klasse *Auto* zugeordnet werden).

² Diese mentalen Modelle dürfen wohl nicht starr gesehen, sondern sollten dynamisch interpretiert werden, denn insbesondere unter dem Einfluß der Prototypentheorie wird heute die These vertreten, daß Konzepte und deren Kategorien nicht objektiv und fest sind, sondern durch die Wahrnehmungsfähigkeit und Vorstellungskraft der Menschen bestimmt sind. (vgl. zum *Konzeptbegriff* G. Lakoff 1987 sowie zur *Prototypensemantik* Kleiber 1998)

- Ins Lokal gehen
- Platz wählen und hinsetzen
- Speisekarte ansehen
- Speise wählen und bestellen
- Essen
- Zahlen
- Lokal verlassen

Die Vermutung ist nun, daß in einem Script bestimmte AKTANTEN (VARIABLEN) per ‚default‘ mental aktiviert werden. So ist es z.B. klar, daß in einem Restaurant Ober bedienen, so daß der Satz *Er ging ins Restaurant und bestellt ein Bier.* impliziert, daß ein Kellner zur Aufnahme der Bestellung vorhanden ist. Per default vorhandene Aktanten können daher, wie bereits oben angesprochen, mit bestimmtem Artikel eingeführt werden:

Er setzte sich an den Tisch. Der Kellner kam sofort mit der Karte.

Auch Schemata verfügen wahrscheinlich über solche Variablen: Werden bestimmte Sachverhalte sprachlich nur angedeutet – etwa mit dem Lexem *Auto* –, reicht dieses Stichwort aus, um den gesamten Wissenskomplex, der im Schema *Auto* enthalten ist (*Tanken, Radkappe, Fahrer* etc.) aufzurufen:

„So enthält das Schema etwa das Wissen, aus welchen Teilen ein Auto sich zusammensetzt, welche verschiedenen Typen unterschieden werden können und in welchen Handlungen es benutzt werden kann. Je mehr Erfahrungen mit Autos vorliegen, desto differenzierter wird das Schema ausfallen. Das Auto-Schema kann als allgemeines Modell oder Grundstruktur eines Realitätsbereiches aufgefaßt werden.“

(Ballstaedt u.a.. 1981, S.28).

Im Hinblick auf definite NPs gilt auch hier das bereits oben für Scripts festgestellte:

Ich war auf der Autobahn unterwegs, als der Motor zu stottern anfang.

Durch den Begriff *Autobahn* wird das Schema *Auto* (bzw. das Script *Auto fahren*) aktiviert, welches standardmäßig als Bestandteil die Variable *Motor* bereithält. Daher kann *Motor* auch mittels einer definiten NP eingeführt werden.

Obwohl es durchaus empirische Evidenz für die Schematheorie gibt, ist deren kognitive Plausibilität nicht unumstritten¹. Insgesamt spricht Schwarz 1992

¹ Problematische Aspekte der Schematheorie sind aufgeführt in Schwarz 1992 S.156 f.

jedoch von einem „psychologisch plausiblen Rezeptionsmodell“ (Schwarz 1992 S.156) Obwohl diese Untersuchung nicht explizit Aspekte der Textverarbeitung untersucht, geht sie dennoch implizit von einem theoretischen Modell von Textverstehen und Textproduktion aus, das sich an die Schematheorie anlehnt. Die Grundthesen im einzelnen lauten hierbei:¹

- Textverarbeitung findet sowohl textgeleitet als auch wissensgeleitet statt.
- Textverarbeitung ist ein aktiver, konstruktiver kognitiver Prozeß.
- Neben der textuellen Information und bestehenden Wissensstrukturen des Lesers spielen Interessen und Zielsetzungen desselben eine große Rolle.
- Beim Textverstehen baut der Leser eine kohärente Wissensrepräsentation über die im Text dargestellten Sachverhalte auf.
- Textverstehen inkludiert mehrere Ebenen einer mentalen Repräsentation:

„Textinformationen [werden] mit dem bestehenden Vorwissen vernetzt. Es werden dabei sog. holistische mentale Repräsentationen in Form kognitiver Schemata und mentaler Modelle aufgebaut.“

(Tergan 1993 S. 16)

- Textverstehen wird auch von sogenanntem metakognitiven Wissen beeinflusst, bei dem der Leser über seine eigenen Denkprozesse reflektiert.

Das Analysemodell für Hypertexte (siehe Kapitel 14) ist, wie gesagt, primär als textgeleitet zu verstehen. Integriert wurden lediglich einige wenige Schnittstellen, die Erklärungsmöglichkeiten für autoren- und lesergeleitete Strategien des Kohärenzaufbaus bieten – wie etwa mentale Schemata. Daher soll nun abschließend ein kurzer Blick auf die mögliche Rolle von Schemata beim Produktions- und Rezeptionsprozeß von Hypertexten geworfen werden: Eventuell können Schemata dem Hypertext-Autor beim Anlegen von Links behilflich sein. Dies in zweifacher Hinsicht: Einerseits könnten Teilschemata Aufschluß über sinnvolle Gliederungen eines komplexeren Schemas geben, welches über mehrere Knoten verteilt werden soll. Andererseits hängen Zugriff auf Schemata und Möglichkeiten der Inferenzbildung stark vom individuellen Vorwissen des Lesers ab. Anders gesagt: Die Ausdifferenziertheit eines Schemas korrespondiert mit der Wissenstiefe vom Leser (bzw. Autor) in dieser Domäne:

¹ Die Aufstellung erfolgt in Anlehnung an Tergan 1993 S. 16.

Eine Anweisung ist es zum Beispiel: wenn Sie einer Variablen einen Wert zuweisen, wie oben in Beispiel in der Anweisung \$Zahl = 42;.

(Münz 1997 b [tgca.htm] Abschnitt *Anweisungen notieren*)

Je nach Vorwissen können beim Lesen des Begriffes *Anweisung* mehrere Variablen (im Sinne der Schematheorie) und/oder Teilschemata aktiviert werden. Im obigen Beispiel würde bei einem Programmierer wohl das Teilschema *Variable deklarieren* beim Lesen von *Anweisung* mit aktiviert werden. Die Variable *Variable* müßte also nicht definitorisch eingeführt werden. Programmieranfängern ist aber eventuell nicht klar, daß *Variablen* zumeist Bestandteile von *Anweisungen* sind – oder, mit den Begriffen der Schematheorie ausgedrückt, daß das Schema *Anweisung* über die Variable (bzw., je nach Definitionstiefe der jeweiligen Schemastruktur, über das Teilschema) *Variable* verfügt. Es böte sich an, wie im obigen Beispiel auch geschehen, diese Variablen per Link in einen separaten Knoten auszulagern. Für den Programmierer, der diese Schemastruktur bereits im Langzeitgedächtnis hat, ist dieser Link wohl eher redundant, da er auf das Explizitmachen der Variable *Variable*, wie es im folgenden Beispiel geschieht, wohl verzichten kann:

Variablen sind Speicherbereiche, in denen Sie Daten, die Sie im Laufe Ihrer Programmprozeduren benötigen, speichern können.

(Münz 1997 b [tgcb.htm] Abschnitt *Skalare – einfache Variablen*)

Eventuell können somit auch schematheoretische Aspekte als Argumente für personalisierte Hypertexte gelten und möglicherweise auch Hinweise zur Ausdifferenzierung bzw. Tiefe der Knoten- und Verknüpfungsstruktur personen(gruppen)bezogener Varianten von Hypertexten liefern.

9.3 Intentionalität

Beim nächsten Textualitätskriterium handelt es sich um die INTENTIONALITÄT. Nach den eher textzentrierten Kriterien der Kohäsion und Kohärenz ist die Intentionalität das erste in einer längeren Reihe der leserzentrierten Kriterien.

Das Prinzip der Intentionalität besagt, daß ein Text von seinem Produzenten immer mit der Absicht verfaßt wird, eine bestimmte Wirkung beim Leser zu erzielen. Diese Intention zieht sich durch den gesamten Text und ist Grund dafür, daß der Autor einen „kohäsiven und kohärenten Text bilden will“ (de Beaugrande/Dressler 1981 S. 8).

Vater kritisiert dieses Kriterium in zweifacher Hinsicht. Erstens bemängelt er die Definition von Intentionalität durch zwei andere Textualitätskriterien:

„Zum einen sind Kohäsion und Kohärenz ja – wie es bei allen Textualitäts-Kriterien sein sollte – unabhängige Kriterien, die nicht wiederum Teil eines anderen Kriteriums sein können. Zum anderen kommen sehr wohl Fälle vor, in denen der Textproduzent nicht Kohäsion und/oder Kohärenz intendiert.“

(Vater 1992 S. 51)

Desweiteren weist Vater darauf hin, daß Intentionalität im Sinne von de Beaugrande und Dressler nicht ausschließlich ein Kriterium für Textualität sondern vielmehr eine allgemeine Voraussetzung von Kommunikation jeder Art sei.

Auf Hypertexte – und besonders auf deren Links bezogen –, muß gesagt werden, daß ohne die Intention des Autors, einen kohärenten Text zu erzeugen, die Links wohl auch nicht an solchen Stellen gesetzt werden würden, an denen sie nicht auch einen Beitrag zum Aufbau von Kohärenz liefern können. Dasselbe gilt auch für die prinzipielle Sequenzierung des ‚Rohtextes‘ in eine bestimmte Knoten-Struktur.

Die Frage, ob ein Hypertext eine ganz bestimmte Intention verfolgt, ist nicht leicht zu beantworten. Die meisten Hypertexte sind zweifelsohne mit einer bestimmten Intention verfaßt worden. Schwieriger einzuschätzen sind ‚offene‘ Hypertext-Systeme, die wenig Einschränkungen im Hinblick auf Inhalt und Umfang der enthaltenen Informationen machen. Zu denken wäre hier etwa an freie Literaturportale o.ä. Auch die

"Vision des kulturumspannenden, ganze Literaturen beinhaltenden Hypertextes, wie sie Nelson [...] vertritt, läßt eine Intention, das heißt, eine selektive Aufnahme von Information in den Hypertext, nicht mehr erkennen."

(Hammwöhner 1993, S.25)

Intentionalität liegt im Kleinen auf alle Fälle dann vor, wenn der Hypertext-Autor GUIDED TOURS durch einen Hypertext vorbereitet, die den Leser in einer bestimmten Sequenz durch diesen geleiten.

9.4 Akzeptabilität

Als weiteres leserzentriertes Textualitätskriterium ist die AKZEPTABILITÄT eines Textes zu nennen: AKZEPTABILITÄT bezeichnet die Erwartungshaltung des Lesers, einen kohärenten Text vorzufinden, der für ihn

„[...] nützlich oder relevant ist, z.B. um Wissen zu erwerben [...] Diese Einstellung [bzw. Erwartungshaltung; Anmerkung von O.H.] spricht auf Faktoren an wie Textsorte, sozialen oder kulturellen Kontext und Wünschbarkeit von Zielen. Hier könnten wir die Aufrechterhaltung von Kohäsion und Kohärenz auch als Ziel des Textrezipienten betrachten, insofern er selbst Material beisteuert oder Störungen, wenn erforderlich, überwindet.“

(de Beaugrande u. Dressler 1981 S. 9)

Dieses Kriterium bürdet dem Autor die Aufgabe auf, so etwas wie den ‚prototypischen‘ Leser zu antizipieren: Der Autor sollte also das mögliche Zielpublikum vor Augen haben und Rezeptionsziele, Erwartungen sowie Lesegewohnheiten der Leserschaft beim Verfassen des Textes berücksichtigen.

Vater 1992 führt auch hier mehrere Kritikpunkte an: Seiner Meinung nach, handelt es sich bei Akzeptabilität, ähnlich wie bei Intentionalität, eher um

„[...] eine allgemeine Voraussetzung für erfolgreiches Kommunizieren als ein Kriterium für Textualität. Zudem ist Akzeptabilität in starkem Maße subjektiv.“

(Vater 1992 S. 52)

Vater weist ferner darauf hin, daß Akzeptabilität –, die er selbst eher als ein Bemühen um dieselbe ansieht, wie dies etwa beim Ziehen von Inferenzen seitens des Lesers der Fall ist, – bei der subjektiven Textrezeption durch den Leser eine bedeutende Rolle spielt.

Wie verhält es sich nun mit der Akzeptabilität von Hypertexten? Die Rezeptionsreihenfolge der Knoten ist in der Regel ja im Voraus nicht bekannt, der Autor kann also nicht den prototypischen Leser antizipieren.

Hypertext bietet im Prinzip die Möglichkeit, verschiedene typische Lesergruppen zu berücksichtigen. Die Navigationsfreiheit eines Hypertextes kann diesen für die unterschiedlichsten Leser akzeptabel machen: Im Hypertext SELF des Corpus’ sind etwa fünf vom Autor definierte Lesepfade vorhanden, wobei sich der erste („Wie fange ich an?“) an reine Anfänger, der letzte an angehende Profis („Wie werde ich zum Profi?“) wendet (SELF tp.htm).¹

¹ Dies ist zugegebenermaßen eine triviale Möglichkeit, einen Hypertext für mehrere Lesergruppen akzeptabel zu machen. Ein Buch kann freilich auch in verschiedene Teile für verschiedene Anforderungen gegliedert werden. Technisch wäre es jedoch möglich, einen Hypertext interaktiv wesentlich mehr an den Leser anzupassen: Denkbar ist hier etwa eine Abfrage nach dem Kenntnisstand des Lesers. Nach Eingabe der entsprechenden Daten würden dann automatisch nur die für die Zielgruppe relevanten Knoten bzw. Links eingeblendet.

Hammwöhner 1993 bemerkt, daß sich aufgrund der relativ jungen Tradition von Hypertexten noch keine festen „Präferenzen und Lesegewohnheiten des Rezipienten“ herausgebildet haben (Hammwöhner 1993 S. 25). Aus diesem Grund wird wohl beim Design von Hypertexten auf Metaphern wie Bücher oder Karteikarten zurückgegriffen, um die traditionellen Erwartungen und Gewohnheiten des Lesers in das neue Medium zu übertragen (siehe hierzu auch Kapitel 17.2).

Allgemein äußert sich Hammwöhner zur Akzeptabilität von Hypertext wie folgt:

„Bisher bezieht Hypertext seine Akzeptanz wohl am ehesten aus dem schnellen Zugriff auf relevante Information, von der ausgehend einige wenige Verweise effizient verfolgt werden können, wie das bei technischen Handbüchern der Fall ist. Hier ist die Gefahr des Orientierungsverlusts noch relativ gering. Sollen jedoch längere assoziative Ketten über Links verfolgt werden, so ist Strukturinformation unverzichtbar.“

(Hammwöhner 1993 S. 25)

Laut Hammwöhner könnte also Strukturinformation die Akzeptabilität von Hypertexten erhöhen: Ein solches Anreichern mit Strukturinformation wird im Teil C am Beispiel der typisierten Links untersucht.

9.5 Informativität

Die INFORMATIVITÄT eines Textes – und damit seine Effizienz – ergibt sich einerseits aus der Menge der enthaltenen neuen, dem Leser bisher nicht bekannten Informationen und andererseits aus dem „Ausmaß der Erwartetheit bzw. Unerwartetheit“ (de Beaugrande u. Dressler 1981 S. 10) derselben.

Vater 1992 begreift für die Informativität eines Textes die „Erwartetheit bzw. Unerwartetheit von Zeichen aus einem dem Rezipienten bekannten Zeicheninventar“ als konstituierend (Vater 1992. 56). Meiner Meinung nach ist dies jedoch der Definition von de Beaugrande u. Dressler bereits inhärent.

Inwieweit Hypertexte die Informationswünsche des Lesers befriedigen können, ist abhängig davon, ob die gewünschte Information in der Menge der Knoten vorhanden ist – und ob diese Knoten mit vertretbarem Aufwand erreicht werden können. In vielen Hypertexten wird der Leser hierbei von Volltextsuchmaschinen unterstützt, so etwa auch in diversen Hypertexten des Corpus’.

Aufgrund der nicht-linearen Ausrichtung und der oft immensen Menge an vorhandenen Knoten werden Hypertexte in der Regel selektiv und nicht vollständig rezipiert. Dadurch steigt die Chance, daß ein und dieselbe Hypertext-Basis für den Leser auch bei wiederholter Lektüre mittels eines völlig neuen Lesepfads erneut informativ bleibt.

Unter Umständen wird die Informativität eines Hypertextes auch dadurch erhöht, daß aufgrund der technischen Möglichkeiten häufig und schnell neue Informationen einer vorhanden Menge von Knoten hinzugefügt werden können. Zur Steigerung der Informativität sollte der Leser den Status der jeweiligen Knoten jedoch kennen: Dies kann einerseits allgemein mit Hilfe eines Datumstempels erfolgen oder aber personalisiert, das bedeutet, das Hypertext-System speichert für jeden Leser, welche Knoten von diesem bereits besucht wurden und zeigt dies auch an – etwa durch typisierte Knoten und/oder Links (vgl. auch Hammwöhner 1993, S.26).

9.6 Situationalität

Mit dem nächsten Textualitätskriterium, der SITUATIONALITÄT, bezeichnen die Autoren das Maß an Relevanz, über welches ein Text in der jeweiligen Kommunikationssituation verfügt, d.h., ob er den spezifischen Informationsbedarf in dieser Situation befriedigen kann. Vater bringt hierfür ein plastisches Beispiel:

„Situationalität liegt auch dann vor, wenn Studenten, die eine an einem germanistischen Seminar angebotene Vorlesung über Morphologie besuchen, kaum erwarten, etwas über Knochenbau zu hören, wogegen andererseits in einer Morphologie-Vorlesung für Mediziner kaum die Struktur von Wörtern zur Diskussion stehen dürfte.“

(Vater 1992, S.58).

Im Hinblick auf die Situationalität sieht Hammwöhner 1993 Hypertexte im Vorteil gegenüber traditionellen Texten:

„Während konventionelle Texte aufgrund ihrer starren Form nur eingeschränkt auf die Situation des Lesers und seinen Informationsbedarf eingehen können, liegt hier eine der Stärken von Hypertexten. [...] Durch Kontextualisierung von (Verweis-) Information kann der Benutzer zu der für ihn relevanten Information geführt werden.“

(Hammwöhner 1993 S. 26)

Natürlich bieten auch traditionelle umfangreichere Texte, z.B. Bücher, die Möglichkeit, Teile selektiv und situationsabhängig zu lesen. Hierfür gibt es standardisierte Zugriffshilfen wie Inhaltsverzeichnisse oder Indices.

Personalisierte Hypertexte könnten das situationsabhängige Lesen jedoch auf eine neue Ebene bringen, denn sie

„bieten dem Leser die Möglichkeit, aus der Flut der angebotenen Informationen diejenigen auszublenden, die für ihn erwiesenermaßen irrelevant sind. So wird der Experte aus einem technischen Handbuch sicherlich diejenigen Verweise ausblenden, die zu begrifflichen Erläuterungen führen.“

(Hammwöhner 1993 S. 26)

Um dieses automatische Ausblenden bzw. Einblenden situationsabhängig zu automatisieren, wäre die konsistente Link- bzw. Knoten-Auszeichnung mittels eines kontrollierten Vokabulars von Nöten (dies am besten auf Ebene der semantischen Tiefenstruktur, siehe hierzu auch Kapitel 15).

9.7 Intertextualität

Beim letzten der Textualitätskriterien nach de Beaugrande und Dressler 1981 handelt es sich um die INTERTEXTUALITÄT. Hierdurch wird ein einzelner Text in einen übergreifenden Kontext eingebettet, es werden also Bezüge zu anderen Texten aufgebaut. De Beaugrande und Dressler versuchen mit diesem Kriterium,

„[...] die Abhängigkeiten zwischen Produktion bzw. Rezeption eines gegebenen Textes und dem Wissen der Kommunikationsteilnehmer über andere Texte zu bezeichnen.“

(de Beaugrande/Dressler 1981 S. 188)

Ohne Intertextualität wären bestimmte Texte, wie etwa Parodien, oft nicht zu verstehen, sofern sie Anspielungen enthalten, die über den ursprünglichen Text hinausgehen. Vater führt hierfür etwa die Kombination *scharz-weiß-tot* an, deren parodistischer Gehalt nur durch die Kenntnis von *schwarz-weiß-rot* möglich ist (vgl. Vater 1992 S. 58f.). Manche INTERTEXTUALISMEN werden so oft verwendet, dass sie zu geflügelten Wörtern werden (z.B. *die Gretchenfrage stellen*). Intertextualismen, die ad hoc gebildet werden, können sich direkt auf die Akzeptabilität des Textes auswirken: So kann der Satz *Er wollte doch auch ein Soloalbum schreiben* in einer Buchrezension nur von denen korrekt verstanden werden, die die Anspielung auf das kommerziell erfolgreiche Werk *Soloalbum* des Autors Stuckrad-Barre kennen.

Bezieht man Intertextualität auf Hypertexte, so kann man dieses Kriterium in mehrfacher Weise auslegen:

Interpretiert man etwa die einzelnen Knoten eines Hypertextes als selbständige

Texte, dann kann der Hypertext als ein hochgradig intertextuelles Gebilde angesehen werden. Intertextualität findet dann in den Links auf andere Knoten ihren Ausdruck. Die Bezüge zu anderen Texten (= Knoten) würden so explizit sichtbar gemacht werden. In dieser Arbeit werden Knoten jedoch als – wenn auch zum Teil oft relativ eigenständige – Teiltexthe einer größeren Einheit HYPERTEXT gesehen. Hierfür sprechen neben einer übergeordneten Funktion des gesamten Hypertextes in der Regel auch designtechnische bzw. paratextuelle Mittel, mit denen ein gemeinsames Erscheinungsbild der Knoten eines Hypertextes erreicht wird. Meiner Auffassung nach kann jedoch von Intertextualität dann gesprochen werden, wenn mittels externer Links auf Knoten anderer Hypertexte verwiesen wird. Solche Intertextualismen können durch Links sehr leicht explizit gemacht und somit auch leicht aufgelöst werden:

Erste theoretische Überlegungen führten zum System Memex.

Ist etwa das Lexem *Memex* im obigen Beispiel als Link zum einschlägigen Aufsatz *As we may think* von Vannevar Bush realisiert, so kann der Leser den impliziten Intertextualismus auf das Skriptum von Bush, in dem die fiktive Hypertext-Maschine *Memex* vorgestellt wird, explizit durch Anklicken des Links nachvollziehen. Meiner Meinung nach handelt es sich hierbei tendenziell um etwas anderes als bei Zitaten oder Verweisen auf weiterführende Literatur: Durch den direkten Link baut der Autor eines Hypertextes nicht nur eine gedankliche Brücke zu Teilen von anderen Hypertexten, denn: Er verweist nämlich nicht nur auf andere Texte und benutzt diese etwa als Argumentationshilfen, wie es in wissenschaftlichen Arbeiten durch Zitieren üblich ist, sondern er bietet dem Leser an, den aktuellen Lesepfad in einem anderen Hypertext fortzusetzen – was freilich zum Verlassen des ursprünglichen Hypertextes führt. Diese durch Links explizit gemachten Schnittstellen zu anderen Hypertexten fasse ich in dieser Arbeit als direkte und explizite, hypertext-spezifische Intertextualismen auf.¹ Gestützt wird das durch de Beaugrande, der sein 20 Jahre altes Textualitätskriterium INTERTEXTUALITÄT 2001 im Hinblick auf Hypertexte betrachtet:

¹ Dies bedeutet nicht, daß Hypertexte völlig frei von impliziten Intertextualismen wären. So könnte etwa der Begriff *Memex* aus obigem Beispiel in einem Fachartikel zu Hypertext auch ohne expliziten Link auf den Ursprungstext erscheinen, da der Autor von einem Fachpublikum erwartet, diesen Terminus (und den zugehörigen Aufsatz) zu kennen. Da jedoch gerade im WWW das Instrument der extra-hypertextuellen Links existiert, wird dies in der Regel auch zum Explizieren dieser Zusammenhänge benutzt.

„Still, these older instances of intertextuality [d.h. wie diese in Printmedien vorkommt, O.H.] do not logically implicate hypertextuality. Without some compelling motive, few readers of a scientific research report actually ran to the library to read the cited references. Now that technology offers us relatively effortless and immediate access from text to text, some novel considerations arise. Research reports are starting to replace their references with direct access to the source being referred to; and doing so can significantly enhance the interest and relevance of making references at all. The essentially monological status of the report can yield to a plurality of discourses and voices, and so heighten the solidarity and vitality of the research community.“

(De Beaugrande 2001, Abschnitt: *H. The advent of 'hypertechnology'*)

9.8 Zusammenfassung

Abschließend soll nun eine Wertung der einzelnen Kriterien erfolgen und deren Aufnahme in das Textlinguistische Analysemodell für Hypertexte (TAH) in Kapitel 14 kritisch besprochen werden: KOHÄSION und KOHÄRENZ bilden die Grundlage des grammatischen, semantischen und pragmatischen Zusammenhalts des Textes und ermöglichen erst das Gelingen der kognitiven Verarbeitung. Kohäsion als Grundlage von Kohärenz wird im TAH im Rahmen einer Untersuchung der thematisch-strukturellen Aspekte auf mehreren Ebenen des Hypertextes behandelt (siehe Kapitel 14.3).

Aus den im Abschnitt INTENTIONALITÄT genannten Gründen sehe ich genauso wie Vater 1992 dieses Kriterium als eine eher allgemeine Konversationsmaxime an, die im TAH unter kontextuellen Aspekten berücksichtigt wird.

Meiner Meinung nach handelt es sich bei der AKZEPTABILITÄT eines Textes nicht um ein Textualitätskriterium im engeren Sinne. Da jedoch die Frage nach der Akzeptabilität eines Hypertextes für mehrere potentielle Lesergruppen durch Personalisierung oder eben typisierte Links eine Frage der Hypertext-Forschung ist, sollte dieses Textualitätskriterium in Analysen integriert werden, die primär auf Fragen der Personalisierung von Hypertexten zielen. Im Fallbeispiel des Teils C dieser Arbeit ist dies jedoch nicht der Fall.

Die Frage nach der INFORMATIVITÄT steht in engstem Zusammenhang mit der der AKZEPTABILITÄT: In der Regel steuert das Maß an Informativität auch die Akzeptabilität eines Textes. In relevanten Untersuchung sollten diese beiden Fragen daher möglichst eng zusammen behandelt werden. Das ebenfalls verwandte Kriterium der SITUATIONALITÄT untersucht, angewendet auf Hypertexte, in welchen Situationen welche Lesepfade für welche Lesergruppen relevant (= informativ) und damit auch akzeptabel sind. Mit Hilfe von Aufgaben könnten in empirischen Untersuchungen solche Situationen geschaffen und die

Reaktionen der Leser anhand der Lesepfade und durch Befragung ausgewertet werden. Die beispielhafte Anwendung von TAH im Teil C dient nicht der Beantwortung solcher Fragen.

Die Einwände machen deutlich, daß eine theoretische Untersuchung ohne eine gewisse Anzahl von Testpersonen lediglich Kohäsion, Kohärenz sowie explizit gemachte Intertextualität eines Textes mit relativer Sicherheit bestimmen kann. Ein auch nur in Ansätzen objektives Beurteilungsverfahren der verwenderzentrierten Kriterien von Beaugrande und Dressler dürfte sich als schwierig zu erstellen erweisen.¹ Dies darf bei möglichen Untersuchungen und beim späteren Beurteilen der Ergebnisse nicht aus den Augen verloren werden.

Fazit: Eine Analyse mit TAH sollte also die Kriterien KOHÄSION, KOHÄRENZ sowie INTENTIONALITÄT einbeziehen. Aus oben bereits genannten Gründen fließt auch die Untersuchung der expliziten Intertextualität in das Analysemodell ein. Die Untersuchung der anderen leserzentrierten Kriterien würde ein breiteres Setting verlangen, um herauszufinden, welche Knoten, Links und/oder Lesepfade für welche Personen in welchen Situationen relevant, informativ und akzeptabel sind. Solche Fragestellungen verfolgt bis heute eher die psychologische Hypertext-Forschung. TAH als primär textzentriertes Analysemodell stößt hier an seine Grenzen, kann jedoch als Grundlage für empirische Forschungen dienen, da eine Klärung der textuellen Verhältnisse notwendig ist, bevor weitere Variablen wie *Situation* und *Benutzergruppe* eingeführt werden.

¹ Ähnliches stellt Labarta Postigo 1996 fest, als sie die Brauchbarkeit der Textualitätskriterien für ihre Analyse von Softwaredokumentationen prüft. Sie sieht Schwierigkeiten, „weil wir die Kompetenzen der Leser nicht kennen“ und weil sich sämtliche Kriterien – laut Labarta Postigo mit Ausnahme von Kohäsion und Kohärenz – „nicht am Text allein“ festmachen lassen (Labarta Postigo 1996 S. 57).

10 Dimensionen einer textlinguistischen Analyse

Nachdem in den letzten Kapiteln der kommunikativ-funktionale Textbegriff dieser Arbeit sowie die wichtigsten Kriterien der Textualität sprachlicher Gebilde vorgestellt worden sind, sollen nun die textuellen Dimensionen eingeführt werden, die bei der praktischen Analyse eine Rolle spielen.

10.1 Allgemein

Nachdem bereits die allgemeine Arbeitsdefinition von TEXT aus Brinker 1997 bezogen wurde, wird ebenso dessen Methodik zur Bestimmung von TEXTSTRUKTUR, TEXTFUNKTION und TEXTSORTE übernommen. Brinkers Stärke liegt darin, daß er mittels eines „konsistenten textlinguistischen Beschreibungsapparat[s] auf integrativer Basis“ (Labarta Postigo 1996) die Textstruktur in Abhängigkeit von der Textfunktion sehr gut zu erklären vermag. Auch überzeugt seine Vorgehensweise, wenn er von den nichtsprachlichen KONTEXTUELLEN Faktoren eines Textes ausgehend, sich über KOMMUNIKATIVE FUNKTIONEN zu sprachlichen STRUKTUREN ‚hinunter arbeitet‘. Aus den kontextuellen, kommunikativen und strukturellen Merkmalen leitet er wiederum die Textsorte ab. Diese einzelnen Dimensionen einer textuellen Analyse werden im folgenden vorgestellt.

10.2 Kontextuelle Bedingungen

Spricht man vom KONTEXT eines Textes, so müssen hier zwei Dimensionen unterschieden werden: Die erste meint die Einbettung eines Textes in den Kontext anderer Texte. Diese Dimension wird von de Beaugrande und Dressler als INTERTEXTUALITÄT bezeichnet und gilt den Autoren als eines der sieben textkonstituierenden Kriterien. Dieser textuelle Kontext wurde in Kapitel 9.7 bereits näher behandelt. Texte sind jedoch auch stets in bestimmte situative und gesellschaftliche Kontexte eingebettet. Diese zweite kontextuelle Dimension erstreckt sich auf die KOMMUNIKATIONSSITUATION, die sich wiederum durch die KOMMUNIKATIONSFORM sowie den HANDLUNGSBEREICH eines Textes beschreiben läßt.

10.2.1 Kommunikationsform

Zur medialen Gebundenheit der KOMMUNIKATIONSFORM bemerkt Brinker 2001:

„Die Kommunikationssituation wird entscheidend durch das Medium bestimmt, das zur Übermittlung von Texten eingesetzt wird. Wir können im wesentlichen fünf Medien unterscheiden: Face-to-face-Kommunikation, Telefon, Rundfunk, Fernsehen und Schrift.“¹

(Brinker 2001 S. 138)

Die Hypertexte des Corpus' sind medial an das WWW gebunden, welches – zumindest in den westlichen Ländern – mittlerweile als Massenmedium angesehen werden kann. Im Gegensatz zu anderen Massenmedien kann im WWW jeder Privatmann mit vergleichsweise geringem technischen und finanziellen Aufwand zum Anbieter von Information werden. Begreift man Zeitungen, Rundfunk und Fernsehen als ASYNCHRONE MEDIEN (kleine Anzahl von Produzenten vs. große Anzahl von Rezipienten), so ist das WWW ein tendenziell SYNCHRONES MEDIUM – jeder Leser kann mit geringem Aufwand zum Publizisten werden.

Die Form der Informationsdarbietung ist primär eine schriftliche. Die Anzahl multimedialer Elemente wächst jedoch beständig, man kann beim WWW in zunehmendem Maße von einem Schmelztiegel unterschiedlicher Medien sprechen. Dies um so mehr, als bereits zahlreiche Radiosender online erreichbar und die ersten Fernsehsender im WWW auf Sendung sind. Elektronische Post und Telefonie ergänzen den Medienmix. Im Fokus dieser Arbeit stehen jedoch primär textuell realisierte Hypertexte.

Die Kommunikationsform des WWW ist die WEBSITE, oder kurz SITE. Sie beherbergt in der Regel einen Hypertext. Die Summe aller Websites im WWW bildet das derzeit größte Hypertext-Netz der Welt. Auf die Unterschiede von Hypertexten und Hypertext-Netzen (und E-Texten) wurde bereits in Kapitel 5 eingegangen. Die unterschiedlichen Situationen der Kommunikation bedingen deren Ausformungen hinsichtlich der Kommunikationsrichtung, des (vorhandenen oder nicht vorhandenen) direkten Kontakts der Kommunikationspartner sowie hinsichtlich der Sprache (gesprochen vs. schriftlich).

Die Kommunikationsform von WWW-Hypertexten ähnelt sehr der von Zeitungsartikeln und Büchern: Die Richtung der Kommunikation ist mono-

¹ Die Kommunikationssituation *Computer* bzw. *WWW* wird hier nicht aufgeführt.

logisch, der Kontakt zwischen den Kommunizierenden in der Regel zeitlich und örtlich begrenzt. Die Kommunikationsform ist an die geschriebene Sprache gebunden.¹

Die Kommunikationsform bietet einen Rahmen, innerhalb dessen sich verschiedene Textfunktionen entfalten können. Sie darf nicht mit der Textsorte verwechselt werden:

„Für die Kommunikationsformen ist charakteristisch, daß sie – im Unterschied zu den Textsorten – allein durch situative bzw. mediale Merkmale definiert, in kommunikativ-funktionaler Hinsicht also nicht festgelegt sind.“

(Brinker 2001 S. 139)

Die Kommunikationsform beeinflußt in starker Weise die Struktur eines Textes. Brinker verweist hier etwa auf die unterschiedlichen Textstrukturen von Anzeigen-, Rundfunk- und Plakatwerbung, hervorgerufen durch differierende mediale Realisierung.

10.2.2 Handlungsbereich

Der HANDLUNGSBEREICH ist neben der Kommunikationsform als zweiter Teil der Kommunikationssituation zu interpretieren:

„Die Kommunikationssituationen, die den Rahmen für Textsorten bilden, sind bestimmten gesellschaftlichen Bereichen zugeordnet, für die jeweils spezifische Handlungs- und Bewertungsnormen gelten.“

(Brinker 2001 S. 139)

Laut Brinker existieren zwei Möglichkeiten, um zu einer Klassifikation der Handlungsbereiche zu gelangen: Man unterteilt hinsichtlich inhaltlicher Kriterien (etwa Handlungsbereiche des Rechts, der Musik, der Religion etc.) oder bzgl. des Rollenverhältnisses der Kommunikationspartner.

Brinker verwirft die erste Möglichkeit mit dem Hinweis auf den nicht ausreichenden Stand der Forschung. Daher sei es nicht möglich, inhaltlich definierte Handlungsbereiche nach Regeln und sozialen Normen, die innerhalb der Bereiche Geltung haben, zu gliedern.

¹ HYPERMEDIA hingegen offerieren zusätzliche Möglichkeiten: Durch Audiointegration kann geschriebene Sprache durch gesprochene ergänzt bzw. ersetzt werden.

Für interaktive Hypertexte bzw. Hypermedia ist darüber hinaus eine scheinbar dialogische Kommunikationsrichtung charakteristisch. Der Dialog entwickelt sich freilich (anders als etwa beim Telefongespräch) innerhalb der engen Grenzen, die der Autor im Hinblick auf den aktuellen Kontext gesetzt hat: nämlich entlang der von ihm gesetzten Hypertext-Links.

Brinker unterscheidet daher im Bezug auf die Rolle, die die Teilnehmer einer Kommunikation einnehmen, und zwar in einen PRIVATEN, einen OFFIZIELLEN sowie einen ÖFFENTLICHEN HANDLUNGSBEREICH. WWW-Hypertexte können als Medien der Massenkommunikation gesehen und damit dem öffentlichen Handlungsbereich zugeordnet werden.¹

10.3 Textfunktion

Nach Interpretation der kontextuellen Faktoren begibt sich die Analyse mit der Untersuchung der TEXTFUNKTION auf die textinterne Ebene im engeren Sinne.

Brinker nimmt den Begriff der SPRACHLICHEN HANDLUNG als theoretische Grundlage zur Bestimmung der Textfunktion. Grundgedanke ist hierbei die Ansicht der kommunikationsorientierten Textlinguistik, daß nicht einzelne Sätze, sondern SPRECHHANDLUNGEN (SPRECHAKTE), die durch ihre Äußerung vollzogen werden, Grundlage der Kommunikation seien (vgl. auch Austin 1962 und Searle 1969). Sprechakte bestehen aus mehreren, simultan vollzogenen TEIL-SPRECHAKTEN (LOKUTION, PROPOSITION, ILLOKUTION, evtl. PERLOKUTION).

Für die Textfunktion ist dabei der Teilakt der Illokution von besonderer Bedeutung. Die Illokution gibt Hinweise, „wie die Proposition aufzufassen ist“, bezeichnet also die „kommunikative Funktion“ der Sprechhandlung (Bußmann 1990 S. 726). Diese Illokution wird oft mittels sprachlicher bzw. kontextueller ILLOKUTIONSINDIKATOREN signalisiert und kann durch diese in einer Analyse ermittelt werden. Die direkteste Art der Signalisierung leisten performative Verben, durch deren Verwendung direkt Handlungen vollzogen werden, so etwa in der Mahnung eines Anwalts:

„Hiermit fordere ich Sie zum unwiderruflich letzten Mal auf, Ihre Rechnung vom 29.2. dieses Jahres zu bezahlen.“

(Bußmann 1990 S. 324)

¹ Prinzipiell sind Hypertexte aber bezüglich des Handlungsbereichs relativ offen. So können beispielsweise Kunden einer Online-Agentur durch einen hypertextuellen Vertrag navigieren und eine Bestellung tätigen bzw. ein Geschäft abschließen (=offizieller Handlungsbereich). Ebenso kann auf einem hypertext-basierten Intranet-Server ein bestimmter Hypertext nur sehr wenigen Personen per Paßwort zur Verfügung gestellt werden (= ‚privater‘ Bereich).

In Analogie zur Illokution, die den „Handlungscharakter einer Äußerung festlegt“ bestimmt laut Brinker die Textfunktion den „Kommunikationsmodus des Textes“. Daher kommt er zu folgender Definition:

„Der Terminus *Textfunktion* bezeichnet die im Text mit bestimmten, konventionell geltenden, d.h. in der Kommunikationsgemeinschaft verbindlich festgelegten Mitteln ausgedrückte Kommunikationsabsicht des Emittenten. Es handelt sich also um die Absicht des Emittenten, die der Rezipient erkennen soll, sozusagen um die Anweisung (Instruktion) des Emittenten an den Rezipienten, als was dieser den Text insgesamt auffassen soll, z. B. als informativen oder als appellativen Text.“

(Brinker 2001 S. 95)

Fraglich ist jedoch, ob der Autor tatsächlich stets den Wunsch hat, daß der Leser die Funktion des Textes erkennen möge. Im Rahmen des Fallbeispiels in Teil C wird sich bei bestimmten Hypertext-Knoten des Corpustextes SPORT1 zeigen, daß der Autor dem Leser eine informative Funktion ‚vorspielt‘, während der Knoten jedoch eine primär appellative Funktion aufweist.

10.4 Textstruktur

Die Struktur eines Textes ist stets in enger Abhängigkeit von der Funktion zu sehen und zu analysieren. TEXTSTRUKTUR faßt Brinker auf

„[...] als Gefüge von Relationen, die zwischen den Sätzen bzw. den Propositionen als den unmittelbaren Strukturelementen des Textes bestehen und die den inneren Zusammenhang, die Kohärenz des Textes bewirken.“

(Brinker 2001 S. 21)

Vom sprachsystematischen Ansatz ausgehend wird bei der Analyse der Struktur der einzelne Satz als textuelle Grundeinheit verstanden.¹ Darauf aufbauend soll auf grammatischer- und thematischer-struktureller Ebene die Bedingtheit der Kohärenz ermittelt werden. Als für Kohärenz und Textverstehen entscheidende Bedingung sieht Brinker die THEMATISCHE ORIENTIERUNG eines Textes an, die sich zumeist durch verschiedene Formen grammatischer bzw. semantischer Wiederaufnahmestrukturen ausdrückt. Eine Analyse dieser Wiederaufnahmestrukturen soll letztendlich die Entfaltung des Themas aufzeigen.

¹ Ausführlich geht Brinker auf die Schwierigkeiten einer Satzdefinition und die damit verbundenen Probleme bei der Verwendung dieses Terminus' innerhalb einer Textanalyse ein. Eine Präzisierung versucht Brinker durch das Unterscheiden der Begriffe SATZ (syntaktische Struktureinheit) und PROPOSITION (semantische Struktureinheit) zu erreichen. Wichtig ist festzuhalten, daß zwischen diesen Teilgrößen keine 1:1-Entsprechung besteht. Insgesamt sieht Brinker dieses Thema, das er als „mikroanalytische Aspekte der Textanalyse“ bezeichnet, für die Textanalyse als solche in einer eher „untergeordnete[n] Rolle“ (Brinker 1997 S. 22 f.). Dieser Aussage ist nur zuzustimmen.

Das Thema ist oftmals durch Überschriften von Texten und Teiltextrn direkt signalisiert. Ist dies nicht der Fall, ist in unübersichtlichen Fällen eine Bottom-Up-Untersuchung der THEMATISCHEN ENTFALTUNG nötig. Zur Analyse dieser thematischen Bedingungen der Textstruktur zieht Brinker das THEMA-RHEMA-KONZEPT¹ der Prager Schule sowie das MAKRO- und SUPERSTRUKTURKONZEPT (siehe Kapitel 12) heran. In Kapitel 13 wird für diese Zwecke zusätzlich das MODELL DER REFERENTIELLEN BEWEGUNG eingeführt.

10.5 Textsorte

In einem diachronen Prozeß, auf den hier nicht näher eingegangen wird, bilden sich aus funktionalen und strukturellen Gebrauchsmustern mit der Zeit TEXTSORTEN aus. Ein solcher Prozeß ist momentan im jungen Medium WWW zu beobachten, in dem sich langsam Textsorten wie etwa Werbebanner oder interaktive Formulare ausformen.

Sowohl die sprachsystematisch als auch die kommunikations- bzw. handlungstheoretisch ausgerichtete Textlinguistik hat versucht, eine Texttypologie anhand voneinander abzugrenzender TEXTSORTEN bzw. Klassen von Textsorten zu erstellen. Brinker bezeichnet den handlungstheoretischen Ansatz als „erfolgsversprechender“ und gibt eine Definition, die auf diesem Standpunkt fußt:

„Textsorten sind konventionell geltende Muster für komplexe sprachliche Handlungen und lassen sich als jeweils typische Verbindungen von kontextuellen (situativen), kommunikativ-funktionalen und strukturellen (grammatischen und thematischen) Merkmalen beschreiben.[...] sie besitzen zwar eine normierende Wirkung, erleichtern aber zugleich den kommunikativen Umgang, indem sie den Kommunizierenden mehr oder weniger feste Orientierungen für die Produktion und Rezeption von Texten geben.“

(Brinker 2001 S. 135)

Basiskriterium für eine Differenzierung der Textsorten ist laut Brinker die Textfunktion. Auf diese Weise unterscheidet er fünf verschiedene Klassen:

- (1) Informationstexte
- (2) Appelltexte
- (3) Obligationstexte

¹ „Nach dieser Konzeption läßt sich ein Satz von seinem *Mittelungswert* her gesehen in zwei Teile gliedern, in das *Thema* als dem *Ausgangspunkt der Aussage* und das *Rhema* als den *Kern der Aussage*.“ (Brinker 1997 S. 48 [Auszeichnung im Original]) Das THEMA ist die bekannte oder irgendwie erschließbare Information, über die mittels des RHEMAS etwas Neues mitgeteilt wird (Zur Schwierigkeit, diese Begriffe exakt zu definieren siehe auch Bußmann 1990 S. 784 f.).

(4) Kontakttexte

(5) Deklarationstexte

Innerhalb dieser Klassen sind nun weitere Unterscheidungen zu treffen, um die in den einzelnen Textklassen enthaltenen Textsorten auszudifferenzieren. Unterscheidungskriterien sind hier laut Brinker vor allem Kommunikationsform, Handlungsbereich, Art des Themas, Form der thematischen Entfaltung sowie Beschreibung der sprachlichen und nichtsprachlichen Gestaltungsmittel.

Durch Erfahrung im Umgang mit Texten ist es möglich, „Prototypen von Textsorten“ zu schaffen, woraus ein intuitives „textsortenspezifische[s] Textmusterwissen“ resultiert, welches wiederum einen „routinierten Alltags-Umgang mit bestimmten Textsorten“ ermöglicht (Linke u.a. 1991 S. 253). Dieses TEXTMUSTERWISSEN unterstützt insofern die Textverarbeitung, als daß man bei einem als Vertreter einer Textsorte erkannten Text (z.B. einem Zeitungsartikel) von bestimmten Voraussetzungen ausgehen kann, welche die Textverarbeitung erleichtern.¹ So gesehen unterstützt Textmusterwissen die Verarbeitung, denn:

„Einzelne Textelemente müssen nicht in jedem Einzelfall aus sich heraus interpretiert und mit den umgebenden Textelementen in Zusammenhang gebracht werden, sondern können vor dem Hintergrund gegebener Muster als Repräsentanten erwartbarer Einheiten ausgedeutet und somit auch rasch kohärent verstanden werden.“

(Linke u.a. 1991 S. 254)

Allgemein muß jedoch mit Brinker 2001 konstatiert werden:

„Bisher existieren lediglich eine Reihe von z.T. recht verschiedenartigen Ansätzen zum Problem der Textsorten. Von der Aufstellung einer geschlossenen und in sich stimmigen Texttypologie ist die Textlinguistik noch weit entfernt.“

(Brinker 2001 S. 130)

Im Rahmen der vorliegenden Untersuchung nimmt das Problem der Textsorten jedoch allgemein eine eher untergeordnete Rolle ein. Die sich dennoch stellende grundlegende Frage ist aber, ob HYPertext als eigene Textsorte zu sehen ist.

Hypertexte unterscheiden sich von traditionellen Texten zumeist bezüglich des medialen Kontextes sowie des Einsatzes nichtsprachlicher Gestaltungsmittel.

¹ Etwa daß bei einem Zeitungsartikel die grundlegende Information in der Regel in den ersten Sätzen zu finden ist. Dies gilt zumindest für seriöse Tageszeitungen. In Boulevardblättern findet sich immer mehr ‚Infotainment‘: Dort ist – um Spannung aufzubauen – der Kern der Nachricht am Ende des Artikels zu finden.

Letzteres meint vor allem das Setzen expliziter Verweise sowie die Integration bestimmter multimedialer Elemente, die im Printmedium so nicht realisierbar wären.

Der Einsatz kohäsiver sprachlicher Mittel unterliegt aufgrund der modularen Struktur gewissen Einschränkungen, die in Kapitel 9.1 angesprochen wurden. Bei der Themenwahl ist der Hypertext-Autor ebensowenig eingeschränkt wie der Produzent eines ‚normalen‘ Textes. Gibt er die konkrete Navigationssequenz allerdings nicht durch eine GUIDED TOUR streng vor, ist sein Einfluß auf die Entfaltung des Themas sehr begrenzt. Sie hängt dann vom jeweiligen, frei wählbaren Lesepfad ab.

Im WWW sind mit großer Häufigkeit Informations- (Nachrichten, Berichte, Dokumentationen, Sachbücher etc.) und Appellationstexte (Werbung) zu finden, gelegentlich auch Texte mit Obligationsfunktion (Verträge). Während Kontakttexte wie Kondolenzen oder Danksagungen den Weg in das neue Massenmedium seltener finden, sind Deklarationstexte (Testament, Ernennungs-urkunde) nahezu gar nicht vertreten.

Das WWW bietet also scheinbar Texten mit unterschiedlichsten Funktionen eine Heimat. Gilt, wie Brinker vorschlägt, die Textfunktion als Hauptcharakteristikum für Textsorten, so kommt man zu dem Schluß, daß es sich beim Corpus um eine Sammlung hypertextueller Realisierungen von bereits in traditionellen medialen Kontexten ausgeprägten Textsorten handelt. Hypertext ist in diesem Sinne nicht als Textsorte zu verstehen!

Eine elektronische Hypertext-Realisierung muß vielmehr als medialer Kontext gesehen werden, dessen Beschränkungen bzw. Möglichkeiten den Rahmen vorgeben, in dem sich ein bestimmtes Textthema abhängig von der jeweiligen Textfunktion zum Repräsentanten einer bestimmten Textsorte entfalten kann.

11 Paratextuelle Elemente

Im letzten Kapitel wurden die kontextuellen und textinternen Dimensionen einer textlinguistischen Analyse vorgestellt. In diesem Kapitel werden die ‚Grenzbereiche‘ von traditionellen Texten und Hypertexten definitorisch im Rahmen einer Beschreibung des sogenannten PARATEXTES eingeführt.

11.1 Traditionelle paratextuelle Elemente von Texten

Texte treten in mehrfacher Hinsicht nicht isoliert auf. Einerseits ist der einzelne Text stets eingebunden zu sehen in eine offene oder geschlossene Sequenz von kommunikativen Äußerungen, die sich jeweils in Texten manifestieren. Diese Sequenz wird häufig als DISKURS bezeichnet. Die Gesamtheit aller Diskurse, die für eine bestimmte Anzahl von Menschen oder auch für die Gesellschaft insgesamt, im Rahmen einer gewissen Thematik oder für eine wissenschaftliche Disziplin relevant ist, wird entsprechend als DISKURSUMIVERSUM bezeichnet.

Andererseits ist ein Text zumeist eingebettet in ein größeres Gebilde, bestehend aus dem eigentlichen Text und Erscheinungen der Peripherie – wie Überschriften, Vorworte, Danksagungen, Marginalien, Anmerkungen, Fußnoten etc. Diese PARATEXTUELLEN Elemente erleichtern die Orientierung in gedruckten Texten und ermöglichen eine schnellere Aufnahme der präsentierten Information. Das Verhältnis zwischen Diskursuniversum und paratextuellen Elementen beschreibt Genette 1989 wie folgt:

„Der Paratext ist also jenes Beiwerk, durch das ein Text zum Buch wird und als solches vor die Leser und, allgemeiner, vor die Öffentlichkeit tritt. Dabei handelt es sich weniger um eine Schranke oder eine undurchlässige Grenze als um eine *Schwelle* oder [...] um ein ‚Vestibül‘, das jedem die Möglichkeit zum Eintreten oder Umkehren bietet; um eine ‚unbestimmte Zone‘ zwischen innen und außen, die selbst wieder keine feste Grenze nach innen (zum Text) und nach außen (dem Diskurs der Welt über den Text) aufweist [...]“

(Genette 1989 S. 10)

Flankiert werden solche Metainformationen durch ein vom Autor bzw. Verleger bewußt gewähltes Layout des Textes. Das Layout betrifft sowohl Papierform, -größe und -art als auch typographische Merkmale des Textes wie Schriftfamilie, -größe oder Zeilenabstand:

„The linguistic means to realise non-linear structures in texts are usually accompanied and supported by the page layout when it comes to the presentation of the text. We take a broad view of page layout: besides the placement of text blocks and pictures on the page, we are concerned with typographical issues such as the choice of type, spacing, leading etc. All these layout resources work hand in hand with linguistical resources to provide meta-information

about the text, which fit together to a structural overview (see Norrish 1987). Such an overview has three main functions. It not only allows for the communication of non-linear discourses but also serves as an access tool for its non-linear, selective consumption and creates expectations concerning the content of the document on the part of the reader, thus influencing the decision to read the document at all, for example.“

(Reichenberger u.a. 1995)

Neben den oben aufgeführten paratextuellen Elementen im engeren Sinne sowie dem eigentlichen textuellen Layout führt Genette 1989 noch weitere Größen in der Peripherie des Textes an – etwa das gewählte Format, die Reihe, in der ein Buch erscheint, die Aufmachung des Umschlages und der Titelseite etc. Diese Größen faßt er unter dem Begriff VERLEGERISCHER PERITEXT zusammen, welcher mit den oben aufgeführten paratextuellen Elementen im engeren Sinne den GESAMT-PERITEXT eines Textes bildet.

Desweiteren erkennt Genette neben dem Diskursuniversum, in das ein Text eingebettet ist, Elemente, die nicht mehr dem Paratext und noch nicht dem Diskursuniversum angehören. Diese Elemente bezeichnet er mit dem Begriff EPITEXT. Der Epitext wiederum gliedert sich in einen öffentlichen Bereich (Beiträge auf Kolloquien, spätere Kommentare des Autors etc.) sowie einen privaten, etwa Briefwechsel, Mündliche Mitteilungen des Autors etc. PARATEXT nach der Definition von Genette 1989 läßt sich also auf folgende Formel bringen:

„Paratext = Peritext + Epitext“

(Genette 1989 S. 13)

Der private Epitext ist für diese Arbeit nicht von Interesse. Der öffentliche hingegen umfaßt Phänomene, die bereits unter dem Terminus INTERTEXTUALITÄT subsumiert wurden. Der Epitext ist also eher als Größe der Diskurswelt denn als Bestand des Paratextes anzusehen. Aus diesem Grund wird die Untersuchung von dem, was Genette als EPITEXT bezeichnet, im Rahmen der Analyse der Intertextualität erfolgen.¹

11.2 Paratextuelle Elemente von Hypertexten

Das Layout eines Hypertextes wird mit dem Begriff SCREENDSIGN beschrieben: Screendesign umfaßt hierbei das konventionelle Textdesign aber auch Fragen nach Gestalt und Anordnung von eventuell vorhandenen interaktiven

¹ EPITEXT ist freilich enger gefaßt als INTERTEXTUALISMUS, da letzterer Begriff nur intertextuelle Phänomene beschreibt, die als Reaktion auf den eigentlichen Text entstanden sind.

Elementen, SCREENDESIGN kann aber auch bedeuten, daß die Erfassung des Hypertextes durch eine METAPHER erleichtert wird: So könnte etwa der Hypertext als Schreibtisch visualisiert werden, an dem per Mausklick verschiedene Aktionen ausgeführt werden können (etwa Aktenordner öffnen o.ä.). Neben diesen designspezifischen Elementen können bei Hypertexten auch NAVIGATIONS-KOMPONENTEN (siehe Kapitel 6.4) zum Paratext gezählt werden.

Der Paratext eines Hypertextes setzt sich also aus folgenden Größen zusammen:¹

- Navigationselemente des Browsers (= anwenderspezifischer Peritext): Die dem Anwender zur Verfügung stehenden Navigationselemente hängen von der verwendeten Software ab und können in Anzahl und Ausprägung stark variieren. Es kann jedoch im Rahmen dieser Arbeit nicht auf jede mögliche spezifische technische Ausstattung Rücksicht genommen werden, als Standard soll hier die in Kapitel 6.4.8 vorgestellte Browser-Software gelten.
- Screendesign (= verlegerischer Peritext): Hier sind die oben beschriebenen allgemeinen Aspekte des Layouts, globale Interaktionen sowie eventuelle hypertext-weite Metaphern gemeint.
- Navigationselemente des Hypertextes (= vom Autor bestimmter Peritext): Die vom Hypertext-Autor in die HTML-Seiten integrierten Navigationselemente müssen bei der Beschreibung des Gesamt-Hypertextes berücksichtigt werden.
- Gliederung (Peritext im engeren Sinn): Begreift man den gesamten Hypertext als einen Text, so spiegelt die Knoten-Struktur die vom Autor gewünschte inhaltliche Gliederung dieses Textes zum Teil wider. Die Knotenüberschriften können somit mit Teilüberschriften traditioneller Texte verglichen werden. Diese Analogie ist freilich nur zum Teil zutreffend, da der Autor eines Hypertextes ja nur eine eingeschränkte Kontrolle über die Rezeptionssequenz seitens des Lesers hat.
- Verlinkte Hypertexte (Beiträge, Diskussionen etc.) über den aktuellen Hypertext (= öffentlicher Epitext): Der verlinkte Peritext – in dieser weiten Definition – gehört meiner Meinung nach nicht mehr zum Text sondern bereits zum Diskursuniversum. Jedoch hat der Autor durch explizite Links deutlich gemacht, daß er den verlinkten externen Text im Bezug auf den

¹ In Klammern stehen die von Genette 1989 eingeführten Termini, die in etwa auf die hier besprochenen hypertextuellen Verhältnisse passen.

aktuellen Hypertext als sehr wichtig betrachtet. Bei den unten erwähnten, nicht verlinkten Beiträgen hat er dies (bewußt oder aus Unkenntnis über die Existenz dieser Beiträge) nicht getan. Links auf verlinkte peritextuelle Erscheinungen gehören zum Aufgabenbereich der vorliegenden Arbeit und werden im Rahmen der INTERTEXTUALITÄT untersucht.

- Nicht verlinkte Hypertexte (Beiträge, Diskussionen etc.) über den aktuellen Hypertext (=ebenfalls öffentlicher Epitext): Diese intertextuellen Erscheinungen sind ebenfalls im Diskursuniversum des Hypertextes angesiedelt. Da sie nicht mit Hilfe von Links an den aktuellen Hypertext gekoppelt sind, fallen sie nicht in den Aufgabenbereich dieser Arbeit.

Der eigentlichen Textanalyse vorgelagert sind also Untersuchungen der Navigationsmöglichkeiten des Browsers, des Designs, der hypertext-internen Navigationsmöglichkeiten sowie der Grobgliederung des Gesamt-Hypertextes. Andere Aspekte des Peritextes im engeren Sinn, wie Fußnoten, Anmerkungen, Querverweise etc. sind in Hypertexten mit internen Links realisiert und fallen somit in den Kernbereich der Untersuchung, d.h., sie werden auf der Ebene *Link* des textlinguistischen Analysemodells für Hypertexte bearbeitet (siehe Kapitel 14.3).

12 Makrostruktur und Makrostrukturanalyse

In unübersichtlichen Fällen bringt eine Top-Down-Analyse der thematisch-strukturellen bzw. funktionalen Verhältnisse eines Textes eventuell nicht die gewünschten Ergebnisse. Dann könnte ergänzend eine Bottom-Up-Analyse für ‚kritische‘ Teiltexthe bzw. Knoten eingesetzt werden. Dieses und das nächste Kapitel stellen daher formale Werkzeuge vor, mit deren Hilfe unklare thematische und funktionale Verhältnisse in komplexeren Texten geklärt werden können. Im Rahmen des Textlinguistischen Analysemodells für Hypertexte (siehe Kapitel 14) bieten sich diese Werkzeuge als Hilfsmittel zur Analyse thematisch-struktureller Aspekte an. Ein erstes mögliches Instrumentarium für eine solche Untersuchung ‚von unten‘ ist die MAKROSTRUKTURANALYSE.

12.1 Allgemeines: Textthema und Makrostruktur

Bevor die eigentliche MAKROSTRUKTURANALYSE eingeführt wird, soll noch kurz das grundsätzliche Verhältnis von MAKROSTRUKTUR und TEXTTHEMA angesprochen werden. Textthema wird bei Brinker 2001 definiert als

„Kern des Textinhalts, wobei der Terminus *Textinhalt* den auf einen oder mehrere Gegenstände [...] bezogenen Gedankengang eines Textes bezeichnet.“

(Brinker 2001 S. 55)

Manchmal ist das Textthema explizit in einem Teil des Textes, etwa in Überschrift oder Schlagzeile, angegeben. Ist dies nicht der Fall, so muß es aus dem TEXTINHALT abstrahiert werden. Eine Möglichkeit, das Thema aus dem Textkörper abzuleiten, ist die von van Dijk eingeführte Makrostrukturanalyse. Die Methode ist nicht unumstritten.¹ Brinker faßt die Kritik wie folgt zusammen:

„Die Kritik betrifft einmal Form und Ableitung der Makrostruktur selbst, zum anderen das Problem, wie aus der semantischen Tiefenstruktur durch textuelle Operationen (Transformationen) die Oberflächenstruktur der Texte generiert werden kann, schließlich die Frage, wie die Anwendung der Makroregeln im einzelnen zu erfolgen hat, um zur Makrostruktur des betreffenden Textes zu gelangen; umstritten ist auch van Dijks Postulat von der kognitiven Relevanz seines Konzepts.“

(Brinker 2001 S. 53)

¹ Siehe hierzu Gülich/Raible 1977 S. 272 f., Quasthoff 1980 S. 39 f. sowie als Zusammenfassung dieser Positionen Brinker 2001 S. 53

Das Ergebnis der Analyse ist im Idealfall das Textthema selbst, dargestellt als „Makroproposition auf einem bestimmten Abstraktionsniveau“ (van Dijk 1980 a S. 50). Die MAKROSTRUKTUR¹ eines Textes setzt sich nun aus einer „macroproposition“ und einem „macrospeech act“ zusammen und stellt die textglobale semanto-pragmatische Struktur dieses Textes dar (van Dijk 1980 b S. 27 u. 43). Aus den auf oberflächen-textueller MIKROSTRUKTUREBENE enthaltenen Propositionen bzw. Sprechakten läßt sich eine tiefenstrukturelle MAKROPROPOSITION bzw. ein tiefenstruktureller MAKROSPRECHAKT (d.h. Illokution des Textes) ableiten.²

Dieses Ableiten der Makrostruktur geschieht mittels von van Dijk eingeführter MAKROREGELN.³ Die grundlegenden Regeln hierbei sind DELETION, GENERALIZATION und CONSTRUCTION.

„DELETION: „It deletes all those propositions of the text base which are not relevant for the interpretation of other propositions of the discourse [...]

[...] the rule operates on all kinds of so-called *irrelevant details* of a description, that is, details that do not contribute to the construction of a theme or topic.“

GENERALIZATION: „[We] abstract from semantic detail in the respective sentences by constructing a proposition that is conceptually more general.“

CONSTRUCTION: „In this rule propositions are, so to speak, taken together by substituting them, as a joint sequence, by a proposition that denotes a global fact of which the micropropositions denote *normal components, conditions, or consequences* [...]“

(van Dijk 1980 b S. 46 f.)

¹ Neben dem Terminus MAKROSTRUKTUR führt van Dijk den Begriff SUPERSTRUKTUR ein:

„[...] a superstructure is the schematic form that organizes the global meaning of a text.“

Eine solche Superstruktur ist als „conventional schemata“ aufzufassen, das aus „functional categories“ und „rules“ besteht, welche die Kategorien ordnen bzw. miteinander kombinieren. Eine Superstruktur ist also „eine Art *Textform*, deren Gegenstand, Thema, d.h.: Makrostruktur, der *Textinhalt* ist.“ Siehe hierzu van Dijk 1980 a S. 128 u. van Dijk 1980 b S. 128 f.

² Obwohl „der generative Bezugsrahmen nur angedeutet wird“ (Brinker 2001 S. 52), steht das Prinzip gedanklich der GENERATIVEN TRANSFORMATIONSGRAMMATIK nahe: Der Textproduktionsprozeß verläuft TOP-DOWN. Der Autor geht von einem Textthema (oberste Makroproposition) aus und entfaltet die inhaltlich-thematische Struktur schrittweise. Der Textrezeptionsprozeß hingegen ‚regeneriert‘ (BOTTOM-UP) die vom Autor angelegte Struktur, indem er (unbewußt) mittels Makroregeln aus den Propositionen Makropropositionen bis hin zum Textthema bildet.

Hierbei ist freilich zu berücksichtigen, daß die Verarbeitung von Texten nicht allein textuell determiniert ist, sondern in großem Umfang von Vorwissen, Einstellung und Intention des Lesers sowie von dessen Einschätzung der Kommunikationssituation abhängt. (siehe hierzu auch Kapitel 10.2)

³ In van Dijk 1980 a wird auf Seite 45 zusätzlich noch die Regel SELEKTIEREN eingeführt, die ebenso wie DELETION folgende abstrakte Form hat: $\langle a,b,c \rangle \rightarrow \langle b \rangle$. Meiner Meinung nach erfolgen Deletion und Selektion stets zusammen: Wichtige Information wird von unwichtiger getrennt, die unwichtige danach getilgt.

Van Dijk 1980 a demonstriert die Anwendung der Makroregeln an folgendem Beispieltext:

„S1 Peter beschloß, diese Jahr zum Wintersport zu fahren.
S2 Bisher war er immer nur im Sommer im Urlaub nach Italien gefahren, aber er wollte jetzt auch mal Ski fahren lernen, und auch die Gebirgsluft erschien ihm sehr gesund.
S3 Er ging in ein Reisebüro, um Prospekte zu holen, so daß er sich aussuchen konnte, wo er am liebsten hin wollte.
S4 Österreich zog ihn eigentlich am meisten an.
[...]"

(van Dijk 1980a S. 51)

Durch sukzessive Anwendung von Makroregeln gelangt van Dijk nun zu einer Makroproposition dieses Teiltexes. So ist etwa die Proposition *beschließen(Peter, dieses Jahr zum Wintersport fahren)* in S1 per Deletion eliminierbar, da sie lediglich eine semantisch implizite Voraussetzung für *fahren_zum(Peter, Wintersport)* und für die weiteren Propositionen nicht relevant ist.¹ S2 gibt Hintergrund und Motivation für *fahren_zum(Peter, Wintersport)* sowie die folgenden Propositionen in S3. S2 kann somit ebenfalls mittels Deletion getilgt werden. Die Propositionen in S3 könnten via Generalisation auf die allgemeinere Proposition *informieren_über(Peter, Reiseziel)* gebracht werden, da die Art und Weise des Einholens der Information nicht relevant ist. Aber auch diese generalisierte Proposition kann in Hinblick auf die unten gebildete Makroproposition lediglich als vorbereitende Handlung gesehen und damit getilgt werden. Auch die Proposition *anziehen(Peter, Österreich)* kann als mentale Einstellung per Deletion getilgt werden. Das Reiseziel *Österreich* wird per Construction als Ortsargument in die Proposition *fahren_zum(Peter, Wintersport)* eingebaut, so daß die dreistellige Makroproposition *fahren_zum(Peter, Wintersport, Österreich)* entsteht.

12.2 Exemplarische Makrostrukturanalyse eines Hypertextes

Eine BOTTOM-UP-ANALYSE würde sich einerseits insbesondere für umfangreiche, stark modulare Hypertexte eignen, deren Textumfang und thematische Struktur vom Leser nicht auf einen Blick erfaßt werden kann: Ausgehend von den extrahierten Makropropositionen der Knoten könnte eine Makroproposition

¹ In der Regel geht der Beschluß, etwas zu tun, der Handlung voraus. Oftmals ist eine Handlung aber auch eine direkte Reaktion auf Sinnesreize, so etwa bei *Reflexartig zog ich die Hand von der heißen Oberfläche zurück*. Durch DELETION der Proposition *heiß sein(Oberfläche)* würde eventuell zuviel Information verloren gehen, da ich meine Hand ja auch aus anderen Gründe von der Oberfläche zurückgezogen haben könnte.

des gesamten Hypertextes abgeleitet werden. Die Komplexität des van Dijk'schen Analysesystems läßt die Untersuchungsarbeit jedoch andererseits wiederum gerade bei umfangreichen Gebilden ins Unermeßliche steigen. Im Rahmen von TAH könnte es aber dennoch an bestimmten Stellen notwendig werden, die Makroproposition desjenigen Teiltextes eines Knotens zu ermitteln, dessen thematische Struktur nicht auf einen Blick ‚von oben‘ erfaßbar ist. Im folgenden sollen exemplarisch die zwei Knoten *Buchvorstellung* und *Seebold* des teilweise bereits bekannten Beispiel-Hypertextes auf deren Makrostrukturen hin untersucht werden.¹

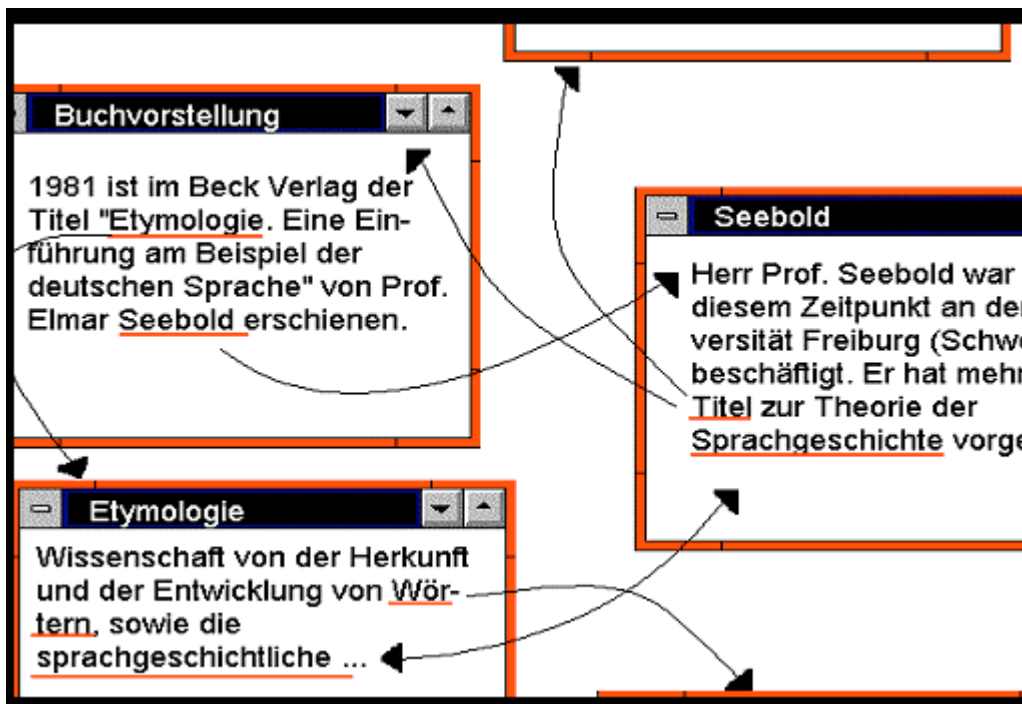


Abbildung 5: Link-Struktur eines einfachen Hypertextes

Die Verhältnisse erläutern die folgenden beiden Tabellen:

| | |
|----------------|--|
| Text: | <i>1981 ist im Beck Verlag der Titel „<u>Etymologie</u>. Eine Einführung am Beispiel der deutschen Sprache“ von <u>Prof. Elmar Seebold</u> erschienen.</i> |
| Propositionen: | <p>P1: <i>erscheinen</i>(<i>der Titel</i>)</p> <p>P2: <i>sein</i>(<i>der Titel, <u>Etymologie</u> [...]</i>)</p> <p>P3: <i>sein</i>(<i>Ort, Beckverlag</i>)</p> <p>P4: <i>sein</i>(<i>Autor, Prof. [...] <u>Seebold</u></i>)</p> |

¹ In den folgenden Tabellen werden Ausgangspunkte von Verweisen unterstrichen, Zielpunkte *kursiv* gekennzeichnet.

| | |
|------------------------|--|
| Makroproposition: | M1: <i>erscheinen(der Titel</i> ¹) |
| Schlüsselwort / Titel: | Buchvorstellung |

Tabelle 2: Makrostrukturanalyse von Knoten *Buchvorstellung*

| | |
|------------------------|---|
| Text: | <i>Herr Professor Seebold war zu diesem Zeitpunkt an der Universität Freiburg (Schweiz) beschäftigt. Er hat mehrere <u>Titel</u> zur Theorie der <u>Sprachgeschichte</u> vorgelegt.</i> |
| Propositionen: | P5: <i>beschäftigt_sein(Herr Professor Seebold, an der Universität [...])</i> P6: <i>vorgelegen(er², <u>Titel</u>)</i> |
| Makroproposition(en): | M2: <i>beschäftigt_sein(Herr Professor Seebold, an der Universität [...])</i> M3: <i>vorgelegen(er, <u>Titel</u>)</i> |
| Schlüsselwort / Titel: | Seebold |

Tabelle 3: Makrostrukturanalyse von Knoten *Seebold*

Das Thema des Knotens *Buchvorstellung* wird durch das Schlüsselwort der Titelzeile bereits genannt. Tatsächlich drückt die Propositionssequenz P1 bis P4 das Konzept *Buchvorstellung* aus und kann mittels M1 zusammengefaßt werden. Da die Propositionen P2, P3 und P4 lediglich weiterführende Informationen zu P1 liefern, können diese per Deletion getilgt werden.

M1 wiederum kann als eine konkrete Realisierung der übergeordneten Makroproposition M3 des Knotens *Seebold* ausgedrückt werden (*Mehrere Bücher wurden veröffentlicht, eines davon hat den Titel Etymologie, ein anderes den Titel X...*).³ Anzunehmen ist eine Argumentüberlappung (*Titel*) zwischen M1, M3,...MX und damit eine zunehmende Dominanz der M1,...MX-integrierenden Makroproposition M3. Daher kann die Makroregel Deletion auf M2 angewendet werden. Hierarchisch am höchsten ist nun M3, die das Textthema darstellt. Eine mögliche Überschrift des (hypothetischen) Gesamt-Hypertextes könnte daher *Die Veröffentlichungen von Herrn Prof. Seebold* lauten.

Fazit: Oben wurde die Makroproposition der einzelnen Knoten ermittelt. Anschließend ist untersucht worden, ob bestimmte Knoten eventuell durch Makroregeln getilgt oder in eine allgemeinere Makroproposition eines anderen Knotens integriert werden können. Im beschriebenen Beispiel erhält man als Ergebnis ein mögliches Textthema des Beispiel-Hypertextes. Allerdings ist anzumerken, daß die Kriterien, nach denen die Makroregeln anzuwenden sind,

¹ Syntaktisch ist der Ausdruck *Etymologie [...]* als enge Apposition zu *Titel* innerhalb der NP *der Titel Etymologie [...]* aufzufassen. Die Wirkung der NP nach außen wird von dem Ausdruck *der Titel* bestimmt, an dem auch die morphologische Markierung festgemacht wird (Zu den Schwierigkeiten bei der Unterscheidung von Kern und Apposition in NPs mit definitivem Artikel vgl. Eisenberg 1994 S. 256).

² Argumentüberlappung mit *Herr Professor Seebold*.

³ Wenn es sich um die Vorstellung mehrerer Titel des Autors handelt, ist das Schlüsselwort *Buchvorstellung* für diesen Knoten zu weit gefaßt.

nicht unabhängig vom – eigentlich durch sie zu ermittelnden – Ergebnis der Untersuchung sind: Welche Makroproposition beispielsweise nun spezieller als eine andere ist und daher von der allgemeineren getilgt werden darf, hängt vom Kontext – also auch vom Textthema bzw. der höchsten Makroproposition der Hierarchie – des gesamten Hypertextes ab. Aus diesem Grund, und wegen der unverhältnismäßigen Komplexität der Theorie, ist diese Methode für die Analyse ganzer Hypertexte als weniger ergiebig zu bewerten. Im Hinblick auf die Verweisstruktur des Hypertextes könnte eine Analyse jedoch durchaus partiell Sinn machen:

So wäre es etwa möglich, durch M2 und M3 per Construction die Makroproposition *haben(Seebold, wissenschaftlichen_Lebenslauf)* für den Knoten *Seebold* zu bilden. Ein Argument dieser Makroproposition ist der Text des Ausgangspunktes des Verweises *Seebold* im Knoten *Buchvorstellung*. Durchaus vorstellbar wäre hier ein typisierter Link, der Auskunft über das Link-Ziel gibt, etwa *Informationen zum Lebenslauf von Prof. Seebold*. Eine Makrostrukturanalyse des Zielknotens (bzw. des in Frage kommenden Teiltexes dieses Zielknotens) könnte vermutlich Hilfestellung beim Verfassen dieser (freien) Typisierungstexte geben.

Ein Punkt, der für die Untersuchung von Verweisen wichtig sein könnte, wird von der Makrostrukturanalyse nicht explizit untersucht: Der Wechsel von Referenten innerhalb einer Sequenz von Propositionen. Aus diesem Grund soll im nächsten Kapitel eine weitere Bottom-Up-Analysemethode vorgestellt werden, die den Fokus genau auf diesen Punkt setzt: das MODELL DER REFERENZIELLEN BEWEGUNG.

13 Referenzstruktur und referentielle Bewegung

Im letzten Kapitel wurde als wichtige Einschränkung der Makrostrukturanalyse die fehlende Beschreibung von eventuellen Referenzwechseln in Propositionssequenzen genannt. Das MODELL DER REFERENTIELLEN BEWEGUNG, welches in diesem Kapitel eingeführt wird, füllt genau diese Lücke.

Vorab muß jedoch erst geklärt werden, inwieweit die REFERENZSTRUKTUR von Texten überhaupt Rückschlüsse auf die thematische Entfaltungen geben kann. Aus diesem Anlaß, wird zunächst einmal der Terminus REFERENZ definiert.

13.1 Referenzbeziehungen in Texten

Der linguistische Begriff der REFERENZ ist umstritten. Traditionell bedeutete er die Beziehung zwischen sprachlichem Ausdruck und dem Phänomen der außersprachlichen Wirklichkeit, auf das der Ausdruck referiert. In der Sprechakttheorie ist die Referenz ein Teilakt der Proposition und steht für die Bezugnahme des Sprechers auf Außersprachliches mittels sprachlicher und nichtsprachlicher Mittel; damit ist

„[...] Referenz ist nicht eine Beziehung zwischen einem Wort (mit einer bestimmten wörtlichen Bedeutung) und einem Gegenstand (=Individuum), sondern zwischen einem Sprecher und einem Gegenstand [...]"

(Meibauer 2001 S. 19)

In aktuelleren kognitiven Theorien (z.B. Jackendoff 1983) meint REFERENZ nicht mehr die Bezugnahme auf Referenten der realen Welt, sondern auf Phänomene in einem mentalen Konzeptsystem des Bewußtseins.

Als allgemeine REFERENZTYPEN nennt Vater GEGENSTANDSREFERENZ, EREIGNISREFERENZ¹, ZEITREFERENZ und ORTSREFERENZ (vgl. Vater 1994 S. 116 f.).

Bei Referenzbeziehung in Texten kann es sich laut Vater um KOREFERENZ, PRODUKT-PRODUZENT-BEZIEHUNGEN, PRÄDIKAT-ARGUMENT-BEZIEHUNGEN, TEMPORALREFERENZ-RELATIONEN und LOKALISIERUNGSRELATIONEN handeln (vgl. Vater 1994 S. 133f.). Koreferenzbeziehungen können weiter nach dem jeweiligen Grad der REFERENZIDENTITÄT untergliedert werden in TOTALE

¹ Umstritten ist, ob lediglich auf Gegenstände oder auch auf Ereignisse referiert werden kann. Vater weist darauf hin, daß die Verwendung von Ereignisreferenz sich offenbar durchsetzt und verweist dabei auf die Arbeiten von van Dijk 1980a und de Beaugrande/Dressler 1981 (Vater 1994 S. 116).

REFERENZIDENTITÄT (z.B. zwischen Anapher und zugehörigem Antezedenten), PARTIELLE REFERENZIDENTITÄT (z.B. zwischen *Hans* und *sein Kopf*; Vater 1994 S. 133) oder ÜBERLAPPENDE REFERENZ (z.B. zwischen *die Jungen* und *die älteren Kinder* bei Bezugnahme auf eine Gesamtmenge *die Kinder*; Vater 1994 S. 133).

Die Produkt-Produzent-Beziehungen, die laut Vater „den Koreferenz-Beziehungen verwandt sind“ (Vater 1994 S. 138) lassen sich eventuell auch als spezielle Form einer partiellen Referenzidentität (z.B. zwischen *die Softwarefirma* und *deren Programme*) auffassen.

13.2 Modell der referentiellen Bewegung

Zur Untersuchung der Referenzstruktur setzt Vater u.a. das MODELL DER REFERENTIELLEN BEWEGUNG (vgl. Klein u. von Stutterheim 1987 u. 1991 und Vater 1994 S. 144) ein. Auch im Rahmen der vorliegenden Arbeit wird im Hinblick auf thematisch-strukturelle Aspekte auf dieses Modell zurückgegriffen. Zunächst soll das Modell jedoch allgemein vorgestellt werden: Nach Meinung von Klein und von Stutterheim dient ein Text dazu, eine implizite oder explizite Frage zu beantworten. Diese Frage wird QUAESTIO genannt (vgl. Klein u. von Stutterheim 1987 S. 163). Die Beantwortung der Quaestio wird durch die Gesamtheit der Äußerungen eines Textes geleistet.

„Jede einzelne Äußerung enthält ein Gefüge von Angaben zu Ort, Raum, [Zeit; Anmerkung von O.H.], Handlung, Personen usw. – zu verschiedenen semantischen Bereichen oder, wie wir hier sagen, Referenzbereichen. Die folgende Äußerung behält einen Teil dieser Informationen bei und führt andere neu ein. Diese Entfaltung der Information von Äußerung zu Äußerung bezeichnen wir als *referentielle Bewegung*.“

(Klein u. von Stutterheim 1987 S. 166)

Eine einzelne Äußerung faßt die besagten Angaben zu einer Proposition¹ zusammen. Referentielle Bewegung geschieht zwischen einer Proposition und der ihr folgenden. Hauptsächliche Intention der Autoren ist, die Rolle der Quaestio bei der Spezifizierung einzelner Propositionen sowie bei der referentiellen Bewegung zwischen diesen Propositionen zu ermitteln: Die Weiterführung von Information referentieller Domänen bzw. die Einführung neuer Information kann auf verschiedene Arten erreicht werden, wie folgendes

¹ Vereinfachend kann in Anlehnung an Klein u. von Stutterheim eine Proposition in einen „inneren Kern“, welcher die Angaben (im allgemeinsprachlichen Sinn) zu Person und Handlung beinhaltet, sowie einen „äußeren Kern“ mit Angaben zu Raum und Zeit unterteilt werden. Die Autoren führen als weiteren Referenzbereich den der modalen Angaben ein, der aber für unsere Zwecke vernachlässigt werden kann; vgl. Klein u. von Stutterheim 1987 S. 170 f.

Schaubild eines Analysemodells der referentiellen Bewegung veranschaulichen soll.¹

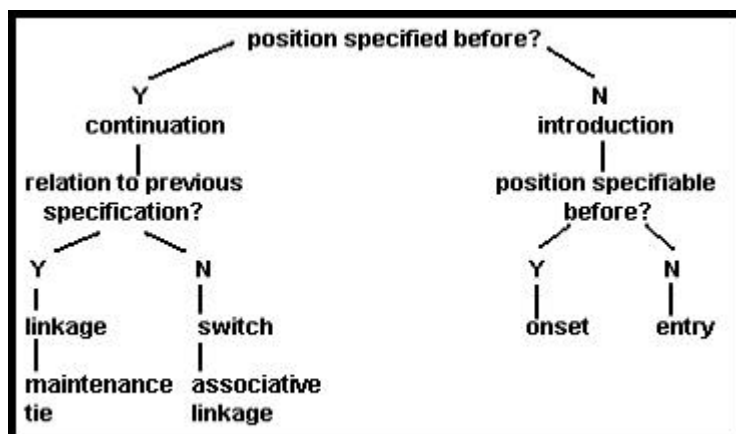


Abbildung 6: MODELL DER REFERENTIELLEN BEWEGUNG

13.3 Referentielle Bewegung in Hypertexten

Eine Anwendung des Modells erfolgt nun bei der Analyse des einfachen Beispiel-Hypertextes aus Abbildung 5 von Kapitel 12.

Untersucht werden diesbezüglich exemplarisch die folgenden Sätze der Knoten *Buchvorstellung*, *Seebold* und *Etymologie*. Als Beispiel einer möglichen Lese-reihenfolge wird die Sequenz *Buchvorstellung* – *Seebold* – *Etymologie* angenommen.²

1981 ist im Beckverlag der Titel „Etymologie. Eine Einführung am Beispiel der deutschen Sprache“ von Professor Elmar Seebold erschienen. I

Herr Professor Seebold war zu diesem Zeitpunkt an der Universität Freiburg (Schweiz) beschäftigt. Er hat mehrere Titel zur Theorie der Sprachgeschichte vorgelegt. I

[Etymologie ist die] Wissenschaft von der Herkunft und der Entwicklung von Wörtern, sowie die sprachgeschichtliche ...

¹ Übernommen aus: Vater 1994 S. 146

² Relevante Links werden unterstrichen, deren Ziele kursiv dargestellt (siehe auch den Abschnitt *Typographische Konventionen*).

| Knoten (Proposition) | Person | Ereignis | Ort | Zeit |
|----------------------------------|---|--|---------------------------------------|----------------------------|
| Buchvorstellung (1) ¹ | Titel „ <i>Etymologie</i> [...]“ | <i>Erscheinen</i> | <i>im Beck Verlag</i> | <i>1981</i> |
| Referentielle Bewegung: | maintenance ² | switch | ++++ ³ | associative linkage |
| Buchvorstellung (2) | <i>Prof. Elmar Seebold</i> _i | [<i>schreiben</i>] | [0] | [0] ⁴ |
| Referentielle Bewegung: | maintenance | switch | switch (bzgl. <i>im Beck Verlag</i>) | maintenance (von 1981) |
| Seebold (1) | <i>Herr Prof. Seebold</i> _i | <i>beschäftigt sein</i> | <i>an der Universität Freiburg...</i> | <i>zu diesem Zeitpunkt</i> |
| Referentielle Bewegung: | maintenance | switch | (maintenance) ⁵ +++++ | associative linkage |
| Seebold (2) | <i>Er</i> _i | <i>vorlegen (Titel zur Sprachg.)</i> | [0] | [0] |
| Referentielle Bewegung: | switch | (switch) | +++++ | switch |
| Etymologie | [<i>Etymologie</i>] | [<i>ist die</i>] <i>Wissenschaft</i> [...] | [0] | [0] |

Tabelle 4: Referentielle Bewegung in einem Hypertext

¹ Der referentielle Bereich *Person* des Knotens *Buchvorstellung* erweist sich als etwas problematisch. Entweder man folgt der obigen Propositionanalyse und sieht die NP *Prof. Elmar Seebold* als Teil der übergeordneten PP lediglich als Erweiterung der eigentlichen Kernproposition mit dem Referenten *Titel* oder man setzt zwei Propositionen an: *erscheinen(Titel)* sowie *schreiben(Prof. Elmar Seebold, Titel)*. Die Überlegung ist also, daß eine grundlegende Voraussetzung für eine Buchveröffentlichung die Tatsache ist, daß dieses Buch bereits geschrieben ist. Das allgemeine Wissen um diese Tatsache macht eine sprachliche Erwähnung derselben unnötig. "Eine Proposition ist [aber] ein theoretisches Konstrukt, das zur Benennung und Darstellung der mentalen Struktur eingeführt wurde [...]", drückt also eher vorsprachlich kognitive als sprachliche Bedeutung aus (Sowarka u.a. 1983, S. 69). Damit geht auch einher, daß laut Klein u. von Stutterheim das explizite Referieren im Ermessen des Autors liegt, d.h. wenn dieser davon ausgehen kann, daß eine bestimmte Referenzbeziehung vom Leser aufgrund anderer Wissensquellen erkannt wird, diese im Text nicht unbedingt mittels sprachlicher Elemente explizit gemacht werden muß.

² Klein u. von Stutterheim 1987 sehen auch das Konzept einer „partiellen Beibehaltung“ vor. Ähnlich könnte auch die Beziehung von *Titel (...)* zu *Prof. Elmar Seebold* als partielle Beibehaltung etwa im Sinne einer PRODUKT-PRODUZENT-RELATION interpretiert werden.

³ Das Zeichen „++++“ bedeutet, daß der jeweilige Referenzbereich in der folgenden Proposition nicht spezifiziert wird.

⁴ Implizit ist der Zeitpunkt des Schreibens mit *früher als* bzw. *gleich 1981* festgelegt.

⁵ Denkbar wäre, daß sämtliche von Prof. Seebold erschienenen Titel an der Universität Freiburg (Schweiz) vorgelegt wurden. In diesem Fall wäre eine MAINTENANCE der Lokalreferenz anzusetzen. Spielt die Universität Freiburg für die Vorlage der Titel allerdings keine Rolle (und dies ist anzunehmen), so muß für die Lokalreferenz von Seebold (2) eine Nullposition [0] angenommen werden. Referentielle Bewegung fände somit hier nicht statt.

Die referentielle Bewegung aller Referenzbereiche¹ der im Text auftretenden Propositionen zeigt die oben stehende Tabelle und wird im folgenden näher beschrieben.

Personenreferenz

Man stellt fest, daß die Personenreferenz innerhalb der ersten beiden Knoten unverändert bleibt, diese Knoten also bezüglich der referentiellen Domäne *Person* in beliebiger Reihenfolge gelesen werden könnten: Die Personenangabe der ersten Proposition des zu Beginn gelesenen Knotens ist in jedem Fall rhematisch, wird mit einer INTRODUCTION realisiert und im folgenden Knoten thematisch mittels koreferenter NP bzw. Anapher wiederaufgenommen. Die Fragen *position specified before?* und *relation to previous specification* können mit *Ja* beantwortet werden. So läßt sich folgern, daß es sich bei der referentiellen Bewegung der Personenangabe um eine Beibehaltung (MAINTENANCE) handelt. Als kohäsive Mittel dienen hier Anapher und Rekurrenz.²

Beim Übergang vom zweiten zum dritten Knoten ändert sich jedoch die Personenreferenz: Mittels eines SWITCH wird von *Er* zu *Etymologie* gewechselt. Wirft man nun aber einen Blick auf den Link-Text (*Sprachgeschichte*) des Knotens *Seebolds* und vergleicht diesen mit der Personenreferenz der ersten Proposition im Knoten *Etymologie*, so erkennt man, daß hier eine Maintenance mittels der synonymen Begriffe *Sprachgeschichte* und *Etymologie* vorliegt. Der dominierende Referenzträger hat sich dennoch geändert, da der Link-Text

¹ In dem von Klein u. von Stutterheim vorgestellten Modell wird als Referenzbereich einer Proposition auch die Dimension *Vorgänge, Zustände, Ereignisse* eingeführt, die hier abgekürzt als *Ereignis* übernommen wird. Allerdings herrscht, wie bereits erwähnt, prinzipielle Unstimmigkeit darüber, ob lediglich Gegenstände und Personen als Referenten fungieren oder auch Ereignisse. Nach allgemeiner Definition ist eine PROPOSITION ein Teilakt einer SPRECHHANDLUNG (Äußerung) und setzt sich aus weiter oben bereits beschriebener REFERENZ (Bezugnahme auf Objekte, Dinge, Konzepte, Personen der außersprachlichen Wirklichkeit) und PRÄDIKATION (Aussage über den Referenten) zusammen. Alle übrigen Referenzbereiche stellen Erweiterungen der eigentlichen Proposition dar. Für die Analyse der referentiellen Bewegung wird aber die Terminologie von Klein u. von Stutterheim beibehalten.

² Vater weist darauf hin, daß es sich bei MAINTENANCE auch um eine Paraphrasierung handeln kann. Als Beispiel gibt er das Paar *Strauß – der Wiener Walzerkönig* an. Eine derartige Verknüpfung mittels Paraphrase über Knotengrenzen hinweg könnte allerdings zu Kohärenzverlust beim Leser führen. Hier spielen KONTEXT und WELTWISSEN des Lesers entscheidende Rollen: Navigiert der Leser durch einen Hypertext mit dem Titel *Richard Strauß*, so ist anzunehmen, daß er um die Koreferenz von *Strauß* und *der Wiener Walzerkönig* weiß. Ist der Knoten, dessen erste Personenangabe *der Wiener Walzerkönig* ist, aber in einen kontextuell breiter angelegten Hypertext (etwa mit dem Titel *Musik der Neuzeit*) enthalten, und verfügt der Leser nicht über das Wissen, daß Strauß in Wien Walzer komponiert hat, so ist der Bezug von *der Wiener Walzerkönig* zu *Strauß* schwierig herzustellen.

Eine MAINTENANCE mittels Rekurrenz scheint hingegen auch über Knoten-Grenzen hinweg stets kohärent zu sein.

Sprachgeschichte im Knoten *Seebold* nicht die Personenreferenz herstellt sondern lediglich untergeordnetes Argument der Ereignisreferenz *vorlegen* ist: *vorlegen(er, mehrere Titel zur Theorie der Sprachgeschichte)*. Dennoch ist die Sequenz bezüglich der Personenreferenz als kohärent zu betrachten. Es darf vermutet werden, daß der im Link-Text vorhandene Ausdruck eine Verschiebung des dominierenden Referenzträgers (= die Personenreferenz) andeutet. *Sprachgeschichte* als Link-Text signalisiert also, daß *Sprachgeschichte* (bzw. ein synonyme oder anderweitig koreferente Ausdruck) im Ziel-Knoten als Personenreferenz eingeführt werden wird. Daß *Sprachgeschichte* im aktuellen Knoten lediglich Argument der Ereignisreferenz ist, spielt hierbei scheinbar keine Rolle.

Ereignisreferenz

Die Ereignisreferenz verschiebt sich jeweils durch Switch-Beziehungen. Der ständige Wechsel innerhalb der Domäne *Ereignis* ist in Zusammenhang mit der dominierenden Funktion Maintenance der Domäne *Person* zu sehen. Es scheint eine ständig wechselnde Prädikation des fast immer gleichen Referenten vorzuherrschen. Diese Feststellung wird im nächsten Kapitel im Zusammenhang mit der Rolle der Quaestio zu erklären versucht.

Ortsreferenz

Bezüglich der Ortsreferenz läßt sich lediglich aussagen, daß sie durch Switch von Knoten zu Knoten wechselt. Sie ist im Beispiel auch eher von untergeordneter Bedeutung.

Zeitreferenz

Die Phrase *Zu diesem Zeitpunkt* ist als Maintenance zu *1981* zu verstehen. Als deiktisches Mittel zu Knotenbeginn kann sie nur in der Lesereihenfolge *Buchvorstellung – Seebold* korrekt interpretiert werden.

Eine Veränderung des zeitlichen Bezugs läßt ferner eventuell auf einen Wechsel von Haupt- zu Nebenstruktur schließen. Die Formen *hat vorgelegt, ist erschienen* referieren auf einen bestimmten Zeitpunkt in der Vergangenheit. Nimmt man als Textthema etwa *Wissenschaftliche Publikationen von Prof. Seebold* an, so ist als Zeitform für die Hauptstruktur die Vergangenheit anzusetzen. Elemente der Nebenstruktur (etwa die ergänzende Definition von

Etymologie im gleichnamigen Knoten) stehen dagegen im Präsens.¹

Vorläufiges Fazit: Introductions von Referenten können offenbar problemlos zu Beginn eines Knotens erfolgen. Wechselt der Referent einer bereits eingeführten Domäne zu Beginn eines per Verweis angesteuerten Knotens, so ist dies mit einer Introduction vergleichbar und wird sich in den meisten Fällen als kohärent erweisen.² Ob eine Maintenance als kohärent interpretiert wird, hängt von deren Realisierung ab: Wird die Beibehaltung mittels Rekurrenz realisiert, so wird diese auch zu Knotenbeginn als kohärent gelten dürfen. Aus besagten Gründen (siehe hierfür Kapitel 9.1) ist eine Maintenance mittels Pronomen zu Knotenbeginn hingegen für manche Lesevarianten nicht akzeptabel und führt zu Kohärenzverlust.

13.4 Die Rolle der Quaestio

Eine mögliche Einschränkung der Aussagen des oberen Abschnittes macht das jetzige Kapitel: Es wurde gezeigt, daß eine zu Knotenbeginn plazierte alleinstehende Anapher in der Regel den Aufbau von Kohärenz verhindert. Die von Klein u. von Stutterheim eingeführte QUAESTIO (= TEXTFRAGE) kann eventuell helfen, diesen Sachverhalt differenzierter zu betrachten.

Ist die Quaestio einer Erzählung bei Klein u. von Stutterheim etwa *Was passierte mit der Person P zum Zeitpunkt T?*, so lassen sich auch für den vorliegenden Hypertext Formulierungen einer Quaestio auf Knoten- sowie auf Hypertext-Ebene finden. Ist dies geschehen, so kann man feststellen, daß in vorliegendem Beispiel die Quaestio eines Textes in Verbindung mit dem Grad an Kohärenz bezüglich der Domäne *Person* steht. Die Beantwortung der Quaestio sollte sich normalerweise in der Überschrift oder Schlagzeile manifestieren. Ausgehend von der Quaestio der einzelnen Knoten kann man für den vorliegenden Beispiel-Hypertext mehrere mögliche Überschriften finden:

¹ Bei Definitionen der Art *Etymologie ist ...* handelt es sich in gewisser Hinsicht um ‚zeitunabhängige‘ Phänomene, die Formulierungen der Form *Morgen wird Etymologie ... sein* oder *Gestern war Etymologie...* sind schwer vorstellbar. (Wobei freilich nicht behauptet werden will, daß solchartige Begriffe im Laufe der Zeit keinen Bedeutungswandel erfahren können. Sätze wie *Um 1970 verstand man unter Etymologie das folgende...* sind daher durchaus denkbar.)

² Diese Aussage ist natürlich nicht allgemein gültig. Wichtig ist das allgemeine Verhältnis zweier Sätze verschiedener Knoten, die hintereinander in einer Navigationssequenz gelesen werden. Ist das Verhältnis etwa ein kontrastives (*Er schlief. / Sie hingegen verrichtete die ganze Arbeit.*), so können die Sätze nur in einer bestimmten Reihenfolge vollständig korrekt interpretiert werden.

- (1) „Der Titel Etymologie [...] des Beck Verlags“
- (2) „Veröffentlichungen/Curriculum vitae von Prof. Seebold“ oder
- (3) „Was ist Etymologie?“

Mit der Überschrift werden eindeutig Hauptstruktur und Nebenstrukturen des Hypertextes vorgegeben, was wiederum Auswirkungen auf weitere referentielle Domänen, wie etwa die Zeitreferenz, hat. Wird als Quaestio *Was ist Etymologie?* angenommen, so ist nicht vorauszusetzen, daß der Leser eine zu Knotenbeginn plazierte Anapher mit dem Antezedenz *Prof. Seebold* eines anderen Knotens in Verbindung bringen kann. Anders verhält es sich allerdings bei einer Quaestio, die ihren Ausdruck in einer Überschrift wie *Veröffentlichungen/Curriculum vitae von Prof. Seebold* findet: Hier ist das Erkennen der Referenzidentität zwischen den besagten Elementen der Domäne *Person* deutlich einfacher, denn: Die referentielle Domäne *Person* ist direkt in der Überschrift signalisiert und ändert sich wahrscheinlich nur wenig. Dagegen wird innerhalb der Domäne *Ereignis* mittels Switch eine ständig wechselnde Zuordnung von Eigenschaften (= Prädikation) des Referenten der Domäne *Person* bewirkt.

Fazit: Wird das maßgebende Element einer referentiellen Domäne bereits in der Überschrift thematisiert und eventuell mittels Maintenance als Referent gehalten – dient der Text also hauptsächlich zur Präzisierung dieses Elements –, so ist dessen Referenz so eindeutig feststellbar, daß selbst eine Anapher zu Knotenbeginn den kohärenten Aufbau der Gesamtstruktur nicht in größerem Maß negativ beeinflussen kann.¹

¹ Durchgehende Maintenance eines Elements einer bestimmten referentiellen Domäne (hier: *Person*) geht zumeist mit Introductions in anderen Domänen einher (hier besonders: *Ereignis*). Durch den Übergang von *Er* der Proposition *Seebold* (2) zu [*Etymologie*] der Proposition *Etymologie* mittels einer Switch-Beziehung manifestiert sich (unter den aktuellen Prämissen) der Übergang von Haupt- zu Nebenstruktur. Hypertexte bieten aufgrund ihres modularen Aufbaus gute Möglichkeiten, Elemente von Haupt- und Nebenstruktur logisch zu trennen.

14 Zusammenfassung: Das Analysemodell TAH

Im folgenden werden die in den vorherigen Kapiteln vorgestellten textlinguistischen Größen und Methoden zu einem konsistenten TEXTLINGUISTISCHEN ANALYSEMODELL FÜR HYPERTEXTES verschmolzen, welches innerhalb dieser Arbeit unter der Abkürzung TAH verwendet werden soll.

14.1 Ausrichtung und Vorteile des Analysemodells

„Ich vertrete in diesem Aufsatz die Ansicht, dass kein neuer Textbegriff benötigt wird, dass begriffliche Differenzierungen ausreichen, um Hypertexte als textuelle Gebilde mit ganz spezifischen Eigenschaften an einen pragmatisch und funktional fundierten Textbegriff anzubinden. Dadurch wird es möglich, die Kategorien, die in der Diskussion um Textualität, um die Struktur und die gesellschaftliche Funktion von Texten entwickelt wurden, sowie die Erkenntnisse der empirischen Textproduktions- und rezeptionsforschung [sic!] zu nutzen, um einer Mythologisierung von ‚hyper‘-Eigenschaften vorzubeugen, die den Blick auf Möglichkeiten und Grenzen der neuen Lese- und Schreibtechnologie verstellt.“

(Storror 2000 Abschnitt: *Einleitung*)

Dieser Aussage kann nur zugestimmt werden. Jeder Text ist zuallererst eingebettet in ein Geflecht sozialer, situativer, kontextueller und medialer Faktoren. Ein darauf basierender linguistische Textbegriff, wie er in Kapitel 8.3 erarbeitet und in dieser Arbeit verwendet wird, geht darüber hinaus von einer strikten Trennung von Funktion und Struktur aus. Die Struktur wird hierbei maßgeblich von der Funktion beeinflusst. Hieraus leitet sich das Primat der Textfunktion ab, welches als fundamentales Merkmal auf jeder Ebene des Analysemodells bemerkbar wird. Ein auf diese Weise „pragmatisch und funktional fundierter Textbegriff“ ist auch für die Analyse von Hypertexten zu verwenden.

Aus diesen eben angesprochenen Grundsätzen leitet sich die Vorgehensweise des TAH ab: Es handelt sich um eine Top-Down-Untersuchung, die von den kontextuellen, konventionellen und intertextuellen Aspekten sowie der globalen Textfunktion, Text- und Paratextstruktur des Hypertextes ausgehend sich über die Ebene der einzelnen Knoten knoten-intern bis auf Absatz- und Satzebene ‚herunterarbeitet‘. Der Hypertext selbst wird als funktionales Ganzes gesehen, von welchem ausgehend die Analyse zu den „konstituierenden Einheiten und Strukturen“ (Brinker 2001 S. 150) zu erfolgen hat.

Zielt die Hypertext-Analyse auf spezielle Aspekte ab, wie in unserem Falle etwa auf die der typisierten Links (siehe Teil C), so muß das TAH um entsprechende

Ebenen erweitert werden. Ein solches ‚Modul‘ zur Untersuchung von Links auf Knoten-Ebene wird in Tabelle 7 vorgestellt. Auch hier müssen strukturelle Aspekte stets in ihrer Wirkung auf die Funktion beschrieben werden. Das Grundmodell des TAH kann um beliebige Module erweitert werden, etwa für sekundäre semiotische Systeme wie Bild- oder Tonelemente der Knoten. Je nach Ausrichtung der praktischen Untersuchung können dann die Schwerpunkte jeweils verschieden gesetzt werden. Im Rahmen des sich anschließenden Fallbeispiels werden hauptsächlich Beschreibungsaspekte der Ebenen *Hypertext* und *Link* bearbeitet, die Beschreibungsaspekte *Knoten*, *Absatz* und *Satz* werden nur selektiv, wo notwendig, besprochen. Eine anders ausgerichtete Untersuchung aber, die beispielsweise das Verhältnis von Knoten-Funktionen hinsichtlich der Gesamt-Textfunktion analysiert, würde die unterschiedlichen Beschreibungsaspekte in anderer Weise gewichten. TAH ist insofern offen für fast alle textlinguistischen Fragestellungen, die an Hypertext gerichtet werden können.

Besonders zu bemerken ist, daß sich TAH sowohl zur Analyse der Hypertext-Struktur als auch zu der einzelner Lesepfade eignet – die Einteilung in Ebenen, Aspekte und Kategorien ist für beide Perspektiven relevant!

Insgesamt könnte man TAH auch – im Hypertext-Jargon bleibend – als ‚Fischaugenmodell‘ beschreiben: Die Darstellung graphischer Übersichten in Browsern geschieht manchmal aus der Perspektive einer sogenannten FISH EYE VIEW (vgl. Gloor u. Saxer 1990 S. 190f.). Dies ist eine Darstellungsform, welche lokale Details und globale Zusammenhänge in den Vordergrund rückt, während entferntere Einzelheiten der Struktur unterschlagen werden: Wie spezifisch auch immer die Details untersucht werden – im Fallbeispiel dieser Arbeit: die typisierten Links –, die kontextuellen, funktionalen und grobstrukturellen Aspekte des Gesamt-Hypertextes dürfen dabei in ihrer konstituierenden Kraft nicht aus den Augen verloren werden.

14.2 Einschränkungen

Auf eine wichtige Einschränkung muß hingewiesen werden: In Kapitel 9 wurden einige kognitive Grundzüge menschlicher Textverarbeitung angerissen. Diese Vorgänge sind sowohl textgeleitet als auch lesergeleitet. Hierbei wird aus textuellen Informationen und bereits vorhandenen Wissensstrukturen des Lesers eine – für jeden Leser andere – mentale Repräsentation des Textes konstruiert.

In einem analytischen aber auch synthetischen Prozeß wird somit Wissen aus Texten konstruiert. Hierbei passieren äußerst komplexe kognitive Prozesse, darunter perzeptive, syntaktisch- bzw. semantisch-analytische und Zugriffe auf bereits vorhandene kognitive Strukturen beim Leser. All diese Prozesse laufen (mehr oder weniger) simultan und stehen im Verhältnis einer gegenseitigen, ständigen Rückkopplung. Diese Aspekte versuchen interdisziplinäre Theorien des Textverstehens zu klären. Der folgende, rein textlinguistische Ansatz ist bescheidener – in seiner Beschreibungskraft aber sehr tiefgehend: Aufbauend auf dem oben vorgestellten Textbegriff geht es dem TAH ‚nur‘ um die Formulierung deskriptiver Analysekriterien für bestimmte Texte, Hypertexte nämlich. Es wird stets auf der textuellen Ebene verblieben, lediglich einige ‚Schnittstellen‘ zu leserzentrierten Aspekten werden angeboten: Hierbei geht es hauptsächlich um kohäsive Lücken im Text, die durch elaborative Inferenzen seitens des Lesers geschlossen werden müssen.

Auf die prinzipielle Offenheit von TAH für die Integration in einen weiter gefaßten kognitiven Ansatz der Textverarbeitung wird im Abschnitt *Ausblick* dieser Arbeit jedoch nochmals ausdrücklich hingewiesen.

Ein weiterer wichtiger Punkt ist, daß die unten angeführte Tabellenform der einzelnen Ebenen und Kategorien des TAH eventuell eine Abgeschlossenheit suggerieren könnte, wo diese in solcher Form nicht vorhanden ist. Oder, um mit Brinker 2001 zu sprechen, der sein Analysemodell für traditionelle Texte ebenfalls als Tabelle visualisiert:

„Die einzelnen Beschreibungsebenen und –aspekte sind bei der Textanalyse zwar genau zu unterscheiden, nicht aber voneinander zu isolieren. Zwischen situativem bzw. medialem Kontext, kommunikativer Funktion (Textfunktion), thematischem Aufbau und sprachlich-grammatischer Strukturierung von Texten bestehen komplexe Beziehungen“.

(Brinker 2001 S. 152)

Dies gilt selbstredend auch für Hypertexte. Wer die Trennung der Kategorien und Ebenen aufgibt, verzichtet auf die analytische Schärfe einer solch modularen Problembehandlung. Wer sich ihr ganz unterwirft, wird ‚betriebsblind‘. Hier das richtige Maß zu finden, ist die Kunst der textlinguistischen Analyse. Der folgende Abschnitt bietet eine tabellarische Übersicht der einzelnen Ebenen, Beschreibungsaspekte und Analysekatoren von TAH. Das meiste sollte nach den Ausführungen der letzten Kapitel selbsterklärend sein. Eventuell kritische Punkte werden in den anschließenden Anmerkungen besprochen.

14.3 Überblicksdarstellung von TAH

| Ebene | Beschreibungsaspekt | Analysekategorie (+ evtl. Unterkategorien) | |
|------------------------|---|---|---|
| Hypertext | Kontextuelle Aspekte | Kommunikations-situation | - Kommunikationsform - Handlungsbereich - Autor (Person u. Intention) - Zielgruppe(n) - Nötiges Welt- u. Fachwissen |
| | Kommunikativ-funktionale Aspekte | Gesamt-Textfunktion | |
| | Konventionelle Aspekte | - Textsorte - (Textsortenwissen) - (Schemataaktivierung durch Textsorte) | |
| | Strukturelle bzw. paratextuelle Aspekte | - Knoten- und Link-Anzahl - Knoten- und Link-Typen (soweit vom Autor definiert) | |
| | | Anwenderspezifischer Peritext | Browser-Software |
| | | Verlegerischer Peritext | - Screendesign - Metaphorik |
| | Allgemeiner Peritext | - Gliederung - Hypertext-interne Navigation - Definierte Lesepfade - Glossare - Suchmaschinen | |
| Intertextuelle Aspekte | - Implizite Einbettung in einen größeren textuellen Komplex aufgrund der Kommunikationssituation - Explizite Intertextualismen durch extra-hypertextuelle Verweise | | |

Tabelle 5: Das TEXTLINGUISTISCHE ANALYSEMODELL FÜR HYPERTEXTTE (TAH) – Ebene *Hypertext*

| Ebene | Beschreibungsaspekt | Analysekategorie (+ evtl. Unterkategorien) | |
|-----------------------|----------------------------------|---|--|
| Knoten | Kommunikativ-funktionale Aspekte | Knoten-Textfunktion | Verhältnis zur Gesamt-Textfunktion |
| | Thematisch-strukturelle Aspekte | - Thema - Themenentfaltung | - Makroproposition - Referentielle Bewegung - Isotopieebenen |
| | Grammatisch-strukturelle Aspekte | - Kohäsion als Grundlage für Kohärenz - Mögliche kohäsive Lücken (füllbar durch intendierte und elaborative Inferenzen?) | |
| Absatz bzw. Abschnitt | Kommunikativ-funktionale Aspekte | Absatz-Textfunktion | Verhältnis zur Knoten-Textfunktion |
| | Thematisch-strukturelle Aspekte | - Thema - Themenentfaltung | - Makroproposition - Referentielle Bewegung - Isotopieebenen |
| | Grammatisch-strukturelle Aspekte | - Kohäsion als Grundlage für Kohärenz - Mögliche kohäsive Lücken (füllbar durch intendierte und elaborative Inferenzen?) | |
| Satz | Kommunikativ-funktionale Aspekte | Satzfunktion | Verhältnis zur Absatz-Textfunktion |
| | Thematisch-strukturelle Aspekte | - Thema - Themenentfaltung | - Makroproposition - Referentielle Bewegung - Isotopieebenen |
| | Grammatisch-strukturelle Aspekte | - Kohäsion als Grundlage für Kohärenz - Mögliche kohäsive Lücken (füllbar durch intendierte und elaborative Inferenzen?) | |

Tabelle 6: Das TEXTLINGUISTISCHE ANALYSEMODELL FÜR HYPERTEXTTE (TAH) – Ebene *Knoten*

| Ebene | Beschreibungsaspekt | Analysekategorie (+ evtl. Unterkategorien) | |
|--------------|--|---|---|
| Link | Aspekte des Bezugsbereichs | <ul style="list-style-type: none"> - Aspekte des Ausgangs-Knotens, die mit Hilfe des Links selektiert werden - Aspekte des Ziel-Knotens, die mit Hilfe des Links verknüpft werden - Deixis | |
| | Kommunikativ-funktionale Aspekte | Link-Funktion | <ul style="list-style-type: none"> - Relation, die der Link zwischen Funktion der Ausgangsressource und der Funktion der Zielressource etabliert - Illokutionäre Beziehung, die über die Knotengrenze hinweg etabliert wird |
| | Thematisch-strukturelle Aspekte | Themenentfaltung | <ul style="list-style-type: none"> - Thematische Einbettung des Link-Textes in die Ausgangsressource - Entfaltung des Themas, die mittels des Links über die Ressourcengrenze hinaus geht - Propositionale Beziehung, die über die Knotengrenze hinweg etabliert wird |
| | Grammatisch-strukturelle Aspekte | Kohäsion | <ul style="list-style-type: none"> - Kohäsive Einbettung des Link-Textes (bzw. Link-Ankers) in die Ausgangs-Ressource - Grammatische Strukturen, die durch den Link über die Ressourcengrenze hinaus wirken - Mögliche kohäsive Lücken beim Übergang zwischen den Ressourcen (füllbar durch intendierte und elaborative Inferenzen?) |
| | Paratextuelle und metasprachlicher Aspekte | Signalisierung der Existenz des Links, Link-Funktion bzw. Link-Rhetorik | <ul style="list-style-type: none"> - Semiotische Hilfsmittel jeder Art, mit denen der Link in irgendeiner Art kenntlich gemacht oder näher spezifiziert wird - Relation zu Beschreibungsaspekten, auf die sich die Spezifizierung bezieht |
| | Browser-spezifische Aspekte | Traversal-Verhalten | <ul style="list-style-type: none"> - Verhaltensweise der Browser-Software - Bezugnahme auf kommunikativ-funktionale Aspekte - Bezugnahme auf paratextuelle und metasprachliche Aspekte |

Tabelle 7: Das TEXTLINGUISTISCHE ANALYSEMODELL FÜR HYPertexte (TAH) – Ebene *Link*

14.4 Anmerkungen zu den Untersuchungsebenen

Die prinzipielle Ausrichtung von TAH wurde eingangs bereits beschrieben. Die einzelnen Aspekte und Dimensionen wurden ausführlich in den vorangegangenen Kapiteln erläutert. Es sind daher nur noch einige wenige zusätzliche Anmerkungen zu machen, die in den folgenden Abschnitten – im Stile einer ‚Fußnotenkollektion‘ – lose verbunden aufgeführt werden.

14.4.1 Anmerkungen zur Ebene *Hypertext*

Kontextuelle Aspekte

Zur Beschreibung der Kommunikationsform und des medialen Rahmens der Corpus-Texte muß beim elektronischen Medium auch eine Analyse der Zugangsvoraussetzungen erfolgen. Es sollten hierfür der Grad nötiger Erfahrung im Umgang mit Computern nicht unerwähnt bleiben, da der elektronische Hypertext doch andere Anforderungen an den Leser (der typischerweise auch USER genannt wird) stellt, als dies etwa bei der Lektüre eines gedruckten Buches der Fall ist. Auch unter finanziellen Aspekten ist die mediale Realisierung von Bedeutung.

Strukturelle und paratextuelle Aspekte

Die strukturellen und paratextuellen ASPEKTE sind allesamt in ihrer Rolle zur Herstellung von Kohärenz und damit letztendlich zur Vermittlung der Textfunktion des Hypertextes zu beschreiben. Dies gilt insbesondere auch für die Analyse des Screendesign: Nicht ästhetische Merkmale sondern funktionale müssen auch hier im Vordergrund stehen.

14.4.2 Anmerkungen zu den Ebenen *Knoten, Absatz und Satz*

Auf Knoten-Ebene wird das Thema häufig durch eine Überschrift erster Ordnung signalisiert. Wo dies nicht der Fall ist oder die Überschrift das Thema augenscheinlich nicht korrekt wiedergibt, kann zusätzlich eine Analyse der Makrostruktur bzw. der referentiellen Bewegung erfolgen.

Auf den Ebenen *Knoten, Absatz* und *Satz* finden sich auch mögliche ‚Schnittstellen‘ von textinternen zu textexternen Phänomenen: Zu denken ist hier an kohäsive Lücken, die nur durch Inferenzen auf Leserseite gefüllt werden

können, oder an die Struktur der im Text vorkommenden Isotopieebenen, die mit den kognitiven Schemata der Leser in Einklang zu bringen sind.

14.4.3 Anmerkungen zur Ebene *Link*

Aspekte des Bezugsbereichs

Die Ebene *Link* bezieht sich einerseits auf den LINK-TEXT (bzw. den LINK-ANKER, der durch Zeichen eines sekundären semiotischen Systems, z.B.: Piktogramme, Farbe etc. dargestellt wird). Der Link-Text ist somit Teil des Ausgangs-Knotens und wird primär dort auf der Ebene *Knoten* (bzw. den Sub-Ebenen *Absatz* und *Satz*) untersucht.

Andererseits bezieht sich die Ebene *Link* auf die Relation, welche aufgrund dieses Links zwischen Ausgangs- und Ziel-Knoten etabliert wird. Zuerst werden hierbei Aspekte des Bezugsbereichs des Links im Ausgangs-Knoten sowie im Ziel-Knoten besprochen. Es geht um die Frage, welche Teile bzw. welche Aspekte von Knoten miteinander verknüpft werden. Zur Verdeutlichung: In einer Lesesequenz verknüpft beispielsweise ein navigationsspezifischer Link – etwa mit dem Link-Text *weiter* – den gesamten Ausgangs-Knoten mit dem gesamten Ziel-Knoten. Demgegenüber verknüpft ein Link vom Typ *Definition* lediglich einen kleinen Teil des Ausgangs-Knoten, nämlich den zu klärenden Terminus – der zumeist auch als Link-Text dient – mit dem Ziel-Knoten (bzw. einem Teil desselben). Solche Fragestellungen sind von eher strukturellem Charakter.

Mit ihnen kann eine Analyse des Links in seiner Funktion als deiktischer Ausdruck einhergehen: Allgemein kann jeder Link als ein Mittel der Deixis gesehen werden, da er auf andere sprachliche (oder auch nichtsprachliche) Zeichen innerhalb desselben oder eines anderen Knotens verweist. Insofern können Links mit deiktischen Adverbialausdrücken wie *dort* oder *oben* verglichen werden. Links können nun darauf hin untersucht werden, ob diese deiktische Komponente im Vordergrund steht oder nur implizit vorhanden ist. Zur Veranschaulichung diene abermals das obige Beispiel: Navigations-Links existieren primär dafür, um auf andere Stellen im Hypertext zu ‚deuten‘. Inhaltlich oder funktional definierte Links deuten natürlich implizit auch auf andere Stellen im Hypertext, ihre primäre Aufgabe ist jedoch die Verdeutlichung der inhaltlichen oder funktionalen Aspekte – wie dies beim oben angeführten Beispiel *Definition* der Fall ist.

Es können auch noch zusätzliche Aussagen über die Perspektive der Verknüpfung gemacht werden: nämlich, ob der Link eher die funktional-kommunikativen Aspekte der Knoten verknüpft – wie in obigem Beispiel *Definition* gezeigt – oder eher eine Verknüpfung der thematisch-strukturellen Aspekte herstellt, wie dies bei Links der Fall ist, die eine bestimmte Überschrift eines Übersichts-Knotens als Link-Text und einen Inhalt-Knoten als Ziel haben. Stehen kommunikativ-funktionale Verknüpfungsaspekte im Vordergrund, kann man von ILLOKUTIONÄREN LINKS sprechen, bei Dominanz thematisch-struktureller, inhaltlicher Aspekte von PROPOSITIONALEN LINKS (siehe hierzu auch Bucher 1999 S.24).

Thematisch-strukturelle Aspekte

Gestaltet sich die Analyse der thematisch-strukturellen Aspekte, also der Themenentfaltung, über die Knoten-Grenze hinweg als schwierig, so kann auch hier das Modell der referentiellen Bewegung als Instrumentarium eingesetzt werden. Untersuchungen, die sich hauptsächlich mit dem Problem der Kohärenz bei der Navigation in Hypertexten beschäftigen – dem momentan populärsten Gebiet der hypertext-linguistischen Betätigung –, werden sich wohl hauptsächlich auf diesen Beschreibungsaspekt von TAH stützen. Vorteil des TAH ist auch hier die klare Abgrenzung der Kategorien. Eine Beschreibung des Kohärenzaufbaus kann – von thematisch-strukturellen Aspekten ausgehend – wiederum Rückschlüsse zulassen auf kommunikativ-funktionale, kohäsive und paratextuelle Bereich.

Paratextuelle und metasprachliche Aspekte

Paratextuelle und metasprachliche Aspekte von Links beziehen sich auf das semiotische System, mit denen Link-Typen kenntlich gemacht und spezifiziert werden; also etwa, ob Piktogramme oder Mouse-Over-Texte zur Visualisierung verwendet werden. Diese Punkte können im Zusammenhang mit den browser-spezifischen Aspekten des Links untersucht werden.

Teil C: Fallbeispiel – Typisierte Links

Methodisches Hauptziel des Fallbeispiels in Teil C ist ein erster Einsatz des TAH anhand einer konkreten Fragestellung um herauszufinden, wie sich die schematischen Beschreibungsaspekte und Analysekategorien mit der Wirklichkeit komplexer Hypertexte decken – und besonders auch, welche Verhältnisse mit diesem Analysemodell nicht zu erfassen sind.

Als Fragestellung wurde die Problematik der sogenannten TYPISIERTEN LINKS gewählt. Typisiert sind Links dann, wenn sie auf irgendeine Art und Weise markiert sind und dadurch etwas über Funktion, Inhalt, Struktur o.ä. des Ziel-Knotens in Relation zum Ausgangs-Knoten aussagen. In der Literatur werden typisierte Links durchweg positiv bewertet, etwa von Storrer 2001c:

"Selektivität und Individualisierung bedeutet natürlich auch, dass Hypertext-Nutzer stets aufs Neue entscheiden müssen, welchen der auf einem Modul angebotenen Links sie verfolgen möchten. Surprise-Links, also Links mit ungewissem Ziel, können in spielerischen Hypertexten ihren Reiz ausüben. Für WWW-Nutzer, die möglichst schnell möglichst viele oder ganz spezifische Informationen zu einem Thema abrufen möchten, ist es hingegen äußerst hilfreich, zu wissen, wohin ein Link führt, welcher Aspekt des Link-Ursprungs mit welchem Aspekt des Link-Ziels verknüpft ist und was den Autor dazu bewogen hat, den Link anzubieten. [...] Je besser dies deutlich wird, um so schneller kann der Nutzer erkennen, welche der Links für seinen Bedarf relevant sind."

(Storrer 2001c Abschnitt *Linkexplikation: Aufgabe und Gestaltungsoptionen*)

Erschöpfende Ergebnisse zu Art und Anzahl dieser typisierten Links liegen jedoch noch nicht vor, insofern besteht hier noch Forschungsbedarf für mehrere Disziplinen. Diese textlinguistische Arbeit untersucht nun, ob und wie hypertextuelle Links der Corpus-Hypertexte mit zusätzlicher Information angereichert sind und um welche Art von Information es sich dabei handelt. Die unterschiedlichen Typen werden hierbei auf passende Beschreibungsaspekte von TAH abgebildet und mit dessen Analysekategorien beschrieben.

Im einzelnen wird hierbei wie folgt vorgegangen:

Für die argumentative Progression dieser Untersuchung eher als ergänzend, inhaltlich aber als äußerst wichtig anzusehen ist das ziemlich technisch ausgelegte nächste Kapitel 15. Hier werden heutige und mögliche zukünftige Implementierungen vorgestellt, mit denen das Konzept der typisierten Links im WWW in die Praxis umgesetzt werden kann. Dort geschieht auch eine ‚technische Verortung‘ des Fallbeispiels (siehe Kapitel 15.1.3): Es sollen keine typisierten Relationen auf einer, wie auch immer gearteten, Ebene der

TIEFENSTRUKTURELLEN WISSENSREPRÄSENTATION untersucht werden. Ansätze und Technologien, die in diese Richtung gehen, werden lediglich kurz angesprochen. Als Untersuchungsebene dieser Arbeit wird vielmehr die OBERFLÄCHENSTRUKTUR von Hypertexten festgelegt – also das, was der Leser tatsächlich sieht.

Nach dieser grundsätzlichen technischen Einordnung der textlinguistischen Analyse des Fallbeispiels, werden die konkreten Implementierungen dargestellt: Nachdem bereits bei der Vorstellung des Untersuchungsgegenstandes die grundlegende Philosophie von HTML angesprochen wurde, sollen in Kapitel 15.2 nun die Aspekte dieser Sprache vertieft werden, die für das Setzen und Typisieren von Links relevant sind. Denn: Bedingt ist die Art der Oberflächenstruktur von den Möglichkeiten der dahinter stehenden Markup-Sprache. Am Rande kommen auch neuere Aspekte des LINKING im Rahmen des Link-Modelles XLINK von XML zur Sprache. In einem nächsten Schritt werden die konkreten Link-Implementierungen der Corpus-Hypertexte sowie eine Auswahl möglicher zukünftiger Umsetzungen erläutert.

Nach diesen technischen Vorüberlegungen geht die Arbeit zu den inhaltlichen über: Hierfür werden einige allgemeine Vorschläge für Vokabularien zur Link-Typisierung vorgestellt (siehe Kapitel 16). Dabei handelt es sich um relativ heterogene Ansätze aus Informationswissenschaft und Linguistik. Manche wurden allgemein zur Beschreibung der inhaltlichen Verbindungen von Teiltexten entwickelt, andere speziell auf die Aufgabe der Link-Typisierung hin optimiert. Gemein haben die Ansätze, daß sie alle von der Sekundärliteratur schon einmal zur Typisierung von Links vorgeschlagen oder in Hypertext-Systemen bereits umgesetzt wurden. Als Abschluß dieser Vorüberlegungen werden die Link-Vokabularien ausgewählt, die in die praktische Untersuchung in Kapitel 16.4 einfließen sollen.

Die praktische Analyse selbst geht nun wie folgt vor: Nach einer allgemeinen und link-spezifischen Untersuchung der Hypertexte mit TAH werden diese Vokabularien auf einige Knoten der Corpus-Hypertexte angewendet. Am Ende steht ein Vergleich der proprietären und der allgemeinen Typisierungen. Als Ergebnis werden erste Vorschläge geliefert, in welchem Rahmen sich ein allgemeiner Katalog von Link-Typen zur Auszeichnung der Oberflächenstruktur von Hypertexten bewegen könnte. Im Rahmen dieses knappen und zusammenfassenden ‚rhetorischen Teils‘ wird versucht, die Ergebnisse der Untersuchung in ‚Regieanweisungen‘ für Hypertext-Autoren umzusetzen.

15 Technische Voraussetzungen: Links im WWW

Bevor zu den eigentlichen inhaltlichen Vorüberlegungen zur Erstellung möglicher Vokabularien typisierter Links übergegangen werden kann, muß, wie gesagt, zuerst die ‚technische Fixierung‘ der Analyse auf die Ebene der Oberflächenstruktur von Hypertexten erfolgen. Grundsätzliche Unterscheidungsmöglichkeiten der einzelnen Ebenen von Hypertext-Implementierungen stellt das nächste Kapitel vor.

Es ist zwar wünschenswert, jedoch für das Verständnis der textlinguistischen Analyse nicht unbedingt notwendig, sich sämtliche Kapitel zu den unterschiedlichen Implementierungen zu Gemüte zu führen. Dem technisch weniger interessierten Leser ist jedoch auf alle Fälle die Lektüre des Kapitels 15.2.1 naheulegen, da die Hypertexte des Corpus‘ in HTML codiert sind, und daher ein Grundverständnis der Prinzipien dieser Sprache doch hilfreich ist.

15.1 Datenhaltung, Zwischenschicht und Oberflächenstruktur

In diesem Kapitel wird der Problembereich der praktischen Untersuchung auf der Ebene der OBERFLÄCHENSTRUKTUR eines HTML-Hypertextes lokalisiert, welche nach semantischen, layout- und browser-spezifischen Aspekten des TAH analysiert wird.

15.1.1 Von HTML zum Semantic Web¹

HTML als Sprache der meisten aktuellen Hypertexte im WWW (siehe Kapitel 4.2) war ursprünglich als reine Strukturbeschreibungssprache konzipiert. Von diesem Grundsatz wurde jedoch bald abgewichen, indem immer mehr Elemente zur physischen Textauszeichnung bereits in der Frühphase der Sprachentwicklung hinzugenommen wurden.²

¹ In diesem Kapitel werden lediglich die Grundlinien der Entwicklung von HTML, über XML hin zum sogenannten SEMANTIC WEB angesprochen. Die Technologien selbst werden in Kapitel 15.2 eingeführt.

² Dies führt dazu, daß die Zwischenschichten des untenstehenden Modells in der Praxis nicht exakt getrennt sind.

Eine zusammenfassende Darstellung der Ziele und Entwicklung der ersten HTML-Versionen befindet sich im WWW unter <http://www.huberoliver.de/historisch.html>. Von dort findet man auch leicht Zugang zu den Primärtexten aus der ‚WWW-Prähistorie‘. Die grundlegenden Prinzipien und Begriffe der Makrosprache HTML werden in Kapitel 15.2 eingeführt. Die einschlägigen Begriffe sind zusätzlich im Glossar aufgeführt.

Die relativ neue Sprache XML (siehe Kapitel 15.2.2) dient hingegen unter anderem dazu, Struktur und Darstellung von Dokumenten strikt zu trennen: Ein XML-Dokument enthält die semistrukturierten Daten, die Darstellung wird ausschließlich von STYLE-SHEETS übernommen. Da XML-Elemente vom Autor selbst benannt werden können – HTML schreibt dagegen ein festes Vokabularium benannter Elemente vor –, kann dies mit – für den menschlichen Leser – ‚sinnvollen‘ weil bedeutungstragenden Lexemen erfolgen. Computer können aber mit diesen Bedeutungen per se erst einmal nichts anfangen. Die Initiative SEMANTIC WEB¹ des W3C arbeitet momentan mit Nachdruck daran, Netzwerke von Ressourcen in Netze von Bedeutungen zu transferieren, welche auch von Computern ‚verstanden‘ werden können. Zu diesem Zweck hat das W3C die formale Sprache RDF geformt. Dieses RESSOURCE DESCRIPTION FRAMEWORK ist eine in XML definierte Sprache zur Beschreibung von Metadaten der Ressourcen:

“Meaning is expressed by RDF, which encodes it in sets of triples, each triple being rather like the subject, verb and object of an elementary sentence. These triples can be written using XML tags. In RDF, a document makes assertions that particular things (people, Web pages or whatever) have properties (such as ‘is a sister of’, ‘is the author of’) with certain values (another person, another Web page).”

(Berners-Lee u. Lassila 2001 Abschnitt *Expressing Meaning*)

Um sicherzustellen, daß Leute, die dasselbe ‚meinen‘ auch dasselbe ‚sagen‘, werden gemeinsame kontrollierte Vokabularien verwendet, die bestimmte Wissensbereiche beschreiben – sogenannte ONTOLOGIEN. Für Erstellung, Austausch sowie das Referenzieren solcher Ontologien gibt es formale Vorschläge wie die in SGML formulierte HTML-Erweiterung SHOE² oder das in RDF formulierte DAML+OIL³.

15.1.2 Die Ebenen einer Link-Typisierung

Die semantische Auszeichnung von Ressourcen, und damit auch die von Links, kann auf mehreren Ebenen stattfinden:⁴

¹ Information im WWW unter <http://www.w3.org/2001/sw/>.

² Informationen im WWW unter <http://www.cs.umd.edu/projects/plus/SHOE/spec.html>.

³ Informationen im WWW unter <http://www.w3.org/TR/daml+oil-reference>.

⁴ Die Begriffe TIEFENSTRUKTUR und OBERFLÄCHENSTRUKTUR sind hierbei nicht aus der Terminologie der Generativen Transformationsgrammatik entnommen. Es handelt sich auch nicht um breit eingeführte Termini der Informatik. Vielmehr wurden die Begriffe in dieser Arbeit aufgrund ihrer intuitiven Eingängigkeit zur Visualisierung der Ebenen von Hypertext-Systemen verwendet.

- (1) TIEFENSTRUKTUR
- (2) ZWISCHENSCHICHT
- (3) OBERFLÄCHENSTRUKTUR

Vereinfacht gesagt, bezieht sich der Begriff TIEFENSTRUKTUR auf die Datenhaltung der semantischen Information von Knoten und Relationen – die konkrete Implementierung derselben ist an dieser Stelle eher nebensächlich. Auf dieser Ebene könnte die Umsetzung des zukünftigen Semantic Web erfolgen.

Mit dem Begriff OBERFLÄCHENSTRUKTUR wird hingegen auf die eigentliche Benutzeroberfläche referiert, also auf ‚alles, was der Benutzer tatsächlich sieht‘. Hiermit sind die semantischen, layout-spezifischen und die browser- bzw. navigationsspezifischen Aspekte der Hypertext-Oberfläche gemeint.

Mit dem Begriff ZWISCHENSCHICHT ist der HTML-Quellcode gemeint, der – mit Hilfe von Style-Sheets – für die Anzeige und das Browser-Verhalten auf Ebene der Oberfläche verantwortlich ist. Das Übersetzen vom Quellcode in den sichtbaren Hypertext wird hierbei von der Browser-Software übernommen. Folgendes Schaubild¹ soll die Verhältnisse visualisieren:

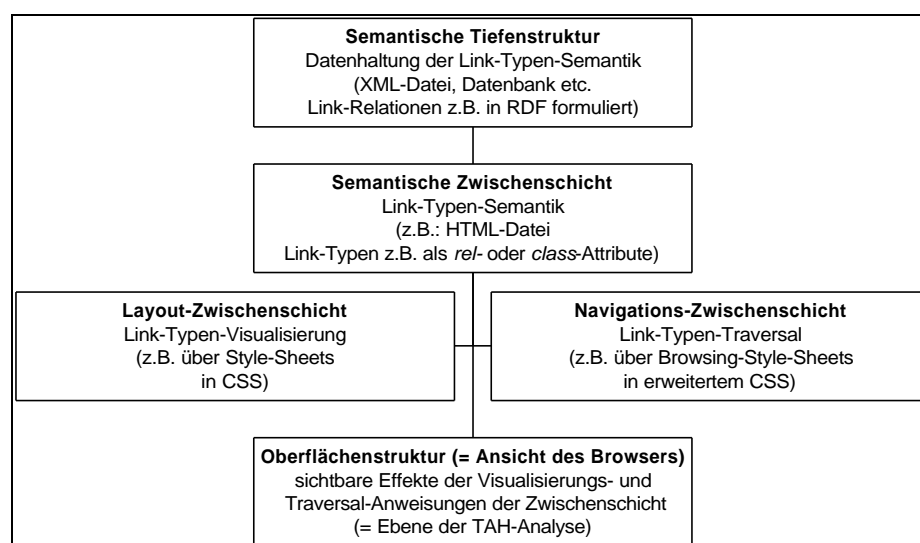


Abbildung 7: Ebenen der Typisierung

Ein komplexes Hypertext-System, das mit semantischer Information angereichert ist, könnte eine Link-Typisierung also wie folgt vornehmen:² Auf Ebene der Tiefenstruktur sind Relationen semantisch mittels eines kontrollierten

¹ Als ein Ergebnis der praktischen Untersuchung in Kapitel 17 wird eine Modifizierung der hier abgebildeten Zwischenschicht vorgeschlagen (siehe hierzu Abbildung 31).

² Die verwendeten Technologien sind hierbei lediglich als beispielhaft zu verstehen. Auf eine Darstellung der Transformationstechniken, etwa XSLT, wird verzichtet.

Vokabulars ausgegliedert und in einer formalen Semantikbeschreibungssprache (etwa RDF oder XML-TOPIC-MAPS, siehe Kapitel 15.2.5) formuliert. Diese Information wird dann der semantischen Zwischenschicht ‚vererbt‘ – wobei hier Veränderungen bzw. Vereinfachungen stattfinden können (siehe unten). Die layout- und browser- bzw. navigationsspezifischen Style-Sheets bekommen ihrerseits wiederum diese Information ‚vererbt‘ und sorgen, in Abhängigkeit vom Link-Typ, für eine bestimmte Darstellung des Links – bzw. ein bestimmtes Navigationsverhalten beim Klick auf den entsprechenden Link auf der Ebene der Hypertext-Oberfläche.

Zu bemerken ist, daß nicht jedes Hypertext-System über sämtliche der oben dargestellten Ebenen zur Link-Typisierung verfügt. Die obige Darstellung ist der vereinfachte Entwurf eines Hypertext-Systems im Sinne einer möglichen technologischen Umsetzung eines SEMANTIC WEB (siehe unten). Die Hypertexte des ‚Corpus‘ verfügen lediglich über die Ebenen ZWISCHENSCHICHT und OBERFLÄCHENSTRUKTUR.¹

Eben Gesagtes war sehr komprimiert dargestellt und verlangt daher nach weiterer Explikation der einzelnen Ebenen. Dies soll nun geschehen, Hauptfrage ist hierbei, auf welcher Ebene eine TAH-Untersuchung von typisierten Links in HTML-Hypertexten stattfinden soll. Zur Beantwortung müssen die Konzeptionen der einzelnen Ebenen genauer betrachtet werden:

ad (1): Die Ebene, die ich als SEMANTISCHE TIEFENSTRUKTUR bezeichne, ist in etwa mit dem vom W3C geplanten SEMANTISCHEN WEB vergleichbar: Mit Hilfe von kontrollierten Bezeichnern und Verknüpfungen aus standardisierten Ontologien werden Ressourcen und deren Beziehungen (die Hyper-Links entsprechen können, dies aber nicht müssen!) typisiert. Allgemein formuliert: Das Semantische Web ist so etwas wie die WISSENSREPRÄSENTATION der im WWW adressierbaren Ressourcen. Intelligente Software könnte aus dieser Information dann Schlußfolgerungen ziehen. Hierbei wäre etwa an logische Schlußfolgerungssprachen wie PROLOG zu denken, mit denen die Aussagen von RDF einfach in Klauseln zu implementieren und durch Regeln zu verbinden und abzufragen wären: so etwa die Beispielsrelation *is sister of* aus dem etwas weiter oben angeführten Zitat von Berners-Lee u. Lassila 2001:

¹ Diese Aussage bezieht sich nur auf die Art und Weise, wie Typen von Links in den Hypertexten gespeichert werden! Eine ‚tiefenstrukturelle‘ Datenhaltung ist durchaus bei einigen der Corpusvertreter vorhanden, so werden etwa die HTML-Seiten von SPORT1 komplett aus Datenbanken generiert.

```
is_sister_of(RESSOURCE1, RESSOURCE2):-  
  is_parent_of(RESSOURCE3, RESSOURCE1),  
  is_parent_of(RESSOURCE3, RESSOURCE2).
```

In einem solchen zukünftigen semantischen Web wäre es dann auch prinzipiell möglich, von den Beziehungstypen auf Ebene dieser Wissensrepräsentation – also der semantischen Tiefenstruktur – die Beziehungstypen für eine andere Ebene abzuleiten, nämlich für die der semantischen Zwischenschicht. Daß hierfür eine Eins-zu-Eins-Übernahme bei der Transformation jedoch wohl kaum sinnvoll wäre, wird im nächsten Abschnitt dargestellt.

ad (2): Die Trennung von semantischer, Layout- sowie Navigations-Zwischenschicht ist in den Corpus-Hypertexten zumeist nicht konsequent durchgeführt (siehe oben), die semantische Zwischenschicht (= der HTML-Quellcode) ist daher durchsetzt von Angaben zu Layout und Browser-Verhalten.

Die erste Teilstruktur der Zwischenschicht ist der HTML-Quellcode, in dem semantische Informationen der Link-Typen abgelegt sein können. Der Quellcode selbst kann hierbei aus einer tiefenstrukturellen Datenbank erzeugt sein. Ein naheliegender Gedanke wäre nun, die Link-Typen schlicht aus der semantischen Tiefenstruktur abzuleiten. Diese ‚Ernte‘ könnte aber unter Umständen zu reichhaltig ausfallen. Noch ist es zu früh vorherzusagen, wie sich die Ontologien im Netz entwickeln werden. Zu mutmaßen ist jedoch, daß sehr komplexe und äußerst heterogene, da domänenspezifische Systeme mit Termini und Beziehungen auf Ebene der semantischen Tiefenstruktur entstehen werden. Als Beispiel sei hier nur die etwas exotische, in SHOE formulierte *Bier-Ontologie* genannt, die Relationen wie *brews(Brewery, Beer)*, *awardedAt(Award, Festival)* oder *sponsors(Association, Festival)*¹ definiert.

In Repräsentationen der Tiefenstruktur machen solche stark domänenspezifischen Relationen natürlich Sinn, immerhin sollen ja in Logik programmierte Software-Agenten dadurch ‚Wissen‘ erwerben; der menschliche Leser dürfte aber mit ständig wechselnden Relationstypen beim Navigieren in Hypertexten kognitiv deutlich überfordert sein. Ich plädiere somit für ein kleines Set allgemeiner Link-Typen, das für eine möglichst große Anzahl von Hypertexten ‚Sinn‘ für den Leser macht. Was mit dieser Forderung natürlich

¹ Informationen im WWW unter <http://www.cs.umd.edu/projects/plus/SHOE/onts/beer1.0.html>.

nicht ausgeschlossen werden soll, ist eine Transformation der domänenspezifischen Relationen der Tiefenstruktur in genau ein solches kleines Set allgemeiner Link-Typen, welche der Browser dann dem Leser anzeigt. Zu diesem Zweck muß jedoch erst ein Inventar solcher allgemeiner Link-Typen erstellt und dieses theoretisch und empirisch auf dessen Plausibilität hin geprüft werden. Wie diese Links dann in zukünftigen Systemen implementiert werden, ist erst einmal zweitrangig.

Zusammenfassend kann gesagt werden: Auf Ebene der semantischen Zwischenschicht von typisierten Links wird codiert, welche Typen für welche konkreten Links verwendet werden. Das Layout der einzelnen Typen wird idealiter dann auf der Ebene der Layout-Zwischenschicht (etwa mit in CSS formulierten Style-Sheets) festgelegt, während die Aktionen, die von der Browser-Software bei Klick auf einen bestimmten Link-Typen ausgeführt werden, auf der Ebene der Navigations-Zwischenschicht festgelegt sind. Eine Kopplung von semantischer und browser-spezifischer Information könnte man sich auf Ebene der Zwischenschicht(en) etwa so vorstellen: Der Ziel-Knoten eines Links des Typs *Definition* wird stets in einem neuen Fenster geöffnet, während ein Klick auf einen Link des Typs *Hintergrundinformation* den Inhalt im ursprünglichen Fenster darstellt. Die Beziehung von Semantik und Travesierung könnte hierbei in sogenannten BROWSING-STYLE-SHEETS festgelegt werden (siehe auch Bry u. Kraus 2002).

ad (3): Die OBERFLÄCHENSTRUKTUR ist der für den Benutzer sichtbare Teil des Hypertextes, welcher durch die Browser-Software aus den Elementen der Zwischenschicht erzeugt wurde. Auf dieser Ebene stellen sich Fragen nach Art und Anzahl der sichtbaren Link-Typen sowie nach deren (eventuell unterschiedliche) Gestaltung. Die Signalisierung des Link-Typs kann durch Mouse-Over-Texte, farbige Markierungen oder Piktogramme geschehen. Denkbar wären aber auch ‚exotischere‘ Varianten wie Sprachausgabe oder ähnliches. Die Oberflächenstruktur ist Ziel einer Analyse mit TAH, mehr hierzu folgt im nächsten Abschnitt.

15.1.3 Die Oberflächenstruktur als Problembereich der Untersuchung

Berners-Lee zufolge könnte das zukünftige SEMANTIC WEB den Zugang zu Information in noch größerer Weise revolutionieren, als es bereits das WWW getan hat:

„Wenn HTML und das Web dafür gesorgt haben, daß alle Onlinedokumente wie ein riesiges Buch aussehen, werden RDF-, Schema- und Schlußfolgerungssprache dafür sorgen, daß alle Daten auf der ganzen Welt wie eine immense Datenbank aussehen.“

Berners-Lee 1999a S. 270

Das heutige WWW würde sich damit in eine riesige, semistrukturierte Datenbank verwandeln. Informationen aus einer Datenbank gewinnt man durch Anfragen. Solche Anfragen zu erleichtern, ist Ziel des Semantic Web. Auf der semantischen Tiefenstruktur greifen die Werkzeuge des Information Retrievals: die Suchmaschinen. Das Semantic Web wird die Recall- und Relevance-Werte dieser Maschinen aller Voraussicht nach deutlich verbessern.

Eine textlinguistische Untersuchung mit TAH beschäftigt sich aber mit dem Lesen von Hypertexten, nicht mit der Informationsgewinnung aus Hypertext-Basen mittels Suchanfragen oder SOFTWARE-AGENTEN: Sie befaßt sich demnach nicht mit der semantischen Tiefenstruktur. Auch die Zwischenschicht, des HTML-Quellcodes sowie die zugehörigen Style-Sheets sind nicht primär Gegenstand dieser Untersuchung: Zwar werden Möglichkeiten besprochen, auf dieser Ebene semantische Informationen von Links zu codieren (samt layout- und navigationsspezifischen Zuordnungen via Style-Sheets) – aber nur im Hinblick auf das sich dem Leser tatsächlich bietende Erscheinungsbild des Hypertextes samt dessen Links! In diesem Sinne sind sämtliche technischen Erläuterungen der nächsten Kapitel zu verstehen.

Gegenstand der Arbeit ist also die Oberflächenstruktur der Corpus-Hypertexte. Hier kann wiederum differenziert werden in eine semantische, layout- und browser-spezifische Oberflächenstruktur – jeder dieser Aspekte der Hypertext-Oberfläche wird wiederum von einem oder mehreren der Beschreibungsaspekte des TAH erfaßt.

Daß zu irgendeinem Zeitpunkt aus den speziellen Relationen der Wissensrepräsentation Relationen für Zwischenschicht und semantische Oberfläche abgeleitet werden, ist anzunehmen: Eine doppelte Codierung der Relationen in Tiefenstruktur und Zwischenschicht wäre nämlich redundant und erschwerte die Datenkonsistenz. Insofern kann die Prognose gegeben werden, daß die in der Arbeit vorgeschlagenen Relationen im zukünftigen semantischen Web wohl eher bei der Transformation von Tiefenstruktur zur Zwischenschicht eine Rolle spielen, wenn die spezielleren Relationen der Wissensrepräsentation durch allgemeinere Relation ersetzt werden (können), welche die kognitive Textverarbeitung des Lesers unterstützen. Eine auf Ebene der semantischen

Zwischenschicht autonom stattfindende Codierung wird es in solchen Systemen wohl nicht mehr geben.¹ Genau auf diese Weise codieren jedoch die meisten heutigen HTML-Hypertexte die Information, die für eine Link-Typisierung nötig sind.

In den nächsten Kapiteln werden die hier nur umrissenen Technologien HTML und XML etwas genauer eingeführt werden. Dies ist aus folgendem Grund notwendig: Zwar betrachtet eine TAH-Analyse Tiefenstruktur und Zwischenschicht von Hypertexten mehr oder weniger als ‚Blackbox‘, befaßt sich also ausschließlich mit der Oberflächenstruktur; dennoch ist ein Grundverständnis der Realisierung solcher HTML-Hypertext-Systeme sinnvoll, um zu erkennen, daß die Hypertext-Oberfläche im Browser nicht im ‚luftleeren‘ Raum existiert: Technische Prämissen besonders der Zwischenschicht – also der Markup-Sprache HTML – bedingen nämlich zum großen Teil das Erscheinungsbild der Oberfläche, im konkreten Beispiel etwa, ob und wie Links dort typisiert erscheinen können.

15.2 Linking in HTML und XML

Im letzten Kapitel wurden die möglichen Ebenen von Hypertext-Systemen vorgestellt. Dabei wurde festgehalten, daß die Hypertexte des Corpus’ lediglich den Quellcode der Zwischenschicht sowie eine semantische, layout-spezifische und browser-spezifische Oberflächenstruktur, nicht aber eine angebundene semantische Tiefenstruktur besitzen. Hauptaugenmerk wird daher auf die Technologien auf Ebene der Zwischenschicht gerichtet sein, die zur Generierung der Oberfläche der Corpus-Hypertexte nötig sind. Besonders relevant sind hier die Markup-Sprache HTML sowie – mit Abstrichen – die zugehörige Style-Sheet-Sprache CSS.

Ein kurzer Ausblick wird aber auch auf die Technologien gegeben, welche für die immer wichtiger werdende semantische Tiefenstruktur relevant sind: Gemeint sind hier hauptsächlich die Meta-Markupsprache XML, das zugehörige Relationsmodell XLINK sowie die XML-Anwendungen RDF und XML-TOPIC-MAPS.

¹ Die Semantik der Zwischenschicht und damit die der Oberfläche würde, wie beschrieben, ‚on the fly‘ bei der Transformation erzeugt werden

15.2.1 HTML als SGML-Anwendung

HTML steht für HYPERTEXT MARKUP LANGUAGE¹. Das bedeutet, daß mittels HTML Hypertexte in einer standardisierten Form für das WWW aufbereitet werden können. Um die Fähigkeiten von HTML richtig deuten zu können, muß zuerst eine andere Sprache kurz vorgestellt werden: die STANDARD GENERALIZED MARKUP LANGUAGE (SGML), eine formale METASPRACHE zur Definition anderer Sprachen.² Dies geschieht, indem eine GRAMMATIK für die zu definierende Sprache entwickelt wird. Eine solche Grammatik wird als DOCUMENT TYPE DEFINITION (DTD) bezeichnet.³ Diese DTD deklariert die erlaubten ELEMENTE sowie deren Kombinationsmöglichkeiten. Eine vereinfachte (und formal unvollständige!) Grammatik zur Beschreibung einer kleinen Teilmenge möglicher Sätze des Deutschen könnte wie folgt in einer SGML-DTD formuliert werden:⁴

```
<! ELEMENT SATZ -- (NP, VP)>
<! ELEMENT NP -- (ARTIKEL, NOMEN)>
<! ELEMENT VP -- (VERB, NP)>
<! ELEMENT VERB -- (#PCDATA)>5
<! ELEMENT NOMEN -- (#PCDATA)>
<! ELEMENT ARTIKEL -- (#PCDATA)>
```

DTDs stellen also ähnliche Ersetzungsregeln zur Verfügung wie Phrasenstrukturgrammatiken:

¹ Die ‚offizielle‘ Seite für HTML ist die des W3C [<http://www.w3.org/MarkUp/>]. Hier befinden sich neben den Spezifikationen und Grammatiken der unterschiedlichen HTML-Versionen auch umfangreiches Lernmaterial sowie weiterführende Links. Im deutschsprachigen Raum stellt Münz 2001 so etwas wie eine Standardstartseite dar, um den Umgang mit HTML zu erlernen [<http://selfhtml.teamone.de/>].

² Das W3C bietet eine umfangreiche Einstiegsseite für SGML an [<http://www.w3.org/MarkUp/SGML/Overview.html>]. Den meiner Meinung nach besseren Überblick gewährt jedoch die SGML-Sektion der Oasis-Seiten [<http://www.oasis-open.org/cover/general.html#overview>]. Die wohl umfangreichste Seite zu XML bieten die *XML Coverpages* [<http://xml.coverpages.org/>]. Hier finden sich zahlreiche Tutorials und sämtliche relevanten Links zum Thema.

³ Die Begriffe GRAMMATIK und DTD werden im folgenden synonym verwendet. Lobin 2000 gibt einen guten, linguistisch orientierten Einstieg in die Problematik des Designs von DTDs mit SGML und XML.

⁴ Die folgenden Beispiele sind in leicht veränderter Form entnommen aus Witt 1999 S. 123 f.

⁵ PCDATA steht für PARSED CHARACTER DATA, d.h. für beliebigen Text. Bei einer tatsächlichen DTD-Grammatik müßte freilich gewährleistet sein, daß für die terminalen Symbole (etwa NOMEN) auch nur die passende kategoriale Füllung (z.B. *Mann*, *Auto*) erlaubt sein darf. Dieser einfache DTD-Auszug soll aber lediglich das grundlegende Prinzip demonstrieren.

SATZ -> NP + VP
 NP -> ARTIKEL + NOMEN
 VP -> VERB + NP
 ARTIKEL -> Lexem
 NOMEN -> Lexem
 VERB -> Lexem

Die DTD kann also als Automat zur Produktion von Sätzen mit einer bestimmten Grammatik aufgefaßt werden. Folgende unmittelbare Konstituenten (bzw. terminale Symbole) werden nun exemplarisch als Vokabular zugefügt:¹

NOMEN -> *man* | *ball*
 ARTIKEL -> *a* | *the*
 VERB -> *hit* | *took*

Mit diesem Vokabular und der DTD-Grammatik wäre es also möglich, einen Satz wie *the man hit the ball* zu produzieren. Begreift man nun jedes Element der DTD als ‚benannten Container‘ mit fest definiertem Inhalt, so ließe sich der Satz auch wie folgt visualisieren:

```
<SENTENCE>
<NP>
  <ARTIKEL>the</ARTIKEL>
  <NOMEN>man</NOMEN>
</NP>
<VP>
  <VERB>hit</VERB>
  <NP>
    <ARTIKEL>the</ARTIKEL>
    <NOMEN>ball</NOMEN>
  </NP>
</VP>
</SENTENCE >
```

Hierzu einige Bemerkungen: Wird ein Element der DTD verwendet, so geschieht das immer in der folgenden ‚Container-Form‘ *<Elementname>Elementinhalt</Elementname>*². Formal wichtig ist hierbei, daß der

¹ Der senkrechte Strich steht hierbei für logisch *oder*.

² Der ‚Container‘ selbst, also das Klammerpaar *<elementname></elementname>*, wird in der Informatik meistens als TAG, ELEMENT oder ETIKETT bezeichnet. Meiner Meinung nach aber visualisiert die Bezeichnung *Container* den tatsächlichen Sachverhalt besser: In einem großen Container *Sentence* sind zwei kleinere Container *NP* und *VP* enthalten. Der Container *NP* enthält beispielsweise wiederum den Container *NOMEN*, welcher nun nicht mehr weiter verschachtelt ist, sondern als Inhalt ein Lexem enthält. In den meisten Kontexten dieser Arbeit wird dennoch der Begriff ELEMENT benutzt, da er in den meisten deutschen Publikationen zum Thema auch verwendet wird.

Inhalt stets von zwei mit dem Elementnamen benannten Klammerpaaren eingerahmt ist und daß diesem Namen in der zweiten Klammer ein Schrägstrich vorangestellt ist. Je nach Festlegung in der DTD-Grammatik kann jeder Container entweder textuelle Daten – in unserem Fall also Lexeme – oder aber weitere Container enthalten (wie z.B. *NP* und *VP*).¹

Auch wenn nur ein Minimum der Möglichkeiten einer formalen Metasprache wie SGML in obigen Ausführungen angedeutet wurde, sollte das Prinzip offensichtlich sein: Eine für alle zugängliche² Grammatikdefinition regelt die Menge aller möglichen gültigen Sätze (bzw. Dokumente), die mit dieser Grammatik erzeugt werden können.

Im Falle von HTML passiert nun das folgende: Mittels SGML wird eine Grammatik für HTML definiert. Hier wird genau festgelegt, welche Elemente verwendet werden und wie diese verschachtelt sein dürfen. Ein Beispiel hierfür: Auf HTML-Seiten finden sich oft zwei Sorten von Listen, eine numerierte (Element *ol*) sowie eine nicht numerierte Variante (Element *ul*). Jede Liste kann einen Eintrag oder beliebig viele Einträge enthalten (Element *li*). In der SGML-DTD ist dies wie folgt ausgedrückt:

```
<!ELEMENT (OL|UL) - - (LI)+>3
```

Auf diese einfache Weise⁴ sind sämtliche HTML-Elemente in der DTD definiert:

Das Haupt-Element (auch ROOT-ELEMENT) einer HTML-Datei heißt praktischerweise `<html>`. Es enthält das Kopf-Element `<head>` sowie das Körper-Element `<body>`. Im Kopfelement werden METAANGABEN im weiteren und engeren Sinn vermerkt – etwa Angaben zu Titel, Autor, Schlüsselwörter u.ä.

¹ Prinzipiell kann der Inhalt eines Containers auch multimedialer Natur sein.

² Dieser Zugang wird über eine GRAMMATIK-DEKLARATION ermöglicht, die in jeder INSTANZ, d.h. in jedem konkreten Dokument dieser Grammatik vorkommt. So verweist etwa das Element `<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN" "http://www.w3.org/TR/REC-html40/strict.dtd">` zu Beginn eines HTML-Dokumentes auf die Grammatik der Sprache HTML 4.0.

³ Der senkrechte Strich steht hierbei wieder für das logische *oder*, das Pluszeichen ist ein Platzhalter, ein sogenannter REGULÄRER AUSDRUCK, mit der Bedeutung *ein oder beliebig viele Elemente li*.

⁴ Die gesamte DTD ist freilich ungleich komplexer. Für jedes Element sind nämlich noch ATTRIBUTE definiert: Attribute sind zusätzliche Angaben, die ein Element enthalten kann. Ein Attribut besteht stets aus einem Namen sowie einer Wertzuweisung. Mit Hilfe des Attributs *href* kann etwa einem Link-Element *a* die Adresse des Zielknotens mitgeteilt werden: `Verweis` (siehe unten).

– sowie Relationen zu anderen Dateien etabliert (siehe unten). Der im Browser sichtbare Knoten-Inhalt steht im Körper-Element. Der eigentlichen Datei vorangestellt ist die Referenz auf eine SGML-Grammatik. Eine einfache Datei könnte also wie folgt aussehen:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN"
"http://www.w3.org/TR/REC-html40/strict.dtd">
<html>
<head>
<title>Text des Titels</title>
<meta name="author" content="Huber O.">
</head>
<body>
<h1>Erste Überschrift erster Ordnung</h1>
<p>Einfacher Textparagraph mit <a href="andere_datei.html">Hyperlink</a> </p>
<h1>Zweite Überschrift erster Ordnung</h1>
<p>Zweiter Textparagraph mit <img src = "Bild_datei.gif" alt = "Bild_text"/> <a/> </p>
<p><b>Zweiter Textparagraph in Fettdruck</b><a/> </p>
</body>
</html>
```

Dieses Beispiel soll nun zur einführenden Beschreibung der grundlegenden HTML-Strukturierung eines Textes dienen.¹ Auf den ersten Blick gut zu sehen ist, daß neben Elementen, die für die logische Gliederung der Datei bzw. des Textes zuständig sind (z.B. `<head>`, `<body>`, `<h1>`², `<p>`³), Elemente zur Textformatierung (z.B. ``⁴), zur Einbettung multimedialer Elemente (z.B. ``⁵) sowie zum Setzen von Hyperlinks (`<a>`) vorhanden sind.

Dieses letztgenannte Link-Element steht durch die Thematik des Fallbeispiels naturgemäß im Vordergrund. Sämtliche Hypertext-Links haben in HTML einen einheitlichen Aufbau. Hierbei spielt das Link-Ziel keine Rolle: Es kann sich in

¹ Im folgenden werden einige Elemente der Sprache HTML mit einer Kurzerläuterung eingeführt. Dies sollte genügen, um den weiteren Ausführungen zu folgen. Eine auch nur in Ansätzen umfassende Einführung in HTML wäre an dieser Stelle fehl am Platz. Für weitere Informationen sollte das Befehlsverzeichnis von Münz 2001 konsultiert werden (Münz 2001 [navigation/html.htm] Abschnitt *Kurzreferenz: HTML*).

² Überschrift-Element erster Ordnung.

³ `<p>` ist ein Absatz-Element, das weitere Elemente sowie Rohtext enthalten kann.

⁴ Element zur Fettformatierung von Text.

⁵ `` bietet die Möglichkeit mittels des Attributs `src` ein Bild in die HTML-Datei zu integrieren. Bei `` handelt es sich um ein LEERES ELEMENT. Der Begriff ist auf den ersten Blick etwas verwirrend, da bei der Anzeige der Datei im Browser das Bild angezeigt wird, der Container scheinbar eben doch über Inhalt verfügt. Da das Bild-Element jedoch im Quellcode kein schließendes Element hat und nicht der Schematik `<Elementname>Elementinhalt</Elementname>` entspricht, wird es als *leeres Element* bezeichnet.

der gleichen oder einer anderen Datei der Homepage, auf einer anderen Homepage oder auch lokal auf dem eigenen Rechner befinden. Stets gilt das folgende Schema für einen einfachen Verweis: `Verweistext`. Als Ziel kommen hierbei entweder ein gesamter Knoten oder sogenannte SPRUNGMARKEN in Betracht. Sprungmarken werden in HTML ebenfalls mit dem Element `<a>` definiert. Mittels des FRAGMENTIDENTIFIZIERERS `#` kann im Attribut `href` eines Links auf die Sprungmarke verwiesen werden. Zum Beispiel: Das Link-Ziel ist die Sprungmarke `` in der Datei `meine_datei1.html`. Von einer anderen Datei aus kann nun auf diese Sprungmarke wie folgt verwiesen werden: ``. Umfangreichere Möglichkeiten, z.B. das Verlinken von fest definierten Ausschnitten von Dokumenten, bietet erst das XML-Linking-Modell XLINK (siehe unten).

Neben den oben beschriebenen Hypertext-Links, die mit `<a>` formuliert werden, bietet HTML noch die Möglichkeit, Links im Sinne von *einfacher Relation zwischen Dokumenten* zu etablieren. Hierfür wird das Element `<link>` benutzt: Es kann lediglich im Kopfbereich einer HTML-Datei verwendet werden und etabliert keine ausführbaren Hypertext-Links! Mit `<link>` können Beziehungen eines Dokumentes zu anderen Dokumenten oder zu Style-Sheets beschrieben werden. Die Einbindung von Style-Sheets wird weiter unten kurz angesprochen, für die Belange des Fallbeispiels relevanter ist jedoch die semantische Bezugnahme zwischen HTML-Dokumenten mit Hilfe von `<link>` und den zugehörigen Attributen `rel` und `rev`. Hierauf wird in Kapitel 16.3.3 nochmals zurückgekommen.

Desweiteren fällt sofort auf, daß die logische Textstruktur nur bedingt durch HTML-Elemente ausdrückbar ist: So kann mit dem Element `<h1>` lediglich dargestellt werden, daß es sich um eine Überschrift erster Ordnung handelt. Daß aber beispielsweise die obige Textzeile `<h1>Zweite Überschrift erster Ordnung</h1>` tatsächlich die zweite Überschrift der ersten Überschriftenhierarchie ist, läßt sich mit HTML nicht codieren.¹ Hierzu wäre etwa die Konstruktion eines ‚Zählerattributes‘ notwendig, was folgende –

¹ Erschwerend kommt hinzu, daß von vielen Web-Designern Mittel zur logischen Auszeichnung zum Zweck der physischen Textdarstellung ‚mißbraucht‘ werden: Die meisten Browser stellen nämlich Textteile, die beispielsweise als Überschrift formatiert sind, fett und in großer Schrift dar. Wird diese Möglichkeit auf Textteile angewendet, die logisch keine Überschriften darstellen, so ist eine spätere automatische Strukturanalyse, die sich ja auf die Hierarchie der HTML-Elemente stützt, nahezu unmöglich. Auch um dies zu verhindern, strebt das W3C mit der Entwicklung der Formatierungssprache CASCADING STYLE-SHEETS (CSS) eine strikte Trennung von Struktur und Layout an. Nähere Informationen zu CSS sind auf der Seite des W3C erhältlich [<http://www.w3.org/Style/CSS/>].

hypothetische – Konstruktion ergeben würde:

```
<h1 counter="2">Zweite Überschrift erster Ordnung </h1>
```

Zusammenfassend kann HTML also als eine Sprache zur groben strukturellen Auszeichnung von Text bezeichnet werden. Mittel zur Formatierung des Layouts sind hingegen ausreichend vorhanden, wie folgendes Beispiel zeigt:¹

```
<p><font size = „12pt“ face = „Avantgarde“ color = „red“><b><i>Text</i></b></font></p>
```

Mit obigem Beispiel wird der mit `<p>` eingeschlossene Absatz in der Schriftart Avantgarde, Größe 12 Punkte, roter Farbe sowie fett und kursiv dargestellt. Jeden Absatz einzeln mit diesen Layoutangaben zu versehen, wird jedoch schnell umständlich. Aus diesem Grund können globale, das Layout betreffende Auszeichnungen für die einzelnen HTML-Elemente in einem Style-Sheet angegeben werden. Wird dieses Style-Sheet an den Anfang der Datei gesetzt, werden von modernen Browsern sämtliche Elemente im gewünschten Layout dargestellt.²

```
<html>
<head>
<style type = "text/css">
  p { font-family:Avantgarde,Arial; font-color:red; font-style:italic; font-size:12pt; font-weight:bold}
</style>
</head>
<body>
<p>Text, der mit dem obigen Style-Sheet formatiert wurde</p>
<p> Noch ein Text, auf den das obige Style-Sheet einwirkt </p>
</body>
</html>
```

Die beiden Beispiele sollten verdeutlichen, daß es zwar umfangreiche Möglichkeiten der Einflußnahme dafür gibt, wie ein Element auf dem Bildschirm dargestellt werden soll. Eine Aussage, was ein Element als Inhalt hat und wie sich dieser Inhalt zu anderen Inhalten verhält, ist jedoch nur sehr beschränkt über die Grobstruktur möglich. Textteile, die sich einer solchen Grobeinteilung

¹ *Font* gilt inzwischen als veraltet. Layout-Angaben werden heute zumeist mit den bereits oben eingeführten STYLE-SHEET-SPRACHEN gemacht.

² Bei tatsächlichen Projekten wird in der Regel ein Weg gegangen, der die Arbeit noch mehr erleichtert: Es wird ein globales Style-Sheet definiert und in sämtlichen Dateien des Projektes ein Link darauf gelegt, etwa in der Form `<link rel = stylesheet type = "text/css" href="meine_style_sheet_datei.css">`.

entziehen, sind mit reinem HTML überhaupt nicht logisch, hierarchisch, funktional und/oder semantisch in den Gesamttext einer HTML-Datei einzuordnen. Dies gilt so auch uneingeschränkt für Hypertext-Links, die durch das Element `<a>` in den Text eingebettet werden. Lediglich ‚Hilfskonstruktionen‘ mittels CSS und einer Programmiersprache wären hier denkbar.¹ Daß beim ursprünglichen Design von HTML jedoch durchaus an eine mögliche Typisierung von Links durch logische, funktionale und/oder semantische Information gedacht wurde, zeigt das Kapitel 16.3.3.

15.2.2 HTML als XML-Anwendung

Die letzte mit SGML definierte Sprachversion von HTML war HTML 4. Die Nachfolgeversion nennt sich XHTML² und ist mit XML³ definiert. XML wiederum kann als vereinfachte Version und damit als echte Teilmenge von SGML verstanden werden. Ebenso wie SGML ist XML eine METASPRACHE zur Definition anderer Sprachen. Da XML die meisten Vorzüge von SGML nicht jedoch dessen Komplexität geerbt hat, scheint sich diese Sprache momentan auf

¹ Per se bieten auch Style-Sheets keine Möglichkeiten zur logischen, funktionalen und/oder semantischen Codierung von Text(teilen) oder Links an. Jedoch können mit Hilfe der Style-Sheet-Sprache CSS in gewissem Umfang ‚Hilfslösungen‘ zur semantischen Auszeichnung von Elementen konstruiert werden: In CSS lassen sich nämlich Klassen definieren, die dann über genau festgelegte Layoutmerkmale verfügen. Diese Klassen könnten auch semantisch motiviert sein. So kann etwa die Klasse *Erläuterung* mit bestimmten Layouteigenschaften in einem Style-Sheet definiert werden. In einer konkreten HTML-Datei werden dann Links, deren Ziele auf Dateien verweisen, welche semantisch tatsächlich die Relation *Erläuterung des Link-Ursprungs* darstellen, mit der entsprechenden CSS-Klasse ausgezeichnet. Konkret sieht das dann etwa so aus:

```
...Einleitungstext <a class="Erlaeuterung" href="meine_erlaeuterung.html">Beginn des  
Erläuterungstextes </a> Eventuell weiterer Einleitungstext ...
```

Mit Hilfe einer Programmiersprache wie *JavaScript* wäre es nun möglich, eine Verhaltensweise für sämtliche Elemente der Klasse *Erläuterung* zu erstellen, etwa daß eine maussensitive Hilfe mit dem Inhalt *Eine ‚Erläuterung‘ hierzu erhalten Sie, wenn Sie diesem Link folgen!* erscheint, wenn der Benutzer die Maus über das Element bewegt. Auch könnte eine CSS-Klasse mit einem ganz spezifischen Verhalten der Browser-Software verknüpft werden. Durch die SELEKTOREN-FUNKTION von CSS sind auch geschachtelte Klassen realisierbar. Auf diese Feinheiten wird hier nicht eingegangen, Information dazu findet sich jedoch im WWW [<http://www.w3c.org/TR/css3-selectors>].

² Einen guten Einstieg auch in diese Thematik bietet die Seite des W3C [<http://www.w3.org/TR/xhtml1/>]. Weitere Informationen und Tutorials finden sich auch unter <http://www.xhtml.org/>.

³ Die beste Adresse zur Information über XML sind die Seiten des W3C [<http://www.w3.org/XML>], wobei zum ersten Einstieg der Kurzlehrgang *XML in 10 points* besonders gut geeignet ist [<http://www.w3.org/XML/1999/XML-in-10-points>].

vielen Gebieten durchzusetzen.¹

Mit XML lassen sich nicht nur textuelle Dokumente sondern auch andere Formen von (semi)strukturierten Daten mit Hilfe von Grammatiken (DTD oder SCHEMATA) beschreiben. Mit XML wurde nicht nur die neue Version von HTML, XHTML, definiert, sondern weitere zunehmend einflußreicher werdende Sprachen wie die SVG (SCALABLE VECTOR GRAPHICS)², MATHML (MATHEMATICAL MARKUP LANGUAGE)³, SMIL (SYNCHRONIZED MULTIMEDIA INTEGRATION LANGUAGE)⁴ oder WML (WIRELESS MARKUP LANGUAGE)⁵. Die große Stärke einer generischen Sprache ist die Möglichkeit, Elemente und Attribute den eigenen Bedürfnissen nach zu definieren. Folgendes Beispiel zeigt ein XML-Dokument samt externer Grammatik in DTD-Syntax:⁶

```
<!ELEMENT nachricht (titel,text,datum,redakteur)>
<!-- Eine Nachricht besteht aus Titel, Text, Datum und Redakteur -->

<!ELEMENT titel (#PCDATA)>
<!-- Der Titel enthält den Titeltext -->

<!ELEMENT text (#PCDATA)>
<!-- Der Text enthält den Nachrichtentext -->

<!ELEMENT datum (#PCDATA)>
<!-- Das Datum enthält die Datumsangabe -->

<!ELEMENT redakteur (#PCDATA)>
<!-- "redakteur" enthält die Angabe zum Redakteur -->
```

¹ Im Mittelpunkt steht hier das MULTIPLE PUBLIZIEREN ein und desselben XML-Dokumentes für verschiedene Ausgabemedien, etwa für den Drucker, den PC, einen Taschencomputer oder ein Mobiltelefon. Aber auch die Kommunikation zwischen heterogenen Datenbanken kann über eine einheitliche Umwandlung nach XML erfolgen. XML nimmt hier eine Vermittlerrolle ein, in der Informatik spricht man daher auch von MIDDLEWARE.

² Informationen hierzu im WWW [<http://www.w3.org/Graphics/SVG/>]

³ Informationen hierzu im WWW [<http://www.w3.org/Math/>]

⁴ Informationen hierzu im WWW [<http://www.w3.org/AudioVideo/>]

⁵ Informationen hierzu im WWW [<http://www.wapforum.org/>]

⁶ Das Beispiel ist angelehnt an Münz 2001[xml\intro.htm] Abschnitt XML, DTDs und Validierung.

```
<nachricht>
<titel>XML wird jetzt auch in SELFHTML beschrieben!</titel>
<text>
Nach langer Pause ist eine neue SELFHTML-Version erschienen. Eine der wichtigsten
Neuerungen ist der XML-Teil, in dem die neue Definitionssprache für Auszeichnungssprachen
behandelt wird. Der Anwender lernt, wie er Regeln definiert, und Beispiele
ergänzen die Beschreibungen.
</text>
<datum>20.10.2001</datum>
<redakteur>Ferdinand Schreiberling</redakteur>
</nachricht>
```

(Grammatik und Beispieldokument der Klasse *nachricht*)

Das Grundprinzip sollte durch dieses Beispiel ausreichend illustriert sein. Das Verhältnis von Grammatik und Dokument-Instanz wurde bereits bei der Einführung von SGML erläutert.

Sämtliche¹ mit SGML formulierten DTDs für HTML wurden mit XML neu definiert. Will ein Hypertext-Autor an Stelle von HTML 4 nun XHTML verwenden, so ist der Umstellungsaufwand gering²: Anstelle der Referenz auf eine SGML-Grammatik erfolgt zuerst eine XML-DEKLARATION³ sowie die Referenz auf eine XML-Grammatik, eine DTD (bzw. ein SCHEMA; siehe unten):

¹ Bei der Einführung einer neuen HTML-Version stellt das W3C in der Regel drei Versionen der DTD zur Verfügung: eine, die die Neuerungen streng umsetzt (*strict*), eine Übergangsversion (*transitional*) und eine, die spezielle Rahmenkonstruktionen innerhalb von HTML-Dateien, sogenannte FRAMES, erlaubt (*frameset*). Der genaue Aufbau dieser DTDs ist für die vorliegende Arbeit nicht von näherer Bedeutung. Für Interessierte seien jedoch die Internet-Adressen der entsprechenden Grammatiken von XHTML 1.0 aufgeführt:

XHTML-Strict [<http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd>]

XHTML-Transitional [<http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd>]

XHTML-Frames [<http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd>]

² Freilich sind einige wichtige Details vom Programmierer zu beachten: so müssen etwa die Namen der Elemente stets klein geschrieben werden. In HTML 4 spielte die Groß- und Kleinschreibung hingegen keine Rolle. Ferner müssen in XHTML sämtliche Elemente geschlossen sein, dies ist besonders relevant bei leeren Elementen. So wird etwa ein Zeilenumbruch in HTML 4 mit `
` eingefügt. In XHTML muß die Schreibweise `
</br>` bzw. vereinfacht `
` verwendet werden.

In HTML 4 war auch eine nicht korrekte Verschachtelung von Elementen erlaubt, etwa `fett kursivfett auskursiv aus`. Dies ist in XHTML nicht mehr möglich, die entsprechende Zeile muß wie folgt lauten: `fett kursiv kursiv aus`.

³ Die XML-Deklaration teilt dem Browser lediglich die verwendete XML-Version sowie den Zeichensatz des folgenden Dokumentes mit. Momentan ist diese Deklaration noch nicht Pflicht, sie wird jedoch empfohlen.


```
<?xml version="1.0" encoding="iso-8859-1"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/2000/REC-xhtml1-20000126/DTD/xhtml1-strict.dtd">
```

Ansonsten bekommt der Hypertext-Autor zum jetzigen Zeitpunkt von der Umstellung hin zu XML kaum etwas mit, aber die Zukunft wird mit ziemlicher Sicherheit einschneidende Modifikationen bringen: Künftig wird XHTML nämlich in Module aufgeteilt, so daß XHTML-Programmierer sich im Grunde ‚ihre‘ DTD selbst zusammenstellen können.¹ Dies kann mit den vom W3C zur Verfügung gestellten Modulen als auch – im Prinzip – mit Hilfe eigener Ergänzungsgrammatiken geschehen.² Um eventuelle Namensgleichheiten von Elementen unterschiedlicher Grammatiken zu vermeiden, besitzt jede Grammatik einen eigenen NAMENSRAUM, der gesondert deklariert werden muß:

```
xmlns="http://www.w3.org/1999/xhtml" xml:lang = "de" lang = "de"
```

Auf diese Weise wird beispielsweise der Standard-Namensraum *XHTML* für ein XHTML-Dokument festgelegt. Sämtliche Elemente (in sämtlichen Kombinationsmöglichkeiten) der XHTML-Grammatik dürfen nun verwendet

¹ Grundlage hierfür ist die sogenannte MODULARISIERUNG von XHTML. Philosophie und technischer Hintergrund dieser Entwicklung sind im *Modul-Framework of XHTML* nachzulesen [<http://www.w3.org/TR/2000/WD-xhtml-modularization-20000105/>]. Zu erwähnen ist, daß DTDs hierbei von sogenannten SCHEMATA abgelöst werden. Ein Schema ist wie eine DTD eine Grammatik zur Beschreibung einer Sprache. Jedoch gibt es bei der Benutzung von Schemata einige Vorteile: So ist eine DTD etwa in der sogenannten ERWEITERTEN BACKUS-NAUR-FORM (EBNF) verfaßt, eine in der Informatik weit verbreitete Form zur Beschreibung formaler Sprachen, während Schemata in XML selbst beschrieben werden können. Dies verringert den Aufwand für Neueinsteiger erheblich, da eben nur eine formale Metasprache erlernt werden muß und mit dieser dann sowohl Grammatiken als auch Dokumente beschrieben werden können. Philosophie sowie technische Grundlagen von Schemata zur Modularisierung von XHTML sind im Entwurf *Modularization of XHTML™ in XML Schema* des W3C nachzulesen [<http://www.w3.org/TR/2001/WD-xhtml-m12n-schema-20010322/>].

² Um Mißverständnissen vorzubeugen: Auch in Zukunft wird sich nicht jeder Designer einer Homepage mit dem Aufbau einer Privatgrammatik beschäftigen müssen. Im Normalfall werden die vom W3C zur Verfügung gestellten modularen XHTML-Grammatiken ausreichen. Bei Bedarf können sich jedoch Organisationen oder große Firmen eigene, ihren Bedürfnissen angepaßte Grammatiken schaffen. Denkbar (und bereits in Entwicklung) sind hier Grammatiken zur Beschreibung von Geschäftsvorgängen im ELEKTRONISCHEN HANDEL. Im Prinzip kann jedoch jeder Privatmann mit eigenem Server auch eine eigene Grammatik definieren und in seinen Dokumenten deklarieren und verwenden. Jeder andere, der über die Internet-Adresse der Grammatik verfügt, kann diese per NAMENSRAUM-DEFINITION (siehe unten) in eigene Dokumente einfügen. (Die neue formale Sprache ist dann jedoch einer Gefahr ausgesetzt, mit der auch natürliche Sprachen zu kämpfen haben: dem Aussterben durch Nichtverwendung. Die Gemeinschaft der ‚Sprecher‘ von Metasprachen im Netz wird also selbst entscheiden, welche der entstehenden Sprachen sich tatsächlich durchsetzen wird. Vollkommen durchgesetzt hat sich ein neues Sprachmodul wohl erst dann, wenn die großen Browser-Hersteller dieses implementiert haben werden.)

werden. Zusätzlich können jedoch weitere Namensräume deklariert werden. Im untenstehenden Beispiel wird etwa der fiktive Namensraum *mein_namensraum*, der unter der fiktiven Adresse *http://mein_server/mein_namensraum_schema* zu finden ist, importiert:

```
xmlns:mein_namensraum="http://mein_server/mein_namensraum_schema"
```

Sämtliche Elemente des Schemas *mein_namensraum_schema* sind nun im XHTML-Dokument mit dem NAMENSRAUM-PRÄFIX *mein_namensraum* benutzbar. Definiert das Schema etwa ein Element, das denselben Namen hat wie ein Element aus dem XHTML-Namensraum (beispielsweise das Absatz-Element *p*), so sind durch das Präfix Verwechslungen ausgeschlossen:

```
<body>
<p>
Standard Absatz in XHTML
</p>
<mein_namensraum:p>
Importiertes Element p aus mein_namensraum
</mein_namensraum:p>
</body>
```

Durch diese Modularisierung ist es also prinzipiell möglich, eigene Elemente und Attribute mittels der Deklaration eines Namensraumes in ein XHTML-Dokument einzufügen. Dies wäre eine Möglichkeit, auf der Ebene der semantischen Zwischenschicht von XHTML-Hypertexten das im Rahmen dieser Arbeit zu entwickelnde Vokabular zur Typisierung von Links in bestehende Dokumente einzufügen.¹ Ein Link könnte dann etwa durch Befehle des fiktiven Namensraums *typo* wie folgt typisiert werden:

¹ Hierzu einige Anmerkungen: Wenn mittels einer Namensraumdeklaration bestehende Hyperlinks mit Attributen zu deren Typisierung ausgezeichnet werden, so hat das zunächst freilich nur Auswirkungen auf der Ebene der semantischen Zwischenschicht des Dokumentes. Das bedeutet, die typisierten Links sind lediglich im Quelltext zu sehen, ein Standard-Browser kann mit den Link-Auszeichnungen nicht umgehen, da diese ja nicht implementiert sind. Hierfür müßte zunächst eine Hilfsanwendung geschrieben werden, die dem Browser mitteilt, wie er mit den typisierten Links auf den Ebenen der layout- und browser-spezifischen Oberfläche umzugehen hat.

```
<a href="meine_ziel_datei.html"
  typo:ausgangsdokuments = "Kurzangabe des Inhalts des Ausgangsdokuments"
  typo:zieldokuments = "Inhaltsangabe des Zieldokuments"
  typo:relation = "Relation zwischen Ausgang und Ziel, z.B. Generalisierung">
```

15.2.3 Das Hyperlink-Modell von XML: Xlink

Das Link-Modell von HTML ist, wie bereits erwähnt, äußerst limitiert: Es lassen sich lediglich unidirektionale Links formulieren, Ziel eines Links kann lediglich ein ganzer Knoten oder eine Sprungmarke innerhalb dieses Knotens sein. Bereiche oder andere Elemente können nicht als Ziele definiert werden. Ein weiteres Problem ist, daß Links lediglich von Knoten ausgehen können, auf die man Schreibrechte hat, da der Link innerhalb des Ausgangsknoten notiert sein muß.

Der für unsere Belange gewichtigste Nachteil ist jedoch die Tatsache, daß zwar mit Hilfe der *rel* und *rev*-Attribute Link-Typisierungen prinzipiell mit Einschränkungen möglich wären (siehe Kapitel 16.3.3), diese aber von den Browser-Herstellern nicht unterstützt werden. Direkt schafft auch XLink hier keine Abhilfe: Mit bestimmten Attributen von XLink (siehe Kapitel 15.2.3) ließen sich zwar Links mit typisierten Auszeichnungen versehen. Das ‚Mapping‘ dieser semantischen Information im Quellcode der Zwischenschicht mit Layout und Browser-Verhaltensweisen müßte dann aber erst, wie bei HTML, von externen Style-Sheets (und/oder Anwendungen) übernommen werden.

Das Linking-Modell von XML schafft jedoch für viele der anderen oben genannten Probleme Abhilfe: XLink kann als Modell zur Beschreibung von Relationen von RESSOURCEN¹ gesehen werden. Diese Relationen können Hyperlinks sein, müssen dies aber nicht. XLink kann Links beschreiben, die beliebig viele Ressourcen involvieren. Die Beschränkung von HTML auf einen Ausgangs- sowie einen Ziel-Knoten ist also aufgehoben, es ließe sich somit etwa ein Link auf mehrere Versionen eines Dokumentes ausdrücken. Desweiteren verfügt XLink durch XPOINTER und XPATH über mächtige Möglichkeiten, um Ressourcen oder Teile von Ressourcen zu adressieren. Die Aspekte des

¹ Unter einer RESSOURCE (engl. RESOURCE) versteht man jedes mittels einer URI adressierbare Objekt. Eine Ressource kann ein Hypertext-Knoten, ein Teil-Knoten, ein Bild, ein Teil eines Bildes o.ä. sein. Bezieht man den Begriff auf einen ganzen KNOTEN, so sind die beiden Termini synonym zu verwenden. Zur Beschreibung von Teil-Knoten werden in der Sprache XPOINTER formulierte Ausdrücke verwendet, die sich der Pfadlokalisierungssprache XPATH bedienen. Hierauf wird im Rahmen dieser Untersuchung nicht eingegangen, der interessierte Leser findet jedoch genug Material auf den Seiten des W3C (XPath: <http://www.w3.org/TR/xpath>, XPointer: <http://www.w3.org/TR/xptr>).

Bezugsbereichs von Links im Ausgangs- sowie Ziel-Knoten – ein wichtiger Beschreibungsaspekt von TAH auf der Ebene *Link* – könnten somit syntaktisch formal beschrieben werden.

Ein weiterer Vorteil ist, daß Links nicht im Ausgangs-Knoten gespeichert werden müssen. Dies macht den Aufbau von leicht wartbaren externen LINK-DATENBANKEN möglich. Auf diese Weise ist es auch machbar, ein Hyper-Link-Netz über Ressourcen zu legen, für die man selbst keine Schreibrechte hat: Hierdurch ließen sich eigene Anmerkungen zu fremden Hypertexten beisteuern oder persönliche Sichten ein und desselben fremden Dokumentes für mehrere Benutzer erstellen.

Auf weitere technische Einzelheiten soll hier nicht eingegangen werden, der interessierte Leser sei jedoch nochmals auf die oben angegebenen einschlägigen Seiten des W3C verwiesen. Die Syntax von XLink als eine Möglichkeit, typisierte Links zu implementieren, wird in Kapitel 15.2.5 kurz vorgestellt.

15.2.4 Implementierung typisierter Links im Corpus'

In diesem Kapitel werden aktuelle technische Mittel zur Implementierung typisierter Links vorgestellt, die auch in den Corpus-Hypertexten verwendet werden.

Das Kapitel 15.2.5 stellt hingegen Implementierungen vor, die sich momentan noch in einem Entwicklungsstadium befinden (bzw. deren Spezifikationen vor kurzem vorgestellt wurden).¹ Mit dem Konzept der XML-TOPIC-MAPS wird beispielhaft eine Möglichkeit der Umsetzung eines tiefenstrukturellen semantischen Netzes kurz vorgestellt, aus welchem heraus Link-Typen (der Zwischenschicht und) der Hypertext-Oberfläche ‚vererbt‘ werden könnten.

Typisierung durch das *title*-Attributs des *a*-Containers

Die erste hier beschriebene Möglichkeit ist eine Codierung der Link-Typen auf der Ebene der semantischen Zwischenschicht des HTML-Quellcodes, d.h. die Semantik des Links ist bereits im Quelltext und nicht erst im Browser wie bei den beiden weiter unten vorgestellten Varianten sichtbar: Neuere Browser-Versionen machen es möglich, das Attribut *title* zur Visualisierung von

¹ Freilich werden Implementierungen in XML und mit XML-Topic-Maps von Tag zu Tag zahlreicher. Die große Mehrheit der WWW-Hypertexte besteht jedoch noch immer aus (statischen oder dynamisch erzeugten) HTML-Knoten.

typisierten Links einzusetzen.¹ Der Text dieses Attributs eines Elements `<a>` wird beim Positionieren der Maus auf den Link eingeblendet. Eine mögliche Vorgehensweise wäre nun, das Vokabular, daß eigentlich für die *rel*- und *rev*-Attribute vorgesehen ist, zusätzlich für das *title*-Attribut zu verwenden:

```
<p>
  Als nächstes wenden wir uns den
  <a title="next" rel="next" href="typisiert.html">
    typisierte Links
  </a> zu.
</p>
```

Im Browser würde der Text des *title*-Attributs dann so angezeigt werden:

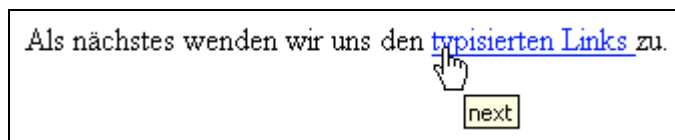


Abbildung 8: Typisierung durch ‚Mißbrauch‘ des *title*-Attributs

In den Corpus-Hypertexten SELF und SELF8 sind die Piktogramme (siehe unten) der typisierten Links (teilweise) mit dem Attribut *title* versehen, die Elemente `<a>` jedoch nicht. Der Mouse-Over-Text wird daher nicht beim Link-Text sondern lediglich beim zugehörigen Piktogramm angezeigt.

Typisierung durch Piktogramme

Bei der Typisierung durch Piktogramme handelt es sich – ebenso wie bei der weiter unten vorgestellten Variante durch farbige Auszeichnung – um Typisierungen auf der Ebene der layout-spezifischen Oberflächenstruktur eines HTML-Hypertextes: Zwar ist ein Piktogramm oder eine farbige Markierung im Quelltext niedergelegt (oder wird per externem Style-Sheet den Link-Typen zugewiesen), die Semantik ist aber im Gegensatz zu der oben beschriebenen Auszeichnung mittels des Attributs *title* erst im Browser selbst sichtbar! Die Semantik von Piktogrammen ist somit auch nicht maschinenlesbar.² Die

¹ Für Browser, die diese Funktionalität nicht besitzen, böten sich Hilfslösungen in der Form an, die weiter oben bereits angesprochen wurde: Mit Style-Sheets werden für die Link-Typen eigene Klassen definiert, die dann über JavaScript gesteuert ebenfalls den gewünschten Mouse-Over-Text anzeigen.

² Um dies zu erreichen, müßte das Alternativtext-Attribut (*alt*) des Bild-Containers (*img*), mit dessen Syntax Piktogramme in HTML eingebettet werden, den Namen des Link-Typen enthalten, für den das Piktogramm steht. Etwa für ein Piktogramm mit der Bedeutung *Beispiel*: ``

Verwendung von Piktogrammen ist eine weitere Möglichkeit, Links in HTML zu typisieren, ohne auf die Beschränkungen der Syntax Rücksicht zu nehmen. Hierbei wird vor den eigentlichen Link-Text ein Symbol bzw. ein Icon eingefügt, das auf den Typ des Links hinweisen soll (so in SELF, SELF8, SPORT1, GRAMMIS).



Abbildung 9: Link-Typisierung durch Piktogramme

Manchmal treten Piktogramme auch als alleinige Link-Anker ohne zugehörigen Link-Texte auf (etwa in SPORT1; siehe hierzu auch Kapitel 17.3.4). In Kapitel 16.3.3 werden verschiedene Link-Typen von Piktogramm-Kollektionen beispielhaft anhand der Texte SELF und SELF8 vorgestellt.

Typisierung durch farbige Kennzeichnung

In GRAMMIS sind Links zum Teil farbig gekennzeichnet, um zu signalisieren, daß die jeweiligen Ziel-Knoten einem bestimmten Hypertext-Bereich angehören. (Die farbige Kennzeichnung wird dort flankiert durch eine zusätzliche Einblendung eines Mouse-Over-Textes, dessen Inhalt im *title*-Attribut des *a*-Containers codiert ist.) Auch die farbige Codierung erfolgt auf der Ebene der layout-spezifischen Oberflächenstruktur¹, jedoch ist offenbar geplant, eine semantische Tiefenstruktur einzuführen.²

15.2.5 Mögliche zukünftige Implementierungen

Die folgenden Abschnitte beziehen sich explizit nicht auf die technischen Realisierungen der Corpus-Hypertexte. Es sind jedoch interessante Spezifikationen und Implementierungen, die für den gesamten Themenkomplex *Text im digitalen Medium* in nächster Zukunft einen großen Bedeutungszuwachs erfahren dürften. Aus diesen Gründen wurden die weiter oben bereits abgehandelte Einführung in XML und der nun folgende kleine Ausblick auf

¹ Auch die Semantik einer farbigen Link-Auszeichnung ist nicht maschinenlesbar, hierfür bräuchte es zusätzliche Information auf Quellcode-Ebene, etwa über das oben beschriebenen Attribut *title* (bzw. *class*).

² Ein solches Unterfangen unternimmt Roman Schneider in seiner sich in Arbeit befindlichen Dissertation. Das geplante Vorhaben hat er im WWW zusammengefaßt [<http://www.ids-mannheim.de/grammis/roman/diss.html>].

Einsatzmöglichkeiten von XML-Anwendungen zur Link-Typisierung (und mehr) in diese Arbeit integriert.

Typisierung durch die Attribute *role* und *title* von XLink

Die design-philosophischen Vorzüge von XLink im Vergleich zum Link-Modell von HTML wurden bereits erläutert. Für die konkrete Realisierung von typisierten Links auf der Ebene der semantischen Zwischenschicht stehen zwei Konstrukte zur Verfügung: SIMPLE LINKS und EXTENDED LINKS¹. Simple Links sind in etwa mit den Links in HTML zu vergleichen, die mit einem Element `<a>` realisiert sind.

Aber: Ein solcher Link kann jedem XML-Element hinzugefügt werden, er ist also nicht auf bestimmte Elemente beschränkt, wie dies in HTML mit *a* und *rel* der Fall ist! Die Syntax für die einfache Verlinkung eines Elements sähe mit XLink wie folgt aus:

```
<Abschnitt1
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xlink:type="simple"
  xlink:href="abschnitt2.xml"
  xlink:role=" http://mein_server/meine_link_typen/fortsetzung.txt"
  xlink:title="Fortsetzung in Abschnitt2"
  xlink:show="new"

  xlink:actuate="onRequest">
</Abschnitt1>
```

Zur Erläuterung: Durch Verweis auf den Namensraum *xlink* ist der Import von Link-Attributen möglich, die einem beliebigen Element beigelegt werden können: *type* gibt an, daß es sich um einen einfachen Link handelt, *href* verweist auf die URI der Ziel-Ressource. Die semantische Codierung erfolgt mit den Attributen *role* und *title*; *role* hat als Wert die Adresse einer Datei, die eine maschinenlesbare Semantik des Link-Typs enthält. Das Attribut *title* hingegen enthält eine für den menschlichen Leser verständliche Beschreibung des Link-Typs. Auf diese Weise kann also eine menschen- und maschinenlesbare

¹ An dieser Stelle sollen lediglich die Simple Links eingeführt werden. Für unsere Belange sind vor allem die möglichen Typisierungen interessant. Die wahre Überlegenheit von XLink wird jedoch erst beim Betrachten der Extended Links offenbar. Nur mit Ihnen ist es etwa möglich, Ausgangs-Knoten in Links einzubinden, für die man keine Schreibrechte hat, und somit externe Link-Datenbanken aufzubauen. Mit Extended Links ist es ferner machbar, von den Typen der beteiligten Ressourcen ausgehend Rückschlüsse auf die Typen der Links zu ziehen, die zwischen diesen Ressourcen etabliert werden.

Typisierung von Links erfolgen, die auf Ebene der semantischen Zwischenschicht mit XLink formuliert wurden.

Typisierung im Rahmen einer XML-Topic-Map

TOPIC-MAPS (TM) stellen ein Implementierungsbeispiel einer semantischen Tiefenstruktur für Ressourcen dar.¹ TM sind eine den eigentlichen Ressourcen untergeordnete Tiefenstruktur, die Verknüpfungen thematisch miteinander zusammenhängender Konzepte und/oder Dokumente erlaubt. Sie dienen der Vereinfachung von Strukturierungen sehr großer Ressourcenmengen, da sie eine ressourcenunabhängige Möglichkeit zur Beschreibung von Informationen anbieten. Das Prinzip und die Syntax sollen hier kurz vorgestellt werden.

TM sind in XML codierte Dokumente, die drei wichtige Elemente zur Indexierung und Beschreibung von Daten beinhalten. TOPICS, OCCURRENCES und ASSOCIATIONS. Diesen Elementen können die Attribute TOPIC TYPE, OCCURRENCE ROLE TYPE und ASSOCIATION ROLE TYPE zugeordnet werden. Auf komplexere Konzepte, wie SKOPUS-BESTIMMUNG, PUBLIC-SUBJECTS und FACETS, soll hierbei nicht eingegangen, sondern lediglich auf die angeführte Literatur verwiesen werden.

Ein Topic kann nun so ziemlich alles sein: Eine Stadt (*Köln*), eine Person (*Richard Wagner*), ein Konzept (*Freiheit*) oder ein Vorgang (*essen*). Was genau ein Topic ist, hängt stets von der konkreten Anwendung ab. Jedes Topic ist nun einem bestimmten Topic-Type zugeordnet: Das Topic *Köln* ist demnach vom Topic-Type *Land* etc. Folgende Abbildung soll diese Verhältnisse visualisieren:

¹ Die Spezifikation sowie weitere Informationen zu TM finden Sie online [<http://www.topicmaps.org>]. TM stellt eine Alternative zu RDF dar.

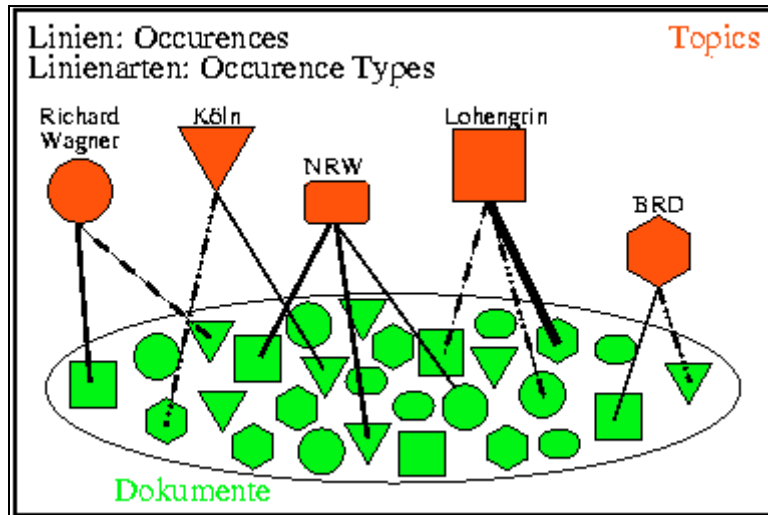


Abbildung 10: TOPIC MAPS – TOPICS UND OCCURRENCES¹

Topic Types sind in XML wie folgt definiert:²

```
<topic id="tt-stadt">
  <topname>
    <basename>
      Land
    </basename>
  </topname>
</topic>
```

Topic Occurrences Types werden formal so beschrieben:³

```
<topic id="or-webpage">
  <topname>
    <basename>
      Web Page
    </basename>
  </topname>
</topic>
```

Aus diesen beiden Typ-Definition lässt sich nun etwa das Topic *BRD* konstruieren, dem eine WWW-Seite als Occurrence zugeordnet ist:

¹ Die Abbildungen in diesem Kapitel sind aus einem Vortrag von Sven Hartmeier entnommen. Online im WWW unter: http://www.techfak.uni-bielefeld.de/~shartmei/Hypertext/html/Folien/TM_000.html.

² Das Kürzel *tt* steht für *Topic Type*. Die syntaktischen Feinheiten sind hier allerdings eher nebensächlich.

³ Das Kürzel *or* steht hierbei für *Occurrences Relation*.

```
<topic id="t-brd" types="tt-staat">
  <topname>
    <basename>
      BRD
    </basename>
    <dispname>
      Bundesrepublik Deutschland
    </dispname>
  </topname>
  <occurs type="or-webpage">
    http://www.brd.de/
  </occurs>
</topic>
```

Ein tiefenstrukturelles Topic kann also Verknüpfungen (Topic Occurrences) mit konkreten – auch heterogenen – Ressourcen (in der Abbildung als *Dokumente* bezeichnet) im WWW haben, die für das jeweilige Topic relevant sind. Vorhandene Ressourcen werden durch die Topic Occurrences typisiert, da die Occurrences selbst durch Occurrence Role Types typisiert sind, und diese Typisierung an die konkreten Ressourcen ‚vererben‘. Das Modell der Topic Maps gibt hier keine festen Typen vor, denkbar wären sowohl medial motivierte Typen (*Text, Bild, Video, Ton* etc.) oder semantisch motivierte (*Kommentar, Zitat* etc.).

Verknüpfungen können aber nicht nur zwischen Topics und Ressourcen sondern auch innerhalb der semantischen Tiefenstruktur, also von Topic zu Topic, realisiert werden. Eine solche Verknüpfung nennt man TOPIC ASSOCIATION. Auch Topic Associations könnten typisiert werden, und zwar durch sogenannte ASSOCIATION TYPES. Im folgenden Beispiel könnte etwa ein Association Typ *ist gelegen in* für die Topic Association vom Topic *Köln* zum Topic *Nordrhein-Westfalen (NRW)* – bzw. von *NRW* zu *BRD* – angelegt werden.

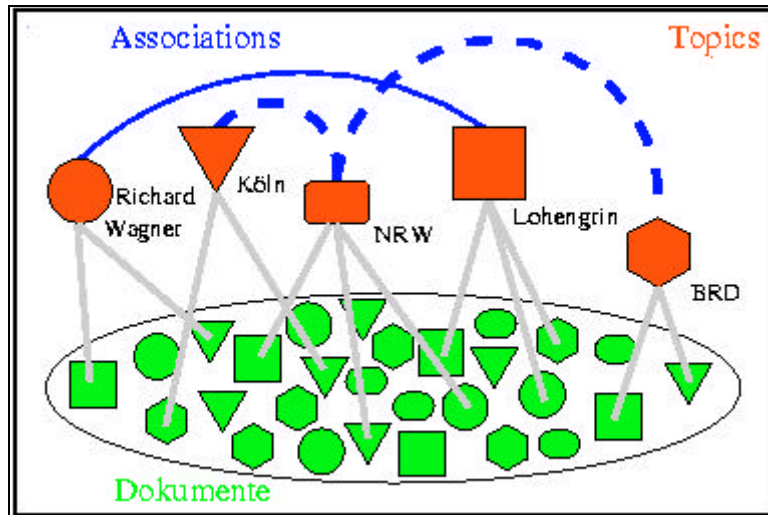


Abbildung 11: TOPIC MAPS – TOPIC ASSOCIATIONS

TM stellen eine Möglichkeit dar, das WWW in ein semantisches Netz zu verwandeln, oder korrekter, diesem eine semantische Tiefenstruktur hinzuzufügen. Vorteil der Topic Maps ist hierbei, daß sie sich auf Standards wie XML, XLink und URIs stützen und durch sogenannte PUBLIC SUBJECTS ‚Schnittstellen‘ zu standardisierten Ontologien und ISO-Codes bieten. Wenn auf solche allgemein bekannten Bezeichner zurückgegriffen wird, ist eine Austauschbarkeit und Mischbarkeit unterschiedlicher Topic Maps zu einem semantischen Netz möglich.

In einer solchen tiefenstrukturellen Auszeichnung liegt offensichtlich eine mögliche Zukunft des WWW: Nur somit wäre es nämlich machbar, die wachsende Zahl semantisch unstrukturierter Ressourcen in den Griff zu bekommen. Ein weiterer Vorteil der Topic Maps ist, daß eine völlig unabhängige Tiefenstruktur aufgebaut wird: Topics und Associations können aufgebaut werden, ohne daß dafür (bereits) passende Ressourcen vorhanden sein müssen!

Meiner Meinung nach, werden Link-Typisierungen an der Oberfläche ihre Bezeichner (via Zwischenschicht) aus der Tiefenstruktur ‚vererbt‘ bekommen. Eine Eins-zu-Eins-Übernahme dürfte hier aber kaum sinnvoll sein (siehe hierzu auch Kapitel 15). Insofern sind die Ergebnisse dieser Untersuchung der Oberflächenstruktur auch für zukünftige Implementierungen nützlich, die die hier erarbeiteten Link-Typen im Rahmen einer Topic Map wohl kaum umsetzen werden: Die Typen sind zu allgemein, als daß sie zur Repräsentation von Wissen dienen könnten. Die Transformation von Tiefen- zu Oberflächenstruktur

jedoch ist eine Wandlung von Präsentation zu Exploration – und genau hier werden meiner Meinung nach die vielen komplexen Typen des semantischen Netzes in wenige, für die Lektüre kognitiv plausible Link-Typen konvertiert. An einem solchen Inventar oberflächenstruktureller Typen wird in dieser Untersuchung gearbeitet. Auf spezielle Probleme der Konversion von Tiefenstruktur zur Zwischenschicht bzw. von Zwischenschicht zur Oberflächenstruktur wird nicht, bzw. nur dann eingegangen, wenn die Problematik relevant für Aspekte der Hypertext-Oberfläche ist.

16 Vokabularien typisierter Links

Im Rahmen der theoretischen Vorüberlegungen für die Anwendung des TAH zur Untersuchung von typisierten Links wurden im letzten Kapitel die heutigen und möglichen zukünftigen technischen Grundlagen zum Linking im WWW geschildert.

In den folgenden Kapiteln wird nun eine Auswahl an Inventaren von Relationstypen eingeführt, die für sehr heterogene Belange in unterschiedlichen Bereichen erstellt worden sind: Die hier vorgestellten Vokabularien älterer Hypertext-Systeme waren zumeist für die Datenhaltung, also die semantische Tiefenstruktur, des Hypertext-Systems konzipiert (siehe Kapitel 16.3.2). Hingegen sind die Möglichkeiten der Link-Typisierung mit Hilfe des Attributs *rel* in HTML auf der Ebene der Zwischenschicht zu lokalisieren (siehe Kapitel 16.3.3).

Vorab werden jedoch zuerst zwei linguistische Verfahren zur Typisierung von Verknüpfungsstrukturen traditioneller Texte vorgestellt (siehe Kapitel 16.2).

Als erster Schritt soll zunächst auf die allgemeinen Schwierigkeiten hingewiesen werden, die sich beim Aufbau solcher Inventare prinzipiell stellen.

16.1 Problemstellung

Prinzipiell wäre es möglich, für jeden Hypertext eigene proprietäre Link-Typen zu vergeben. Dies macht aber für die menschliche Textverarbeitung wohl kaum Sinn: Unterscheiden sich die Link-Typen von Hypertext zu Hypertext, so muß der Leser das den Typen zugrundeliegende System bei jedem Hypertext aufs neue begreifen. Aus diesem Grund ist es sinnvoll, sich auf eine begrenzte Anzahl von TERMINI zur Beschreibung von Links zu einigen. Eine TERMINOLOGIE wiederum ist nichts anderes als ein exakt definiertes, sich dadurch von der allgemeinen Sprache unterscheidendes geordnetes

KONTROLLIERTES VOKABULAR.¹ Einige solcher kontrollierter Vokabularien werden in den nächsten Kapiteln eingeführt.

16.2 Linguistische Ansätze

In den folgenden Kapiteln werden zwei genuin linguistische Ansätze zur Beschreibung von Relationen zwischen Textteilen vorgestellt: Es handelt sich hierbei um die RELATION ELEMENT THEORY und die RHETORICAL STRUCTURE THEORY. Die Grundkonzepte und Relationstypen beider Ansätze werden kurz eingeführt, um anschließend auf eine mögliche Anwendbarkeit zur Typisierung von Hypertext-Links hin untersucht zu werden.

16.2.1 Relation Element Theory

Der erste vorgestellte linguistische Ansatz ist die RELATION ELEMENT THEORY (RET). Grundlage der RET ist die Annahme, daß komplexe Relationen zwischen Wörtern stets durch ein SET ELEMENTARER RELATIONEN auszudrücken sind (Chaffin u. Hermann 1984 sowie Chaffin u. Hermann 1986). Eine beispielhafte Analyse der Lexempaare *alt-jung* sowie *tot-lebendig* soll diese Aussage illustrieren. Die Lexeme *alt* und *jung* stehen in einem Verhältnis der Antonymie (*contrary* in der Schreibweise der RET): Es herrscht also eine gegensätzliche Beziehung vor, auf einer imaginären Skala sind jedoch Zwischenstufen zu finden:

Alt ist er noch nicht, richtig jung aber auch nicht mehr.

Ein kontradiktorisches Verhältnis zeichnet hingegen die Lexeme *tot* und *lebendig* aus (RET: *contradictory*). Die Bedeutungen dieser Wörter schließen sich strikt aus, die Negation des einen ergibt automatisch die Bedeutung des anderen (*nicht lebendig = tot*, *nicht tot = lebendig*). Zwischenstufen sind hier nicht möglich.² Die unterschiedlichen Relationen von *alt-jung* bzw. *tot-lebendig* resultieren also aus dem Gegebensein bzw. Nichtgegebensein einer Gradzierbarkeit. Folgerichtig setzt die RET hier zwei elementare Dimensionen an: *Cont* für die Paraphrase *Word1 and Word2 can be qualified, the dimension is*

¹ Ein kontrolliertes Vokabular kann nun auf verschieden Arten organisiert sein: FLACH, TAXONOMISCH, als THESAURUS oder ONTOLOGISCH. Die meisten der in den folgenden Kapiteln vorgestellten Vorschläge für Link-Typisierungen liegen als flachen Liste bzw. als einfache Taxonomie vor.

² Lediglich im übertragenen Sinne: *Ich war den ganzen Tag unterwegs, jetzt bin ich halbtot.*

gradable sowie *Dis* für *Word1* and *Word2* cannot be qualified, the dimension is not gradable.

Die komplexe Relation *Antonymie* (*contrary*) führt die RET auf folgende elementare Relationen¹ zurück:

Dim(Bip(Sym),Cont)

Alt und *jung* existieren also in der gleichen Dimension (Dim), ihre Referenten sind gleich weit (auf unterschiedlichen Seiten) vom Mittelpunkt dieser Dimension entfernt (Bip(Sym)). Die Dimension ist fortlaufend (Cont = Continuous). Dies unterscheidet die Gesamrelation *contrary* von *contradictory*: Hier können (bei Gleichheit der drei weiteren oben beschriebenen elementaren Relationen) lediglich diskrete Positionen (Dis) innerhalb einer Dimension eingenommen werden, eben entweder *tot* oder *lebendig*:

Dim(Bip(Sym),Dis)

Elementare und komplexe Relation der RET

Elementare Relationen:

Nach dieser kurzen Einführung zeigen die folgenden Tabellen sämtliche mit RET formulierbaren elementaren Relationen. In RET finden sich hierfür fünf ‚Familien‘. Jede Familie ist im folgenden samt ihrer Mitglieder aufgeführt.²

| Familie 1: ELEMENTS OF INTENSIONAL FORCE | | |
|--|------|--|
| Relation | Abk. | Beschreibung ³ |
| Denotative | Den | W1 und W2 teilen denotative Bedeutung. |
| Connotative | Con | W1 konnotiert W2. |

Tabelle 8: RET-Familie 1: ELEMENTS OF INTENSIONAL FORCE

| Familie 2: DIMENSIONAL ELEMENTS | | |
|---------------------------------|------|---|
| Relation | Abk. | Beschreibung |
| Dimension | Dim | W1 und W2 teilen gemeinsame Dimension. |
| Unilateral position | Unip | W1 und W2 sind auf der gleichen Seite (einer Dimension bzgl. des Mittelpunkts dieser Dimension). |
| Bilateral position | Bip | W1 und W2 sind auf gegenüberliegenden Seiten (einer Dimension bzgl. des Mittelpunkts dieser Dimension). |
| Symetrical position | Sym | W1 und W2 sind jeweils gleich weit vom Mittelpunkt entfernt. |

¹ Die Dimensionen *Dim*, *Bip* und *Sym* sind in der untenstehenden Tabelle 9 erläutert.

² Die Tabellen sind entlehnt aus Chaffin u. Herrmann 1986. Die englische Terminologie der Relationen wurde beibehalten, Beispiele sowie Beschreibungen sind eingedeutscht und teilweise lediglich sinngemäß wiedergegeben.

³ *W1* und *W2* sind Abkürzungen für *Wort 1* und *Wort 2*.

| Familie 2: DIMENSIONAL ELEMENTS | | |
|--|-------------|---|
| Relation | Abk. | Beschreibung |
| Continuous | Cont | W1 und W2 sind qualifizierbar, die Dimension ist graduierbar. |
| Discrete | Dis | W1 und W2 sind nicht qualifizierbar, die Dimension ist nicht graduierbar. |
| Dichotomous | Dich | Regel Wenn nicht W1, dann W2 und umgekehrt ist anwendbar. |
| Spatial | Spa | W1 ist räumlich das Gegenteil von W2. |
| Vector | Vec | W1 ist bzgl. der Richtung das Gegenteil von W2. |

Tabelle 9:RET- Familie 2: DIMENSIONAL ELEMENTS

| Familie 3: ELEMENTS OF AGREEMENT | | |
|---|-------------|--|
| Relation | Abk. | Beschreibung |
| Inclusion | Inc | W1 ist semantisch oder physisch in W2 enthalten. |
| Overlap | Over | W1 und W2 sind semantisch ‚ähnlich‘. |
| Intersection | Int | W1 ist semantisch in W2 enthalten. |
| Unilateral inclusion | Unil | W2 enthält sämtliche Merkmale von W1, nicht jedoch umgekehrt. |
| Bilateral inclusion | Bil | W1 entspricht sematisch W2. |
| Attribute | Att | W2 ist ein Attribut von W1; W1 ‚ist ein‘ W2 im Sinne einer Hyperonymiebeziehung. |

Tabelle 10: RET-Familie 3: ELEMENTS OF AGREEMENT

| Familie 4: PROPOSITIONAL ELEMENTS | | |
|--|-------------|---|
| Relation | Abk. | Beschreibung |
| Event | Evt | W1 und W2 sind via eines Ereignisses miteinander verbunden. |
| Action | Act | W1 und W2 sind via einer Handlung miteinander verbunden. |
| Agent | Agt | W1 ist der Verursacher einer Handlung. |
| Object | Obj | W2 ist das Objekt einer Handlung. |
| Instrument | Inst | W2 ist das Instrument, das für eine Handlung benutzt wird. |

Tabelle 11: RET-Familie 4: PROPOSITIONAL ELEMENTS

| Familie 5: ELEMENTS OF PART-WHOLE INCLUSION | | |
|--|-------------|---|
| Relation | Abk. | Beschreibung |
| Partitive Inclusion | Partive | W2 ist physisch in W1 enthalten. |
| Locative Inclusion | Loc | W2 befindet sich räumlich in W1. |
| Attachment | Attach | W2 ist W1 hinzugefügt. |
| Social | Soc | W2 ist Teil von W1 aufgrund einer Vereinbarung. |
| Homogenous | Homog | Verschiedene Instanzen von W2 sind austauschbar. |
| Component | Comp | W2 ist ein Bestandteil von W1. |
| Property | Prop | W2 ist Eigentum von W1. |
| Possession | Possession | W1 verfügt über eine Eigenschaft oder ein Attribut. |

Tabelle 12: RET-Familie 5: ELEMENTS OF PART-WHOLE INCLUSION

Komplexe Relationen:

Aus den oben angeführten elementaren Relationen konstruieren Chaffin und Herrmann nun 31 komplexe Relationen, die sich wiederum zu fünf Familien gruppieren lassen:

| Komplexe Familie 1: CONTRAST | | |
|-------------------------------------|-----------------------------------|--|
| Komplexe Relation | Konstituierende Relationen | Beispielspaar |
| Contrary | Dim(Bip(Sym), Cont) | <i>alt – jung, glücklich – traurig</i> |
| Contradictory | Dim(Bip(Sym), Dis) | <i>tot – lebendig</i> |
| Reverse | Dim(Bip, Dich, Vec) | <i>kaufen – verkaufen</i> |
| Directional | Dim(Bip, Dich, Rop) | <i>rechts – links</i> |
| Incompatible | Dim(Bip) | <i>glücklich – deprimiert</i> |
| Asymmetrical contrary | Dim(Bip, Cont) | <i>heiß – kalt</i> |
| Pseudoantonym | Dim(Bip, Cont) | <i>glauben, daß – leugnen, daß</i> |
| Attribute similar | Int(Over(Att, Dis)) | <i>Rechen – Gabel</i> |

Tabelle 13: Komplexe RET-Familie 1: CONTRAST

| Komplexe Familie 2: SIMILARS | | |
|-------------------------------------|-----------------------------------|---------------------------------|
| Komplexe Relation | Konstituierende Relationen | Beispielspaar |
| Synonymity | Int(Inc(Bil)) | <i>Auto – Wagen</i> |
| Dimensional similar | Int(Over, Dim(Unip)) | <i>lachen – lächeln</i> |
| Necessary attribute | Int, Poss, Con | <i>Turm – hoch</i> |
| Invited attribute | Inc, Poss, Con | <i>Messer – scharf</i> |
| Action subordinate | Int(Inc(Unil)) | <i>reden – Vorlesung halten</i> |

Tabelle 14: Komplexe RET-Familie 2: SIMILARS

| Komplexe Familie 3: CLASS INCLUSION | | |
|--|-----------------------------------|--------------------------|
| Komplexe Relation | Konstituierende Relationen | Beispielspaar |
| Perceptual subordinate | Int(Inc(Unil)) | <i>Tier – Löwe</i> |
| Functional subordinate | Int(Inc(Unil)) | <i>Werkzeug – Hammer</i> |
| State subordinate | Int(Inc(Unil)) | <i>Emotion – Freude</i> |
| Activity subordinate | Int(Inc(Unil)) | <i>Spiel – Schach</i> |
| Geographical subordinate | Int(Inc(Unil)) | <i>Land – Spanien</i> |
| Place | Inc(Partive(Poss,Loc)) | <i>Europa – Spanien</i> |

Tabelle 15: Komplexe RET-Familie 3: CLASS INCLUSION

| Komplexe Familie 4: CASE RELATIONS | | |
|---|-----------------------------------|---------------------------|
| Komplexe Relation | Konstituierende Relationen | Beispielspaar |
| Agent-action | Evt(Act, Obj) | <i>Hund – bellen</i> |
| Agent-instrument | Evt(Agent, Inst) | <i>Jäger – Hund</i> |
| Agent-object | Evt(Agent, Obj) | <i>Bäcker – Brot</i> |
| Agent-recipient | Evt(Act, Obj) | <i>sitzen – Stuhl</i> |
| Action-instrument | Evt(Act, Inst) | <i>schneiden – Messer</i> |

Tabelle 16: Komplexe RET-Familie 3: CASE RELATION

| Komplexe Familie 5: PART-WHOLE | | |
|---------------------------------------|---|--------------------------|
| Komplexe Relation | Konstituierende Relationen | Beispielspaar |
| Functional object | Inc(Partive(Attach, Comp, Prop, Poss)) | <i>Baum – Blatt</i> |
| Collection | Inc(Partive(Homo, Prop, Poss)) | <i>Wald – Baum</i> |
| Group | Inc(Partive(Homo, Prop, Poss, Soc)) | <i>Chor – Sänger</i> |
| Ingredient | Inc(Partive(Comp, Prop, Poss, Loc)) | <i>Pizza – Tomate</i> |
| Functional location | Inc(Partive(Attach, Comp, Prop, Poss)) | <i>Haus – Wohnzimmer</i> |
| Organisation | Inc(Partive(Attach, Comp, Prop, Poss, Soc)) | <i>Armee – Division</i> |
| Measure | Inc(Partive(Homo)) | <i>Stunde – Minute</i> |

Tabelle 17: Komplexe RET-Familie 5: PART-WHOLE

Probleme einer möglichen Anwendung der RET zur Typisierung von Links

Zuerst die Probleme allgemeiner Art: Über die KOGNITIVE PLAUSIBILITÄT der ELEMENTAREN RELATIONEN der RET soll an dieser Stelle nicht diskutiert werden. Die Vermutung ist jedoch, daß am Ende einer solchen Diskussion – ähnlich wie bei der über elementare semantische Merkmale – das bekannte Brecht-Zitat stehen könnte: „Es geht auch anders, aber so geht es auch.“

Auch über interne Probleme der Relationskategorisierung gäbe es einige Anmerkungen zu machen, hierfür sei das folgende nur exemplarisch genannt: Bei den elementaren Relationen werden Beziehungen zwischen den Bedeutungen der Lexeme und Beziehungen zwischen deren Referenten vermischt. Beispielhaft sind hierfür Überlappungen bei den Relationen *Inclusion*, *Intersection* und *Partative Inclusion*. Darüber hinaus sind einige Relationen so allgemein gehalten, daß diese eigentlich nichts mehr aussagen, so etwa *Denotative*. Ferner ist zu bemerken, daß die Granularität der Relationen äußerst unterschiedlich ist: So sind etwa Relationen wie *Dimension* auf einer Stufe mit semantischen Merkmalen zu sehen, während andere, wie etwa *Bilateral Inclusion* oder *Component*, auf der Ebene der semantischen Beziehungen anzusiedeln sind. So ist etwa auch die elementare *Bilateral Inclusion* meiner Meinung nach bereits mit der semantischen Beziehung der Synonymie gleichzusetzen. Chaffin u. Hermann setzen hierfür jedoch eine komplexe Relation *Synonymity* an, die sich aus dem Cluster *Inclusion*, *Bilateral Inclusion* und *Intersection* zusammensetzt. Meiner Meinung nach ist dies nicht nur künstlich komplex gehalten und theoretisch zumindest diskussionswürdig – sondern macht dies eine praktische Anwendbarkeit der RET auf der Ebene der elementaren Relationen sehr schwierig.

Chaffin u. Hermann 1986 führen zur Stützung ihrer Theorie mehrere Studien an, in denen Probanden gefragt werden, ob zwischen einem Lexempaar eine bestimmte Relation besteht oder nicht. Sind etwa das Paar *Amsel* und *Vogel* sowie die Fragestellung *Ist Amsel in der Klasse Vogel enthalten?* gegeben, so können die Probanden mit *ja* oder *nein* antworten. Auf diese Weise soll das Verständnis der Relation durch die Versuchspersonen verifiziert werden.

Man muß jedoch bemängeln, daß die Untersuchung lediglich Aussagen darüber macht, ob die Probanden einem Lexempaar eine bestimmte Relation zusprechen. Sie zeigt nicht, ob die Versuchspersonen diese Relationen konsequent und selbständig zwischen Lexempaaaren spezifizieren können.

Auch ist nicht klar, woher Chaffin und Hermann ihre 31 Relationen beziehen, ob diese etwa ‚a priori‘ angenommen werden oder aber im Hinblick auf bestimmte Aufgabentypen abgeleitet wurden.

Die wichtigste Einschränkung ist jedoch, daß Chaffin u. Hermann ihre Relationen lediglich auf Lexempaare nicht aber auf Paare von Texteinheiten beziehen. Eine Eins-zu-eins-Übertragung auf Teiltex-te bzw. Hypertext-Knoten ist also auf alle Fälle zunächst einmal fraglich und soll im folgenden Abschnitt besprochen werden.

Die bis jetzt erwähnten Probleme waren theorieintern, jedoch gibt es auch spezifische Schwierigkeiten bei der Anwendung der RET für eine Link-Typisierung: Neben den oben bereits angeführten allgemeineren Problemen der RET ergeben sich noch spezifische für eine mögliche RET-Link-Typisierung. Zuerst stellt sich die Frage, welche Ebene der Relationen zur Typisierung verwendet werden soll. Die Elementaren Relationen sind wohl oftmals zu feinkörnig – obwohl es, wie oben beschrieben, auch tendenziell komplexere elementare Relationen gibt, die vergleichbar mit den semantischen Beziehungen *Synonymie* oder *Hyponymie* sind –, als daß sie allein Beziehungen zwischen verlinkten Textteilen beschreiben könnten. Hierfür bräuchte es Cluster von Relationen – und nichts anderes sind ja die von Chaffin u. Hermann beschriebenen komplexen Relationen.

Ein Problem der komplexen Relationen ist jedoch ihre Anzahl von 31. Für die praktische Arbeit der Typisierung auf Ebene der semantischen Oberfläche von Hypertexten ist eine solche hohe Zahl unterschiedlicher Relationstypen natürlich ungeeignet. Aber auch die Art mancher Relation macht deutlich, daß RET für Lexempaare und eben nicht für die Beschreibung semantischer Relationen zwischen Textteilen konzipiert war: So ließen sich etwa mit den komplexen Relationen der Familie 4 (Case-Relations) wohl gut Semantiken von Verben samt ihrer Ergänzungen beschreiben, zur Klassifizierung von Hypertext-Links dürften sie jedoch nur in ganz ausgefallenen Kontexten nützlich sein (Agent-Action: *Prozessor – berechnen*, Action-Instrument: *zeigen – Maus*). Ebenso scheinen auch die Part-Whole-Relationen eher zur Beschreibung der Hyponymie-Struktur von Substantiven im (elektronischen) Lexikon geeignet zu sein denn zur Typisierung von Links (Hyperonymie: *Markup-Sprache – HTML*, Kohyponymie: *Perl – PHP*). Denkbar wäre aber eine RET-Beschreibung der Relation zwischen dem Link-Text sowie den Schlüsselwörtern des Ziel-Knotens, die etwa aus den Metaangaben ausgelesen werden könnten.

Eine vergleichende Betrachtung von RST (siehe nächstes Kapitel) und RET

ergibt, daß bereits auf den ersten Blick hin Relationen der RST wie *Antithesis*, *Summary*, *Background* etc. besser zur Link-Typisierung geeignet zu sein scheinen als die Gegenbeispiele *Agent – Recipient*, *Functional Subordinate* oder *Incompatible* der RET.

Trotz aller Kritik: Die RET stellt ein breites Inventar fein- und grobkörniger Relationen zur Verfügung, jedoch scheint sie eher zur wissensrepräsentativen Codierung von Ressourcen auf der Ebene der semantischen Tiefenstruktur geeignet zu sein (siehe Kapitel 15) Bei der praktischen Untersuchung in Kapitel 17 soll sie daher nicht berücksichtigt werden.

16.2.2 Rhetorical Structure Theory

Als zweiter linguistischer Ansatz zur Typisierung von Textteilen wird in diesem Kapitel die RHETORICAL STRUCTURE THEORY (RST)¹ vorgestellt. Die RST stellt ein formales Werkzeug zur Beschreibung der thematisch-funktionalen Entfaltung eines Textes zur Verfügung. In der Computer Linguistik hat die RST besonders im Bereich der automatischen Textgenerierung weite Verbreitung gefunden.²

Die Theorie basiert auf der völligen Relationierung von Texteinheiten – auf einem „Netzwerk von Relationen zwischen Informationseinheiten“ (Lobin 1999 S. 160):

„Zwei Teile eines Textes sind danach nicht deshalb miteinander verbunden, weil es zwischen ihnen einen Link gibt oder die Verbindung mit bestimmten sprachlichen Mitteln (z.B. X, weil Y) ausgedrückt wird, sondern weil es eine inhaltliche Verbindung, eine Relation eines bestimmten Types zwischen den Textteilen gibt.“

(Lobin 1999 S. 160)

Kernaussage ist, daß mit einem begrenzten Inventar von RELATIONEN ‚Textteile‘ (= „spans of text“ Mann 1999a) zueinander in Beziehung gesetzt werden, so daß

¹ Eingeführt wurde die Theorie von Mann u. Thompson 1987. Eine kritische Würdigung findet sich bei Rösner u. Stede 1993. Online unter <http://www.sil.org/linguistics/rst/> (= Mann1999) stehen neben Beispielanalysen, Forschungsberichten und Referenzprojekten auch eine gute Einführung in die Thematik (Mann 1999a) sowie eine Auflistung und Erläuterung sämtlicher rhetorischer Relationen (Mann 1999b) zur Verfügung.

² Lobin 1999 schätzt die RST ein als „[...] das einzige verbreitete und in Sprachverarbeitungssystemen erfolgreiche deklarative Modell zur Strukturierung von Diskursrelationen [...]“ (Lobin 1999 S.160). Als beispielhafte Implementierung wäre etwa der *Intelligent Labelling Explorer* der Universität Edinburgh zu nennen [<http://cirrus.dai.ed.ac.uk:8000/ilex/>]. Im Online-Manual dieses Systems ist folgendes als Erklärung für die allgemeine Akzeptanz von RST angeführt: „Its popularity is perhaps best attributed to a combination of features: the emphasis on a functional conception of relations; the carefully presented set of relation definitions; the simply stated structural theory.“ (Knott 1998)

der eine Textteil eine spezifische, relative Rolle für den anderen Textteil erfüllt. Auf diese Weise wird die Kohärenz eines Textes hergestellt.

Die Dimension dieser Textteile ist nicht eindeutig festgelegt, in Beispielsanalysen umfassen sie im ersten Arbeitsschritt Teilsätze bzw. Sätze – als kleinste Einheiten, die funktionale Integrität besitzen – und fassen in den darauffolgenden Arbeitsschritten die ermittelten Relationen zum Teil als ‚Makro-Relationen‘ zusammen. Durch diese rekursive Relationsfindung entsteht eine Bottom-Up-Baumdarstellung der Relationsstruktur der Textes.¹

Eigenschaften der Relationen

Nach Knott 1998 weisen die Relationen der RST folgende Eigenschaften auf:

- Eine Relation nimmt auf den propositionalen Inhalt der betroffenen Textteile Bezug und macht zugleich eine Aussage über die Intention des Autors beim Produzieren dieser Relation.
- In der RST wird keine klare Aussage über die Beziehung einer Relation zu einer möglichen linguistischen Bezugseinheit gemacht. Vielmehr wird darauf hingewiesen, daß bestimmte Relationen gar nicht über solche Bezugseinheiten verfügen (müssen).
- Eine Relation verfügt in der Regel über einen NUKLEUS und einen SATELLITEN: „rhetorical relations also convey information about which span is more central to the writer's purposes“ (Knott 1998 Abschnitt *The Underlying theory*). Der funktionalen Denkweise der RST folgend, hat der Autor (auf jeder Ebene des Analysebaums) zentrale Ziele (= Nukleus) und weniger zentrale bzw. das Nukleus-Ziel unterstützende Nebenziele (= Satellit).
- Eine Relation muß man sich stets in folgenden Rahmen eingebettet vorstellen: „It is plausible to the analyst that it was plausible to the author that ... “ (Mann 1999b).

¹ In letzter Konsequenz müßte der oberste Knoten des Baumes dann die oberste Relation des Textes beinhalten. Man kann hier durchaus Ähnlichkeiten zum methodischen Vorgehen von van Dijk sehen: Während in der Makrostruktur-Theorie mittels spezieller Makro-Regeln eine Makro-Proposition ermittelt wird, versucht die RST von den Relationen der funktionalen Basiseinheiten eines Textes ausgehend durch rekursives Vorgehen so etwas wie eine Makro-Relation zu finden (siehe Kapitel 12 sowie van Dijk 1980a und van Dijk 1980b).

Typologie der Relationen

Die Relationstypen wurden, wie gesagt, von Mann u. Thompson 1987 eingeführt und sind für die RST noch immer maßgebend¹, dem Originalset wurden jedoch einige wenige neue Relationen hinzugefügt. Mann 1999b unterscheidet zwischen PRESENTATIONAL RELATIONS, SUBJECT-MATTER RELATIONS und MULTINUCLEAR RELATIONS. Knott 1998 veranschaulicht den Unterschied zwischen den Presentational Relations und den Subject-Matter Relations:

“SUBJECT-MATTER relations have as their effect that the reader *recognize* the relation in question; while PRESENTATIONAL relations have as their effect to *increase some inclination* in the reader.”

(Knott 1998 Abschnitt *The Underlying theory* [alle Auszeichnung im Original])

Nach Mann 1999b existieren die in folgenden Tabellen² aufgelisteten Presentational Relations und Subject-Matter Relations.

| Presentational Relation | Beispiel |
|-------------------------|---|
| Antithesis | <i>Dies ist ein Zeichen von Jobarmut, nicht Faulheit.</i> |
| Background | <i>Das Verhältnis Sprache und Denken gilt noch immer als unklar. Das folgende Buch bemüht sich um eine Einführung in ...</i> |
| Concession | <i>Jede Regel hat ihre Ausnahme, aber hier scheint es sich um einen Fehler zu handeln.</i> |
| Enablement | <i>Gib mir die Nummer, dann rufe ich an.</i> |
| Evidence | <i>Der Minister ist zurückgetreten, soeben erreichte uns eine entsprechende Meldung der dpa.</i> |
| Justify | <i>Er mußte es tun, weil sein Vater es so wollte.</i> |
| Motivation | <i>Bitte helfen Sie der Caritas. Ihre Spende kommt sofort den Bedürftigen zu gute.</i> |
| Preparation | <i>Dieses Buch handelt von Syntax. Unter Syntax versteht man ...</i> |
| Restatement | <i>Wir hatten nicht mit einer solchen Reaktion der Öffentlichkeit gerechnet. Tatsächlich war die öffentliche Meinung auch schwer vorauszusagen.</i> |
| Summary | <i>Es gibt Presentational, Subject-Matter und Multinuclear Relations, insgesamt als drei Typen von Relationen.</i> |

Tabelle 18: Presentational Relations der RST

| Subject-Matter Relation | Beispiel |
|-------------------------|---|
| Circumstance | <i>Ich bin in die Kälte gegangen, ohne mir die Mütze aufzusetzen.</i> |
| Condition | <i>Wenn du dir die Mütze aufsetzt, wirst du nicht krank.</i> |
| Elaboration | <i>Ich bin krank. Gestern war ich ohne Mütze in der Kälte.</i> |
| Evaluation | <i>Du warst ohne Mütze draußen und bist jetzt krank? Da bist du wohl selber schuld!</i> |
| Interpretation | <i>Die Partei X hat mit 51% gewonnen. Ein überraschend knappes Resultat.</i> |
| Means | <i>Ich putze den Boden mit dem neuen Wischmob.</i> |
| Non-volitional Cause | <i>Ich arbeite heute nicht, weil ich krank bin.</i> |

¹ Hierzu äußert sich Mann: „The set of relations is in principle open, but the set above, defined in Mann & Thompson 1988 (which is currently the definitive paper on RST) has proven effective for many purposes.“ (Mann 1999a Abschnitt *More Relations*)

² Nuklei sind dort im Fettdruck dargestellt.

| Subject-Matter Relation | Beispiel |
|-------------------------|---|
| Otherwise | <i>Sie müssen sich an die Regeln halten, sonst müssen Sie gehen.</i> |
| Purpose | <i>Um nicht krank zu werden, setze dir deine Mütze auf.</i> |
| Solutionhood | <i>Kaputte Hände vom Spülen? Hier hilft das neue Spülmittel X!</i> |
| Unconditional | <i>In keinem Fall ist der Autor verantwortlich, auch wenn dies von der Presse fälschlicherweise behauptet wird.</i> |
| Unless | <i>Ich arbeite morgen wieder, außer es geht mir schlechter.</i> |
| Volitional Cause | <i>Ich arbeite heute nicht, weil ich nicht mag.</i> |
| Volitional Result | <i>Die Polizei mußte den Verkehr kontrollieren, tausende Urlauber brachen gleichzeitig in die Ferien auf.</i> |

Tabelle 19: Subject Matter Relations der RST

Zusätzlich treten noch Multinuclear Relations auf, bei denen die Rollenverteilung zwischen Nukleus und Satelliten nicht eindeutig ist:

| Multinuclear Relation | Beispiel |
|-----------------------|---|
| Contrast | <i>Die erste Halbzeit war äußerst langweilig, die zweite hingegen wurde sehr spannend.</i> |
| Joint | <i>Ich bin ohne Mütze spazierengegangen und habe mich dabei erkältet.</i> |
| List | <i>(Man nehme:) 200 Gramm Mehl, ¼ Liter Milch, 2 Eier ...</i> |
| Restatement | <i>Wir hatten nicht mit einer solchen Reaktion der Öffentlichkeit gerechnet. Tatsächlich war die öffentliche Meinung auch schwer vorauszusagen.</i> |
| Sequence | <i>Gestern klingelte das Telefon ständig. In einer Tour kamen neue Mails. Oft klingelte jemand an der Tür.</i> |

Tabelle 20: Multinuclear Relations der RST

Auf den ersten Blick fällt auf, daß eine Relation, nämlich *Restatement*, doppelt aufgeführt ist, einmal als Presentational und nochmals als Multinuclear Relation. Mann 1999b meint hierzu, daß bei einem Presentational *Restatement* der Nukleus der Relation zentraler für die Absichten des Autors sei. Bei einem Multinuclear *Restatement* seien die beiden in Relation zueinander stehenden Textteile gleichwertig im Bezug auf die Absichten und Ziele des Autors. Die Unterscheidbarkeit erscheint mir aber in der Praxis schwer durchführbar zu sein, das zeigt auch die Analyse „The Zero Population Growth Text“ auf der RST-Homepage (Mann 1999 [zpg\index.htm]).

Jede Relation ist definiert über ein Bündel von CONSTRAINTS, die entweder den Nukleus und den Satelliten jeweils einzeln betreffen oder aber Bedingungen an die Kombination der beiden stellen. Zur Verdeutlichung ist im folgenden jeweils ein Beispiel für die drei Relationstypen aufgeführt.¹

¹ Die folgenden drei Tabellen sind übernommen aus Mann 1999b Abschnitt *RST Relation Definitions*. R steht jeweils für *reader*, W für *writer*, N für *Nukleus* und S für *Satellit*.

Relation *Evidence* (Presentational Relation)

| Relation Name | Constraints on either S or N individually | Constraints on N + S | Intention of W |
|---------------|--|---|------------------------------|
| Evidence | on N: R might not believe N to a degree satisfactory to W on S: R believes S or will find it credible | R's comprehending S increases R's belief of N | R's belief of N is increased |

Tabelle 21: Relation *Evidence* als Beispiel einer PRESENTATIONAL RELATION der RSTRelation *Condition* (Subject Matter Relation)

| Relation Name | Constraints on either S or N individually | Constraints on N + S | Intention of W |
|---------------|---|--|---|
| Condition | on S: S presents a hypothetical, future, or otherwise unrealized situation (relative to the situational context of S) | Realization of N depends on realization of S | R recognizes how the realization of N depends on the realization of S |

Tabelle 22: Relation *Condition* als Beispiel einer SUBJECT MATTER RELATION der RSTRelation *Contrast* (Multinuclear Relation)

| Relation Name | Constraints on each pair of N | Intention of W |
|---------------|---|--|
| Contrast | no more than two nuclei; the situations in these two nuclei are (a) comprehended as the same in many respects (b) comprehended as differing in a few respects and (c) compared with respect to one or more of these differences | R recognizes the comparability and the difference(s) yielded by the comparison is being made |

Tabelle 23: Relation *Contrast* als Beispiel einer MULTINUCLEAR RELATION der RST

Die Relation *Evidence* ließe sich etwa wie folgt paraphrasieren: Der Sachverhalt / der Gedanke / die Einstellung die sich im Nukleus ausdrückt, wird nach Ansicht des Autors vom Leser nicht im gewünschten Maß akzeptiert. Der Autor geht aber davon aus, daß der Sachverhalt / der Gedanke / die Einstellung, die er im Satelliten ausdrückt, vom Leser auf alle Fälle geglaubt wird (bzw. es sich hierbei um einen objektiven Beweis handelt). Begreift und akzeptiert der Leser nun den Satelliten, so steigert dies nach Ansicht des Autors auch dessen unbewußte Akzeptanz des Nukleus.

Bei der Relation *Condition* hingegen soll der Leser nach Intention des Autors bewußt erkennen, daß der Nukleus direkt von der Erfüllung der Bedingung abhängt, die im Satelliten ausgedrückt ist. Vergleichbar ist die bi-nukleare Relation *Contrast*: Nach Intention des Autors erkennt der Leser bewußt die Vergleichbarkeit der beiden Nuklei. Der Vergleich selbst bringt dann die Unterschiedlichkeit der Sachverhalte / Gedanken / Einstellungen zum Ausdruck.

Weitere Beschreibungsebenen

Das zentrale Konstrukt von RST sind die oben dargestellten Rhetorischen Relationen. Eine solche Relation wird jedoch (nach der ursprünglichen Theorie) nicht direkt auf Textteile angewandt, sondern im Rahmen einer SCHEMA-APPLIKATION, die ihrerseits wiederum von einem SCHEMA abgeleitet ist.¹ Den Zusammenhang zwischen Relationen, Schemata und Schema-Applikationen erklärt Knott wie folgt:

„In RST, relations are not mapped directly onto texts; they are fitted onto structures called **schema applications**, and these in turn are fitted to text. Schema applications are derived from simpler structures called **schemas**.“

(Knott 1998 Abschnitt *The Underlying theory* [alle Auszeichnungen im Original])

Im folgenden erläutert Knott, wie mit Hilfe dieser drei Konstrukte eine hierarchische rhetorische Struktur eines Textes aufgebaut werden kann:

„A **rhetorical structure tree** is a hierarchical system of schema applications. A schema application links a number of consecutive spans, and creates a complex span which can in turn be linked by a higher level schema application. This enables tree structures to be built--it is a central claim of RST that the structure of every coherent discourse can be described by a single rhetorical structure tree, whose top schema application creates a span encompassing the whole discourse.“

(Knott 1998 Abschnitt *The Underlying theory* [Auszeichnung im Original])

Warum der Aufbau eines rhetorischen Strukturbaums lediglich mit Hilfe von Schema-Applikationen und nicht unter bloßer Verwendung der rhetorischen Relationen möglich scheint, ist nicht einleuchtend. Bezeichnend ist, daß in aktuellen einführenden Texten zur RST auf Schemata und Schema-Applikationen lediglich mit Bemerkungen folgender Art eingegangen wird:

„A second level of structures, called *schemas*, is not presented here.“

(Mann 1999a Abschnitt *Structures*)

Da meiner Meinung nach eine tiefere Einsicht in die Notwendigkeit von Schemata und Schema-Applikationen durch die Vertreter der Theorie nicht vermittelt werden kann und diese Konzepte auch in deren eigenen Analysen offensichtlich nicht verwendet werden, wird im Rahmen der vorliegenden Arbeit von der direkten Anwendbarkeit der Relationen auf den konkreten Text ausgegangen.

¹ Die Termini SCHEMA und SCHEMA-APPLIKATION haben nichts mit der in Kapitel 9.2.2 eingeführten SCHEMATHEORIE der kognitiven Linguistik gemein.

Analysemodell von RST

Lobin 1999 präsentiert auf Seite 161 eine einfache Beispielsanalyse mit Hilfe von RST. Gut zu erkennen ist dort, wie eine Bottom-up-Analyse durchgeführt wird, die zur Relation *cause* für den Gesamttext führt. Auf unterster Ebene werden in diesem Fall Beziehungen zwischen Satzteilen (3 in Relation zu 4) als Relationen dargestellt, um dann diese RST-strukturierten Textteile auf höherer Ebene wieder als Bestandteile einer neuen Relation zu verwenden (3+4 in Relation zu 5 bzw. 3+5 in Relation zu zwei usw.):

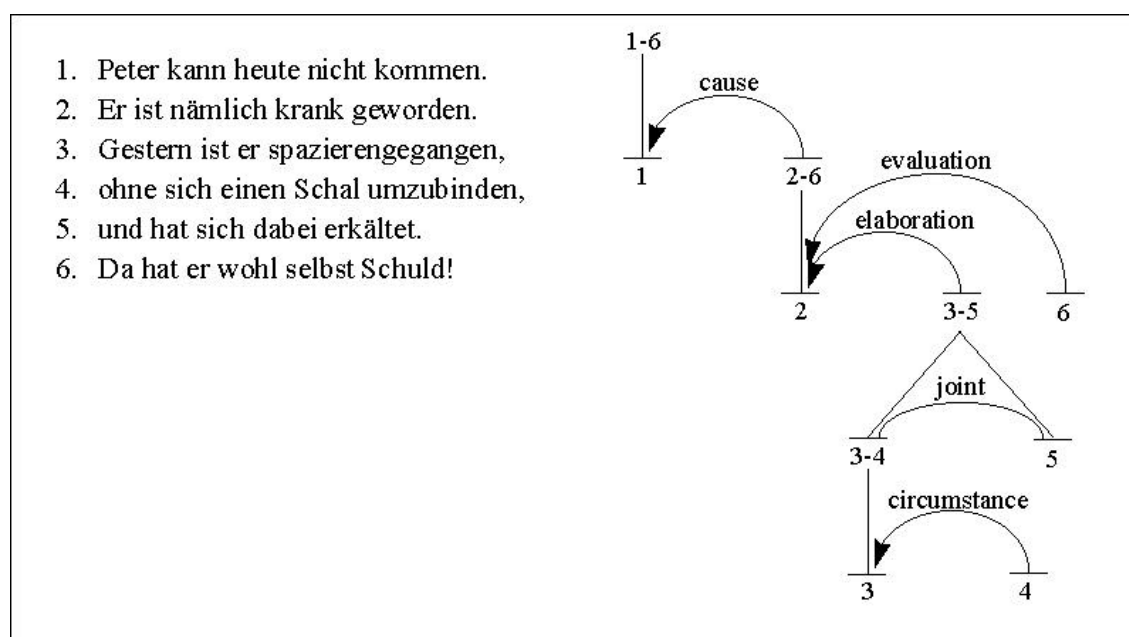


Abbildung 12: Einfach Textanalyse mit RST

Kritische Würdigung der RST

Im folgenden werden einige mögliche Probleme erwähnt, die entweder theorieintern begründet sind oder aber bei einer möglichen RST-Analyse von Hypertexten auftreten könnten.

Unklare Rolle der SCHEMATA:

Ein erster Kritikpunkt ist die oben bereits angesprochene etwas verwirrende Einführung der Konstrukte SCHEMATA sowie SCHEMA-APPLIKATIONEN in Mann u. Thompson 1987. Jedoch tragen sowohl in aktuellen Beiträgen (Mann 1999a oder Lobin 1999) sowie in den zahlreichen praktischen Analysen von Mann 1999 lediglich die rhetorischen Relationen zum Aufbau des Strukturbaumes bei. Daher können wohl auch in der vorliegenden Untersuchung die rhetorischen

Relationen ohne zur Hilfenahme der Schema-Applikationen auf das Corpus angewendet werden.

Subjektivität auf mehreren Ebenen:

Bei der Arbeit mit der RST darf nie die Tatsache aus den Augen verloren werden, daß man in starkem Maß mit Subjektivität konfrontiert ist: Die Relationen sowie das Vokabular zur Bezeichnung derselben sind von Mann und Thompson wohlüberlegt jedoch nach ihrem subjektiven Empfinden ausgesucht. Der diese RST-Relationen verwendende Analysierende macht wiederum subjektive Aussagen über Annahmen des Autors, die dieser bzgl. des Wissensaufbaus beim potentiellen Leser gemacht haben soll.¹ Zur Mächtigkeit des RST-Vokabulars selbst können wohl erst im Laufe der praktischen Anwendung Aussagen gemacht werden. Für die RST spricht auf alle Fälle ihre lange Anwendungsgeschichte und eine ‚gewisse Eingängigkeit‘ der gewählten Relationen auf den ersten Blick.

Unklare Unterscheidung zwischen Subject Matter und Presentational Relations: Etwas vage ist meiner Meinung nach auch die strikte Unterscheidung zwischen Presentational Relations und Subject Matter Relations: In der Praxis werden höchstwahrscheinlich Fälle des Übergangs dieser beiden Kategorien auftreten, bei denen es schwer sein könnte zu entscheiden, ob der Autor nun den Leser durch von diesem erkennbare Subject Matter Relations überzeugen oder den Leser subtiler mit Presentational Relations beeinflussen will.

Hier tritt eine weitere Problematik der RST zu Tage: Die Vermischung von mehrheitlich rhetorisch-funktional definierten und einigen eher logisch ausgerichteten Relationen, etwa *Joint* oder *List*. Diese Inkonsistenz ist bei der Verwendung der RST zu beachten, kann man mit ihr eben wahrscheinlich eher – im Hinblick auf die rhetorische Funktionalität der Link-Struktur – die Entfaltung der Kohärenz eines Lesepfades beschreiben, als eine logische Struktur der Knoten aufbauen. Hierfür benötigte man zusätzliche Relationen wie *Enthaltensein*, *Teil von* oder *Ersetzung durch* etc.

¹ Hier sei nochmals Mann selbst zitiert: „It is plausible to the analyst that it was plausible to the author that ... ” (Mann 1999b). Mann sieht diese Konstruktion als Hilfestellung, um sich als Benutzer der RST über den subjektiven bzw. objektiven Charakter der Untersuchung im Klaren zu sein: „Because of its specification of the role of the observer, it provides a basis for statements about the objectivity and subjectivity of the analysis.“ (Mann 1999a) Ob die bloße Einbettung einer rhetorischen Relation in den obigen Rahmen nun ein mächtiges Werkzeug zur Überprüfung der Subjektivität darstellt, muß jedoch eher bezweifelt werden.

Übertragung der RST von Texten auf Hypertexte:

Ein möglicherweise schwerwiegendes Problem ist die ungeklärte Anwendung der RST auf Hypertexte. Die RST wird traditionell für die Analyse von Texten benutzt, Hypertexte enthalten neben textuellen meist jedoch auch eine gewisse Anzahl multimedialer Elemente. Lobin 1999 spricht sich in diesem Zusammenhang für eine „hypermedial erweiterte Version von RST“ aus:

„Grundsätzlich sollen die Relationen in der hypermedial erweiterten Version von RST zwischen allen Typen von Informationseinheiten – Phrasen, Sätze, Textsegmente[n], Überschriften, Bilder[n] und Bildteilen – möglich sein und dabei insbesondere auch modalitätsübergreifende Verbindungen zu etablieren erlauben. Eine Grafik kann dann z.B. als die Erläuterung eines Textsegments verstanden werden oder ein Textsegment als eine Bewertung eines Bildes.“

(Lobin 1999 S. 162)

Während auch bei der ursprünglichen Version der RST ein gewisses Ausweichen auf die Frage zu bemerken ist, was denn die konstituierenden Einheiten des Textes seien, zeigen doch die Analysen von Mann 1999, daß auf unterster Analyseebene Propositionen die Rolle der ‚funktionstragenden Einheiten‘ spielen. Mit Hilfe der oben dargestellten Bottom-up-Prozedur werden in weiteren Analyseschritten dann Sätze und Textteile als Funktionsträger aufgebaut.

Auch Lobin 1999 sieht Propositionen als konstituierende Einheiten der Analyse an:

„Propositionen sind also deshalb die kleinsten Einheiten der RST-Strukturierung, weil die sie miteinander verknüpfenden Relationen auf unterschiedliche Art realisiert werden können, ohne daß dadurch der eigentliche Gehalt der Proposition berührt wird.“

(Lobin 1999 S. 163)

Mit Hilfe der RST ließen sich somit Relationen zwischen wie auch immer realisierten Bestandteilen von Hypertexten darstellen und somit die Funktionen dieser Bestandteile beschreiben, ohne sich dabei auf die strukturelle Ebene der medialen Realisierung zu begeben. Ein solcher Bestandteil kann, Lobin folgend, „minimales mediales Element“ genannt werden, „was ein Satz, ein Bild, aber auch ein Textsegment oder nur ein Teil eines Bildes sein kann.“ (Lobin 1999 S.163) Eine so ausgelegte RST stellt ein pragmatisch ausgerichtetes Werkzeug zur Analyse der Funktion zur Verfügung, von welcher abgeleitet bei Bedarf natürlich auch eine Analyse der Struktur der medialen Realisierung folgen kann. Allgemein anzumerken ist jedoch, daß die Textlinguistik nicht über geeignete Begriffe und Methoden verfügt, um Funktion und (noch in größerem Maß) Struktur von nichttextuellen Elementen von Hypertexten zu beschreiben!

Abschließend kann festgestellt werden, daß sich die RST nach ersten theoretischen Betrachtungen durchaus zur funktionalen bzw. thematisch-strukturellen Analyse bzw. Typisierung von Links eignen kann. Dies gilt mit den oben gemachten Einschränkungen.

16.3 Hypertext-spezifische Ansätze

Nach den genuin linguistischen Ansätzen der beiden letzten Kapitel werden nun Vorschläge aus der Informationswissenschaft geschildert, die speziell zur Beschreibung von Relationen in Hypertexten konzipiert wurden.

Nach den einführenden Feststellungen des nächsten Kapitels kommt danach eine Auswahl älterer Hypertext-Systeme zur Sprache, die mit ihrer Fähigkeiten zur Link-Typisierung teilweise den heutigen Möglichkeiten von HTML überlegen waren. In einem zweiten Schritt werden dann genau diese Möglichkeiten von HTML geschildert. Es wird sich dabei zeigen, daß die typisierten Links im HTML-Corpus zumeist mit ‚Workarounds‘ gelöst wurden – also nicht mit syntaktischen Mitteln von HTML, die explizit zur Link-Typisierung geschaffen wurden.

16.3.1 Allgemeine Anmerkungen und kleine Pilotstudie

In vielen Publikationen wird dafür plädiert, Hypertexte mit Typisierten Links zu versehen. Wie solche Link-Typen aber beschaffen sein und in welchem Umfang sie auftreten sollen, ist hingegen umstritten.¹ Denkbar wären mindestens drei Herangehensweisen: syntaktisch-formal, semantisch und pragmatisch-rhetorisch.²

Beim Aufbau eines kontrollierten Vokabulars für Link-Typen stellt sich nun die Frage, welche Ausformungen dieses haben soll. Als erste Annäherung an die Thematik wurde eine bescheidene qualitative empirische Pilotstudie unter HTML-Entwicklern (60% der Stichprobe) und ‚normalen‘ Internet-Benutzern

¹ Exemplarisch hierfür: Kuhlen 1991, Schnupp 1992 sowie Gerdes 1997

² Kapitel 6.3.2 beschreibt ausführlich die formalen Möglichkeiten, Links zu unterscheiden. Hier sind Gegensatzgruppierungen wie etwa intra- vs. extra-hypertextuell bzw. uni- vs. bi- vs. multidirektional denkbar. Linguistische Ansätze, die sich mit Textverknüpfungstypen befassen, ziehen natürlich mediumsbedingt solche hypertext-spezifischen Verknüpfungen nicht in Betracht. Links von Hypertext-Systemen, deren Implementierung eine Typisierung zulassen, sind zumeist semantischer und/oder pragmatisch-rhetorischer Natur.

(40%) durchgeführt. Die Stichprobe ist mit 30 Probanden eher klein, die Ergebnisse dürfen daher keinesfalls als repräsentativ angesehen werden. Erste, recht allgemeine Erkenntnisse über mögliche Akzeptanz und gewünschte Ausrichtung typisierter Links sind jedoch bereits ersichtlich.

Die Versuchspersonen erhielten den im Anhang aufgeführten Fragenkatalog. Die erste Gruppe der Fragen zielt hierbei auf die Erfahrung im Umgang mit Hypertext- und Markup-Techniken ab. Die zweite Gruppe fragt explizit nach Erfahrung im Umgang mit typisierten Links. Die dritte Fragengruppe zielt auf die Einschätzungen typisierter Links seitens der Personen ab, die damit bereits zu tun hatten.

Die Umfrage brachte interessante Einzelergebnisse. Typisierte Links sind weitestgehend bekannt (ca. 80 % der Befragten kennen typisierte Links. Bevorzugt wurde auf die Frage, welche WWW-Seiten solche Links einsetzen, u.a. genannt: *www.sport1.de* sowie *www.spiegel.de*) und werden mehrheitlich für sinnvoll erachtet (ca. 70% derjenigen, die typisierte Links kennen). Die Befürworter typisierter Links äußerten sich bzgl. deren Anzahl sehr unterschiedlich: Die Spanne reicht hierbei von fünf bis 25 Typen. Nahezu übereinstimmend (ca. 75%) wurde eine ‚zu große‘ Anzahl als – für das Verständnis der Link-Bedeutung – problematisch erachtet.

Personen, die mit der Nutzung des WWW erfahren sind, können laut Umfrage sehr gut zwischen navigationsspezifischen (z.B. der Typ *Nächstes Kapitel*), inhalts- (*Definition*) und browser-spezifischen Link-Typen (*Pop-Up-Fenster*) unterscheiden; 65% der Erfahrenen äußerten sich in diesem Sinne. Unerfahrene Benutzer bemerken zwar einen Unterschied, können diesen aber nicht konkretisieren. Personen, die mit Markup-Sprachen vertraut sind, plädieren für eine Unterscheidung zwischen Navigations-, Inhalts- und Browser-Verhalten-Links: 70% der HTML Entwickler äußerten sich in diese Richtung. Personen, die sich beruflich mit den Möglichkeiten der semantischen Auszeichnung von Ressourcen beschäftigen (lediglich zwei Personen der Stichprobe) plädieren für eine Link-Auszeichnung, die sich eng an Relationen der Wissensrepräsentation anlehnen (Typ *Teil-von-Relation* u.ä.).

Dieses ‚Brainstorming‘ unter 30 Probanden hat im einzelnen folgende Vorschläge für Link-Typen erbracht:¹

| Link-Typ | Absolute Häufigkeit der Nennungen |
|------------------|--|
| Anwendung | 3 |
| Beispiel | 7 |
| Bild | 9 |
| Definition | 5 |
| Drucken | 7 |
| Email | 20 |
| Extern(er Link) | 15 |
| Fortsetzung | 2 |
| Fußnote | 11 |
| Home | 17 |
| Index | 5 |
| Intern(er Link) | 1 |
| Pop-Up-Link | 18 |
| Programmbeispiel | 5 |
| Suchen | 19 |
| Vertonung | 12 |
| Video | 9 |

Tabelle 24: Ergebnis der Pilotstudie

Auffällig ist, daß computer-spezifische Phänomene wie *Email*, *Drucken*, *Home* oder *Suchen* häufiger genannt wurden als solche, die auch in anderen Medien vorhanden sind (etwa *Fußnote* oder *Bild*). Auch allgemeine, semantische oder logische Relationen zwischen Textteilen bzw. Knoten wurden eher selten genannt.

Diese heuristische Typensammlung ließe sich in folgende Klassen unterteilen:

| Typklasse | Beispiel |
|---|--|
| Typen, die Auskunft über den semantischen Gehalt des Zielknotens geben. | <i>Beispiel, Definition, Anwendung, Fortsetzung, Programmbeispiel, Fußnote</i> |
| Typen, die Auskunft über die mediale Gestalt des Zielknotens geben. | <i>Vertonung, Visualisierung, Video</i> |
| Typen, die Auskunft über die navigationsspezifische Einordnung des Zielknotens in den Gesamtkontext des Hypertextes geben. | <i>Home, Fortsetzung, Index</i> |
| Typen, deren Ziel die Ausführung eines hypertext-externen Programmes ist. | <i>Email, Drucken</i> |
| Typen, die nähere Informationen über die Art des Browser-Verhaltens geben. | <i>Pop-Up-Link</i> |
| Typen, die Informationen darüber geben, ob sich der Zielknoten innerhalb oder außerhalb des aktuellen Hypertextes befindet. | <i>Intern, Extern</i> |

Tabelle 25: Heuristische Typensammlung der Pilotstudie

¹ Es handelte sich hierbei um eine freie Befragung. Synonyme wurden hier – in einer weiten Interpretation – zusammengefaßt.

Bereits auf den ersten Blick ist erkennbar, daß eine einheitliche Taxonomie der Link-Typen wohl kaum aufzustellen ist. Denn einerseits wurden Objekte als Typen vorgeschlagen (*Video*, *Visualisierung* im Sinne von *Grafik*, *Photo* etc.) andererseits Konzepte aus der Rhetorik bzw. den Printmedien wie *Definition*, *Beispiel* oder *Fußnote*. Desweiteren finden sich hypertext-spezifische bzw. edv-spezifische Konzepte wie *Interner Link*, *Externer Link*, *Email*, *Drucken* oder *Programmbeispiel*.¹ Eine mögliche Grobeinteilung wäre nun, sämtliche Links danach zu unterscheiden, ob sie hypertext-intern oder -extern sind. Dann könnten in einem nächsten Schritt mediale Links von Navigations-Links bzw. semantischen Links getrennt werden. Was ist aber, wenn etwa ein Beispiel durch einen Video-Knoten realisiert ist? Soll man dann nach dem medialen Typ, nach dem semantischen oder nach beiden Typen klassifizieren? Wie soll man mit Links umgehen, die nicht nur aus dem Hypertext hinausweisen, sondern deren Aktivierung gar ein neues Programm startet, wie dies etwa beim Link-Typ *Email* der Fall ist?

Diese Fragen können freilich in einer ersten Annäherung nicht gelöst werden. Sie deuten aber an, daß bei genauerer Betrachtung die Problematik noch deutlich komplexer werden könnte.

Nach dieser ersten, praktischen Herangehensweise werden in den nächsten Kapiteln Vorschläge zur Link-Typisierung aus bestehenden Hypertext-Systemen besprochen.

16.3.2 Typisierte Links in älteren Hypertext-Systemen

Die hier vorgestellten Typisierungen sind aus einer deutlich größeren Anzahl von Vorschlägen der Sekundärliteratur ausgewählt worden.² Im folgenden Kapitel soll lediglich ein Eindruck vermittelt werden, wie heterogen die einzelnen Ansätze sind. Aus diesem Grund wird nachstehend ein recht allgemeiner und ein sehr spezieller Typisierungsvorschlag beschrieben.

¹ *Programmbeispiel* wurde lediglich von der Gruppe der HTML-Entwickler genannt. Dieser Link-Typ ist als textsortenspezifisch für Hypertexte aus dem Bereich der technischen Dokumentation zu sehen.

² Einen (für die damaligen Verhältnisse) vollständigeren Überblick gibt Kuhlen 1991.

Duncans allgemeine Typisierung

Wie strittig der gesamte Komplex ist, zeigt das folgende Zitat, in dem sich Schnupp etwas darum ‚windet‘, eine konkrete Aussage darüber zu machen, warum er die Klassifizierung von Duncan 1989 als besonders gelungen ansieht:

„Sofern sich die nötigen oder zweckmäßigen Verweistypen nicht ‚von selbst‘ aus der konkreten Anwendung heraus ergeben, kann als allgemeine, für die meisten zu beschreibenden Sachverhalte brauchbare Klassifizierung ein von *Duncan* [...] stammender *Typenkatalog* empfohlen werden. Dieser unterscheidet die in semantischen Netzen oder auch in Hypertext-Datenbasen üblicherweise vorkommenden *links* nach ihrer *abstrakten Bedeutung* im Informationskontext, das heißt nach den von ihnen repräsentierten, logischen Relationen [...]“

(Schnupp 1992 S. 1991 [alle Auszeichnungen im Original])

Duncan nimmt also eine Typisierung auf der Ebene der Hypertext-Datenbasis – auf der semantischen Tiefenstruktur – vor. Hierbei setzt er folgende allgemeine Relationen an:¹

| Relation | Beschreibung |
|---------------------|--|
| Mengenzugehörigkeit | A ist ein B. B enthält A. |
| Demonstration | A ist ein Beispiel für B. B kann durch A gezeigt werden. |
| Ursache | A verursacht B. B ist ein Ergebnis von A. |
| Verwendung | A verwendet B. B kann von / durch A benutzt werden. |
| Besitz | A besitzt B. B wird besessen von A. |
| Enthaltensein | A ist ein Teil von B. B besteht (nicht unbedingt ausschließlich) aus A. |
| Konsequenz | A folgt aus B. B inkludiert A. |
| Ähnlichkeit | A ist ähnlich zu B. B ist ähnlich zu A. |

Tabelle 26: Link-Typen nach Duncan

Schnupp bemerkt zu Recht, daß lediglich die Relation *Ähnlichkeit* reflexiv ist (Schnupp 1992 S.191). Alle anderen sind gerichtete Relationen. Übernimmt man diese Relationen also unverändert, könnten damit erst einmal lediglich unidirektionale Links typisiert werden, denn

¹ Übernommen aus Schnupp 1992 S.192. Die Beschreibungen der Relationen wurden ins Deutsche übersetzt. Die ursprünglich von Duncan angesetzte Relation *including* teilt Schnupp auf in ein *physisches Enthaltensein* sowie ein *logisches*, welches er als *Konsequenz* benennt.

„[...] die zweite in der Tabelle genannte Relation [ist] im allgemeinen nicht die Umkehrung der ersten sondern eine (etwas) andere Interpretation.“

(Schnupp 1992 S.191)

Meiner Meinung nach, sind die Interpretationen jedoch so nahe an der „Umkehrung“, daß diese Umkehrrelationen durchaus zur Konstruktion bidirektionaler Links verwendet werden können. Lediglich bei der Namengebung der Relationen könnte es Schwierigkeiten geben, so etwa bei der Festlegung der Umkehrrelation zu *Besitz*. Aufgelöst werden könnte dies hier durch Verwendung von Verben als Link-Etikette: *besitzt* vs. *wird besessen*.

Vergleicht man jedoch die von Duncan vorgeschlagenen Relationen mit denen, die in den Corpus-Hypertexten verwendet werden, so finden sich so gut wie keine Überschneidungen: Lediglich die Relation *Demonstration* ließe sich anstelle des Link-Typs *Anzeigebeispiel* von SELF verwenden (siehe auch Kapitel 17.3.1).

Die Probleme liegen darin begründet, daß Duncans Relationen, wie gesagt, die Beziehungen von Knoten auf der Ebene der semantischen Tiefenstruktur beschreiben, die sich nicht automatisch auf die Ebene der Oberfläche übertragen lassen. Duncans Modell soll daher nicht für die praktische Untersuchung in Kapitel 17 verwendet werden.

Author's Argumentation Assistant (AAA)

Der AUTHOR'S ARGUMENTATION ASSISTANT (AAA) ist der Prototyp einer Autorenumgebung zur Produktion argumentativer Hypertexte (vgl. Schuler 1990). Der AAA bedient sich hierbei Typisierter Links. Die Link-Typen beanspruchen jedoch keine universelle Gültigkeit für sämtliche Textsorten, sondern sind auf die Domäne der argumentativen Texte beschränkt. Insofern ist der Ansatz bereits im voraus lediglich für einen kleinen Teilbereich des Corpus' interessant – hier herrschen informative Texte vor –, jedoch bietet der AAA ein sehr gutes Beispiel einer domänenspezifischen Link-Typisierung. Insofern sollten die Relationen des AAA für die Link-Typisierung zumindest in Erwägung gezogen werden: Mit dem AAA lassen sich vier Typen von Knoten festlegen: *issue* (i), *argument* (a), *position* (p) und *fact* (f).

Darauf aufbauend lassen sich folgende Relationen zwischen Knoten ausdrücken:¹

| Relation | Beschreibung |
|--------------------|---|
| serve (A, Z) | Z(i) dient in irgendeiner Form A(i) . |
| replacement (A, Z) | Z(i) ersetzt A(i). |
| suggestion (Z, A) | A(i) ist ein Vorschlag von Z(i, a, p, f). |
| answer (A, Z) | Z(p) ist eine Antwort auf A(i). |
| objection (A, Z) | Z(a) ist Einwand gegen A(p). |
| support (A, Z) | Z(a) stützt A(p). |
| contributes (A, Z) | Z steuert in irgendeiner Form etwas zu A bei. |
| reference (A, Z) | Z(f) ist eine Referenz von A(i, p, a, f). |
| contradicts (A, Z) | Z widerspricht in allgemeiner Art A. |

Tabelle 27: Link-Typen des AUTHOR'S ARGUMENTATION ASSISTANT (AAA)

Bei der Betrachtung eines Knotens aus SELF fällt auf, daß die AAA-Typen und -Relationen durchaus auftreten, etwa in:

„Verweise können in HTML an jeder beliebigen Stelle im Text stehen. Wenn Sie jedoch einmal Text lesen, in dem jedes zweite Wort ein Verweis ist, werden Sie schnell merken, daß dies den Lesefluß ungemein stört. Der Grund dafür ist, daß Verweise immer gleich die Aufmerksamkeit auf sich ziehen und den Leser von seiner eigentlichen Aufgabe, dem geistigen Erfassen des im Text Gemeinten, ablenken. Um so wichtiger ist es, daß Verweise innerhalb des Fließtextes dem Anwender keine Rätsel aufgeben, sondern sofort erfaßbar sind.“

(SELF [tcai.htm] Abschnitt *Ehrliche und aussagekräftige Verweistexte verwenden*)

Eine genaue Analyse soll hier nicht erfolgen, die groben Verhältnisse sind wohl auf den ersten Blick sichtbar: So führt Satz 2 eine *objection* gegen das *issue* aus Satz 1 an. Dieser Einwand wiederum erfährt einen *support* von Satz 3 usw. Wichtig ist nun zu bemerken, daß sich die gesamte Argumentation innerhalb eines Knotens von SELF abspielt. Die Konstellation, daß ein gesamter Knoten zu einem anderen in einer der oben aufgeführten Relationen steht ist selten: Lediglich die Ziel-Knoten, die eine bildliche Darstellung der im aktuellen Knoten besprochenen Fakten beherbergen, könnten als *serve*-, *contribute* oder *support*-Relation gesehen werden.

Als Anwendungsbereich für den AAA ist zweierlei denkbar: Entweder man setzt ihn zur Analyse der Argumentationsstruktur von Texten oder Hypertexten ein – ähnlich wie die RST (siehe Kapitel 16.2.2) – oder aber zur Repräsentation dieser Argumentationsstruktur auf Ebene der semantischen Tiefenstruktur eines Hypertext-Systems. Für eine Oberflächentypisierung der primär informativen

¹ A steht hierbei für *Aktueller Knoten*, Z für *Zielknoten*. In der Spalte *Beschreibung* bezeichnen die Kleinbuchstaben *i*, *a*, *p* und *f* die oben eingeführten möglichen Knoten-Typen. Kann der Zielknoten beispielsweise ein *issue* oder eine *position* sein, so wird dies auf folgende Weise kodiert: Z(i, p).

Hypertexte des Corpus' dürfte der AAA somit ebenso wenig geeignet sein wie die oben vorgestellte allgemeine Typologie von Duncan.

16.3.3 Typisierte Links in HTML

Die Relationen der älteren Hypertext-Systeme haben sich im letzten Kapitel als wenig ergiebig für die Typisierung der Oberfläche aktueller HTML-Hypertexte des Corpus' erwiesen. In dem vorliegenden Kapitel wird nun ein Blick auf die sprachinternen Mittel von HTML zur Link-Typisierung auf Quellcode-Ebene geworfen. Diese Typisierung kann dann – die geeigneten Style-Sheets oder Anwendungen vorausgesetzt – auf die Ebene der semantischen Oberfläche übertragen werden. Darüber hinaus werden Methoden zur Link-Typisierung der layout-spezifischen Oberfläche vorgestellt.

Allgemeines zu Links in HTML

In HTML existieren nach gängiger Meinung keine sprachinternen Möglichkeiten (= Möglichkeiten der semantischen Zwischenschicht), Typisierte Verweise zu setzen:

„Alle Verweise in HTML haben den gleichen Aufbau [...]. Das vereinfacht die Syntax, hat aber auch Nachteile: es gibt nämlich keine Verweise für bestimmte Aufgaben.“

(Münz 1997 b [tcfa.htm] Abschnitt *Allgemeines zu Verweisen*)

Diese Syntax wurde bereits eingeführt (siehe Kapitel 15.2), ein einfacher Link wird in HTML wie folgt realisiert:

```
<a href="Link-Ziel">Link-Text</a>
```

Münz sieht typisierte Links aber ebenfalls als sinnvoll an und rät dem Hypertext-Autor daher folgendes:

"Bei komplexeren Projekten sollten Sie sich überlegen, wie Sie dem Anwender die Vielzahl der Verweise mit unterschiedlicher Bedeutung intuitiv zugänglich machen. Ein sinnvoller Weg ist, den Verweisen kleine Symbolgrafiken voranzustellen [...] Durch aussagekräftige Symbole können Sie dem Anwender sofort signalisieren, um welche Art von Verweis es sich handelt."¹

(Münz 1997 b [tcfa.htm] Abschnitt *Allgemeines zu Verweisen*)

Daß beim heutigen Stand der Technik solche Umwege auf layout-spezifischer Oberfläche zur Typisierung von Links gegangen werden müssen, erstaunt

¹ Auf besagte Art und Weise visualisiert Münz 1997 b unterschiedliche Verweisklassen.

besonders, wenn man sich die programmatischen Anfänge des WWW vor Augen führt. In seinem Aufsatz *Information Management: A Proposal* von 1989 schreibt Berners-Lee über ein geplantes Informationssystem am Forschungszentrum CERN und erläutert, warum sich das HYPERTEXT-KONZEPT dafür als besonders nützlich erweisen könnte:

„[...] a ‘web’ of notes with links (like references) between them is far more useful than a fixed hierarchical system. When describing a complex system, many people resort to diagrams with circles and arrows. Circles and arrows leave one free to describe the interrelationships between things in a way that tables, for example, do not. The system we need is like a diagram of circles and arrows, where circles and arrows can stand for anything. [...]

The arrows which links circle A to circle B can mean, for example, that A... [sic!]
depends on B
is part of B
made B
refers to B
uses B
is an example of B“

(Berners-Lee 1989 Abschnitt *Linked information systems*)

Auch in den ersten *Design Issues* zu HTML (WC3 1990) wird sich für typisierte Links ausgesprochen, eventuell mit dem voreingestellten Wert *untyped* für Links, die vom Autor nicht explizit ausgezeichnet worden sind.

Typisierung mittels der Rel- und Rev-Attribute

Interessanterweise bieten die aktuellen Sprachspezifikationen von HTML tatsächlich Möglichkeiten, bestimmte semantische Zusatzinformationen für Knoten bzw. Links zu geben:

„ [...] authors may insert links in their documents that express other relationships between resources than simply ‘activate this link to visit that related resource’. Links that express other types of relationships have one or more link types specified in their source anchor. The *roles* of a link defined by A or LINK are specified via the rel and rev attributes.“

(W3C 1998)

Das Attribut *rel* typisiert eine Vorwärtsbeziehung des aktuellen Knotens bzgl. des Zielknotens, *rev* hingegen eine Rückwärtsbeziehung. Diese Beziehungen können unterschiedliche Bezugsbereiche von Knoten miteinander in Beziehung setzen: Verwendet man die Attribute in Kombination mit dem Element `<link>`, dann werden Beziehungen zwischen zwei ganzen Knoten ausgedrückt. In Kombination mit dem Element `<a>` hingegen wird lediglich der Bezugsbereich

des Links mit dem Link-Ziel in Beziehung gesetzt.¹ Folgende allgemeine Zusammenhänge können mit Hilfe von *rel* oder *rev* ausgedrückt werden:

| Relation | Beschreibung |
|-------------------|---|
| <i>contents</i> | Verweis zum Inhaltsverzeichnis |
| <i>chapter</i> | Verweis zum Kapitel |
| <i>section</i> | Verweis zum Abschnitt |
| <i>subsection</i> | Verweis zum Unterabschnitt |
| <i>index</i> | Verweis zum Stichwortverzeichnis |
| <i>glossary</i> | Verweis zum Glossar |
| <i>appendix</i> | Verweis zum Anhang |
| <i>copyright</i> | Verweis zur Copyrightangabe |
| <i>next</i> | Verweis zur nächsten Datei (etwa in einer Guided Tour) |
| <i>prev</i> | Verweis zur vorherigen Datei (etwa in einer Guided Tour) |
| <i>start</i> | Verweis zur ersten Datei (etwa in einer Guided Tour) |
| <i>help</i> | Verweis zum Hilfekontext |
| <i>bookmark</i> | Verweis zu einem allgemeinen Orientierungspunkt |
| <i>alternate</i> | Verweis zu einer Datei mit dem gleichen Inhalt wie der der aktuellen, jedoch in einer anderen Dokumentversion |

Tabelle 28: Relationen der *rel*- und *ref*-Attribute

Zu Beginn eines HTML-Knotens können auf diese Art, Beziehungen zu anderen Knoten festgelegt werden:

```
<head>
  <link rel="contents" href="inhalt.htm" title="Inhaltsverzeichnis" />
  <link rel="index" href="stichwort.htm" title="Stichwortverzeichnis" />
  <link rel="glossary" href="glossar.htm" title="Begriffs-Glossar" />
  <link rel="next" href="naechste_seite.htm" title="Nächste Seite" />
  <link rel="previous" href="vorherige_seite" title="Vorherige Seite" />
</head>
```

Jedoch können die Attribute *rel* und *rev* auch in die eigentlichen Hypertext-Links von HTML eingebaut werden:

```
<a href="../../index.htm" rev="contents">Inhalt</a> |
<a href="../naechster_knoten.htm" rel="next">Weiter</a> |
<a href="../letzter_knoten.htm" rel="prev">Zurück</a> |
<a href="../index.htm" rev="chapter">Kapitelanfang</a>
```

Mit Hilfe der Attributs *rel* und *rev* ließen sich also immerhin einige fest definierte Beziehungen codieren. Leider interpretiert jedoch keiner der gängigen

¹ Der Bezugsbereich im Ausgangs-Knoten kann hierbei sehr unterschiedlich sein. Genauere Informationen liefert die praktische Analyse in Kapitel 17. Der Bezugsbereich des Link-Ziels kann mit syntaktischen Mitteln von HTML zweierlei Ausformung haben: entweder den gesamten Ziel-Knoten oder aber lediglich einen Bereich desselben, dessen Anfang per Fragmentidentifizierer im Attribut *href* kenntlich gemacht wurde.

Browser die Container *rel* und *rev*.¹ Im Rahmen unserer Untersuchung muß das vom W3C vorgeschlagene Vokabular der *rel*- und *rev*-Attribute jedoch auf alle Fälle darauf hin geprüft werden, ob es sich für ein allgemeines kontrolliertes Vokabular zur Typisierung von Links eignet. Hierbei fällt zuallererst auf, daß es sich bei den obigen Relationen in ihrer Gesamtheit um hierarchische Angaben handelt, thematische oder semantische Relationen fehlen hingegen.

Typisierung mit Metaangaben nach dem Dublin Core

Eine weitere Möglichkeit auf Ebene der semantischen Zwischenschicht im HTML-Quellcode Relationen zwischen Knoten auszudrücken, bieten die Metaangaben.² Mit Hilfe dieser Angaben werden in erster Linie zusätzliche Informationen für Browser, Server und automatische Indexierer geliefert. So können etwa Autor, Beschreibung und Schlüsselwörter eines Knotens auf folgende Weise in einer HTML-Datei festgehalten werden:

```
<head>
<meta name="description" content="Die Seite handelt von typisierten Links in HTML." />
<meta name="author" content="Oliver Huber" />
<meta name="keywords" content="HTML, Link, Verweis, Typ, typisiert, Meta" />
</head>
```

Eine internationale Expertengruppe hat eine Systematik für Metaangaben erarbeitet, die unter dem Namen DUBLIN CORE (DC) bekannt geworden ist.³ Das System ist leicht zu verstehen und integriert sämtliche wichtigen Angaben, die zu den Eigenschaften eines Dokument gemacht werden können. Das obige Beispiel in der Schreibweise des DC sieht wie folgt aus:

```
<head>
<meta name="DC.Description" content="Die Seite handelt von typisierten Links in HTML." />
<meta name="DC.Creator" content="Oliver Huber" />
<meta name="DC.Subject" content=" HTML, Link, Verweis, Typ, typisiert, Meta " />
</head>
```

Mit der Angabe *DC.Relation* ist es möglich, Beziehungen des aktuellen Knotens zu anderen festzulegen. Folgende Relationen können mit dem offiziellen DC-Vokabular formuliert werden:

¹ Auf einen möglichen ‚Workaround‘ mit Hilfe des Attributs *title* wurde bereits in Kapitel 15.2.4 hingewiesen.

² Metaangaben wurden in Kapitel 15.2.1 bereits kurz vorgestellt.

³ Ausführliche Informationen hierzu bietet die Homepage der Initiative [<http://dublincore.org/>].

| Relation | Beschreibung |
|--------------------------------|---|
| <i>Is Version Of</i> | Aktueller Knoten ist eine bestimmte Version des Zielknotens. |
| <i>Has Version</i> | Zielknoten ist eine bestimmte Version des aktuellen Knotens. |
| <i>Is Replaced By</i> | Zielknoten ersetzt den aktuellen Knoten. |
| <i>Replaces</i> | Aktueller Knoten ersetzt den Zielknoten. |
| <i>Is Required By</i> | Zielknoten benötigt technisch oder inhaltlich in irgendeiner Form den aktuellen Knoten. |
| <i>Requires</i> | Aktueller Knoten benötigt technisch oder inhaltlich in irgendeiner Form den Zielknoten. |
| <i>Is Part Of</i> | Aktueller Knoten ist physikalisch oder logisch Teil des Zielknotens. |
| <i>Has Part</i> | Zielknoten ist physikalisch oder logisch Teil des aktuellen Knotens. |
| <i>Is Referenced By</i> | Zielknoten stellt eine Referenz zum aktuellen Knoten her. Der aktuelle Knoten kann hierbei als Zitat, Danksagung, Gegenposition o.ä. bzgl. des Zielknotens fungieren. |
| <i>References</i> | Aktueller Knoten stellt eine Referenz zu einem Zielknoten her. Der Zielknoten kann hierbei als Zitat, Danksagung, Gegenposition o.ä. bzgl. des aktuellen Knotens fungieren. |
| <i>Is Format Of</i> | Aktueller Knoten ist vom Zielknoten durch Formatierungstechniken in der Weise abgeleitet, daß der Zielknoten zwar keine Interpretation jedoch eine unterschiedliche Repräsentation des aktuellen Knotens ist. |
| <i>Has Format</i> | Zielknoten ist vom aktuellen Knoten durch Formatierungstechniken in der Weise abgeleitet, daß der aktuelle Knoten zwar keine Interpretation jedoch eine unterschiedliche Repräsentation des Zielknotens ist. |
| <i>Is Based on¹</i> | Aktueller Knoten ist eine Ausführung, Ableitung, Übersetzung, Umsetzung bzw. ein Produkt des Zielknotens. |
| <i>Is Basis for</i> | Zielknoten ist eine Ausführung, Ableitung, Übersetzung, Umsetzung bzw. ein Produkt des aktuellen Knotens, der hierfür als Basis dient. |

Tabelle 29: Relationen des DUBLIN CORE

Diese Relationen können auf unterschiedliche Art syntaktisch ausgedrückt werden. Neben der Möglichkeit, DC mittels XML und RDF zu realisieren, können DC-Relationen auch mit Metaangaben in HTML beschrieben werden:

```
<meta name="DC.Relation.IsPartOf" content="(Scheme=URL) http://www.teil.htm" />2
```

Da der Dublin Core mehr als Empfehlung denn als Vorschrift zu verstehen ist, können diesem Vokabular eigene Erweiterungen hinzugefügt werden³:

Specialized communities will need to extend the references relations to express specific types of references or authorial claims.

(Bearman 1997)

Dies impliziert natürlich, daß eine geeignete Implementierung existiert, die mit dem Vokabular umgehen und das Layout der Oberfläche entsprechend mani-

¹ In der aktuellsten Empfehlung der Dublin Core Initiative sind die Relationen *Is Based on* sowie *Is Basis for* nicht mehr aufgeführt [<http://dublincore.org/documents/dcmes-qualifiers/#relation>].

² Hierbei ist anzumerken: Das Standardformat für den Wert von DC-Relationen ist eine URI, benutzt man andere Formate, so muß dies extra erwähnt werden. Die URI muß nicht direkt auf die Ressource (etwa den Hypertext-Knoten) verweisen. Vielmehr empfiehlt der Dublin Core, die Metaangaben in eine gesonderte Datei auszulagern.

³ Eine solche Erweiterung hat der *Deutsche Bildungsserver* entworfen. Mehr hierzu in Diephold u. Richter 1998.

pulieren kann: Denkbar wäre etwa ein Browser, der die Metaangaben ausliest und je nach Relation die Links unterschiedlich markiert, sei es nun farblich, mit interaktivem Text oder durch Piktogramme.

Laut Bearman 1997 lassen sich die möglichen Typen von *DC.Relation* in folgende Gruppen einteilen: *IsPartOf* und *HasPart* können benutzt werden, um Hyperonymie-Relationen zu beschreiben. Verschiedene Entwicklungsstufen von Knoten lassen sich mittels *IsVersionOf* und *HasVersion* voneinander unterscheiden. Referenzbeziehungen, wie etwa Zitate, lassen sich via *References* oder *IsReferencedBy* ausdrücken während *IsBasedOn* und *IsBasisFor* offen sind für kreative Beziehungen im weiteren Sinn: Hiermit können Knoten beschrieben werden, die etwa Übersetzungen, Interpretationen oder sonstige Ableitungen des aktuellen Knotens sind. Die stärksten Bindungsrelationen drücken wohl *Requires* und *IsRequiredBy* aus. Hierbei kann der eine Knoten ohne den in Relation gesetzten inhaltlich oder technisch¹ nicht korrekt angezeigt werden. *Is Format of*, *Has Format*, *Replaces* sowie *Is Replaced by* werden bei Bearman 1997 nicht erwähnt.

Bei einer Bewertung der Relationen fällt zunächst auf, daß diese paarweise auftreten. Daher lassen sich bidirektionale Relationen leicht ausdrücken. Desweiteren ist zu bemerken, daß die Anzahl relativ überschaubar ausfällt. Mit einem Großteil der Relationen lassen sich hauptsächlich ‚formale‘ Eigenschaften von Knoten ausdrücken: welcher Knoten die aktuellere Version ist (z.B. in *Is Version of* oder *Is Replaced by*), welcher Knoten formal einen anderen ‚in irgendeiner Beziehung‘ braucht (z.B. *is Required by*) oder welcher Knoten für ein anderes Ausgabeformat konzipiert wurde (z.B. *is Format of*). Inhaltlich lassen sich lediglich einfache Hyperonymie-Relationen ausdrücken (z.B. *has Part*). *Is Referenced by*, *References*, *is Based on* und *is Basis for* bleiben in ihrer semantischen Aussage etwas vage: Interessant wäre es ja gerade kenntlich zu machen, ob etwa ein Knoten bzgl. eines anderen in einem Zitats- oder aber in einem Kontrapositionsverhältnis steht.

Positiv bemerkt werden muß die breite Anerkennung und Unterstützung, die DC bereits erfahren hat. So hat sich das W3C – auch im Rahmen der Diskussion um

¹ Das DC-Vokabular muß nicht zwangsläufig HTML-Knoten inhaltlich beschreiben, sondern ist offen für Ressourcen jeder Art, die für beliebige Ausgabemedien konzipiert sein können. In einer solchen Multi-Publishing-Umgebung könnten dann Angaben wie
`<meta name="DC.Relation.Requires" content="Browser" />` oder
`< meta name = "DC.Relation.Requires" content = "Printer" />`
zur Unterscheidung der Ausgabegeräte Sinn machen.

RDF und das Semantic Web – für diese Klassifikation ausgesprochen. Das Verhältnis von RDF zum Dublin Core beschreibt das W3C wie folgt:

“One obvious application for RDF is in the description of Web pages. This is one of the basic goals of the Dublin Core [DC] Metadata Initiative. The Dublin Core Element Set is a set of 15 elements believed to be broadly applicable to describing Web resources to enable their discovery. The Dublin Core has been a major influence on the development of RDF. An important consideration in the development of the Dublin Core was to not only allow simple descriptions, but also to provide the ability to qualify descriptions in order to provide both domain specific elaboration and descriptive precision.

The RDF Schema Specification provides a machine-understandable system for defining schemas for descriptive vocabularies like the Dublin Core. It allows designers to specify classes of resource types and properties to convey descriptions of those classes, relationships between those properties and classes, and constraints on the allowed combinations of classes, properties, and values.“

(W3C 2001 Abschnitt 1.1.2. *Simple Web Metadata*)

Im Zusammenspiel mit RDF lassen sich also durch das Dublin Core Relationen zwischen Knoten zugleich menschen- und maschinenlesbar formulieren. Hinzu kommt die Erweiterbarkeit mittels RDF Schemata. Auch aus diesem Grund soll das DC in die nähere Wahl für die praktische Untersuchung gezogen werden. Allerdings sei hier nochmals auf die Bedenken verwiesen, die weiter oben hinsichtlich der stark formalen Ausrichtung der Relationstypen des DC geäußert wurden: Ob diese daher wirklich für eine Beschreibung der Links auf der Ebene der semantischen Oberfläche der Corpus-Hypertexte geeignet sind – darüber wird die kurze praktische Untersuchung in Kapitel 17.3.6 erste Rückschlüsse zulassen.

Hilfskonstruktionen: Typisierte Links durch Piktogramme

PIKTOGRAMME (oder auch SYMBOLE, ICONS bzw. IKONE) wurden allgemein bereits in Kapitel 15.2.4 kurz eingeführt. Hier sollen nun einige beispielhafte Kollektionen solcher Piktogramme vorgestellt werden.

Piktogramme werden, wie gesagt, vor oder nach einem Link eingefügt und sollen so aufgrund ihrer ‚Bildhaftigkeit‘ die Link-Funktion verdeutlichen. Die Links werden also nicht im Quellcode mit einer maschinenlesbaren Typisierung versehen¹, sondern auf der Ebene der layout-spezifischen Oberfläche mit Pikto-

¹ Das leisten die weiter oben eingeführten Typisierungen auf Ebene der semantischen Zwischenschicht mittels der Attribute *rel* und *rev* bzw. des Dublin Core. Eventuelle Konfusion über die Ähnlichkeit der Möglichkeiten kommt lediglich daher, daß strukturelle, semantische, layout-spezifische und navigationsspezifische Zwischenschichten in heutigen HTML-Hypertexten noch nicht so exakt getrennt sind, wie sie es sein sollten und aller Voraussicht nach zukünftig auch sein werden.

grammen typisiert, die für den Menschen ‚Sinn‘ machen sollen. Voraussetzung hierfür ist eine gewisse Eingängigkeit der verwendeten Piktogramme.

Im folgenden werden nun drei Beispiele für Symbolgrafiken im WWW aufgeführt: die Piktogramme der Corpus-Hypertexte SELF und SELF 8 sowie die allgemeine Piktogramm-Kollektion *Qbullets*.

In SELF werden folgende Symbole zur Typisierung benutzt:¹













| Nr. | Symbol | Bedeutung |
|-----|---|--|
| 1 |  | „Geschlossener Ordner – wenn Sie den daneben stehenden Verweis anklicken, gelangen Sie in einen größeren Abschnitt dieses Dokumentes.“ |
| 2 |  | „Geöffneter Ordner – wenn Sie den daneben stehenden Verweis anklicken, gelangen Sie zu einem übergeordneten Bereich des Dokumentes. Mehrere übergeordnete Ebenen werden von links an eingerückt dargestellt.“ |
| 3 |  | „Datei oder Textstelle innerhalb des Dokumentes – wenn Sie den daneben stehenden Verweis anklicken, gelangen Sie zu einer Stelle mit konkreter Information. Das Symbol wird auch bei Querverweisen verwendet.“ |
| 4 |  | Datei, die ein Anzeigebeispiel enthält. ² |
| 5 |  | „Textstelle auf der gleichen Seite weiter unten – wenn Sie den daneben stehenden Verweis anklicken, gelangen Sie zu einer Stelle innerhalb der gleichen Datei weiter unten.“ |
| 6 |  | „Textstelle auf der gleichen Seite weiter oben – wenn Sie den daneben stehenden Verweis anklicken, gelangen Sie zu einer Stelle innerhalb der gleichen Datei weiter oben.“ |
| 7 |  | „Steht am Ende von vielen Seiten und bedeutet: zur vorhergehenden Seite blättern.“ |
| 8 |  | „Steht am Ende von vielen Seiten und bedeutet: zur nächsten Seite blättern.“ |
| 9 |  | „Steht für Email. Wenn Sie den daneben stehenden Verweis anklicken, können Sie eine Email an die betreffende Person absetzen.“ |
| 10 |  | „Steht vor Verweisen ins Internet: das Link-Ziel ist eine hypertext-externe deutschsprachige Web-Seite.“ |
| 11 |  | „Steht vor Verweisen ins Internet: das Link-Ziel ist eine hypertext-externe englischsprachige Web-Seite.“ |
| 12 |  | „Steht vor Verweisen ins Internet: das Link-Ziel ist eine gepackte Datei zum Downloaden.“ |

Tabelle 30: Piktogramme zur Link-Typisierung in SELF

Die Mehrzahl der Typisierungen machen Aussagen zur Struktur des Dokumentes (1-3, 5-8). Die Typen 10, 11 und 12 sind als ‚Grenzsignale‘ von SELF zu verstehen, die darüber hinaus noch zusätzliche Informationen über das externe Link-Ziel bieten (etwa über Sprache oder Dateiformat). Als semantisch motivierte Typen können lediglich Nr. 4 und – mit Einschränkungen – Nr. 9 gesehen werden: Durch den Klick auf einen Email-Verweis verläßt der Leser (vorübergehend) nicht nur den Hypertext sondern auch die Browser-Software: Ein Email-Programm wird – die richtige Konfiguration der Software vorausgesetzt – gestartet. Nach Verfassen und Absenden der Post kann der Leser zum eigentlichen Hypertext zurückkehren.

Ein Link des Typs 4 läßt sich hingegen als semantisch motiviert interpretieren: Ein Klick führt den Leser zu einer Visualisierung der im aktuellen Knoten auf-

¹ Entnommen aus: Münz 1997b [tab.htm] Abschnitt: *Verwendete Symbole*

² Das Symbol wird in der Symbolübersicht von SELF nicht aufgeführt, im Text jedoch benutzt.

geworfenen Problematik.

Es soll nun kurz die Eingängigkeit bzw. Schlüssigkeit der in SELF benutzten Piktogramme anhand eines Beispiels besprochen werden:



Abbildung 13: SELF [tcab.htm] Abschnitt *SGML-gerechten Dokumenttyp angeben*

Schön zu sehen ist hier die Unterscheidung zwischen internen (typisiert mit den Symbolen Nr. 2, 3, 6, 7 und 8) und externen Links (Nr. 11). Fraglich ist jedoch, ob Links, die mit Symbol Nr. 11 typisiert sind, sofort als externe Varianten interpretierbar sind. Im Beispiel ist die Lage relativ klar: Als Link-Text dienen externe URLs – einem mit dem Internet etwas vertrautem Leser müßte bewußt sein, daß er hier durch einen Klick den Text SELF verläßt, da dieser eine andere URL hat. Die Verwendung des Symbols der englischen Flagge für Links, die auf externe englische Hypertexte verweisen, ist deshalb eindeutig, da SELF vollständig in deutscher Sprache vorliegt. Eine Verwendung der deutschen Flagge (Symbol Nr. 10) für eine Verlinkung auf externe deutsche Hypertexte führt, meiner Meinung nach, aus diesem Grund nicht unbedingt bei der ersten Lektüre zu der gewünschten Erkenntnis beim Leser, daß dieser sich mit dem typisierten Link aus dem aktuellen Hypertext ‚herausklickt‘.

Intuitiv zu erfassen sind hingegen die Link-Symbole zur Navigation innerhalb des aktuellen Dokumentes (Nr. 5 und 6). Im Beispiel stimmig wirken auch die Symbole 7 und 8. Diese sind jedoch auf das Ende¹ eines Knotens beschränkt:

¹ Hier kommen sie als Elemente der Navigationsleiste vor.

Stünden sie nämlich im Fließtext, ließen sich ihre Bedeutungen nicht ohne Probleme von denen der Symbole 5 (im Sinne von *weiter unten*) und 6 (*weiter oben*) unterscheiden.

Im unteren Teil der obigen Abbildung ist die Navigationsleiste von SELF zu erkennen. Die geöffneten Ordner zeigen, wo sich der Leser momentan im Hypertext befindet: aktuell also im Kapitel *Allgemeine Regeln* des Bereichs *HTML* des Hypertextes *SelfHTML*¹. Der Name des aktuellen Knotens (*Grundgerüst einer HTML-Datei*) ist hier leider nicht zu erkennen, er könnte jedoch mittels eines Piktogramms Nr. 3 und dem entsprechenden Link-Text neben dem geöffneten Ordner *Allgemeine Regeln* visualisiert werden.

Das Piktogramm Nr. 3 wird im obigen Beispiel aber an anderer Stelle verwendet, und zwar im Fließtext („[...] daß Sie in der HTML-Datei in Frameset definieren“). Hier wird visualisiert, daß ein Klick auf diesen Link einen Knoten von SELF mit textueller Information öffnet.² Mehr Information zur hierarchischen Einordnung des Zielknotens wird jedoch nicht gegeben: Ob sich ein mit Symbol Nr. 3 typisierter Zielknoten innerhalb oder außerhalb des aktuellen ‚Kapitels‘ oder ‚Sinnbereichs‘ befindet, wird nicht angezeigt. Dies könnte durch zusätzliche Informationen zur Einordnung des Link-Ziels in den Gesamt-Hypertext signalisiert werden. Genau dies wird mit den zusätzlichen Piktogrammen in SELF8 versucht:³




| Nr. | Symbol | Bedeutung |
|-----|---|---|
| 1 |  | „Kleines Seitensymbol für Popup-Seiten, z.B. Seiten mit Anzeigebeispielen oder Quellenhinweisen. Wenn Sie den daneben stehenden Verweis anklicken, wird die entsprechende Seite geöffnet, und zwar in einem neuen Browserfenster.“ |
| 2 |  | „Geschlossener kleiner Ordner hautfarben – wenn Sie den daneben stehenden Verweis anklicken, gelangen Sie zu einer Übersichtsseite eines größeren Abschnitts im Online-Angebot von SELFHTML aktuell oder im Online-Angebot eines Partnerprojekts von SELFHTML. Das Link-Ziel hat das gleiche Layout wie die Seiten von SELFHTML. Um solche Verweise ausführen zu können, ist eine Internetverbindung erforderlich.“ |
| 3 |  | „Geschlossenes kleines Seitensymbol hautfarben – wenn Sie den daneben stehenden Verweis anklicken, gelangen Sie zu einer Informationsseite im Online-Angebot von SELFHTML aktuell oder im Online-Angebot eines Partnerprojekts von SELFHTML. Das Link-Ziel hat das gleiche Layout wie die Seiten von SELFHTML. Um solche Verweise ausführen zu können, ist eine Internetverbindung erforderlich.“ |

Tabelle 31: Zusätzliche Piktogramme zur Link-Typisierung in SELF8

¹ Die neben „SELFHTML“ angeführte Quickbar dient der schnelleren Navigation, ist aber an dieser Stelle nicht von Belang.

² Ein Übersichtsknoten würde durch ein Symbol Nr. 1 oder 2 angedeutet werden.

³ Folgende Tabelle ist übernommen aus Münz 2001 [editorial\arbeitshinweise.htm] Abschnitt *Verwendete Symbole*. Bitte beachten Sie besonders die Farbbezeichnungen in der Spalte *Bedeutung*, da die farbigen Markierungen der Piktogramme im Schwarz-Weiß-Ausdruck naturgemäß ‚suboptimal‘ zu erkennen sind.

Die Piktogramme 4 (*Anzeigebeispiel*) und 12 (*Gepackte Datei*) von SELF sind hingegen in SELF8 nicht mehr vertreten. Das alte Piktogramm 4 wird in SELF8 faktisch vom neuen Piktogramm 1 abgelöst. Interessant ist hierbei, daß mit der alten Nr. 4 das einzige Piktogramm verschwunden ist, das in obigen Erläuterungen mit *semantisch motiviert* bezeichnet wurde. Ersetzt wird es in SELF8 durch das Symbol 1, welches nähere Angaben über die Darstellung des Zielknotens macht – nämlich, daß dieser in einem neuen Fenster der Browser-Software geöffnet wird. Die mediale Füllung selbst wird hingegen nicht mehr signalisiert, der Zielknoten kann „Anzeigebeispiele oder Quellenhinweise“ (Tabelle 31) enthalten.

Um die neu eingeführten Piktogramme 2 und 3 von SELF8 zu verstehen, muß man etwas mehr über die Struktur dieses Hypertextes erfahren. Was in dieser Arbeit als *SELF8* bezeichnet wird, ist der eigentliche Kern eines kleinen Hypertext-Netzes, das als *SELF-Raum* bezeichnet wird. Hierzu gehören neben wenigen Partnerprojekten, ergänzenden Fachartikeln, einem Nachrichtenbereich auch Einrichtungen zur interaktiven Kommunikation: ein Chat sowie zwei Diskussionsforen. Alle Knoten des Hypertext-Netzes *SELF-Raum* zeichnen sich durch ein gemeinsames Layout sowie eine gemeinsame Makro-Textfunktion aus. Daher kann man hier von einem tatsächlichen Hypertext-Netz sprechen. Mit den neuen farbig differenziert dargestellten Symbolen 2 und 3 werden nun hypertext-netz-interne Links von hypertext-internen (repräsentiert durch die alten Symbole 1 bis 8) sowie hypertext-(netz)-externen Links (alte Symbole 10 bis 12) getrennt.

Im praktischen Teil dieses Fallbeispiels werden die Piktogramme von SELF und SELF8 anhand mehrerer Beispiele genauer untersucht.

Als letztes Exempel einer Piktogramm-Kollektion seien die *Qbullets* erwähnt: Diese folgende Sammlung der Firma *Matterform*¹ wird nicht im Corpus benutzt, stellt aber eine der am weitesten ausdifferenzierten Sammlungen typisierender Piktogramme im WWW dar. Daher und aufgrund des Anspruchs dieser sogenannten *Qbullets*², auf verschiedenen, sehr heterogenen WWW-Sites von Kunden eingesetzt werden zu können, ist diese Piktogramm-Kollektion in die vorliegende Untersuchung mit aufgenommen worden. Hierbei findet auch ein knapper komparativer Vergleich mit den Piktogrammen aus SELF bzw. SELF8 statt. Die *Qbullets* sind in folgende fünf Gruppen eingeteilt:

¹ Online im WWW [<http://www.matterform.com>].

² *Q* in *QBullet* steht hierbei für *Cue* – also für *Hinweis*.





| Navigationssymbole | | |
|--------------------|---|---|
| Nr. | Symbol | Bedeutung |
| 1 |  | Hypertext-externer Link |
| 2 |  | Link auf Stelle im gleichen Knoten weiter oben |
| 3 |  | Link auf Stelle im gleichen Knoten weiter unten |
| 4 |  | Link-Ziel in neuem Fenster |

Tabelle 32: QBullets – Navigationssymbole










| Multimediale Symbole | | |
|----------------------|---|--|
| Nr. | Symbol | Bedeutung |
| 5 |  | Link-Ziel ist ein PDF-Dokument. |
| 6 |  | Link-Ziel ist ein Flash-Dokument. |
| 7 |  | Link-Ziel ist eine Real-Media-Tondatei |
| 8 |  | Link-Ziel ist eine Tondatei |
| 9 |  | Link-Ziel ist eine Filmdatei |
| 10 |  | Link-Ziel ist ein 3-D-Dokument |
| 11 |  | Link-Ziel ist eine Word-Datei |
| 12 |  | Link-Ziel ist eine Excel-Datei |
| 13 |  | Link-Ziel ist eine Power-Point-Datei |

Tabelle 33: QBullets – Multimediale Symbole








| Semantische Symbole | | |
|---------------------|---|--|
| Nr. | Symbol | Bedeutung |
| 14 |  | Link auf eine Definition |
| 15 |  | Link auf eine auf eine 'Kurz Notiz' |
| 16 |  | Link auf 'hilfreiche Hinweise' |
| 17 |  | Link auf Information im weiteren Sinn |
| 18 |  | Link-Ziel enthält ein Formular |
| 19 |  | Link auf eine Suchseite |
| 20 |  | Inhalt des Link-Ziels kann editiert werden |

Tabelle 34: QBullets – Semantische Symbole







| Netzwerkspezifische Symbole | | |
|-----------------------------|---|---------------------------------|
| Nr. | Symbol | Bedeutung |
| 21 |  | Link-Ziel ist verschlüsselt |
| 22 |  | Link-Ziel ist passwortgeschützt |
| 23 |  | Link öffnet Email-Programm |
| 24 |  | Download-Link |
| 25 |  | Link auf FTP-Verzeichnis |
| 26 |  | Link auf eine Newsgroup |

Tabelle 35: QBullets – Netzwerkspezifische Symbole


| Vermischte Symbole | | |
|--------------------|---|--|
| Nr. | Symbol | Bedeutung |
| 27 | NEW | Link-Ziel ist 'neu' |
| 28 | UPDATED | Link-Ziel liegt in einer neuen Version vor |
| 29 |  | Link-Ziel ist nicht ganz ernsthaft zu interpretieren (?) |

Tabelle 36: QBullets – Vermischte Symbole

Die erste Gruppe der Navigationssymbole fällt mit lediglich vier Vertretern etwas klein aus. Symbol 1 signalisiert das Verlassen des Hypertextes. In SELF und SELF8 wird hier noch die Sprache des Ziel-Hypertextes als

Zusatzinformation geliefert. Die Navigation innerhalb des aktuellen Knotens kann mit den Symbolen 2 und 3 typisiert werden, die in dieser Form auch in SELF und SELF8 vorhanden sind. Für die sonstige hypertext-interne Navigation ist das Symbolsystem von SELF und besonders das von SELF8 weitaus ergiebiger (vgl. hierzu die obigen Ausführungen): Mit Hilfe des *QBullet* Nr. 4 kann lediglich signalisiert werden, daß ein Ziel-Knoten in einem neuen Fenster geöffnet wird. Eine weitergehende navigations- bzw. browser-spezifische Typisierung ist nicht möglich.

Die zweite Gruppe, die der medialen Piktogramme, ist deutlich größer. Hiermit können Ziel-Knoten bzgl. ihres medialen Gehalts typisiert werden. Jedoch sind bis auf Nummer 9 und 10 sämtliche Symbole auf die Anwendersoftware hin ausgerichtet, die zur Erstellung des medialen Gehalts des Knotens notwendig war. Sinnvoller wäre es, meiner Meinung nach, jedoch, Knoten nach ihren medialen Inhalten (*Text, Bild, Grafik, Sound, Film*) und nicht nach Dokumenttypen (*Word Dokument* vs. *Adobe PDF, Real Media Movie* vs. *Flash Animation* etc.) zu unterscheiden, da bestimmte Typen von Medien immer von einer Vielzahl von Programmen erzeugt werden können (z.B. Texte von *Word, StarWriter, LaTeX* etc.).

Der Nutzen der semantischen Piktogramme, die die dritte Gruppe stellen, ist ad hoc sehr unterschiedlich zu beurteilen. Sinnvoll erachte ich in jedem Fall Nr. 14 und 15 (*Definition, Fußnote*). Wobei streng betrachtet lediglich *Definition* ein semantisches Symbol im engeren Sinn ist, *Fußnote* hingegen mehr die mediale Realisierung angibt: Eine Definition kann beispielsweise als Pop-Up-Fenster, als Mouse-Over-Text oder eben auch als Fußnote realisiert sein. Die Piktogramme 16 und 17 sollen Links typisieren, deren Ziel-Knoten Hilfestellungen im weitesten Sinne zum aktuellen Knoten geben. Denkbar wäre etwa, daß zu Beginn oder am Ende jedes Knotens eines Hypertextes ein solcher Link wäre, dessen Ziel-Knoten Informationen zur Bedienung des Hypertextes sowie eine Legende der Piktogramme o.ä. enthält. Vergleichbar böte ein mit Piktogramm 19 typisierter Link Zugriff auf die hypertext-internen Suchfunktionen. Nr. 18 (*Formular*) macht meiner Meinung nach keinen Sinn, denn ob der Ziel-Knoten ein Formular enthält oder nicht, darf wohl weder als semantische noch als sonstwie besonders relevante Information gelten. Symbol 20 (*Editierfunktion*) ist zwar auch nicht semantisch motiviert, kann aber in gewissen Kontexten durchaus Sinn machen: Denkbar sind hier besonders dynamische Hypertexte, in welchen der Leser das Aussehen der Seite sowie die Inhalte bestimmter

Textbereiche selbsttätig ändern kann. Hypertexte mit solcher Funktionalität sind aber (noch) die Ausnahme.¹ Als allgemeiner Link-Typ in einem kontrollierten Vokabular scheidet die *Editierfunktion* daher (noch) aus.

In der vierten Gruppe sind die netzwerkspezifischen Piktogramme vereint. Die Typen *Email* und *Download* sind ebenfalls in SELF vorhanden, *Email* auch in SELF8. Die anderen Symbole der Gruppe beziehen sich auf die Protokolle, mit denen die Inhalte der Ziel-Knoten übertragen werden (*Email*, *News*, *FTP*). Je nach Voreinstellung des Benutzers werden hierbei andere Programme, etwa spezielle CLIENTS² für Email, FTP oder Newsgroups, aktiviert – und somit der Hypertext verlassen. In ein begrenztes Typenvokabular wird wahrscheinlich lediglich ein Typ wie *externer Link* aufgenommen werden können, eine Aufzählung sämtlicher Varianten würde wohl zu einer kognitiven Überbelastung besonders bei solchen Lesern führen, die sich nicht so gut mit den unterschiedlichen Internet-Protokollen auskennen.

In einer fünften Gruppe werden die restlichen Piktogramme gesammelt. Die Typen *New* und *Updated* könnten prinzipiell jedem bereits typisierten Link als zusätzliche Metainformation beigefügt werden, denn: Auch wenn sich die Firma *Matterform* nicht explizit zur Kombinierbarkeit ihrer Piktogramme äußert, so ist dennoch vorstellbar, daß – je nach Ausrichtung der Typisierung – semantische oder navigationsspezifische Piktogramme das ‚Gerüst‘ der Typisierung bilden und durch mediale oder eben auch durch Piktogramme der Typen *New* und *Updated* ‚modifiziert‘ werden. In Kapitel 17.4.2 werden auf ähnliche Weise Kerntypen von Links und mögliche ‚Modifizierer‘ vorgeschlagen.

QBullets bieten im Vergleich zu den Piktogrammen aus SELF und SELF8 (aber auch im Vergleich zu den übrigen, bereits vorgestellten Link-Kollektionen) keine deutlichen Vorteile. Im Gegenteil: Die navigationsspezifischen Piktogramme sind eindeutig zugunsten proprietärer Formatbezeichner unterrepräsentiert. Die netzwerkspezifischen Piktogramme machen, meiner Meinung nach, für den normalen Leser wenig Sinn. Lediglich die Typen *New* und *Updated* scheinen eine sinnvolle Ergänzung darzustellen, da sich Inhalte im WWW bekanntlich schnell ändern.

¹ Portale wie *Yahoo!* ermöglichen aber bereits eine inhaltliche sowie layout-spezifische Personalisierung ihrer Seiten per Klick [<http://www.yahoo.com>].

² Das Begriffspaar SERVER/CLIENT ist im Glossar erläutert.

Eine kurze Zusammenstellung der für das Fallbeispiel relevanten Vokabulare im nächsten Kapitel runden diesen Komplex ab, bevor zur eigentlichen praktischen Analyse der Corpus-Hypertexte mit Hilfe von TAH übergegangen wird.

16.4 Zusammenfassung: Geeignete Vokabularien

Einige Texte des Corpus‘ führen, wie gesagt, eine Link-Typisierung auf Ebene der layout-spezifischen Oberflächenstruktur durch, allerdings mit eigenen, proprietären Link-Typen. Dies hat einen, bereits mehrfach erwähnten gewichtigen Nachteil: Der menschliche Leser muß sich bei der Lektüre verschiedener Hypertexte auf wechselnde Typisierungen einstellen. Ein möglicher Ausweg wäre eben, sich auf ein allgemeines Inventar von Link-Typen zu einigen. Im Kapitel 17.3 werden einige der in den letzten Kapitel vorgestellten Vorschläge untersucht, ob und in welchem Ausmaß sie die proprietären Link-Typen der Corpus-Hypertexte nachbilden können.

Im einzelnen werden folgende Typeninventare in die praktische Analyse einbezogen: Typisierungen mittels des Attributs *rel* von HTML, Typisierungen mit dem Dublin Core (Kapitel 16.3.3) sowie solche mit Hilfe der Rhetorical Structure Theory (Kapitel 16.2.2).

Für eine Codierung mit Hilfe des *rel*-Attributs spricht, daß die Werte dieses Attributs von den Schöpfern der Sprache HTML bestimmt wurden: Daß die Attributwerte also von Anfang an für die Typisierung von HTML-Knoten bzw. -Links konzipiert waren, nährt die Hoffnung, daß diese Werte die tatsächlichen Verhältnisse in den Corpus-Hypertexten des WWW beschreiben können. Ähnliches läßt sich aufgrund der mit dem WWW eng verflochtenen Entwicklung des Dublin Core für dessen Relationstypen vermuten. Die Relationen der Rhetorical Structure Theory werden hingegen vorwiegend aufgrund der fächerübergreifenden Popularität der Theorie verwendet.

Einige der vorgestellten Vokabularien wurden nicht in die praktische Analyse aufgenommen. Gründe für die abschlägigen Wertungen wurden bereits bei Besprechung der einzelnen Ansätze genannt. Der eigentlichen Verwendung der ausgesuchten Vokabularien geht die intensive Untersuchung der proprietären Link-Typen der Corpus-Hypertexte mit Hilfe von TAH voraus.

17 Analyse des Corpus' mit Hilfe von TAH

17.1 Vorbemerkung zur Analyse

Die folgende Analyse der Hypertexte des Corpus hat zwei – unterschiedlich gewichtete – Ziele, ein methodisches und ein inhaltliches.

Das methodische steht im Vordergrund: Hierbei soll sich das in Kapitel 14 vorgestellte textlinguistische Analysemodell als methodischer Rahmen in der praktischen Anwendung bewähren – denn keine theoriegeleitete Frage sollte völlig unabhängig von der empirischen Basis behandelt werden.

Die Untersuchung selbst darf aber auch nicht ‚methodischer Selbstzweck‘ sein und ist daher mit einer speziellen inhaltlichen Fragestellung beauftragt: Es werden primär die unterschiedlichen Link-Typen der Corpus-Hypertexte untersucht, danach einige der in Kapitel 16 vorgestellten Link-Inventarien auf das Corpus angewandt und analysiert. Am Ende steht ein Vergleich der proprietären und der allgemeinen Link-Typen. Die Lösung der inhaltlichen Fragestellung muß im Laufe der gesamten Untersuchung jedoch stets in ihrer ‚dienenden‘ Funktion gesehen werden: Sie dient, wie gesagt, primär dazu, das TAH nicht nur theoretisch zu fundieren, sondern das Modell mit einer Fragestellung aus der ‚Wirklichkeit‘ der Hypertext-Forschung zu konfrontieren.

Klar sollte auch sein, daß der praktische Test der Methodik von TAH hier nur exemplarisch, sozusagen als kleine Pilotstudie erfolgen kann. Das Modell muß sich aber erst in einer Reihe von Untersuchungen ähnlicher Zielrichtung als nützlich erweisen, bevor über dessen deskriptive Kraft gesicherte Aussagen gemacht werden können. Erste Stärken und Schwächen sollten aber dennoch bereits im Rahmen dieser praktischen Anwendung sichtbar werden.

Auch die Beantwortung der inhaltlichen Frage nach einem sinnvollen Inventar typisierter Links kann, wie gesagt, hier keinesfalls erschöpfend behandelt werden. Inhaltliche Ergebnisse sollten daher keinesfalls als unumstößlich gesehen, sondern eher so interpretiert werden, als daß sie die zukünftige Ausrichtung von Link-Vokabularien auf der Oberflächenstruktur von Hypertexten andeuten können. Breiter angelegte Untersuchungen zur Thematik der typisierten Links aus Psychologie und Kognitionswissenschaft werden aufgrund der zunehmenden Wichtigkeit des Themas aber wohl kaum lange auf sich warten lassen.

Um die Schritte der praktischen Analyse besser nachvollziehen zu können, dürfte es für den Leser hilfreich sein, ab und an auf die überblicksartige Darstellung des TAH in den Tabellen 5, 6 und 7 des Kapitels 14.3 zurückzugreifen.

17.2 Allgemeine, exemplarische Analyse des Corpus'

Bevor auf die eigentliche inhaltliche Fragestellung nach Art und Signalisierung der typisierten Links eingegangen werden kann, muß zuerst – getreu der Top-Down-Philosophie des TAH – eine allgemeine Analyse der kontextuellen, funktionalen und grobstrukturellen Aspekte der Corpus-Hypertexte erfolgen.

17.2.1 Kontextuelle Aspekte

In einem ersten Schritt werden die technischen Bedingungen des Lesens von WWW-Hypertexten, sowie deren Kommunikationsform und Handlungsbereich besprochen.

Allgemeine Bedingungen des Zugriffs auf WWW-Hypertexte

Conditio sine qua non für die Lektüre eines Buches ist, dieses Buch physisch vor sich zu haben. Um Hypertexte im WWW lesen zu können, muß man Zugriff auf einen Computer mit Zugang zum Internet haben. Während ein Buch ohne Bedienungsanleitung ausgeliefert wird¹, verlangt das Verbinden des eigenen Rechners mit dem Netz gewisse elementare Fähigkeiten im Umgang mit technischen Geräten und Software.² Dazu zählt die Benutzung der Browser-Software: Das Programm muß zuerst lokalisiert und gestartet werden. Dann hat sich der Leser mit der internen Logik des Browsers auseinanderzusetzen: Zuallererst muß er wissen, wo die Adresse des zu lesenden Hypertextes einzugeben ist.³ Um sich erfolgreich mit komplexeren Hypertexten

¹ Komplexe Bücher, besonders technische und wissenschaftliche, verfügen oftmals jedoch über paratextuelle Bereiche, in denen Aussagen über den Aufbau des Buches sowie Lektüreempfehlungen gegeben werden (siehe hierzu auch Kapitel 11).

² Erfolgt der Zugriff auf das WWW von einem bereits eingerichteten Rechner am Arbeitsplatz oder in einem Internet-Cafe, sind solche Fähigkeiten keine unbedingt notwendige Bedingung, um Hypertexte im WWW lesen zu können.

³ Dies mag elementar klingen. Interessant ist jedoch, daß es (mir mündlich überlieferte) Erfahrungsberichte von (in der Computerbenutzung eher weniger versierten) Mitmenschen gibt, deren neuer Computer mit fest installierter Browser-Software geliefert wurde. Die Browser-Software war so konfiguriert, daß die Adresszeile nicht eingeblendet war. Der Kunde ging von einem Defekt des Computers aus, da er nirgendwo die Adresse *http://www.yahoo.de* eingeben konnte.

auseinanderzusetzen, muß er darüber hinaus die Navigationstechniken des Browsers an solchen Stellen verwenden können, an denen die hypertext-internen Navigationsmöglichkeiten nicht ausreichen.

Es ist jedoch anzunehmen, daß sich der elementare Umgang mit Computern zu einer allgemeinen Kulturtechnik entwickeln wird. Das ‚Surfen im Internet‘ dürfte dann kaum mehr intellektuell-technische Bedingungen an den Leser stellen als die Lektüre eines Buches.

Kontextuelle Aspekte im engeren Sinn

Die kontextuellen Aspekte von Hypertexten wurden bereits überwiegend in Kapitel 10.2 geklärt und sollen daher an dieser Stelle nur nochmals kurz zusammenfassend dargestellt werden: Die Hypertexte sind medial an das WWW gebunden¹, die Form der Informationsdarbietung ist primär schriftlich, zunehmend jedoch multimedial. Die Kommunikationsform aller Corpus-Hypertexte ist die WEBSITE: Die Kommunikationsrichtung ist stets monologisch. Auch die durch Hyper-Links ermöglichte, bedingte Interaktion macht es, meiner Meinung nach, nicht möglich, von einer dialogischen Kommunikationsrichtung zu sprechen – zu eingeschränkt ist hierfür die ‚Dialogfähigkeit‘ des Systems, die im Vergleich zum Face-to-Face-Gespräch nur in den engen Grenzen der vom Autor definierten Links erfolgen kann. Der Kontakt zwischen Autor und Leser ist in der Regel zeitlich und räumlich getrennt. SELF, SELF8 und SPORT1 bieten zusätzlich sogenannte CHATS an: elektronische ‚Räume‘, in denen in Echtzeit kommuniziert werden kann. Chats fallen jedoch nicht in den Rahmen dieser Untersuchung, da deren Texte (= Transkripte) keine hypertextuellen Gebilde darstellen, sondern eher mit Mitschnitten von Telefonkonferenzen zu vergleichen sind.²

Sämtliche Hypertexte des Corpus‘ können als Medien der Massenkommunikation gesehen und damit dem öffentlichen Handlungsbereich

¹ Dies bedeutet jedoch keineswegs, daß die Hypertexte nicht zum lokalen Lesen auf den eigenen Computer geladen und gespeichert werden können. Bei SELF u. SELF8 wird dies ausdrücklich angeboten. Die Verlinkungsstruktur wird dann relativ zum Verzeichnis der eigenen Festplatte angegeben, die lokale Lesbarkeit ist damit gewährleistet. Bei den anderen Texten des Corpus‘ müßte technisch eingegriffen werden, um aus der absoluten Adressierung (wie W3C etwa <http://www.w3.org/TR/NOTE-xlink-req>) eine relative zu generieren (etwa *./TR/NOTE-xlink-req* o.ä.). Für solche Belange existieren spezielle Programme.

² Interessant ist hierbei natürlich in erster Linie der ‚quasi-mündliche‘ Stil in solchen – eigentlich an die Schriftlichkeit gebundenen – Chat-Räumen. Vergleichbares ist auch in Email-Kommunikation zu beobachten.

zugeordnet werden, wobei das Verhältnis *Autor-Leser* sehr unterschiedlich ist: Einem potentiellen Massenpublikum auf Leserseite steht auf Autoreseite eine Organisation (W3C), eine Firma (SPORT1), eine Institution (GRAMMIS) oder eine von einer Privatperson geleitete Initiative (SELF, SELF8) gegenüber. Das nötige Fachwissen zur Lektüre wird manchmal durch die Angabe der Zielgruppe direkt signalisiert (SELF, SELF8, GRAMMIS), womit die potentielle Rezipientengruppe eingeschränkt werden soll. Im Text W3C wirkt hingegen die technische Komplexität der einzelnen Knoten als impliziter Filter bzgl. der Rezipientengruppe: Ein Leser, der zufällig auf die Startseite dieses Hypertextes gerät und nicht in die Thematik der Markup-Sprachen eingearbeitet ist, dürfte durch die hohe Anzahl einschlägiger Begriffe sowie einen äußerst technischen Schreibstil leicht ‚abgeschreckt‘ werden. Als weiterer impliziter Filter wirkt (im deutschen Sprachraum) auch die Verwendung der englischen, informationswissenschaftlichen Fachsprache für den gesamten Hypertext W3C¹.

17.2.2 Kommunikativ-funktionale Aspekte

Nach Klärung der kontextuellen Verhältnisse befaßt sich dieses Kapitel auf der Ebene *Hypertext* unter kommunikativ-funktionalen Aspekten mit dem Corpus. Besonders interessant ist in diesem Zusammenhang die Beziehung zwischen informativer Hauptfunktion und appellativer Nebenfunktion mancher Texte.

Zunächst gilt das Interesse dem Corpusvertreter SPORT1: Die Hauptfunktion von SPORT1 ist eine appellative, die Informationsfunktion ist lediglich in einer untergeordneten Rolle zu sehen. Appellativ ist dieser Corpus-Text in zweifacher Hinsicht: Einerseits aufgrund der hohen Zahl von Werbe-Teil-Knoten innerhalb der Platzhalter-Knoten (siehe hierzu Kapitel 17.2.4), die von informativen Texten mit sportspezifischer Thematik flankiert sind. Diese appellativen Teilknoten sind alle mit extra-hypertextuellen Links versehen, die zumeist auf die Internet-Auftritte von Werbeträgern verweisen. Als Textsorte läßt sich hier die elektronische Variante der Anzeigenwerbung, der WERBEBANNER, festlegen. Die Werbebanner selbst fallen nicht in den Bereich dieser Untersuchung. Festzuhalten ist jedoch, daß in einem mit Werbung so stark durchzogenen Hypertext wie SPORT1 die ‚appellative Kraft‘ dieser Werbebanner offensichtlich auch auf die redaktionellen Teil-Knoten durchschlägt: Der Erlös für das Schalten von Werbung im WWW ist stark von der Besucherzahl der

¹ Einige wenige Knoten sind jedoch in anderen Sprachen abrufbar. Die Startseite des deutschen Teil-Hypertextes des W3C lautet: <http://www.w3.org/Consortium/Offices/Germany/uebersetzungen>

jeweiligen Seite abhängig. Insofern haben die Betreiber ein natürliches Interesse, möglichst viele Besucher möglichst lange auf ihren Seiten zu halten – was sich wiederum auf den Stil der informativen Texte durchschlägt.¹ Die in Abschnitt 17.2.4 beschriebenen strukturellen Aspekte unterstützen diese Intention.

Ebenso verfügt SPORT1 über Teilbereiche mit Obligationsfunktion. Hier sind die verschiedenen Online-Shops angesprochen, in denen der Kunde per elektronischem Kaufvertrag Waren erwerben kann, die thematisch mit den Beiträgen des redaktionellen Teils verknüpft sind: Wird etwa über das Spiel eines bestimmten Fußballclubs berichtet, so ist diese Information stets flankiert von Links zu Knoten, auf denen Trikots dieses Vereins zu erwerben sind.

Als nächstes soll ein Blick auf die globale Textfunktion des Corpus-Hypertextes W3C geworfen werden: W3C ist funktional als rein informativ aufzufassen. Auf den Seiten sind keine Werbebanner geschaltet. Lediglich die Verwendung der Suchmaschine *Google* könnte unter Umständen als appellativ interpretiert werden: Die Bevorzugung einer Suchmaschine seitens der Organisation, die für sämtliche WWW-Standards verantwortlich ist, stellt einen nicht unerheblichen Prestigegewinn für die Betreiber der Suchmaschine dar.

Die Hypertexte SELF und SELF8 werden im folgenden Abschnitt gemeinsam behandelt, deren Hauptfunktionen sind informative und instruktive. In Kapitel 17.3 wird nochmals exemplarisch genauer auf die Verschränkung dieser beiden Funktionen eingegangen. Beide Hypertexte verfügen ebenfalls über eine gewisse appellative Nebenfunktion, die in bestimmten Knoten direkt signalisiert wird, zum Beispiel:

„Das Motto – die Energie des Verstehens – ist ernst gemeint. SELFHTML richtet sich nicht an Leute, die mit dem Anspruch antreten, alles fertig vorgesetzt zu bekommen und dann damit prahlen zu können, was sie tolles Eigenes geschaffen haben. Diese Dokumentation ist kein "Ihre-Homepage-in-zehn-Minuten"-Quickie. Das Dokument möchte erreichen, dass seine Anwender verstehen was sie lernen und wissen was sie tun, wenn sie an eigenen Webseiten arbeiten. Es richtet sich an Lernwillige, die bereit sind, sich mit einer faszinierenden Materie auseinanderzusetzen, die aber immer wieder neue Fragen aufwirft, mitunter zur Verzweiflung treibt und viel Zeit zum Entdecken und Ausprobieren erfordert.“

(SELF8 [editorial\hinweise.htm] Abschnitt: *Zielgruppen*)

¹ Eine (höchstwahrscheinlich äußerst interessante) genauere Analyse der Auswirkung eines appellativen Umfeldes auf informative Texte im WWW muß aufgrund der Zielsetzung dieser Arbeit aber leider entfallen.

SELF und SELF8 verstehen sich also nicht als rein informative Hypertexte, sondern als solche mit ‚erzieherischem Auftrag‘, wie dem obigen Zitat zu entnehmen ist. Der Leser soll nicht fertige Konzepte der HTML-Programmierung kopieren, sondern aus Eigeninitiative und Begeisterung sich selbst – self! – in die Problematik von HTML hineindenken. Diese ‚werbende‘ Funktion dient auch dazu, eine Gemeinschaft Gleichgesinnter aufzubauen:

„Die SELFHTML Community, das ist die Gemeinschaft aller, die sich regelmäßig in Foren und Chats rund um SELFHTML herumtreiben und auch vor Live-Treffen nicht zurückschrecken. Damit geben sie dem SELF-Raum sein Gepräge und verleihen ihm erst richtig Leben.“

(SELF8 [editorial\danke.htm] Abschnitt: *an die SELFHTML Community*)

Diese vergleichsweise ausgeprägte appellative Funktion unterscheidet SELF bzw. SELF8 von rein informativen Hypertexten wie etwa W3C.

Der letzte Vertreter des Corpus' ist der Hypertext GRAMMIS: Er verfügt lediglich über eine sehr schwach ausgeprägte appellative Teilfunktion:

„GRAMMIS® ist ein Projekt des Instituts für Deutsche Sprache in Mannheim, das zum Ziel hat, die ach so trockene, oft schwer verständliche deutsche Grammatik leichter zugänglich und, soweit möglich, unterhaltsam und interessant zu machen.“

GRAMMIS [index.html]

„Es muss schon einiges zusammenkommen, bis ein normaler Zeitgenosse zu einer Grammatik greift: ein wichtiges Schreiben vielleicht, womöglich eine Bewerbung, und die Angst sich zu blamieren oder gar für unqualifiziert gehalten zu werden, weil man sich nicht korrekt auszudrücken weiß. Wer das Glück hat, Experten zu kennen, darf auf eine ad hoc Lösung seiner Probleme hoffen [...] Wer auf sich gestellt ist und sich jetzt daran erinnert, dass es ja Grammatiken gibt, in denen man nachsehen kann, wie "es richtig heißen muss", der wird kaum zu einschlägigen Informationen vordringen: Zu abstrakt, zu komplex sind die Regeln, die Sprachen beschreiben, und das liegt nicht einfach daran, dass hier weltfremde Sprachforscher ihr "Fachchinesisch" ausbreiten, jedenfalls nicht in erster Linie.“

GRAMMIS [infologo.html] Abschnitt *Wozu Grammatik? Wozu GRAMMIS?*

Hiermit wird Werbung für die Informativität von GRAMMIS betrieben, die deutlich besser sein soll, als dies bei herkömmlichen Grammatiken der Fall ist. Ansonsten überwiegt eindeutig die informative Textfunktion. Gestützt wird diese Behauptung durch die Abwesenheit von werbenden Elementen sowie den informativen, schmucklosen Schreibstil und die primär explikative thematische Entfaltung.

17.2.3 Konventionelle Aspekte

Nach der überblicksartigen Darstellung der funktionalen Verhältnisse soll der Blick nun auf die konventionellen Aspekte der Corpus-Hypertexte gelenkt werden: In SELF und SELF8 ist die Textsorte jeweils explizit als instruktive und informative „Mischung aus Referenz und Tutorial“ signalisiert (SELF8 [editorial\hinweise.htm] Abschnitt *Zielgruppen*). Ist die Textsorte nicht in vergleichbarer Weise direkt angegeben, so ist für deren Ermittlung hauptsächlich die globale Textfunktion (siehe Kapitel 17.2.2) maßgebend. Insofern können GRAMMIS und W3C als Vertreter der Klasse der informativen Texte gesehen und bzgl. des Typs als *Referenzen* eingeordnet werden. Wobei W3C als ‚reines‘ Nachschlagewerk und Informationsmedium im Wortsinne konzipiert ist. In W3C existieren verschwindend wenig Knoten mit primär instruktiver Funktion. GRAMMIS hingegen ist als grammatische Referenz mit instruktivem Charakter zu bezeichnen, der jedoch deutlich geringer ausgeprägt ist als der von SELF und dem Nachfolger SELF8.

Im Gegensatz dazu kann SPORT1 als Sammlung von Nachrichten, Kommentaren, Werbetexten und ähnlichen Textsorten gelten, welche auch für gedruckte Zeitungen und Zeitschriften prägend sind. Eine saubere Trennung dieser redaktionellen Teile von Werbeelementen ist allerdings nicht vollzogen. Zu letzterem Punkt wird im folgenden unter strukturellen Aspekten noch Ergänzendes zu bemerken sein.

17.2.4 Strukturelle und paratextuelle Aspekte

In diesem Abschnitt wird auf Analysekatoren wie Anzahl und Typen der Knoten, Gliederungsaspekte, Screendesign und browser-spezifische Besonderheiten eingegangen.

Knoten und Links: Anzahl und Typen

Die Anzahl der Knoten von WWW-Hypertexten ist – falls diese nicht vom Autor explizit angegeben wurden – nicht einfach zu ermitteln. Im Falle von SELF und SELF8 ist es hilfreich, daß der Autor ein gepacktes Archiv zum Download des gesamten Hypertextes zur Verfügung stellt. Lokal ist dann mit Suchalgorithmen relativ komfortabel die Anzahl der Knoten und Links zu ermitteln: SELF besteht aus 744 Knoten und 12.450 Links. Die Druckversion erstreckt sich auf 944 Seiten (Münz u. Nefzger 1999). SELF8 hingegen ist auf

1403 Knoten und 22.908 Links angewachsen, die Druckversion ist in zwei Bänden mit über 2000 Seiten erhältlich (Münz 2002).

Zwar bestünde auch für den Hypertext W3C die Möglichkeit, ihn mit speziellen Programmen für den lokalen Gebrauch zu spiegeln: Eine Möglichkeit wäre, mit Hilfe einer kleinen Anwendung, einem sogenannten CRAWLER, sämtliche hypertext-internen Verweise von W3C – beginnend mit der Startseite – zu durchlaufen, um somit online die Anzahl der Knoten und Links zu ermitteln. Dies erzeugt jedoch (unnötigen) Übertragungsverkehr auf einer ohnehin stark frequentierten Site und steht vom Aufwand in keinem Verhältnis zum Ergebnis. Technisch nahezu unmöglich wäre eine wie oben beschriebene Ermittlung der Knoten- und Link-Anzahl von GRAMMIS und SPORT1: Die Navigation in diesen teilweise dynamisch erzeugten Knoten geschieht unter anderem mit Hilfe von in JavaScript programmierten Menüs. Überblick über die Anzahl der Knoten und Links haben hier wohl nur die Redakteure und Programmierer dieser Hypertexte.

Die etwas ausführlichen Erläuterungen zu Problemen der Ermittlung von Knoten- und Link-Anzahl machen abermals die grundlegenden Unterschiede zu gedruckten Texten deutlich: Der Umfang mancher Hypertexte läßt sich nicht einmal mit technischen Tricks erkennen – noch viel weniger aber beim normalen Lesen!

Über die Knoten-Typen lassen sich jedoch folgende zusammenfassende Aussagen treffen:¹ SPORT1 ist der tendenziell am stärksten multimedial ausgelegte Hypertext des Corpus'. Die Knoten von SELF, SELF8, W3C sind primär textuell gehalten – mit eingebetteten Quellcode- und Grafikbeispielen. GRAMMIS verfügt neben hauptsächlich textuellen Knoten auch über einige Tondokumente, Grafiken und Animationen.

Verlegerischer Peritext: Screendesign und Metaphorik

Im folgenden wird kurz das Screendesign der Corpus-Hypertexte angesprochen. Die Dekonstruktion des allgemeinen Screendesigns von WWW-Sites sei jedoch den Disziplinen überlassen, die hierfür über geeignete Begrifflichkeiten und Methoden verfügen. Interessant für eine Analyse mit TAH sind nur die Bereiche, in denen sich die inhaltlichen Fragestellungen mit Aspekten des Designs berühren. Im Falle der Untersuchung typisierter Links ist natürlich

¹ Die Link-Typen werden ausführlich in Kapitel 17.3 behandelt!

besonders das Visualisieren dieser Links mit Mitteln des Screendesigns interessant. Details werden hierzu bei der Untersuchung der Ebene *Link* im Rahmen der metasprachlichen Aspekte besprochen.

An dieser Stelle sei jedoch etwas anderes angesprochen: Hypertext-weit kann nämlich ein metaphorisches Screendesign verwendet werden, daß sich auch auf die Darstellung bzw. Typisierung der Links auswirkt. In SELF und SELF8 ist das der Fall. Diese beiden Hypertexte benutzen die Ordnerstruktur von Windows als Metapher zur Visualisierung der Hierarchie des Hypertextes:



Abbildung 14: Windows-Ordnerstruktur als Metapher für die Knoten-Hierarchie in SELF8

GRAMMIS bedient sich hingegen der Buch- bzw. Bibliotheksmetapher:

„In der grammatischen Bibliothek des Systems findet sich ein Bücherregal mit elektronischen Büchern zu verschiedenen grammatischen Themenbereichen.“

(Storrer 1999 S. 53)

Dieses Zitat bezieht sich allerdings auf eine frühe, im Hypertext-System *ToolBook* erzeugte Version von GRAMMIS. Die Version des Corpus' hingegen bedient sich nicht mehr in so starker Form dieser Metaphorik, lediglich die Namengebung der Teil-Hypertexte von GRAMMIS erinnert an eine Bibliothek: *Terminologisches Wörterbuch*, *Rechtschreib-Wörterbuch* oder *Grammatische Bibliographie*. Dies geschieht jedoch nicht aufgrund einer bewußt aufgebauten Metaphorik sondern aufgrund der Thematik von GRAMMIS als solcher: Ein Wörterbuch ist nun mal ein Wörterbuch, auch wenn es elektronisch umgesetzt wurde, wie dies bei GRAMMIS der Fall ist.

SPORT1 bedient sich der Boulevardzeitungsmetapher. Bucher 1999 beschreibt diesen Typ so:

„Die Homepage dieser Blätter ist mit einer Vielzahl von Links gespickt, teilweise verteilt auf mehrere Frames, so daß sich die Architektur der virtuellen Zeitung nur mühsam erschließen läßt.“

(Bucher 1999 S. 18)

Dieses metaphorische Screendesign beeinflußt die Gesamtstruktur des Hypertextes in hohem Maß und wird daher weiter unten bei der Besprechung von Gliederung und Navigation nochmals ausführlich aufgenommen.

Der Corpus-Text W3C bedient sich keiner Metapher. Das Screendesign kann als ‚klassisch‘ für das WWW bezeichnet werden, da es sich wenig von den Seiten der ersten Generation unterscheidet.¹

Anwenderspezifischer Peritext: Browser-Software

Der anwenderspezifische Peritext kann sich je nach verwendeter Software stark unterscheiden und soll hier, wie bereits erwähnt, nicht en détail untersucht werden. Beispielhaft sei nur eine Beschränkung älterer Browser-Software genannt: Das Anzeigen von Mouse-Over-Texten ist hier nur komplex über Skript-Sprachen zu realisieren (siehe auch Kapitel 16.3.3). Alternativ- bzw. Typisierungstexte von Piktogrammen die mit dem *alt*- bzw. *title*-Attribut realisiert sind, werden mit diesen Programmen nicht angezeigt, mit neueren Versionen des *Internet Explorer* von *Microsoft* oder des *Navigator* von *Netscape* jedoch schon. Statistiken können über die Benutzung der unterschiedlichen Browser empirisch Aufschluß geben.² Darauf wird hier jedoch nicht weiter eingegangen.

Allgemeiner Peritext: Gliederung und Navigationshilfen

Im Rahmen der Beschreibung des allgemeinen Peritextes wird auf die Gliederung der Hypertexte, auf deren Visualisierung sowie auf Hilfen zur Navigation eingegangen.

SELF und SELF8 sind hierarchisch aufgebaut. Die dort beschriebenen Technologien gliedern den jeweiligen Hypertext in verschiedene logische Hauptbereiche, die – da in sich thematisch geschlossen aber doch via Links mit anderen Bereichen von SELF bzw. SELF8 verbunden – auch als ‚Teil-Hypertexte‘ bezeichnet werden können.

Die hierarchische Gliederung von SELF bzw. SELF8 ist relativ zur aktuellen Position in jedem Knoten oben und unten auf der Bildschirmseite angeführt (siehe Abbildung 22 und 23). Desweiteren existiert eine zentrale Übersichtsseite sowie eine zusätzliche Übersichtsseite pro Themenbereich. Außerdem steht dem

¹ WWW-Sites von heutigem Design werden als der vierten oder fünften Generation zugehörig eingeordnet. Genauer über USABILITY, SREEN- und WEB-DESIGN informiert Nielsen 2000.

² Ein Beispiel unter vielen für solche Browser-Statistiken im WWW findet sich unter <http://www.webhits.de/webhits/inetstat.htm>.

Leser eine gut ausgebaute hypertext-interne Navigation zur Verfügung: Mit Hilfe typisierter Piktogramme kann (innerhalb der Logik der hypertext-internen Gliederung) zum vorherigen bzw. zum folgenden Knoten ‚geblättert‘ werden. Zusätzlich gibt es folgende komplexe „Navigationshilfen“ (SELF8 [navigation\index.htm]): Kurzreferenzen, Inhaltsverzeichnis, Syntaxverzeichnis, Stichwortverzeichnis sowie ein Verzeichnis häufig gestellter Fragen.

Diese auch aus traditionellen Medien bekannten Orientierungshilfen werden ergänzt durch eine Volltextsuche sowie eine Quickbar, mit der sich durch die Sinnabschnitte von SELF bzw. SELF8 navigieren läßt:

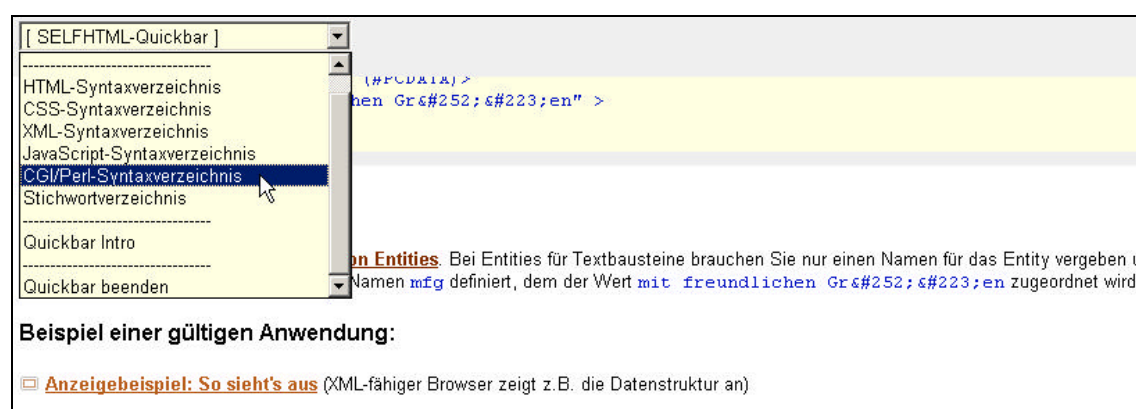


Abbildung 15: Pull-Down-Menü *Quickbar* als hypertext-interne Navigationshilfe von SELF8

Als nächster Corpusvertreter wird der Text SPORT1 im Hinblick auf Aspekte des allgemeinen Peritextes betrachtet: Um den Grobaufbau von SPORT1 zu beschreiben, muß zunächst ein Vorgriff auf die Betrachtungen der Ebene *Knoten* erfolgen: Die Struktur der Knoten in SPORT1 unterscheidet sich grundlegend von der der anderen Texte im Corpus. Während dort den Knoten in der Regel funktionale Integrität zuzusprechen ist, vereint eine im Browser angezeigte Seite von SPORT1 mehrere Textfunktionen – zumeist appellative und informative. Dies wirkt sich auf die Struktur einer Bildschirmseite aus, die ich hier bewußt nicht *Knoten* nenne! Sie ist zumeist derart aufgebaut, daß neben Werbeabschnitten kurze sogenannte TEASER-TEXTE¹ stehen, welche mit einem Pfeil-Piktogramm auf den Knoten verweisen, der den gesamten Text enthält. Dieser Seitenaufbau erinnert an Printausgaben von Magazinen und Zeitschriften,

¹ Der Begriff TEASER stammt aus der Werbung. Mit Teasern wird (in möglichst spektakulären Szenen) zeitlich lange vor dem Kinostart auf einen kommenden Film hingewiesen. Ein Teaser soll ‚Appetit‘ auf den Film machen aber, anders als der später herausgebrachte TRAILER, noch nicht viel von der Handlung vorwegnehmen.

die Online-Auftritte von *Spiegel*, *Fokus* und der *Süddeutschen Zeitung* sind ähnlich konzipiert.¹ Eine Bildschirmseite solcher Hypertexte kann als ‚Platzhalter-Knoten‘ gesehen werden, der eine Art Rahmen für Knoten oder Teil-Knoten darstellt. Die auf diese Weise eingebetteten redaktionellen Teil-Knoten sind als Ausschnitte der eigentlichen Inhalt-Knoten zu verstehen, die den Leser motivieren sollen, mehr von den eigentlichen Inhalt-Knoten sehen zu wollen – und damit länger auf der Seite SPORT1 zu verweilen. Was hier als *Platzhalter-Knoten* bezeichnet wird, nennt sich bei Bucher 1999 „advanced organizer“, der im Falle von SPORT1 in einer nach dem „Dschungelprinzip“ geordneten „Boulevard-Variante“ vorliegt (Bucher 1999 S. 18). Aufgrund dieser Ausformung einer „elektronischen Zeitschrift“ führt Bucher folgende Kritik bzgl. der Link-Struktur an, die in dieser Form auch an SPORT1 zu richten ist:

„Erschwerend kommt hinzu [zu den allgemeinen Schwierigkeiten einer nach dem ‚Dschungelprinzip‘ in chaotischer Form strukturierten elektronischen Zeitschrift, O. H.] daß die Links sehr unterschiedlich gestaltet sind: als Icons, Buttons, Stichwörter, Rubrikenkennzeichnung oder Überschriften. Das Nebeneinander verschiedener Strukturprinzipien behindert bei dieser Gestaltungsstrategie ein zielgerichtetes Navigieren.“

(Bucher 1999 S. 18)

En détail werden in Kapitel 17.3.4 für die Links in SPORT1 konkrete Belege angeführt.

SPORT1 stellt außer einer anklickbaren Gliederung auf der linken Bildschirmseite keine weiteren Navigationshilfen wie Suchfunktionen, Pull-Down-Menüs o.ä. zur Verfügung. Dies liegt wohl hauptsächlich an der Gliederung dieses Hypertextes (siehe oben), die der in Kapitel 17.2.2 beschriebenen Funktion dient: Der Leser soll nicht möglichst schnell zur gewünschten Information gelangen, sondern möglichst lange auf der Site bleiben.

Als nächstes wird der allgemeine Peritext von GRAMMIS betrachtet: Der Grobaufbau von GRAMMIS wird im Hypertext selbst beschrieben. Demnach besteht GRAMMIS aus folgenden informationstragenden Einheiten:²

¹ Online unter <http://www.spiegel.de> (Spiegel), <http://www.fokus.de> (Fokus) und <http://www.sueddeutsche.de> (Süddeutsche Zeitung).

² Bestimmte Teilbereiche von GRAMMIS sind noch nicht fertiggestellt [Stand: 20.02.2002].

- „1. Hypertexte zur Systematischen Grammatik (Kompakttexte für schnellen Überblick über ganze Bereiche, Detailtexte zur gründlichen Information, Vertiefungstexte für die eingehendere Beschäftigung mit bestimmten Fragen
2. eine Datenbank ‚Terminologisches Wörterbuch‘
3. eine Datenbank ‚Grammatisches Wörterbuch‘
4. Bibliografische Datenbanken, insbesondere die ‚Bibliografie zur deutschen ‚Grammatik‘“

(GRAMMIS [information.html] Abschnitt 1 Die globale Struktur)

Die unterschiedlichen Teilbereiche sind farblich voneinander unterschieden. Eine in diesen Farben ausgezeichnete Hauptnavigationsleiste ist stets auf der linken Bildschirmseite sichtbar und visualisiert damit permanent die Grobstruktur des Hypertextes.

Auf die oben dargestellten informationstragenden Einheiten gibt es folgende Zugriffsmöglichkeiten:

- a. über das Inhaltsverzeichnis des virtuelles [sic!] Buchs ‚Systematische Grammatik‘
- b. über das alphabetische Verzeichnis grammatischer Fachausdrücke im ‚Terminologischen Wörterbuch‘, das, soweit entsprechende Texte verfügbar sind, zu Detailtexten der ‚Systematischen Grammatik‘ weiterführt
- c. über das alphabetische Verzeichnis des ‚Grammatischen Wörterbuchs‘, das ebenfalls, soweit entsprechende Texte verfügbar sind, zu Detailtexten der ‚Systematischen Grammatik‘ weiterführt
- d. eine Volltextsuchmaschine in den Texten der Systematischen Grammatik
- e. ein System zur Recherche in den bibliografischen Datenbanken“

(GRAMMIS [information.html] Abschnitt 1 Die globale Struktur)

Mit Hilfe des Pull-Down-Menüs *Übersicht* der im oberen Bildschirmbereich dargestellten Subnavigationsleiste läßt sich beispielsweise zwischen den großen Sinnabschnitten im Teilbereich *Systematische Grammatik* von GRAMMIS navigieren:

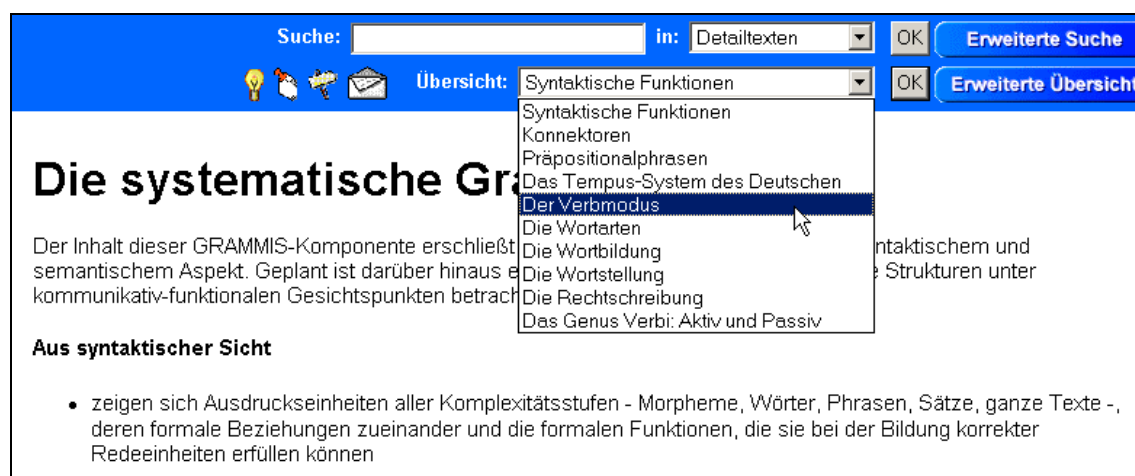


Abbildung 16: Pull-Down-Menü *Übersicht* als hypertext-interne Navigationshilfe von GRAMMIS

Innerhalb dieses Sinnabschnittes läßt sich jeder Knoten als überblicksartiger *Kompakttext* oder als ausführlicher *Detailtext* anzeigen. Zusätzlich steht auch in GRAMMIS eine Volltextsuche zur Verfügung.

Desweiteren enthält die Subnavigationsleiste der systematischen Grammatik vier Piktogramme, die Links auf folgende Funktionen bieten:

- Druckansicht des aktuellen Knotens
- Tips zur Navigation
- Anlegen von Notizen für angemeldete Benutzer¹
- Aufrufen eines vorab gespeicherten Lesepfades für angemeldete Benutzer
- Funktion zum Versand von Emails.

Die anderen Teilbereich von Grammis – Wörterbücher und Bibliographie – verfügen über ähnliche Zugriffsmöglichkeiten via Pull-Down-Menüs oder Suchfelder.

Abschließend soll noch ein kurzer Blick auf den Corpus-Hypertext W3C geworfen werden: W3C stellt ebenfalls auf der linken Bildschirmseite eine Navigationsleiste mit den alphabetisch geordneten ‚Teil-Hypertexten‘ zur Verfügung. Jeder Teil-Hypertext ist hier, wie auch in SELF bzw. SELF8, einer bestimmten Technologie zugeordnet (HTML, XML, CSS o.ä.). Man kann diese Navigation also als inhaltlich-semantisch motiviert beschreiben. Auf der rechten Bildschirmseite existiert zusätzlich noch eine eher pragmatisch orientierte Navigation: Hier besteht die Möglichkeit, auf verschiedene Arten mit dem W3C in Interaktion zu treten („Contac us“, „Get Involved“, „Member Area“ o.ä.). Am oberen Bildschirmrand sind fünf weitere, sehr heterogene ‚Navigations-Links‘ aufgeführt: Abermals die Möglichkeit zur Kontaktaufnahme („Contact us“), inhaltliche Grobkategorien („Activities“, „Technical Reports“), eine Navigationshilfe („Site Index“) sowie ein Link zu Meta-Information über kontextuelle Aspekte des Hypertextes („About W3C“).

¹ Auf die Möglichkeiten der Personalisierung von GRAMMIS wird in dieser Arbeit nicht eingegangen.



Abbildung 17: Navigationsleisten und Suchfunktion in W3C

Zusätzlich existiert noch eine Volltextsuchmaschine der Firma *Google*, mit welcher wahlweise das gesamte WWW oder aber die Site des W3C durchsucht werden kann.

17.2.5 Intertextuelle Aspekte

Der letzte Beschreibungsaspekt der TAH-Analyse auf Ebene des gesamten Hypertextes betrifft die Intertextualität. Als erster Vertreter soll der Text GRAMMIS herangezogen werden, da er hier eine gewisse Sonderstellung einnimmt: GRAMMIS verweist lediglich auf die Hauptseite des *Instituts für deutsche Sprache*, dessen Projekt GRAMMIS auch ist. Von diesen ‚hypertext-netz-internen‘ Verweisen¹ (und einigen wenigen hypertext-externen Links auf den Seiten der Autoren) abgesehen, zeigt sich GRAMMIS geschlossen – fast schon mit einer Hypertext-Insel (siehe Kapitel 5.1) vergleichbar. Dieser Festlegung steht jedoch entgegen, daß von fremden Hypertexten aus durchaus auf Knoten von GRAMMIS verwiesen wird.

Alle übrigen Texte des Corpus’ sind hochgradig intertextuell, d.h. sie verweisen einerseits auf weitere Hypertexte im WWW und werden andererseits aus dem WWW heraus referenziert. Im Rahmen der vorliegenden Untersuchung interessiert lediglich der erste Aspekt, wobei die tatsächliche layout-spezifische

¹ Die jedoch mit dem Mouse-Over-Text *Hyperlink ins WWW* von den GRAMMIS-Autoren als hypertext-(netz-)extern typisiert wird.

Unterscheidung von intra- und extra-hypertextuellen Links in Kapitel 17.3 abgehandelt wird. An dieser Stelle erfolgt vielmehr eine allgemeine Untersuchung der prinzipiellen Organisation von Intertextualität in den Corpus-Texten.

Als erstes Beispiel eines tatsächlich intertextuellen Hypertextes dient W3C: Hier sind extra-hypertextuelle Links nicht von den internen Varianten zu unterscheiden:

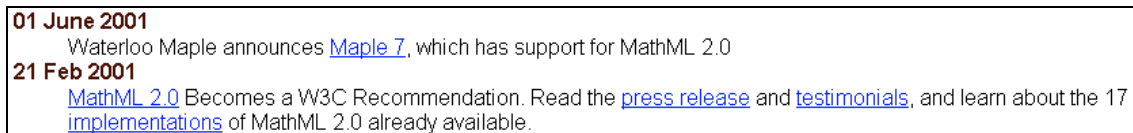


Abbildung 18: Untypisierte Links in W3C [/Math/] Abschnitt *Highlights*

In der Abbildung ist beispielsweise nicht ersichtlich, daß der Link *Maple 7* auf einen externen Hypertext verweist. Interne Links, wie etwa *MathML 2.0* unterscheiden sich nicht von der externen Variante. Lediglich im (dynamisch erzeugten) anwenderspezifischen Peritext ist der Unterschied eventuell sichtbar: Bewegt sich die Maus über den Link, so wird in der Statuszeile mancher Browser die URL des Ziel-Knotens sichtbar. Allerdings hat der Autor, wie schon erwähnt, keinen Einfluß auf die Wahl des Browsers, mit dem der Benutzer den Hypertext liest.

In SELF und SELF8 sind die Verhältnisse anders als in W3C: Im Fließtext der Knoten von SELF stehen ebenfalls sowohl interne und als auch externe Links. Bei Links auf externe Hypertexte wird aber per Piktogramm die Sprache dieser Hypertexte signalisiert:

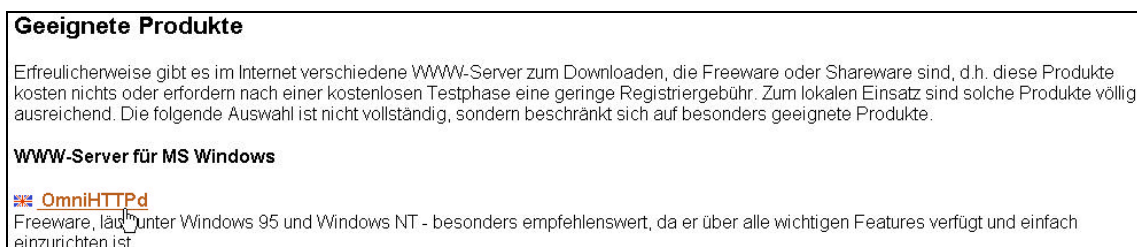


Abbildung 19: Extra-hypertextueller Link in SELF [tgba.htm] Abschnitt *WWW-Server für MS*

Auf diese und andere paratextuellen Aspekte der Typisierung der Link-Varianten wird in Kapitel 17.3.1 näher eingegangen.

Im Unterschied zu SELF sind in SELF8 die extra-hypertextuellen Links in separate Knoten außerhalb von SELF8 – aber innerhalb des Hypertext-Netzes *SELF-Raum* – in das sogenannte „Linkverzeichnis“¹ ausgelagert:

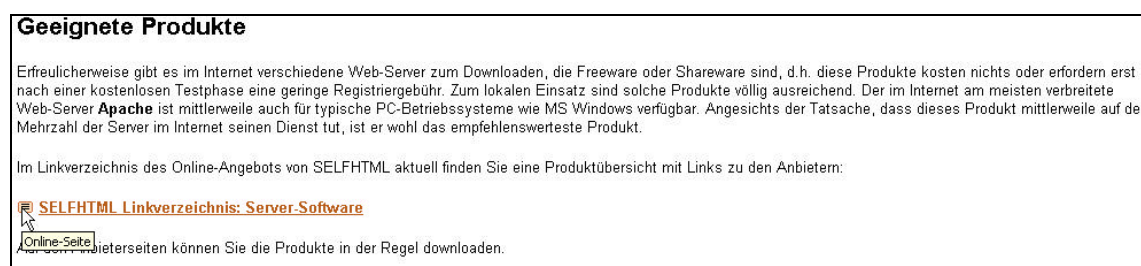


Abbildung 20: SELF8 [cgiperl/intro/webserverpc.htm] Abschnitt *Geeignete Produkte*

Diese Links sind mit Piktogrammen (und Mouse-Over-Texten²) typisiert, die bereits in Kapitel 16.3.3 in HTML besprochen wurden.

SELF und der Hypertext *Selbstaktuell* waren früher eher lose Ergänzungen zueinander, die hauptsächlich durch gemeinsames Screendesign ein erkennbares Hypertext-Netz bildeten. In SELF8 sind diese Zusammenhänge expliziter gemacht, SELF8 ist deutlicher ein Bestandteil des Hypertext-Netzes *SELF-Raum*, dem auch der Hypertext *Selbstaktuell* angehört.

Als letzter der Corpus-Hypertexte muß noch SPORT1 auf die allgemeinen Aspekte seiner Intertextualität hin untersucht werden: Hier sind extra-hypertextuelle Links nicht explizit kenntlich gemacht. Ein möglicher Grund hierfür ergibt sich, wenn man SPORT1 unter kontextuellen und funktionalen Aspekten betrachtet (siehe oben): Die Herausgeber von SPORT1 sind die Fernsehsender *SAT1* und *DSF* sowie die Zeitung *Sport Bild*.³ Verbunden sind diese darüber hinaus mit zahlreichen Vertragspartnern und Werbeträgern. Zusammen wird auf diese Weise ein kommerziell orientiertes Hypertext-Netz geschaffen. Den Herausgebern ist offensichtlich daran gelegen, den Leser (und Kunden!) so lange wie möglich innerhalb dieses Netzes zu halten. Jeder angeklickte Link, der auf die Seite eines Vertragspartners bzw. Werbeträgers führt, bedeutet für SPORT1 finanzielle Vorteile. Insofern wird dem Leser nicht signalisiert, daß er etwa per Klick auf einen Link zu *karstadt.de* den redaktionellen Bereich von SPORT1 verläßt:

¹ Online im Internet unter <http://selfaktuell.teamone.de/links/index.htm>.

² Die Bezeichnung *Online-Seite* macht Sinn, wenn der Leser den gesamten Hypertext als Archiv auf die eigene Festplatte entpackt und somit ‚offline‘ lesbar gemacht hat. Der Mouse-Over-Text weist ihn dann darauf hin, daß sich das Link-Ziel ‚online‘ im WWW befindet.

³ Im Hintergrund steht die *Kirch Intermedia GmbH*. Siehe hierzu auch das Impressum SPORT1 unter: <http://relaunch.sport1.de/coremedia/generator/www.sport1.de/Info/Impressum/Main.html>.

| | |
|--|--|
| <p>Pagelsdorf sagt Köln ab Der Ex-Hamburger entscheidet sich gegen die "Geißböcke", gesteht aber: "Das ist mir sehr schwer gefallen." [E]</p> |  |
| <p>Leverkusens Angst vor dem Absturz Durch die unglückliche Heimpleite gegen Schalke 04 hat Bayer nicht nur die Tabellenführung verspielt. [E]</p> | <p>COMMUNITY ▶ "Vorbild Uli Hoeneß"</p> |
| <p>VIDEO Der 21. Spieltag als multimediales Erlebnis</p>  <p>Die Hatz durch die englische Woche, Bayerns Blamage, Dortmunder Glücksmomente... - bei ran-online sind Sie Programmchef! Durch Q-DSL mit der besten Qualität im Netz. [E]</p> |  <p>Olaf Marschall gab im Chat mit www.fussballprofis.de überraschende Antworten. [E]</p> |
| <p>2. LIGA ▶ Bochum knackt Hannovers Serie</p>  <p>Die Festung ist "eingenommen". Am 21. Spieltag kassierte Hannover 96 die erste Pleite. In einem packenden Duell holte der VfL Bochum den bis dato ungeschlagenen Tabellenführer auf den</p> | <p>SHOP ▶ Jeder will eins... €46.02 Bayern München Trikot karstadt.sport.de</p>  <p>Das Trikot der Bayern - jetzt günstig im Sport1-Shop. [E]</p> |

Abbildung 21: Platzhalter-Knoten in Sport1 [dynamisch erzeugter Knoten] ¹

Ganz im Gegenteil: In SPORT1 wird dem Augenschein nach versucht, eine möglichst undurchsichtige Verlinkung von informativ und appellativ orientierten Knoten zu erreichen. Im obigen Beispiel werden auf diese Weise im eigentlich redaktionellen Teil Links auf Knoten mit Werbung für *Q-DSL*, *karstadt.de* und *iobox* eingebaut. Eine Verquickung informativer und appellativer Piktogramme tut das übrige, um eine vollständig unübersichtliche Mischung aus Informationstexten und Werbung zu erreichen (siehe auch Kapitel 17.3.4).

Nach diesen allgemeinen Betrachtungen der TAH-Ebene *Hypertext* geht das nächste Kapitel intensiver auf die Problematik der Link-Typisierung innerhalb des Corpus' ein.

¹ Die Knoten von SPORT1 werden tagesaktuell dynamisch erzeugt. Insofern ist bei diesem Corpus-Text eine genaue Referenzierbarkeit nicht möglich.

17.3 Analyse der Typisierten Links

Es folgt nun die Untersuchung einzelner Knoten¹ des Corpus² auf der TAH-Ebene *Link*. Der erste Blick gilt hierbei den proprietären Typisierungen der jeweiligen Hypertexte. In einem zweiten Schritt werden die in Kapitel 16.4 ausgewählten allgemeinen Vokabularien exemplarisch als Ersatz für die proprietären Codierungen verwendet. Die hierfür notwendige Auszeichnung der Links erfolgt durch das HTML-Attribut *title* (siehe Kapitel 15.2.4), eine Analyse der paratextuellen und browser-spezifischen Aspekte erübrigen sich damit bei diesem zweiten Schritt.

Das erste der folgenden Beispiele stellt eine TAH-Analyse vor, die auf sämtliche Beschreibungsaspekte der Ebene *Link* eingeht. Es darf daher als exemplarisch für eine Komplettanalyse dieser Ebene gelten. Die darauf folgenden Beispiele hingegen vertiefen lediglich ausgewählte interessante Bereiche der Link-Konstellation der jeweiligen Knoten.

Die Analyse hält sich an die Top-Down-Vorgehensweise des TAH. Vom Bezugsbereich ausgehend wird sich über kommunikativ-funktionale, thematisch- und grammatisch-strukturelle Aspekte zu paratextuellen bzw. metasprachlichen und browser-spezifischen Erscheinungen ‚heruntergearbeitet‘.

17.3.1 Analyse der proprietären Codierung von SELF

Die beiden folgenden Knoten sind aus dem Hypertext SELF entnommen. Schwerpunkte im ersten Beispiel sind die interessanten Aspekte der unterschiedlichen Bezugsbereiche. Im zweiten zu untersuchenden Knoten wird hingegen die thematische Progression durch eine Anwendung der Makrostrukturanalyse geklärt.

Beispiel 1

Der unten dargestellte Knoten beschäftigt sich thematisch mit der Definition von Style-Sheets innerhalb einzelner HTML-Elemente. Die Thematik wird durch

¹ Sämtliche Knoten sind in schwarz-weiß abgebildet. Dies ist insofern problematisch, als daß oftmals farbige Markierungen zur Link-Typisierung verwendet werden. Wo nötig, wurden diese graphischen Besonderheiten ‚in Worte gefaßt‘. Um sich jedoch ein besseres Bild zu machen, sei dem geneigten Leser der Besuch der zitierten WWW-Knoten empfohlen.

² Alle Corpus-Hypertexte mit Ausnahme von W3C – hier kommen keine typisierten Links zur Anwendung – werden praktisch analysiert.

normalen Text, ein Quellcode- sowie ein Ansichtsbispiel erläutert und visualisiert:



Abbildung 22: SELF [tdbc.htm] Abschnitt *Schnellformatierung im Text*

Dem Aufbau von TAH folgend, werden die Links des obigen Knotens zuerst unter dem Aspekt des Bezugsbereichs analysiert.

BEZUGSBEREICH:

Die Bezugsbereiche sind hierbei recht unterschiedlich:

Die Hyper-Links der Navigationsleiste verweisen auf hierarchisch höher stehende Inhaltsknoten: Hier wird also der gesamte aktuelle Knoten mit dem Ziel-Knoten in Bezug gesetzt.

Die Links des ‚Knoten-Inhaltsverzeichnis‘ mit den Link-Texten *Einzelne HTML-Tags formatieren* bzw. *Bereiche mit formatieren* haben als Zielbezug hingegen jeweils Teilbereiche im aktuellen Knoten. Dies ist jedoch nur das Ergebnis der hier erfolgten subjektiven Analyse und nicht etwa durch formale Auszeichnung erkennbar, denn: Auf Ebene des Quellcodes von HTML ist der Zielbereich eines Links nicht exakt zu definieren. Inhaltlich kann jedoch interpretiert werden, daß der Link *Einzelne HTML-Tags formatieren* eine Verknüpfung mit dem gleichnamigen ‚Abschnitt‘ des aktuellen Knotens etabliert.

Die Links im Fließtext, *Absatztypen und Textgestaltung* bzw. *Tabellen*, verknüpfen ebenfalls nur einen kleinen Teil des aktuellen Knotens, nämlich das

jeweilige Konzept, das mit dem Link-Text ausgedrückt wird, mit dem Ziel-Knoten.

Der Link *Anzeigebeispiel: So sieht's aus* verknüpft hingegen nicht nur das im Link-Text ausgedrückte Konzept mit einem bestimmten Bezugsbereich des Ziel-Knotens: Eine naheliegende Interpretation aufgrund der Link-Position wäre, als Bezugsbereich im aktuellen Knoten den ‚Abschnitt‘ *Einzelne HTML-Tags formatieren* anzusetzen. Daß die Verhältnisse jedoch anders sind, ergibt die kontrastive Untersuchung desselben Knotens in SELF8 (siehe Beispiel 2 von Kapitel 17.3.2).

Zum Abschluß der Betrachtung des Bezugsbereiches müssen noch ein paar Worte zur Deixis gesagt werden: Bis auf den Link *Anzeigebeispiel* sind sämtliche Links mit aus gedruckten Texten bekannten deiktischen Mitteln zu vergleichen und ließen sich etwa mit *mehr Information finden Sie dort* paraphrasieren. Darüber hinaus spezifizieren Piktogramme die Links als deiktische Ausdrücke noch näher. So wird etwa zwischen Intra- und Inter-Knoten-Verweisen unterschieden. Außerdem weisen die geöffneten Ordner-Piktogramme in der Navigationsleiste darauf hin, in welchem ‚Sinnbereich‘ (oder strukturell: in welchem Teil-Hypertext) von SELF sich der Leser des aktuellen Knotens befindet.¹

Die geschlossenen Ordner der Links *Absatztypen und Textgestaltung* bzw. *Tabellen* symbolisieren hingegen, daß sich die jeweiligen Zielknoten in anderen Sinnbereichen befinden. Die Link-Struktur kann unter diesen Gesichtspunkten auch als ‚Zeigefeld‘ interpretiert werden, in dem durch Piktogramme auf inhaltliche (bzw. auch strukturelle) Nähe und Ferne der Zielbereiche der deiktischen Ausdrücke hingewiesen wird.

Der Link *Anzeigebeispiel* ist hingegen unter Gesichtspunkten der Deixis anders zu deuten: Piktogramm und Link-Text typisieren das Link-Ziel in inhaltlicher Hinsicht. Auch dieser Link ist freilich mit einem deiktischen Ausdruck zu vergleichen, da auch er auf andere Zeichen im Hypertext ‚hinweist‘. Aber: Hier steht nicht die bloße Existenz eines möglicherweise interessanten Ziel-Knotens im Vordergrund, sondern eindeutig dessen inhaltliche bzw. mediale Gestalt. Dies unterscheidet den Link von den anderen besprochenen.

¹ Mehrere übergeordnete Ebenen werden hierbei durch Einrückung von links symbolisiert.

Im nächsten Schritt der TAH-Analyse gilt es, die kommunikativ-funktionalen Aspekte der einzelnen Links zu ermitteln.

KOMMUNIKATIV-FUNKTIONALE ASPEKTE:

Der Link-Text *Anzeigebeispiel: So sieht's aus* hat eine informative Funktion: Er lenkt die Aufmerksamkeit des Lesers vom (allgemein) informativen Knoten (bzw. Abschnitt) zum (speziell) informativen Knoten des Anschauungsbeispiels. Dies demonstriert sehr schön die Verwendung multimedialer Elemente zur Ausbreitung eines Themas (bzw. zum Konzeptaufbau im kognitiven System des Lesers): Mit sprachlichen Mitteln wird ein Thema allgemein eingeführt. Mit Hilfe von Beispielen wird dieses konkretisiert: Zur Erläuterung dienen ein Quellcode-Beispiel sowie eine Abbildung.¹

Die informativen Einheiten – Fließtext, Quellcode und Anzeigebeispiel – sind insgesamt jedoch als größere Einheit mit instruktiver Funktion aufzufassen. Der Abschnitt ist als ‚Miniaturlehrgang‘ zur Vermittlung des in der Abschnittsüberschrift signalisierten Themas *Einzelne HTML-Tags formatieren* aufzufassen. Insofern ist die Funktion des Links *Anzeigebeispiel: So sieht's aus* nicht als rein informativ zu bezeichnen. Vielmehr liefert er Evidenz und Anschaulichkeit für das im Fließtext Beschriebene bzw. im Quelltext Codierte und ist somit Glied einer instruktiven Sequenz der oben aufgezählten informativen Einheiten.

Bei den übrigen Links des Knotens sind die Verhältnisse anders. Sie informieren darüber, an welcher Stelle die im Link-Text angesprochene Information zu finden ist – sind damit, wie oben bereits erwähnt, reine Mittel der Textdeixis. Es gilt, zwischen der Funktion, die ein Link aufgrund seines bloßen ‚Link-Seins‘ hat, und der eigentlichen inhaltlichen Funktion zu unterscheiden. Die Links der Navigationsleiste sind somit als rein ‚deiktisch-informativ‘ aufzufassen, während der Link *Anzeigebeispiel: So sieht's aus* aufgrund seines Link-Seins ebenfalls deiktische Kraft besitzt, darüber hinaus aber eine inhaltlich-informative Funktion ausübt, die er in den Dienst der oben beschriebenen instruktiven Sequenz stellt.

¹ Die Beispiele sind hierbei aus zwei anderen semiotischen Systemen entnommen worden, die direkt aufeinander bezogen sind. Mit Hilfe des symbolischen Systems, der Markup-Sprache HTML – ergänzt um die Style-Sheet-Sprache CSS, wird via Transformation der Browser-Software eine HTML-Seite erzeugt. Interessant (und für Entwickler von Web-Seiten zum Teil ernüchternd) ist hierbei die Tatsache, daß je nach Browser (in Details) unterschiedliche Abbilder des Quellcodes im Browser-Fenster angezeigt werden.

Nach Klärung der Funktion wird nun auf die Struktur eingegangen. Im Vordergrund stehen hierbei die thematisch-strukturellen Aspekte, die als Bedingung für die im nächsten Abschnitt abzuhandelnden grammatisch-strukturellen zu sehen sind.

THEMATISCH-STRUKTURELLE ASPEKTE:

Der Link *Anzeigebeispiel: So sieht's aus* ist in einer deskriptiven thematischen Entfaltungssequenz zu sehen. Die Abschnittsüberschrift *Einzelne HTML-Tags formatieren* führt hierbei das Thema ein. Der Link selbst macht nun folgendes ‚Angebot‘ an den Leser: Wenn dieser wissen will, wie *es* (aus der verkürzten Form *sieht's*) – das eine Wiederaufnahme des Themas (*Einzelne HTML-Tags formatieren*) ist – sich graphisch darstellt, muß er auf diesen Link klicken. Im Ziel-Knoten werden dann Personen- bzw. Gegenstandsreferenz (*Einzelne HTML-Tags*) sowie Ereignisreferenz beibehalten, allerdings graphisch, also in einem anderen semiotischen System, dargestellt. Navigiert der Leser von dieser graphischen Ansicht zum Ausgangsknoten zurück und nimmt dort die Lektüre nach dem eben benutzten Link wieder auf, so bleibt die Referenzstruktur beibehalten (*Sie können einzelne HTML-Tags innerhalb des HTML-Dateikörpers mit Hilfe von Style-Sheet-Angaben formatieren.*). Vereinfacht gesagt: Der Link – samt Inhalt des Ziel-Knotens – ist in die thematische Entfaltung des aktuellen Knotens eingepaßt. Der Link bietet also lediglich eine ‚thematische Schleife‘ an, nach dessen Durchlaufen der Leser mit zusätzlichem Wissen ausgestattet ist, daß ihn beim Kohärenzaufbau innerhalb der aktuellen Themenprogression unterstützen kann. Insofern unterstützt diese Art der Themenentfaltung die oben erwähnte instruktive Gesamtfunktion des Abschnitts.

Bei den Links *Absatztypen und Textgestaltung* bzw. *Tabellen* sind die Verhältnisse anders: Die Gegenstandsreferenzen *Formatierungsangaben*¹ bzw. *HTML-Tags* werden in der folgenden Satzsequenz beibehalten:

¹ *Formatierungsangaben* wird im zweiten Satz implizit beibehalten, und zwar mit dem Platzhalter *Das*, der die gesamte Proposition *gelten_für(Formatierungsangaben, Geltungsbereich)* wiederaufnimmt.

Die Formatierungsangaben gelten dann genau für den Geltungsbereich des betreffenden HTML-Tags. Das ist zum Beispiel für HTML-Tags zu Absatztypen und Textgestaltung oder für solche zu Tabellen sinnvoll.

(SELF [tdbc.htm] Abschnitt *Schnellformatierung im Text*)

Hierbei ist die referentielle Beibehaltung von *Formatierungsangaben* als stärker zu bewerten, als die von *HTML-Tags*. Dies gilt, weil mit ersterer die Gegenstandsreferenz der Hauptüberschrift des Knotens (*Einzelne HTML-Tags formatieren*) beibehalten wird. Die HTML-Tags, die für die Auszeichnung von Absätzen und Tabellen nötig sind, stellen ontologisch eine Untermenge aller HTML-Tags dar. Die hier aufgeführten Links verweisen also auf einen Unterbereich des Gegenstandes, der zwar referenziell auch in den folgenden Sätzen weitergeführt wird, aber bzgl. der gesamten Referenzstruktur des aktuellen Knotens als untergeordnet angesehen werden kann. Auch aus diesem Grund werden die in den Links *Absatztypen und Textgestaltung* bzw. *Tabellen* thematisierten Bereiche nicht im aktuellen Knoten entfaltet, sondern sind in externe Knoten ausgelagert. Die referentielle Bewegung kann man hierbei als „vage Verknüpfung“ charakterisieren (siehe auch Vater 1994 S. 147). An den Leser wird folgendes Angebot gemacht: Die Links signalisieren den Beginn einer für den aktuellen Knoten nebensächlichen Themenentfaltung.¹ Möchte der Leser diese dennoch verfolgen, kann er die Links dafür benutzen. Dies ist der Unterschied zum oben besprochenen Link des Anzeigebeispiels, dessen Ziel-Knoten thematisch (als Schleife) in die Entwicklung des Hauptthemas eingebunden ist.

Im Folgenden werden die strukturellen Aspekte aus grammatischer Sicht betrachtet.

GRAMMATISCH-STRUKTURELLE ASPEKTE:

Der Link *Anzeigebeispiel: So sieht's aus* ist durch Wiederaufnahme mittels der Anapher *es (sieht's)* mit dem Antezedenz *Einzelne HTML-Tags formatieren* verbunden und somit in die kohäsive Struktur des Knotens eingebunden. Der Link-Text ist ein vollständiger Satz.

Die Link-Texte *Absatztypen und Textgestaltung* bzw. *Tabellen* sind als Konstituenten in folgenden Satz eingebunden:

¹ Hierfür steht auch die Auszeichnung mittels der Ordner-Piktogramme.

Das ist zum Beispiel für HTML-Tags zu Absatztypen und Textgestaltung oder für solche zu Tabellen sinnvoll.

Die Link-Texte fungieren grammatisch-strukturell als Attribute zum Träger der Gegenstandsreferenz *HTML-Tags* (siehe oben). In den Ziel-Knoten werden die jeweiligen Link-Texte als Knoten-Überschriften direkt wiederaufgenommen.

Eine interessante Nebenbeobachtung ist hierbei, daß die Links, die thematisch-strukturell nicht so eng mit dem aktuellen Knoten verknüpft sind, grammatisch-strukturell jedoch enger in den Fließtext eingebunden sind (z.B. *Absatztypen und Textgestaltung* sowie *Tabellen*). Der thematisch-strukturell fester eingebundene Link *Anzeigebeispiel: So sieht's aus* ist grammatisch-strukturell loser verknüpft. Sämtliche Links der Navigationsleiste sind hingegen sowohl thematisch- als auch grammatisch-strukturell eher lose mit dem aktuellen Knoten verbunden. Als letzte Teilaspekte dieses ersten Beispiels werden im folgenden kurz die paratextuellen, metasprachlichen und browser-spezifischen Aspekte angesprochen.

PARATEXTUELLE und METASPRACHLICHE ASPEKTE:

Alle in SELF benutzten Piktogramme wurden bereits ausführlich in Kapitel 16.3.3. vorgestellt. Im aktuellen Beispiel kommen die Piktogramme Nr. 1 bis 5 der Tabelle 30 zur Anwendung, die allesamt hypertext-interne Links typisieren. Sämtliche Links, bis auf *Anzeigebeispiel* dienen primär der Navigation. Deren Piktogramme wurden weiter oben bereits so charakterisiert, daß sie Aussagen über die inhaltliche Nähe bzw. Ferne des Link-Ziels machen.

Als sonstige metasprachliche Auszeichnung ist lediglich der Mouse-Over-Text *Beispiel* des Links *Anzeigebeispiel* vorhanden.

BROWSER-SPEZIFISCHE ASPEKTE:

Sämtliche Links des obigen Beispiels werden im selben Fenster der Browser-Software angezeigt, das heißt, der aktuelle Knoten wird durch den Ziel-Knoten ersetzt. Das Browser-Verhalten gilt für alle Links von SELF – auch für die extra-hypertextuellen.¹ Aufgrund der Eindeutigkeit des Browser-Verhaltens wird dieses auch nicht metasprachlich in irgendeiner Weise signalisiert.

¹ Die einzige Ausnahme sind die Links vom Typ *Email*, die im zweiten Beispiel näher besprochen werden.

SELF8 erfuhr im Bezug auf browser-spezifische Aspekte Veränderungen (siehe Kapitel 17.3.2).

Beispiel 2

Im zweiten Beispiel des Textes SELF wird exemplarisch auf die thematische Progression von Knoten eingegangen, die durch Links der Navigationsleiste miteinander verbunden sind.

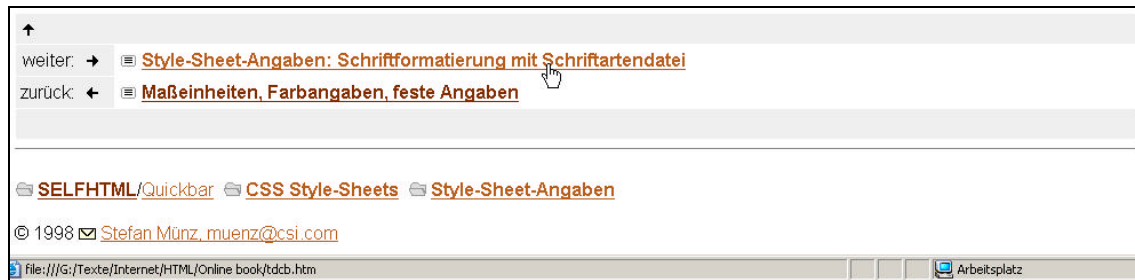


Abbildung 23: SELF [tdca.htm] Abschnitt *Textschatten (text-shadow)*

BEZUGSBEREICH:

Eine detaillierte Untersuchung der Bezugsbereiche unterschiedlicher Link-Typen wurde bereits in Beispiel 1 durchgeführt. Die im aktuellen Beispiel aufgeführten Links dienen im weitesten Sinne der Navigation in der Hierarchie des Hypertextes und verknüpfen jeweils den gesamten aktuellen Knoten mit dem gesamten Ziel-Knoten. Eine Ausnahme ist der Link auf die Email-Adresse (*Stefan Münz muenz@csi.com*). Hier wird nur ein Teilbereich des Knotens – der Link-Text – mit dem externen Programmaufruf eines Email-Programms verknüpft. Dieses Phänomen läßt sich jedoch passender unter dem Aspekt der browser-spezifischen Umgangsweise mit Links auf externe Anwendungen beschreiben (siehe unten).

KOMMUNIKATIV-FUNKTIONALE ASPEKTE:

Hier soll lediglich kurz auf die beiden Links eingegangen werden, die mit *weiter* und *zurück* typisiert sind. Durch diese Links werden zwei Knoten mit jeweils informativer Funktion als Teile einer möglichen Lesesequenz miteinander verknüpft. Welche Funktion genau der Ziel-Knoten bzgl. des aktuellen Knotens einnimmt, wird nicht spezifiziert – die informative Funktion der Links beschränkt sich auf den bloßen Hinweis, daß ein Vorgänger- bzw. Folge-Knoten existiert, deren Thematik durch den Link-Text visualisiert wird und mit der Hauptüberschrift des jeweiligen Ziel-Knotens identisch ist. Aussagen zur Relation selbst werden jedoch nicht gemacht.

THEMATISCH- und GRAMMATISCH-STRUKTURELLE ASPEKTE:

Funktion der Links *weiter* und *zurück* ist die, auf das mögliche Fortschreiten (bzw. Zurückschreiten) innerhalb einer vom Autor definierten Lesefolge hinzuweisen. Es wird also direkte thematische Progression zweier ‚gleichgeordneter‘ Knoten signalisiert. Wirft man aber einen genaueren Blick auf den aktuellen Knoten sowie den mittels des Links *weiter* verknüpften (SELF [tdcb.htm]), so stellt man folgendes fest: Hierarchisch sind die beiden Knoten gleichgeordnet – dies ist auch im Übersichts-Knoten SELF [tdc.htm] deutlich visualisiert. Durch eine Makrostrukturanalyse gelangt man durch geeignetes Anwenden der Makroregeln (siehe Kapitel 12) zu folgenden Makropropositionen: *informieren_über(Textlayout mit CSS)* für den Ausgangsknoten sowie *informieren_über(Externes Speichern von Textlayout mit CSS)* für den Ziel-Knoten. Durch die Makroregel *Generalisieren* kann die spezifischere Makroproposition des Ziel-Knotens in die des aktuellen Knotens integriert werden. Thematisch stellt also der Ziel-Knoten (SELF [tdcb.htm]) eine Spezifizierung bzw. eine Zusatzinformation zum aktuellen Knoten (SELF [tdca.htm]) dar.¹ Durch die Link-Typisierung (aber auch die hierarchische Struktur) von SELF wird dieser Sachverhalt jedoch nicht berücksichtigt.

PARATEXTUELLE und METASPRACHLICHE ASPEKTE:

Die Typen der Links *weiter* und *zurück* werden auf zweifache Weise typisiert: Einmal mit den Begriffen *weiter* und *zurück*, zum anderen durch die Verwendung von Piktogrammen. Anders gesagt: Die Bedeutung der Piktogramme wird durch den davorstehenden Text direkt ausgedrückt. In SELF8 werden diese textuellen Erläuterung im Knoten weggelassen, bzw. als Mouse-Over-Text angeboten.

BROWSER-SPEZIFISCHE ASPEKTE:

Betrachtet man die browser-spezifischen Aspekte, so ist besonders der Link auf die Email-Adresse des Autors interessant (*Stefan Münz muenz@csi.com*): Während sich alle anderen Links von SELF im selben Browser-Fenster öffnen, aktiviert ein Klick auf besagten Email-Link ein externes Programm. Welches Email-Programm dies ist, hängt, wie gesagt, von den Einstellungen des Browsers ab. Insofern könnte man Links, deren Ausführung externe Programme

¹ Zum Teil wird dies auch durch die Hauptüberschriften der beiden Knoten signalisiert (*Style-Sheet-Angaben: Schriftformatierung* bzw. *Style-Sheet-Angaben: Schriftformatierung mit Schriftartendatei*).

öffnet, auch unter Aspekten des anwenderspezifischen Peritextes besprechen. Unabhängig von der Frage nach dem auf dem Rechner des Lesers tatsächlich installierten Email-Programm besteht jedoch das grundsätzliche Problem, daß beim Klick auf einen solchen Link nicht nur der aktuelle Hypertext bzw. das aktuelle Browser-Fenster verlassen, sondern gar zu einem anderen, externen Programm gewechselt wird. Je nach Komplexität des Hypertextes bzw. der leserseitigen Installation sind zahlreiche Links auf solche externen Programme denkbar:¹ Videofilme müssen beispielsweise nicht unbedingt im Browser angezeigt, sondern können auch von externen, funktionsmächtigeren Programmen abgespielt werden. Dasselbe gilt für Tondokumente. Auch können Links auf FTP-Server mit externen FTP-Programmen geöffnet werden. Ob und in welchem Umfang solche Links auf externe Programme typisiert werden sollten, läßt sich wohl nur durch empirische Vergleichsuntersuchungen herausfinden. Eventuell ließe sich aber eine theoretische Unterscheidung danach treffen, ob Links auf externe Programme mit einem Wechsel der Kommunikationsform einhergehen. Bei Links auf Email-Adressen (aber auch zu Chat-Räumen) wäre dies der Fall, bei solchen auf FTP-Sites o.ä. jedoch nicht.

17.3.2 Analyse der proprietären Codierung von SELF8

Bei der Analyse von SELF8 sind besonders die Veränderungen im Vergleich zu SELF interessant.

Beispiel 1

In SELF8 sind nicht nur extra- und intra-hypertextuelle Links mit speziellen Piktogrammen typisiert sondern auch solche Links, die auf Ziele innerhalb des Hypertext-Netzes *SELF-Raum* verweisen.

- | |
|---|
| <ul style="list-style-type: none"> • Die Beschreibungen zu HTML, CSS, JavaScript-Sprachelementen und Dynamischem HTML wurden mehr oder weniger stark aktualisiert. • SELFHTML hat nunmehr eine engere Anbindung an das Online-Angebot von SELFHTML aktuell. In vielen Kapiteln und auf vielen einzelnen Seiten gibt es thematische Verweise zu dem Aktuell-Angebot, das in den letzten Jahren seit der Version 7.0 gewaltig angewachsen ist und längst das "zweite Standbein" von SELFHTML ist. |
|---|

Abbildung 24: SELF8 [editorial\hinweise.htm] Abschnitt *Neu in Version 8.0*

Der Link-Text *SELFHTML aktuell* verweist auf den gleichnamigen externen Hypertext, der sich jedoch innerhalb des SELF-Raumes (siehe Kapitel 16.3.3) befindet und durch dasselbe Layout wie SELF8 gekennzeichnet ist. Die weiteren

¹ Eine Piktogramm-Kollektion, die diese Verhältnisse nachzubilden versucht, ist *Qbullets* (siehe Kapitel 16.3.3).

Links verweisen auf interne Übersichts-Knoten von SELF8. Diese Unterscheidung ist durch die unterschiedliche Farbe des Ordner-Icons kenntlich gemacht – welche in der obigen Schwarz-Weiß-Abbildung leider nicht so gut zu erkennen ist.

Die meisten Leser von SELF8 haben sich – aus Kosten- und Geschwindigkeitsgründen – den Hypertext zum ‚Offline-Lesen‘ auf die lokale Festplatte gespeichert. Insofern symbolisieren die andersfarbigen Piktogramme der Links zu *SelfHTML aktuell* auch die physikalische Grenze zwischen lokalen Dateien und solchen im WWW. Im folgenden werden in knapper Form weitere Beschreibungsaspekte des TAH auf der Ebene *Link* angesprochen.

KOMMUNIKATIV-FUNKTIONALE ASPEKTE:

Die Funktion des Abschnitts *Neu in Version 8.0* des Knotens SELF8 [editorial/hinweis.htm] ist eine informative, man könnte auch sagen, eine meta-informative oder paratextuelle. Der gesamte Knoten SELF8 [editorial] samt seiner hierarchisch tiefer liegenden Knoten ist mit dem Editorial von Printmedien nahezu identisch: Hier wird zusätzliche Information über die zu erwartende Information – bzw. über die Vorgehensweise der zu erwartenden Instruktion – in den eigentlichen Inhaltsknoten gegeben. Die Funktionen dieser möglichen Zielknoten sind informative (bzw. instruktive) – aber eben spezifischerer inhaltlicher Natur. Diese kommunikativ-funktionale Unterscheidung wird durch das verwendete Ordner-Piktogramm in dieser Form nicht signalisiert.

THEMATISCH-STRUKTURELLE und GRAMMATISCH-STRUKTURELLE ASPEKTE:

Das Textthema des Abschnitts wird durch die Überschrift *Neu in Version 8.0* klar signalisiert. Die Entfaltung des Themas geschieht deskriptiv, die Modalität der Themenbehandlung ist sachbetont. Die Texte der oben aufgeführten Links werden allesamt als Überschrift im jeweiligen Zielknoten explizit wiederaufgenommen.

PARATEXTUELLE und METASPRACHLICHE ASPEKTE:

Die vom Ordner-Piktogramm symbolisierte Link-Funktion kann etwa für den Link-Text *HTML* mit folgender Paraphrase wiedergeben werden: *Per Klick kommen Sie auf eine Übersichtsseite für die Thematik ‚HTML‘*. Die Abgrenzung von Übersichtsseiten einerseits und Inhaltsseiten andererseits stand bei der Auswahl des Piktogramms im Vordergrund. Die eigentlich kommunikative Funktion, die oben beschrieben wurde, wird durch das Ordner-Icon nicht

wiedergegeben, da dieses reine Strukturinformation des Hypertextes ‚abbildet‘. Aber auch diese Angaben zur Struktur können als etwas ‚schief‘ bezeichnet werden: Während nämlich das Ordner-Piktogramm in den meisten Fällen auf neue ‚Sinnabschnitte‘ verweist (im obigen Beispiel die Links *HTML*, *CSS*, *JavaScript-Sprachelemente* und *Dynamischem HTML*), verweist der Link *SelfHTML aktuell* auf einen kompletten Hypertext, der sich im selben Hypertext-Netz wie der Hypertext des aktuellen Knotens befindet.

BROWSER-SPEZIFISCHE ASPEKTE:

Die vom Ausgangs-Knoten zu erreichenden Ziel-Knoten gliedern sich strukturell und funktional in den Hypertext SELF8 bzw. in das Hypertext-Netz des SELF-Raumes (Link: *SELFHTML aktuell*) ein. Die Ziel-Knoten sind eigenständige informative Einheiten (also nicht etwa Knoten, die eine Definition des Link-Textes erhalten – was beim obigen Beispiel durchaus denkbar wäre), daher werden sie auch im selben Fenster wie der aktuelle Knoten angezeigt. Im nächsten Beispiel ist dies anders.

Beispiel 2

Der folgende Knoten ist eine neue Version des ersten Beispiels aus Kapitel 17.3.1:



Abbildung 25: SELF8 [css\formate\direkt.htm] Abschnitt *Formatdefinitionen*

BEZUGSBEREICH:

Besonders interessant ist hier die kontrastive Betrachtung des Bezugsbereichs des Links *Anzeigebeispiel: So sieht's aus*: In SELF ist dieser Link direkt nach der Überschrift des Abschnitts plaziert. Damit wird ein Bezugsbereich des Anzeigebeispiels für diesen ganzen Abschnitt signalisiert. Im aktuellen Beispiel ist der Link anders angeordnet: Der Überschrift folgen einige einführende Sätze

zur Thematik. Der nächste logisch-rhetorische Abschnitt ist eine Demonstration des eben Gesagten, dies wird durch den fettgedruckten ‚Abschnittsmarkierer‘ *Beispiel*: ausgedrückt. Die Demonstration wiederum ist in ein Anschauungs- sowie ein dazu passendes Quellcode-Beispiel unterteilt. Der Quellcode wird im selben Knoten dargestellt, das Bildbeispiel hingegen in einem externen Knoten.

Das Anschauungsbeispiel bezieht sich also in erster Linie auf das Quellcode-Beispiel, da es dieses visualisiert. Dies ist im Prinzip auch beim weiter oben besprochenen Beispiel aus SELF der Fall, dort wird jedoch durch die Link-Plazierung direkt unter der Abschnittsüberschrift ein ‚falscher‘ Bezugsbereich signalisiert: Das Anschauungsbeispiel ist eben keine Visualisierung der ganzen im Abschnitt geschilderten Sachverhalte, sondern lediglich die bildliche Auswirkung der Anwendung des Quellcode-Beispiels innerhalb einer Browser-Software:



Abbildung 26: SELF8 [css\formate\anzeige\styleattr.htm] Abschnitt *Die Seite ...*

Da jedoch das Quellcode-Beispiel eine Exemplifizierung des im Fließtext Gesagten darstellt, und das Bildbeispiel wiederum, wie gesagt, die Visualisierung des Quellcodes ist, wirkt sich der ‚falsche‘ Bezugsbereich wohl kaum drastisch auf die Kohärenzbildung beim Leser aus. Dies ist jedoch lediglich eine Vermutung. Der Komplex der unterschiedlichen Bezugsbereiche von Link, Ausgangs-Knoten und Ziel-Knoten kann, meiner Meinung nach, wertvolle Impulse für das Setting empirischer Untersuchungen geben.

17.3.3 Analyse der proprietären Codierung von GRAMMIS

Als nächster Gegenstand einer TAH-Untersuchung dient ein Knoten des Hypertextes GRAMMIS, der die Wortklasse der Konnektoren zum Thema hat.

systematische grammatik
 terminologisches wörterbuch
 grammatisches wörterbuch
 rechtschreib-wörterbuch
 grammatische bibliografie
 über grammis
 grammis PLUS

Konnektoren sind sprachliche Ausdrücke, die Sätze in eine spezifische semantische Beziehung zueinander sein können. Das sind im Deutschen etwa 300 Ausdrücke, die traditionell als Konjunktionen und als bestimmte Subk von **Adverbien** und **Partikeln** beschrieben werden.

1. [Hyperlink zum terminologischen Wörterbuch](#) schön billig.
2. Das Haus ist schön billig, **also** nehmen wir's.
3. Wir haben im Lotto gewonnen, **sodass** wir das Haus kaufen können.
4. Wir müssen mehr sparen, **sonst** können wir das Haus nicht kaufen.
5. Sie haben nur wenig gespart, **trotzdem** wollen sie ein Haus kaufen.
6. Sie haben nichts gespart **und** keine Bank will ihnen Geld leihen.
7. Man muss viel sparen, **vorausgesetzt**, man will ein Haus kaufen.
8. Sie wollen ein Haus kaufen, **weil** sie etwas Geld gespart haben.

Konnektoren im Text

[\[Text einblenden\]](#)

In der GRAMMIS-Einheit "Konnektoren" werden die syntaktischen und semantischen Gebrauchsbedingungen Konnektoren beschrieben.

Dazu wird zunächst eine [Bestimmung des Gegenstands](#) vorgenommen und die definierenden [Merkmale von Konnektoren](#) werden genannt und erläutert.

Abbildung 27: GRAMMIS [grammis_sys.ansicht] Abschnitt *Konnektoren* (Auswahl *Konnektoren*)¹

Die Vermittlung der Inhalte geschieht mit Fließtext, Abbildungen, Beispielen und Belegen. Besonders interessant sind hier die nichtsprachlichen Mittel, mit denen die Typisierung der Links vorgenommen wird, einige Aspekte der thematischen Entfaltung sowie die browser-spezifischen Aspekte der Belege, die als Stretchtexte integriert sind.

BEZUGSBEREICH:

Die Links der links positionierten farbigen Navigationsleiste haben als Bezugsbereiche den gesamten aktuellen Knoten sowie den gesamten Ziel-Knoten. Die beiden Links zum terminologischen Wörterbuch haben als Bezugspunkte die entsprechenden Link-Texte (*Adverbien*, *Partikel*) sowie den gesamten jeweiligen Ziel-Knoten. Der Link *Text einblenden* ist mit einem Ziel-Knoten verknüpft, der als Belegtext einen Artikel der *Süddeutschen Zeitung* enthält, in dem sämtliche Konnektoren annotiert sind. Der Ziel-Knoten wird per Klick als Stretchtext in den aktuellen Knoten eingebunden (siehe unten). Als Bezugsbereich ist der Abschnitt *Konnektoren im Text* anzusetzen, da dieser auch mit einem gleichnamigen Mouse-Over-Text des Links direkt signalisiert wird.

Die Links *Bestimmung des Gegenstandes* sowie *Merkmale von Konnektoren* haben im aktuellen Knoten den Link-Text als Bezugsbereich. Ein Klick bringt den Leser zu einem Detailtext, in dem näher auf das im Link-Text

¹ Per Pull-Down-Menü im oberen Bereich des Knotens (in der Abbildung nicht sichtbar) können bestimmte Subbereiche des aktuellen Bereichs von GRAMMIS ausgewählt und via JavaScript angesteuert werden. Es handelt sich hier also um eine ergänzende Subnavigation der in der Abbildung sichtbaren Hauptnavigationsleiste.

Angesprochene eingegangen wird. Insofern ist der gesamte Ziel-Knoten als Zielbezugsbereich des Links zu interpretieren.

Dem TAH folgend werden nun die kommunikativ-funktionalen Gesichtspunkte untersucht.

KOMMUNIKATIV-FUNKTIONALER ASPEKT:

Die Funktion des Ausgangs-Knotens ist eine informative. Die lila gefärbten Links *Adverbien* und *Partikeln* verweisen auf Knoten des Teilbereiches *Terminologisches Wörterbuch* von GRAMMIS. Die Funktionen dieser Ziel-Knoten sind ebenfalls informative. In Bezug auf den aktuellen Knoten könnte man auch von ‚stützend informativ‘ sprechen: Die Definitionen des Wörterbuches sollen helfen, die Termini im Fließtext – und damit den Fließtext selbst – des aktuellen Knotens besser zu verstehen.

Die Funktion des Links *Text einblenden* ist hingegen eine andere. Per Klick wird ein Beispieltext als Stretchedtext (siehe unten) angezeigt, in dem Belege für Konnektoren markiert sind. Hier liegt meiner Meinung nach eine Mischung aus mehreren Funktionen vor: informativ (Leser wird informiert, daß der jeweilige Beleg im Text ein Beispiel für einen Konnektor ist.) und – als Nebenfunktion – appellativ (Leser soll zu der Erkenntnis gebracht werden, daß in einem normalen Text sehr viele Konnektoren vorkommen und das Thema daher sehr wichtig ist.) Nach dem TOULMINSCHEN RHETORIK-MODELL (Toulmin 1975) oder dem AUTHOR'S ARGUMENTATION ASSISTANT (siehe Kapitel 16.3.2) könnte man hierfür auch die Funktion *Stützung* ansetzen. Diese Ansicht ist freilich diskutierbar, da die thematische Entfaltung (siehe unten) natürlich keine primär argumentative sondern eine deskriptive ist.

Die Links *Bestimmung des Gegenstands* und *Merkmale von Konnektoren* sind blau eingefärbt. Damit wird signalisiert, daß der Ziel-Knoten innerhalb des Bereiches der *Systematischen Grammatik* liegt. Die Funktionen der Ziel-Knoten sind ebenfalls informative. Als Mouse-Over-Text wird die zusätzliche Information *Hyperlink zu einem Detailtext* eingeblendet – also Strukturinformation geliefert. Hierzu muß man wissen, daß für die *Systematische Grammatik* drei Knoten-Arten definiert wurden: *Detailtexte*, *Kompakttexte* und *Vertiefungstexte* (siehe Kapitel 17.2.4).

Im nächsten Abschnitt werden strukturelle Aspekte thematischer und grammatischer Natur besprochen.

THEMATISCH-STRUKTURELLE und GRAMMATISCH-STRUKTURELLE ASPEKTE:

Die knoten-übergreifende Themenentfaltung bei Links, deren Ziel-Knoten im Bereich *Terminologisches Wörterbuch* von GRAMMIS liegen, ist eine deskriptive. Exemplarisch sei hierfür folgender Lesepfad konstruiert:

Das sind im Deutschen etwa 300 Ausdrücke, die traditionell als Konjunktionen und als bestimmte Subklassen von Adverbien ... ||

Adverb

Thematisch verwandte Bezeichnungen: Umstandswort

Zur Wortart Adverb gehören unflektierbare Ausdrücke, deren prototypische Funktion darin besteht,

Prädikate unterschiedlicher Komplexität zu modifizieren... ||

... und Partikeln beschrieben werden.

(Lesepfad: GRAMMIS [grammis_sys.ansicht?v_typ=d&v_id=1182&v_pos=konntext],
GRAMMIS [grammis/grammis_term.eintrag?v_id=140],
GRAMMIS [grammis_sys.ansicht?v_typ=d&v_id=1182&v_pos=konntext])

Die deskriptive Themenentfaltung wird durch einen Klick auf den Link *Adverbien* nicht unterbrochen sondern biegt lediglich in eine Schleife ein, in deren Verlauf die Bedeutung des Terminus' *Adverb* erklärt wird. Hat der Leser sein Informationsbedürfnis befriedigt und ist zum ursprünglichen Knoten zurückgekehrt, wird die Themenentfaltung bruchlos fortgesetzt. Als kohäsives Mittel dient die Wiederholungsstruktur *Adverbien // Adverb*.

Beim Hyper-Link *Text einblenden* sind die Verhältnisse anders, ein Klick bringt hier, wie gesagt, einen Belegtext. Es liegt, je nach Auslegung, eine eher argumentative oder eher deskriptive Themenentfaltung vor.

Nach der deskriptiven Variante handelt es sich beim Inhalt des Ziel-Knotens um eine kohäsiv äußerst lose angebundene thematische Weiterführung. Das Thema *Konnektoren* wird dadurch weiter entfaltet, daß ein Corpustext gezeigt wird, in dem sämtliche Konnektoren graphisch annotiert dargestellt sind. Nimmt man hingegen eine argumentative Variante an, so ist das demonstrative Anzeigen einer großen Menge Konnektoren in einem relativ kurzen Text auch als Stützung für das Argument zu sehen, daß die Beschäftigung mit Konnektoren eine hohe Relevanz für den Leser hat.

Meiner Meinung nach handelt es sich jedoch um eine Fortführung der deskriptiven Entfaltung. Ich sehe die objektsprachlichen Belege als vergleichbar mit graphischen und formalen Beispielen der Hypertexte SELF und SELF8 an. Insofern biegt auch hier die thematische Entfaltung in eine ‚Schleife‘ ein, nach deren Durchlaufen sie wieder im aktuellen Knoten aufgenommen wird.

Der Systematik von TAH folgend, schließt sich nun eine Analyse unter paratextuellen und metasprachlichen Gesichtspunkten an.

PARATEXTUELLE und METASPRACHLICHE ASPEKTE:

In GRAMMIS sind sämtliche Hyper-Links farbig gekennzeichnet. Jede Farbe entspricht einem der sieben Teilbereiche des Gesamt-Hypertextes. Diese Bereiche sind, einer Hauptnavigation gleich, stets auf der linken Seite des Bildschirms eingeblendet, der Leser kann sich also mit einem kurzen Blick informieren, welche Link-Farbe zu welchem Bereich gehört (siehe auch Kapitel 17.2.4).

Darüber hinaus wird (jedenfalls teilweise) ein Mouse-Over-Text mit dem Hinweis auf den Bereich angezeigt, auf den der jeweilige Link verweist.

Abschließend sind noch die browser-spezifischen Aspekte zu besprechen.

Browser-spezifische Aspekte:

Im aktuellen Beispiel sind die drei gängigen Möglichkeiten des Link-Traversals vertreten: Die Links zu weiteren Detailtexten (*Bestimmung des Gegenstands* und *Merkmale von Konnektoren*) werden im selben Fenster dargestellt, die zu Definitionen von Termini (*Adverbien* und *Partikeln*) hingegen in einem Pop-Up-Fenster. Der Link *Text einblenden* wird ebenfalls im selben Fenster angezeigt, ersetzt den aktuellen Knoten aber nicht, sondern ergänzt diesen als Stretchtext mit dem Inhalt des Ziel-Knotens ab der Position des Link-Textes.

Betrachtet man diese browser-spezifischen Aspekte unter Bezugnahme auf thematisch-strukturelle, so läßt sich folgendes feststellen: Die Links zu weiteren Detailtexten stellen Bezüge zu thematisch verwandten Knoten her. Analysiert man die Übergänge mittels des Modells der referentiellen Bewegung (siehe Kapitel 13), kann man für die maßgebenden Referenzdomänen ein Associative Linkage ansetzen. Folgerichtig wird der aktuelle Knoten im Fenster durch den Ziel-Knoten des Links ersetzt. Bei den Links *Adverbien* und *Partikeln* bzw. *Text einblenden* ergibt die Analyse der referentiellen Bewegung hingegen eine Maintenance.

Der Ziel-Knoten ist also dem aktuellen Knoten thematisch ähnlicher. Bei der Themenentfaltung wurde weiter oben bereits festgestellt, daß die letztgenannten Knoten innerhalb der Entfaltungssequenz des aktuellen Knotens zu sehen sind –

somit eine direkte Ergänzung dieses Knotens darstellen. Es erscheint daher sinnvoll, den aktuellen Knoten angezeigt zu lassen, während die zusätzliche Information als (kleines) neues Fenster bzw. als Stretchedtext eingeblendet wird.¹

17.3.4 Analyse der proprietären Codierung von SPORT1

Bei der Analyse des Corpus-Textes SPORT1 stehen funktionale und thematisch-strukturelle Aspekte im Vordergrund. Besonders interessant ist hierbei die funktionale Verflechtung von Information und Appellation, die sich auch auf Kategorien anderer Beschreibungsaspekte auswirkt.

In SPORT1 werden Hyper-Links auf sehr unterschiedliche Arten dargestellt: Neben untypisierten, rein textuell realisierten Links (Links der Navigation, etwa *TV-Programm*, bzw. Links auf ,irgendwie ähnliche Knoten'² am Ende eines Teaser-Textes – im Beispiel der Link-Text *Butch Goring wird neuer Trainer in Frankfurt*) kommen ebenso per Piktogramm typisierte Text-Links (*50. Spieltag in der DEL*) als auch solche Links vor, deren Link-Anker lediglich durch ein alleinstehendes Piktogramm realisiert ist.



Abbildung 28: Piktogramme in SPORT1 [dynamisch erzeugter Knoten]³

¹ Diese Untersuchungen waren ein gutes Beispiel dafür, daß bei allem Bedacht auf eine strenge Differenzierung der einzelnen TAH-Beschreibungsaspekte doch stets auch die übergreifenden Interaktionen berücksichtigt werden müssen.

² Zum Begriff des Knotens im Corpus-Text SPORT1 siehe Kapitel 17.2.4.

³ Die Knoten von SPORT1 werden tagesaktuell dynamisch erzeugt. Insofern ist hier keine genaue Referenzierbarkeit möglich.

BEZUGSBEREICH:

Als erster Aspekt soll kurz der Bezugsbereich der einzelnen Link-Varianten erläutert werden: Die Links der links positionierten Navigationsleiste haben als Bezugsbereich den gesamten aktuellen Knoten sowie den gesamten jeweiligen Ziel-Knoten. Hier werden thematisch-strukturelle Aspekte von Knoten miteinander verknüpft, es geht um mögliche Progressionen innerhalb der Gliederung des Hypertextes.

Links hingegen, die als Piktogramme realisiert sind (also nicht nur mit Piktogrammen typisiert sind), haben im Ausgangs-Knoten, der ja, wie in Kapitel 17.2.4 beschrieben wurde, eher als Platzhalter-Knoten zu charakterisieren ist, als Bezugsbereich den aktuellen Teaser-Text. Im obigen Beispiel etwa, verfügt der Teaser-Text *Kölner Haie setzen Lance Nethery vor die Tür* über einige mit Piktogrammen realisierte Links, die zu ‚irgendwie verwandten‘ Knoten führen (siehe unten). Hier findet also eine Verknüpfung thematisch ähnlicher, kommunikativ-funktional (informativ vs. appellativ) bzw. medial (Interview vs. Video) jedoch verschiedener Knoten statt.

Die Betrachtung der kommunikativ-funktionalen Aspekte ist hier stellenweise sehr aufschlußreich.

KOMMUNIKATIV-FUNKTIONALE ASPEKTE:

In Kapitel 17.2.4 wurde auf die besondere Struktur von SPORT1 hingewiesen: Eine Bildschirmseite beinhaltet einen Platzhalter-Knoten mit einer bestimmten Anzahl von Teil-Knoten. Die redaktionellen Knoten haben hierbei eine informative Funktion, die Werbebanner eine appellative. Aber bereits die Aufteilung der eigentlichen Inhaltsknoten – ein Teilangebot des Inhalts wird auf einer der zahlreichen Übersichtsseiten als Teil-Knoten des Platzhalter-Knotens angeboten, von hier aus kann dann zu den eigentlichen Inhalt-Knotens navigiert werden – unterstützt eine ebenfalls vorhandene appellative Teilfunktion der eher informativen redaktionellen Knoten. Durch diese Informationszerstückelung wird in deutlich höherem Maße als in den anderen Texten des Corpus’ versucht, den Leser innerhalb des Hypertextes zu halten.

Auch die multimediale Aufbereitung eines Inhalts unterstützt indirekt diese Vermischung informativer und appellativer Funktionen: Multimediale Inhalte sind in SPORT1 mit Piktogrammen typisiert. Diese ikonischen Zeichen eines sekundären semiotischen Systems sind die Träger zweier Funktionstypen. Einige Piktogramme stehen für die mediale Realisierung des Zielknotens, welcher ebenso wie der aktuelle Knoten eine überwiegend informative Funktion

hat – die Information wird lediglich mit anderen medialen Mitteln weitergegeben. Andere Links hingegen verweisen von einem Knoten mit informativer Funktion auf einen mit appellativer bzw. obligativer. Folgende Tabelle nennt hierfür einige Beispiele:







| Piktogramm | Bedeutung | Funktion |
|---|---|--|
|  | Ziel-Knoten enthält Video zum Inhalt des aktuellen Knotens. | Information (/ Appell) |
|  | Ziel-Knoten enthält Diaschau zum Inhalt des aktuellen Knotens. | Information |
|  | Ziel-Knoten enthält ein Interview zum Inhalt des aktuellen Knotens. | Information |
|  | Ziel-Knoten enthält (kostenpflichtige,) per Mobiltelefon zugängliche Informationen zum Inhalt des aktuellen Knotens. | Information / Appell |
|  | Ziel-Knoten ermöglicht, eine (kostenpflichtige) Sportwette der Firma <i>oddset</i> abzuschließen, die in Verbindung zum Inhalt des aktuellen Knotens steht. | Appell / (Information) / Obligation ¹ |
|  | Ziel-Knoten ermöglicht, Produkte über das Internet käuflich zu erwerben, die in Verbindung zum Inhalt des aktuellen Knotens stehen. | Appell / (Information) / Obligation |

Tabelle 37: Funktion der Piktogramme in SPORT1

Auch primär informative Piktogramme, etwa die für Videos, sind zum Teil mit einem Mouse-Over-Text versehen, der tendenziell appellativen Charakter hat:



Abbildung 29: Mouse-Over-Text in Sport1 [dynamisch erzeugter Knoten]

Hier wird etwa nicht nur auf einen Video-Knoten hingewiesen, sondern zugleich Werbung für die Fernsehsendung *ran* gemacht (siehe Kapitel 17.2.1).

In einem nächsten Schritt folgt die Strukturanalyse thematischer und grammatischer Aspekte des Beispiels.

THEMATISCH-STRUKTURELLE und GRAMMATISCH-STRUKTURELLE ASPEKTE:

Der thematischen Analyse vorgreifend läßt sich über die Links in SPORT1 unter grammatisch-strukturellen Aspekten folgende Aussage treffen: Eine kohäsive Einbettung des Link-Textes bzw. Link-Ankers in grammatische Strukturen des Ausgangs-Knotens findet kaum statt. Aus diesem Grund lassen sich auch

¹ Der Ziel-Knoten bietet dem Leser die Möglichkeit, eine Sportwette abzuschließen. Geschieht dies, so verpflichtet sich dieser per Online-Vertrag zur Zahlung des Wetteinsatzes. Ähnliches gilt für Links, die auf Seiten des Online-Shops von SPORT1 verweisen.

wenige grammatische Strukturen feststellen, die über Knoten-Grenzen hinweg gehen.

Kohärenz wird lediglich durch thematisch-strukturelle Mittel hergestellt: So folgen den Teaser-Texten zumeist Piktogramm-Links, die auf thematisch ähnliche Knoten verweisen. Die Art dieser Ähnlichkeit zu erschließen, gelingt dem Leser aber oftmals nur mit entsprechendem Vorwissen, Kenntnis über die Struktur von SPORT1 und damit verbundenem Ziehen von Inferenzen. So muß der Leser im obigen Beispiel wissen, daß es sich bei *oddset*¹ um einen Vertragspartner von SPORT1 handelt. Ist dies nicht der Fall, dürfte es wohl zu Kohärenzverlust führen, wenn sich der Leser plötzlich auf einem Knoten mit Sport-Wetten wiederfindet, nachdem er einen Link benutzte, der bzgl. Platzierung und Screendesign einen für den Teaser-Text relevanten Knoten mit informativer Funktion zu verknüpfen schien.

Im folgenden werden nun kurz die unterschiedlichen Ausformungen der referentiellen Bewegungen vom Teaser-Text hin zu den einzelnen Links bzw. deren Ziel-Knoten beschrieben.

Eine Analyse der referentiellen Bewegung des Beispiels in Abbildung 28 ergibt folgendes: Im Fließtext des Teaser *Kölner Haie setzen Lance Nethery vor die Tür* wird die dominante Personenreferenz *Lance Nethery* permanent wieder aufgenommen (*Cheftrainer, der Kanadier, geschasster Coach*). Dies gilt zum Teil auch für die Ereignisreferenz. Am Ende des Fließtextes verweist ein Pfeil-Piktogramm auf den Knoten, in dem der ergänzende Text zum Teaser steht. Dieses Piktogramm kann thematisch-strukturell bzgl. der referentiellen Bewegung als Beibehaltung der Referenzen des Teaser-Textes betrachtet werden. Im Ziel-Knoten wird sowohl die Personen- als auch die Ereignisreferenz wiederaufgenommen.

Anders verhält es sich mit dem textuellen Link *Butch Goring wird neuer Trainer in Frankfurt* sowie den darauf folgenden Piktogramm-Links. Die Personenreferenz wird von *Nance Lethery* zu *Butch Goring* verschoben. Nach dem Modell der referentiellen Bewegung kann hier von einer Assoziativen Linkage gesprochen werden. Kognitiv zu verarbeiten ist diese Verbindung jedoch nur durch geeignete Inferenzen: Der Leser muß wissen, daß Butch Goring Eishockeytrainer ist und Frankfurt über ein Eishockeyteam verfügt.

¹ Dieser Text erscheint als Mouse-Over-Text bei den Piktogrammen *Sportwette*.

Betrachtet man die referentielle Bewegung vom Fließtext hin zu den mit Piktogrammen realisierten Links darunter, so läßt sich feststellen, daß hier wohl nicht mehr von einem Associative Linkage zu sprechen ist. Im obigen Beispiel wird keine der Personen- bzw. Ereignisreferenzen durch die Piktogramme bzw. den zugehörigen Mouse-Over-Text wiederaufgenommen. Mouse-Over-Texte wie *Im Shop: Trikots der DEL* ermöglichen lediglich dem in der Domäne *Eishockey* erfahrenen Leser durch Inferenzziehen den Aufbau eines kohärenten Lesepfades – aber genau für diese erfahrenen, interessierten (und eventuell kaufwilligen) Leser sind ja auch die Ziel-Knoten mit Appell- und Obligationsfunktion gedacht. Thematisch passen diese Ziel-Knoten aber definitiv nicht in die deklarative Entfaltung, die im Teaser-Text begonnen und im Ziel-Knoten des mit dem Pfeil-Piktogramm typisierten Links fortgesetzt wird (siehe oben).

Im folgenden werden nochmals kurz die verschiedenen Visualisierungsmittel der Link-Typisierung zusammengefaßt.

PARATEXTUELLE und METASPRACHLICHE ASPEKTE:

Links, die die primäre Gliederungsstruktur von SPORT1 vermitteln, sind nicht typisiert und auch nicht mit paratextuellen oder metasprachlichen Mitteln auf andere Art kenntlich gemacht. Hingegen sind die Links, die auf ergänzende Knoten von Teaser-Texten hinweisen, mit Piktogrammen typisiert. Oftmals ist der eigentliche Link-Anker auch selbst ein Piktogramm. In den meisten Fällen sind die typisierten Links mit ergänzenden Mouse-Over-Texten versehen.

Die Piktogramm-Typisierung der Links ist primär medial ausgerichtet, kann jedoch auch (kritisch) unter funktionalen Gesichtspunkten betrachtet werden (siehe Tabelle 37).

Als letzte Beschreibungsaspekte seien die browser-spezifischen angesprochen.

BROWSER-SPEZIFISCHE ASPEKTE:

Bis auf wenige Ausnahmen, führen sämtliche Links zu Ziel-Knoten, die im selben Fenster dargestellt werden. Die Ausnahmen sind der Live-Ticker¹, der

¹ Hier werden aktuelle Sportveranstaltungen nahezu in Echtzeit von Redakteuren in schriftlicher Form kommentiert. Beim Live Ticker handelt es sich um keinen Hypertext.

Online-Shop¹ sowie einige Werbe-Knoten. Daß aber bestimmte Knoten mit primär appellativer Funktion im selben Fenster wie die informativen Knoten angezeigt werden, verstärkt die Vermischung von Informations- und Werbe-Knoten.

17.3.5 Versuch einer RST-Codierung des Corpus'

Die nächsten drei Kapitel wenden die allgemeinen Link-Vokabularien, die in Kapitel 16.4 ausgesucht wurden, auf die Texte des Corpus' an. Es wird untersucht, inwieweit diese Vokabularien die dortigen proprietären Codierungen bzw. relativen Verhältnisse nachbilden können. Daß diese Kapitel vom Umfang eher knapp ausfallen, deutet bereits an, daß die Vokabularien für die Corpus-Texte wenig ergiebig waren!

Mit Hilfe der Rhetorical Structure Theory (RST) lassen sich thematisch-strukturelle Aspekte von Texten oder Hypertexten analysieren. In Kapitel 16.2.2 wurde die prinzipielle Anwendbarkeit der RST auf multimediale Hypertexte festgestellt. Knoten können hierfür auf Propositionsebene in minimale mediale Elemente zerlegt werden. Die Relationen, die zwischen diesen Einheiten desselben Knotens oder aber zwischen per Link verknüpften Einheiten verschiedener Knoten herrschen, können mit der RST analysiert werden. Exemplarisch werden nun einige Links aus den Hypertexten des Corpus' mit möglichst passenden Relationen des RST typisiert.

Technisch läßt sich die Typisierung der Corpus-Texte leicht durchführen. Eventuelle Typisierungen per Piktogramm werden hierbei entfernt und durch solche mit Relationen der RST ersetzt. Sichtbar gemacht werden die Link-Typen durch Mouse-Over-Texte, etwa in SELF:

¹ Inwieweit bei Online-Shops, Online-Wettbüros o.ä. überhaupt noch von Hypertexten gesprochen werden kann, ist fraglich: Die Navigation innerhalb eines Shops erfolgt zwar auch über Links. Der Shop selbst ist jedoch deutlich interaktiver als ein ‚normaler‘ Hypertext. Per Klick fügt der Leser bzw. Kunde seinem Warenkorb Artikel hinzu. Meiner Meinung nach kann hier eher von einem ‚normalen‘ Programm gesprochen werden, daß sich der Benutzeroberfläche eines erweiterten HTML-Hypertextes bedient. Die Benutzerschnittstelle solcher Programme könnte wohl eher mit einer modifizierten linguistischen Gesprächsanalyse denn mit einem textlinguistischen Analysemodell für Hypertexte analysiert werden. Zu vermuten ist, daß sich die WWW-Hypertexte der Zukunft immer mehr zu interaktiven Anwendungen entwickeln – und somit die Unterschiede zu traditionellen Texten deutlich größer werden.

Anzeigebeispiel: So sieht's aus

Mit dieser Angabe können Sie für Text einen Schatteneffekt erzeugen.

Beispiel (Style-Sheet-Definition für HTML-Tag im Text):

```
<p style="text-shadow:black; font-size:24pt;">Text</p>
```

Erläuterung:

Mit `text-shadow`: können Sie einen Textschatten erzwingen. Erlaubt sind **Farbangaben** oder der Wert `none` für "keinen Textschatten".

Beachten Sie:

Dieser Befehl, der zur Version 2.0 der CSS Style-Sheets gehört, wird von Netscape und MS Internet Explorer in den Versionen 4.x noch nicht interpretiert.

↑

weiter: [Style-Sheet-Angaben: Schriftformatierung mit Schriftartendatei](#)

zurück: [Maßeinheiten, Farbangaben, feste Angaben](#) Sequenz

Abbildung 30: RST-typisierter Knoten SELF [tdca.htm] Abschnitt *Textschatten (text-shadow)*

Geht man sämtliche typisierte Links von SELF durch (siehe hierzu Tabelle 30), so ist das Ergebnis ziemlich ernüchternd. Außer der Relation *Sequenz* für die Link-Typen *weiter* und *zurück* läßt sich kaum eine der Typisierungen in SELF durch RST-Relationen ersetzen. Zu überlegen wäre lediglich, die Relation *Background* für den Typ *Anzeigebeispiel* zu verwenden. Allerdings ist auch hier die Übereinstimmung eher fraglich.

Grund für den mangelhaften Erfolg einer RST-Codierung von SELF ist wohl primär die Ausrichtung der Typisierung der SELF-Links: In Kapitel 17.3.1 wurde festgestellt, daß die Mehrzahl der Links Aussagen zur Struktur bzw. zu den Navigationsmöglichkeiten machen. Diese Verhältnisse sind mit RST nicht zu beschreiben.

Ein versuchtes Mapping der RST-Relation auf die Link-Typen von SPORT1 bringt eine ähnlich magere Ausbeute. Zwar existieren im Vergleich zu SELF und SELF8 wenig Links, die die Hierarchie des Hypertextes typisieren. Dafür stellen die überwiegend medial ausgerichteten Link-Typen die RST vor unüberwindbare Probleme. *Sequenz* ließe sich auch hier als einzige Relation anwenden, und zwar als Ersatz für die Pfeil-Piktogramme, die einen Link von einem Teaser-Text zum Ziel-Knoten mit dem Volltext signalisieren.

In GRAMMIS ließe sich die RST-Relation *Background* für Links zum terminologischen Wörterbuch oder für Stretchede mit objektsprachlichen Beispielen verwenden. Ein informativer Mehrwert für Leser darf jedoch

bezweifelt werden. Die Verhältnisse in W3C sind vergleichbar, auch hier ließen sich allenfalls partiell die Relation *Sequence* und *Background* verwenden.

Fazit: RST eignet sich nicht zur Typisierung der Links der Corpus-Hypertexte. Dies liegt daran, daß im Corpus so gut wie keine rhetorischen Strukturen durch typisierte Links dargestellt werden. Die rhetorische Entfaltung findet primär auf Knoten-Ebene statt. Ausnahmen sind lediglich die Quellcode- und Anzeigebeispiele in SELF und SELF8 sowie die Belege in Schrift und Ton von GRAMMIS. Aber auch hier legten die Hypertext-Autoren größeren Wert auf eine mediale bzw. browser-spezifische Typisierung als auf eine rhetorische – diese Verhältnisse sind natürlich mit der RST nicht nachzubilden.

17.3.6 Versuch einer Codierung des Corpus' mit Hilfe des Dublin Core

Ebenso wenig wie mit RST gelingt eine Nachbildung der proprietären Link-Typen mit Hilfe des Dublin Core: Mit dessen Relationen (siehe Tabelle 29) lassen sich weder die rhetorisch-funktionalen noch die medialen Aspekte des Ziel-Knotens von Links typisieren.

Es wäre lediglich möglich, rein logisch Relationen wie *Has part* bzw. *Is part of* zu vergeben – das heißt beispielsweise: Der Ziel-Knoten mit einem Anzeigebeispiel ist logisch ein Teil des Ausgangs-Knotens, der den eigentlichen Fließtext enthält. Hiermit ließen sich beispielsweise Tondokumente bzw. Belege in GRAMMIS dem zugehörigen Abschnitt im Fließtext eines Knotens zuordnen. Über die Art der inhaltlichen Zuordnung würde damit aber noch keine Aussage gemacht. Auf diese Weise ließen sich auch die unterschiedlich geschachtelten Strukturen der Navigationsleisten der Corpus-Hypertexte beschreiben. Allerdings ist die dem Leser auf diese Weise vermittelte Zusatzinformation so abstrakt, daß ein Mehrwert dieser Typisierung anzuzweifeln ist. Meiner Meinung nach sollten die Relationen *has part* bzw. *is part of* des DC eher auf Ebene der Wissensrepräsentation realisiert werden.

Auch die Relationen, die mit *Is Required By*, *Requires*, *Is Referenced By*, *References* ausgedrückt werden, sind ebenfalls weniger für den menschlichen Leser als vielmehr für Computer relevant, die ‚Wissen‘ aus einer Wissensrepräsentation ableiten. Die Relationen *Is Basis for* und *Is Based on* sind hingegen wieder so allgemein formuliert (*Knoten A ist in irgendeiner Form Basis für Knoten B*; siehe hierzu die Besprechung dieser Relation in Kapitel 16.3.3), daß sie, obwohl eigentlich zur Typisierung auf Ebene der Oberfläche

geeignet, ebenfalls für den Leser nicht relevant – weil nicht präzise genug – sind: Interessant wäre es ja gerade kenntlich zu machen, ob etwa ein Knoten bzgl. eines anderen in einem Übersetzungs- oder aber in einem Ableitungsverhältnis steht. Ein bloßes *basiert auf* ist hier zu wenig. Dieselbe Kritik läßt sich auch beim Relationenpaar *Is Referenced By* und *References* anbringen. Ein Knoten kann hier für einen anderen als Zitat, Danksagung, Gegenposition oder Ähnliches fungieren. Eine genauere Differenzierung sieht das Dublin Core hier nicht vor.

Interessanter wäre eine DC-Typisierung eventuell für den Hypertext W3C. Hier existieren oftmals verschiedene Versionen von Knoten, eine neue Spezifikation ersetzt eine alte, welche aber dennoch für bestimmte Nutzergruppen relevant und deshalb im Netz verbleibt u.s.w. Diese Abhängigkeiten ließen sich mit den Relationen *IsVersionOf*, *HasVersion*, *Replaces* und *Is Replaced by* ausdrücken. Insgesamt muß jedoch festgestellt werden, daß auch die möglichen Werte der Relationen des Dublin Core zur Typisierung der semantischen Oberfläche von Hypertexten eher schlecht geeignet sind. Die Umwandlung der proprietären Typen des Corpus' ginge, wenn überhaupt, mit einem großen Verlust der von den Link-Typen vermittelten Information einher.

17.3.7 Versuch einer Codierung des Corpus' mit Hilfe des *rel*-Attributs

Die Werte des HTML-Attributs *rel* (bzw. *rev*) stellen das letzte allgemeine Vokabular, das auf seine Verwendung im Corpus getestet werden soll.

Abbildung 23 zeigt das Prinzip, wie in SELF (und auch SELF8) das Verhältnis des aktuellen Knotens zur Gesamthierarchie des Hypertextes visualisiert wird. Links, die zum Navigieren innerhalb dieser Hierarchie verwendet werden, könnten gut mit den Werten *Chapter*, *Section* und *Subsection* des *rel*-Attributs typisiert werden. Mit den zusätzlichen Werten *Contents*, *Index*, *Glossary*, *Next* und *Prev(ious)* ließen sich die gesamte hypertext-interne Navigation sowie die strukturellen Informationen sämtlicher Vertreter des Corpus' typisieren. Eine Typisierung von hypertext-internen vs. –externen Links wäre hingegen mit dem *rel*-Attribut nicht möglich, ebenso wenig wie eine medial orientierte Auszeichnung.

Auch bei GRAMMIS stößt eine Typisierung mittels des *rel*-Attributs an Grenzen. *Definitionen* – die häufigsten Link-Typen in GRAMMIS, können nicht

ausgezeichnet werden. Dasselbe gilt für Tonbeispiele. Auch die in GRAMMIS häufig benutzten Stretch-Texte können nicht kenntlich gemacht werden.

In W3C erwiesen sich die Werte *Chapter*, *Section*, *Subsection*, *Contents*, *Index*, *Glossary*, *Next* und *Prev(ious)* als brauchbar. Der Hypertext ist klar gegliedert, ohne dies jedoch auf der Oberfläche dem Leser deutlich zu visualisieren. Eine Typisierung mit dem Attribut *rel* könnte hier durchaus sinnvoll sein. Besonders dürfte sich auch der Wert *Alternate* eignen, befindet sich doch im Hypertext W3C eine große Anzahl verschiedener Versionen inhaltlich gleicher Knoten. Denkbar wäre auch eine Typisierung solch inhaltlich gleicher Knoten, die in unterschiedlichen Formaten vorliegen.

In SPORT1 könnte lediglich die grobe Navigationsstruktur des Hypertextes typisiert werden. Ansonsten liefe hier eine Typisierung mit dem *rel*-Attribut völlig ins Leere: Die proprietäre, medial und funktional orientierte Typisierung könnte auf diese Weise nicht nachgebildet werden.

Die Werte des Attributs *rel* stellen somit das allgemeine Vokabular dar, welches am besten für die Typisierung der Links des Corpus' geeignet scheint.

In der Endabrechnung erscheinen die Link-Typen der proprietären Vokabulare für den jeweiligen Hypertext dennoch als geeigneter. Im zusammenfassenden nächsten Kapitel wird daher ein Vorschlag für eine eigene allgemeine Menge von Link-Typen gemacht, die sich aus einigen häufig vorkommenden Vertreter allgemeiner Typen sowie beliebten proprietären zusammensetzt.

17.4 Zusammenfassung: Auswertung der Ergebnisse

Im folgenden werden die Ergebnisse der praktischen Untersuchungen aus den letzten beiden Kapiteln nochmals komprimiert dargestellt.

17.4.1 Aussagen zur Methode

Als erstes sind diejenigen Ergebnisse aufgeführt, die das TAH allgemein als Methode betreffen.

Allgemein

Die Aufteilung des Hypertextes in die Ebenen *Hypertext*, *Knoten* und *Link* hat sich als nützlich für die Übersichtlichkeit der Analyse erwiesen. Im Rahmen der Untersuchung wurden die Ebenen *Absatz* und *Satz* lediglich implizit auf *Knoten*-

oder *Link*-Ebene untersucht, eine exakte Unterscheidung war für die Fragestellung nicht von Nöten. Für anders ausgerichtete Arbeiten, die sich etwa primär mit Problemen auf Knoten-Ebene befassen, dürfte diese feinere Differenzierung jedoch relevant sein.

Auch die grundlegende Entscheidung für eine Top-Down-Analyse auf allen Ebenen war sinnvoll, verhinderte sie doch ein ‚Verzetteln‘ und entspricht dem Verständnis von TEXT als komplexem kommunikativ-funktionalen Akt. Daher brachte auch die in den Vordergrund gestellte Analyse der Funktion auf unterschiedlichen Ebenen brauchbare Ergebnisse. Besonders die funktionale Verschränkung des Corpus-Hypertextes SPORT1 hätte bei einer eher strukturell orientierten Analyse in der Form nicht analysiert werden können.

Ebene Hypertext

Kontextuelle Aspekte der Corpus-Texte konnten mit TAH gut beschrieben werden. Durch das Einbeziehen kommunikativ-funktionaler Aspekte wurde festgestellt, daß das WWW Hypertexte mit unterschiedlichen Funktionen beheimatet. Die theoretischen Überlegungen aus Kapitel 10.5 wurden somit durch praktische Analysen bestätigt: Ein Hypertext ist ein medialer Rahmen, in dem sich ein Textthema bestimmt von der jeweiligen Textfunktion zum Repräsentanten einer bestimmten Textsorte entfalten kann.

Ebene Knoten

Die Arbeit beschäftigte sich, wie gesagt, nur insofern mit Analysen auf Knoten-Ebene, als diese zum Beschreiben der Link-Struktur nötig waren. Zwei Probleme, welche aufgrund der Systematik von TAH in den Corpus-Hypertexten bereits gesichtet wurden, sollen hier dennoch kurz angesprochen werden: Die Platzhalter-Knoten in SPORT1 sowie die Stretchtexte in GRAMMIS. SPORT1 zeigte die Problematik, daß eine Bildschirmseite nicht immer einem Knoten entsprechen muß.¹ In der obigen Analyse wurde hierfür

¹ Dies ist natürlich auch immer dann der Fall, wenn der Leser scrollen muß. Darauf zielt der Einwand jedoch nicht ab. Auch geht es nicht um die technische Möglichkeit, in einem Frameset mehrere HTML-Dateien, also Knoten, auf den Bildschirm zu bringen. Es geht vielmehr um den ‚visuellen Eindruck‘ des Lesers. Ein Frameset, das Navigation und Inhalt technisch voneinander trennt, wird vom Leser dennoch aller Wahrscheinlichkeit nach als ein Knoten aufgefaßt. Eine Bildschirmseite von SPORT1 ‚wirkt‘ jedoch disparat (unabhängig von der technischen Realisierung). Auch hier läßt sich die Analogie zu gedruckten Zeitungen ziehen, die den Rahmen für verschiedenartige Teiltex te vorgeben.

der Begriff *Platzhalter-Knoten* eingeführt. Dies ist eine Hilfskonstruktion: Natürlich ließe sich mit den Mitteln von TAH zuerst eine Knoten-Analyse dieses Platzhalters und dann eine der eingebetteten Einzel-Knoten durchführen. Allerdings muß hierbei herausgearbeitet werden, daß dieser ‚Meta-Knoten‘ mehr ist, als die Summe seiner Einzel-Knoten. Eventuell müßte eine neue Ebene *Meta-Knoten* in TAH eingeführt werden, von welcher abhängig dann die Verteilung der Einzel-Knoten interpretiert werden muß. TAH ist insofern auf ‚klassische‘ Hypertexte ausgerichtet, das heißt, ein Knoten entspricht einer Bildschirmseite (Es darf jedoch gescrollt werden!).

Aus diesem Grund ist es auch verständlich, daß eine weitere Realisierungsform komplexer Knoten Schwierigkeiten bereitet – die in GRAMMIS verwendeten Stretchtexte. Natürlich lassen sich diese nach browser-spezifischen Aspekten bzgl. des Verhaltens eindeutig gegen andere Ziel-Knoten von Links abgrenzen. Der strukturelle Status von Stretchtexten läßt sich jedoch auf diese Weise nicht völlig befriedigend beschreiben: In gewisser Hinsicht sind sie ein Teil des Ursprung-Knotens andererseits aber auch als separate Knoten zu sehen. Ich tendiere vorschlagsweise dazu, der Semantik des Begriffes *Stretchtext* (sowie der der zugehörigen Verhaltensweise *embedded*) zu folgen, und somit diese Phänomene als Teile der Ursprung-Knoten zu sehen.

17.4.2 Aussagen zu inhaltlichen Fragestellungen bzgl. der Ebene *Link*

In den folgenden Abschnitten sind die inhaltlichen Ergebnisse aufgeführt, die das TAH bezogen auf die Problematik der Link-Typisierung erbrachte. Zum Schluß werden diese Aussagen benutzt, um damit einen eigenen Katalog typisierter Links vorzuschlagen.

In den Corpusanalysen, in der Umfrage aus Kapitel 16.3.1 sowie beim Betrachten weiterer Hypertexte des WWW¹ ist erkennbar, daß gewisse Typen von Links – ob nun explizit typisiert oder nicht – immer wieder vorkommen. Diese Typen sind aber in ihrer Gesamtheit weder in einer der vorgestellten proprietären Codierungen noch in einem der allgemeinen Vokabularien versammelt. Aus diesem Grund wird ein neues Inventar von Link-Typen vorgeschlagen. Diese Typen werden dann anhand der Beschreibungsaspekte des TAH in Gruppen geordnet.

¹ Unter anderem wurden folgende Hypertext gesichtet: Spiegel Online (www.spiegel.de), Yahoo (www.yahoo.de) und Lycos (www.lycos.de).

Gruppen von LINK-TYPEN

Als erstes werden nun die untersuchten Links den Beschreibungsaspekten des TAH folgend in Gruppen eingeteilt:

Bestimmte Links dienen primär der hypertext-internen Navigation und vermitteln keine zusätzlichen funktionalen oder thematisch-strukturellen Informationen. Diese Links können als direkt ausführbare kohäsive Mittel charakterisiert und mit den kohäsiven Mitteln gedruckter Texte verglichen werden (Inhaltsverzeichnisse, Indices etc.). Im Rahmen von TAH können diese Links auf Ebene des Hypertextes als Teil des allgemeinen Peritextes sowie auf Link-Ebene unter grammatisch-strukturellen Aspekten beschrieben werden. Auf diese Weise ist die erste Gruppe von Links definiert worden – die der NAVIGATIONSSPEZIFISCHEN TYPEN.

Eine zweite Gruppe filterte die Analyse auf Link-Ebene unter funktionalen bzw. thematisch-strukturellen Aspekten des TAH heraus. Zusammenfassend kann man hier von SEMANTISCHEN TYPEN sprechen.

Betrachtet man Links nach browser-spezifischen Aspekten des TAH, so lassen sich Typen bzgl. des TRAVERSAL-VERHALTENS voneinander unterscheiden. Diese Typen stellen die dritte Gruppe dar.

Die vierte Gruppe wird von den MEDIAL DEFINIERTEN LINKS gebildet. Hier scheint sich eine gewisse Einschränkung des TAH zu offenbaren, nämlich dessen Festlegung auf primär textuelle Ressourcen. Werkzeuge zur Analyse dieser nichtsprachlichen Zeichen werden von TAH nicht zur Verfügung gestellt. Im Analysemodell lassen sich zwar grob kommunikativ-funktionale und thematisch-strukturelle Aspekte von multimedialen Knoten beschreiben, eine Untersuchung, die den Besonderheiten der nichtsprachlichen Zeichen gerecht wird, läßt sich hiermit aber nicht erreichen. Durch die Ausrichtung des TAH (Primat der Funktion, Top-Down-Analyse) ließen sich jedoch unterhalb der Ebene *Knoten* weitere Ebenen (auf gleicher Höhe wie *Absatz* oder *Satz*) in das Modell integrieren. Diese Ebenen könnten dann auf die Beschreibung der multimedialen Inhalte von Knoten ausgerichtet sein. Hierfür wäre jedoch der Input anderer Disziplinen, wie etwa der allgemeinen Semiotik, der Kommunikationswissenschaft oder des Informations- bzw. Interfacedesigns, nötig. Eine rein textlinguistische Arbeit kann hier nicht mehr leisten, als einen

auch für multimediale Strukturen offenen pragmatischen Textbegriff zu liefern, und, auf diesen aufbauend, einen offenen, modularen Rahmen eines Analysemodells anzubieten.

Signalisierung der LINK-TYPEN

Ein weiterer Schwerpunkt der praktischen Analyse war die Untersuchung der paratextuellen bzw. metasprachlichen Signalisierung der typisierten Links. Die Information, um welchen Typ es sich bei einem Link handelt, wird in den Texten des Corpus' auf unterschiedliche Weise angezeigt:

In SELF8, SPORT1 und GRAMMIS werden Mouse-Over-Texte zur sprachlichen Auszeichnung von Links benutzt. Diese Mouse-Over-Texte kann man als ‚Metasprache des Links‘ bezeichnen, da sie Aussagen über die Aussage des Link-Textes machen.

In GRAMMIS existiert mit der farbigen Auszeichnung der Link-Texte ein sekundäres semiotisches System. Es handelt sich hierbei um ein symbolisches Zeichensystem, dessen konventionelle Grundlage stets auf der linken Seite eines Knotens aufgelistet ist und damit decodiert werden kann.

Darüber hinaus werden in zahlreichen Texten des Corpus' (SELF, SELF8, GRAMMIS, SPORT1) Links mit Piktogrammen ausgezeichnet. Auch hier handelt es sich um ein sekundäres semiotisches System, um eine Art Metasprache, mit der nichtsprachliche Aussagen über den Link-Text gemacht werden. Jedes Set von Piktogrammen stellt ein (tendenziell) ikonisches Zeichensystem dar. Oft scheinen die Schöpfer dieser Piktogramme der ‚Hinweiskraft‘ ihren ‚Abbildern‘ nicht völlig zu trauen, sonst müßten sie keine Legende der verwendeten Piktogramme in ihren Hypertext aufnehmen.¹

Allgemein stellen wohl nur ikonische Piktogramme mit hoher Bildhaftigkeit einen kognitiven Vorteil dar. Zu denken wäre hier an die vom Betriebssystem

¹ Vergleiche hierzu etwa SELF8 [editorial\arbeitshinweise.htm] Abschnitt *Verwendete Symbole*. Interessant ist, daß die Piktogramme in SELF8 (allgemeinsprachlich) *Symbole* genannt werden; interessant deshalb, denn je weniger die Kraft der Bildhaftigkeit des Piktogrammes ist, desto mehr verliert dieses an ‚Ikonenhaftigkeit‘ und nähert sich einem Symbol an, dessen (hypertext-weite) Konvention extra vermerkt werden muß. Hier sind natürlich ganz deutlich die Unterschiede zwischen fachwissenschaftlicher und allgemeinsprachlicher Verwendung dieser Begriffe zu bemerken!

Windows bekannten, globalen¹ Piktogramme für *Ordner öffnen* oder *Datei speichern*. Piktogramme mit Symbolcharakter, deren Konventionen einer Klärung durch Legenden bedürfen und deren Gültigkeit auf lediglich einen Hypertext beschränkt ist, dürften wohl eher eine kognitive Belastung darstellen. Validität für diese Hypothese müßte aber erst noch von umfangreicheren empirischen Untersuchung erbracht werden.

Ansätze einer Linking-Rhetorik

In Anlehnung an die oben beschriebenen Link-Gruppen werden im folgenden zwei Hauptklassen von Link-Typen vorgeschlagen, die durch zwei Gruppen von Modifizierern ergänzt sein können.

In einer ersten Hauptgruppe werden solche Typen zusammengefaßt, die Informationen über thematische und funktionale Zusammenhänge von Ausgangs- und Ziel-Knoten geben. Vorgeschlagen werden hier die Relationen *Beispiel*, *Hintergrund* und *Definition*. Die Typen dieser ersten Gruppe könnten technisch in einem SEMANTISCHEN STYLE-SHEET gespeichert werden.

Die zweite Gruppe besteht aus Link-Typen, die man als hypertext-spezifische Umsetzungen einiger aus Print-Texten bekannten (paratextuellen) deiktischen Mitteln beschreiben kann. Mit diesen Link-Typen wird die hypertext-interne Navigation realisiert. Die Gruppe besteht aus den Typen *Übersicht*, *neuer Abschnitt*, *extern*, *oben*, *unten*, *vor*, *zurück* und könnte in einem NAVIGATIONS-STYLE-SHEET gespeichert werden.

Im Unterschied zu den weiter unten noch einzuführenden Modifizierern sind die Links der beiden Hauptgruppen für sämtliche Hypertexte des Corpus' relevant.² Es ist zu vermuten, daß in den meisten der im WWW existierenden Hypertexten sowohl thematisch-funktionale als auch navigationsspezifische Zusammenhänge bestehen: Bezogen auf den thematisch-funktionalen Aspekt ist dies abzuleiten aus der Intention des Autors, einen kohärenten Text zu erstellen. Unter Aspekten der Navigation besteht für den Autor die Notwendigkeit, die Erreichbarkeit der einzelnen Knoten zu gewährleisten. Er kann sich hierbei nicht ausschließlich auf die Navigationsmöglichkeiten der Browser-Software verlassen, da er nicht weiß, welchen Browser mit welchen Fähigkeiten der Leser tatsächlich benutzt. Relationen, die ausschließlich medial oder browser-spezifisch modifizierend

¹ *Global* in der ‚Windows-Welt‘!

² Im Text W3C spielen beispielsweise die hier als *Modifizierer* bezeichneten Link-Typen keine Rolle.

wirken, sind in vielen Hypertexten (noch) nicht vertreten. Die strukturelle und thematische Organisation von Hypertexten wird also durch Relationen der von mir angenommenen beiden Hauptgruppen gewährleistet. Medial und browser-spezifische Typen modifizieren diese Relationen lediglich, aus diesem Grund wurden sie von mir *Modifizierer* genannt. Die browser-spezifischen Typen könnten in einem BROWSING-STYLE-SHEET gespeichert werden. Durch Verknüpfung mit einem semantischen Style-Sheet könnte somit etwa semantisch codierten Link-Typen automatisch ein bestimmtes Traversal-Verhalten zugeordnet werden (siehe unten).

Die Link-Typen der beiden Hauptgruppen zeigt folgende Tabelle:¹

| Typ | Relation Ausgangs- zu Ziel-Ressource | Beschreibung der Ziel-Ressource | Ort der Codierung |
|-----------------------------|---|--|--------------------------|
| <i>Beispiel</i> | Ziel-Ressource gibt Beispiel für Ausgangs-Ressource. | - ² | Semantisches Style-Sheet |
| <i>Definition</i> | Ziel- Ressource gibt eine Definition der Ausgangs-Ressource. | - | Semantisches Style-Sheet |
| <i>extern</i> ³ | - | Ziel-Ressource liegt außerhalb des aktuellen Hypertextes. | Navigations-Style-Sheet |
| <i>Hintergrund</i> | Ziel- Ressource gibt Hintergrundinformationen für Ausgangs-Ressource. | - | Semantisches Style-Sheet |
| <i>neuer Sinn-Abschnitt</i> | - | Ziel-Ressource liegt außerhalb des aktuellen Sinnabschnittes. ⁴ | Semantisches Style-Sheet |
| <i>oben</i> | Ziel steht weiter gegen Anfang der aktuellen Ressource. | | Navigations-Style-Sheet |
| <i>Übersicht</i> | - ⁵ | Ziel- Ressource ist ein Übersichts-Knoten (des gesamten Hypertextes). | Navigations-Style-Sheet |

¹ Die Typen sind hierbei nach dem Alphabet geordnet aufgeführt. Die Gruppenzugehörigkeit ist in der Spalte *Ort der Codierung* abzulesen: Semantische Link-Typen werden im Semantischen Style-Sheet, navigationsspezifische im Navigations-Style-Sheet gespeichert.

² *Beispiel* ist keine bloße Beschreibung des Ziel-Knotens, da die zugehörige Proposition zwei Argumente hat: *ist_Beispiel_für(Ziel-Knoten, Ausgangs-Knoten)* und nicht *ist_Beispiel_für(Ziel-Knoten)*.

³ Klar ist, daß signalisiert werden soll, wenn der Leser den aktuellen Hypertext verläßt. Der hypertext-interne Link ist hingegen der Standardfall. Daher ist fraglich, ob diese Link-Eigenschaft mit dem Typ *intern* explizit visualisiert werden sollte.

⁴ Die Beschreibung ist extra vage gehalten. Je nach Beschaffenheit des Hypertextes soll vom Autor selbst entschieden werden, was er unter *Sinnabschnitten* versteht. Es kann sich etwa um Teil-Hypertexte handeln (wie etwa in GRAMMIS), oder aber auch auf Co-Hypertexte in einem Hypertext-Netz verwiesen werden. Letzteres geschieht in SELF8, wobei hier eine noch feinere Unterteilung vorgenommen wird (siehe hierzu Tabelle 31).

⁵ *Übersicht* ist eine reine Beschreibung des Ziel-Knotens, da es sich um die einwertige Proposition *ist_ein_Übersichts-Knoten(Ziel-Knoten)* handelt und eben nicht um eine zweiwertige wie etwa *gibt_Übersicht_zu(Ziel-Knoten, Ausgangs-Knoten)*.

| Typ | Relation Ausgangs- zu Ziel-Ressource | Beschreibung der Ziel-Ressource | Ort der Codierung |
|----------------------------|---|---------------------------------|-------------------------|
| <i>unten</i> | Ziel steht weiter gegen Ende der aktuellen Ressource. | | Navigations-Style-Sheet |
| <i>vor</i> | Ziel ist die vorhergehende Ressource (etwa in einer Guided Tour). | | Navigations-Style-Sheet |
| <i>zurück</i> ¹ | Ziel ist die vorhergehende Ressource (etwa in einer Guided Tour). | | Navigations-Style-Sheet |

Tabelle 38: Vorschlag einer Link-Typologie

Die Signalisierung der zwei Gruppen sollte nach Möglichkeit auf Ebene der Oberfläche mit verschiedenen paratextuellen Mitteln erfolgen. Ein Vorschlag wäre, navigationsspezifische Typen durch Piktogramme und thematisch-funktionale Link-Typen durch Mouse-Over-Texte zu signalisieren. Ausschlaggebend hierfür ist die Annahme einer ‚gewissen Eingängigkeit‘ navigationsspezifischer Piktogramme. Ein gutes Beispiel hierfür ist die Signalisierung der Navigation in SELF8, besonders die durch Pfeil-Piktogramme. Hingegen ließen sich wohl für semantische Beziehungen schwerer Piktogramme finden, die gleich auf den ersten Blick einer Bedeutung zuzuordnen wären. Daher sollten diese Beziehungen eher sprachlich mit Mouse-Over-Texten signalisiert werden. Denkbar wäre auch eine farbige Markierung der einzelnen semantischen Typen, jedoch müßte der Leser dann erst die Farbcodierung internalisieren, was zu einer (hier nicht nötigen) kognitiven Zusatzbelastung führen würde.² Optional lassen sich jedem Link noch mediale bzw. browser-spezifische Modifizierer beifügen, die ihrerseits wieder in separaten Style-Sheets abgelegt sein könnten.

¹ Im Unterschied zur Funktion *Back* der Browser-Software (siehe auch Kapitel 6.4.9) führt ein Link des Typs *zurück* nicht unbedingt zur vorher besuchten Seite! Gleiches gilt auch für Links des Typs *vor*, die einen Schritt vorwärts in einer vom Autor definierten Lesefolge machen, während die Browser-Funktion *Forward* (bei vorheriger Benutzung von *Back*) wieder einen Schritt vorwärts hin zum zuletzt besuchten Knoten macht.

² Ein weiterer Grund spricht gegen eine farbige Markierung semantischer Relationen. Wie in GRAMMIS geschehen, kann es Sinn machen, große Hypertexte in Teil-Hypertexte zu zerlegen und diese einem ‚cooperate design‘ folgend zu gestalten. Hierbei kann die farbige Markierung der Knoten und der Links eine Rolle spielen. Insofern hält man sich diese Markierungsmöglichkeiten besser für solche oder ähnliche Aufgaben frei.

| |
|--------------------------------|
| Eigenschaft¹ |
| Bild |
| Ton |
| Film |

Tabelle 39: Mediale Modifizierer

| |
|--------------------------------|
| Eigenschaft² |
| Neues Fenster |
| Stretchtext |

Tabelle 40: Modifizierer des Browser-Verhaltens

Wie gesagt, die konkrete Implementierung ist für diese Arbeit eher nebensächlich. Ein Vorschlag ist jedoch, technisch eine strikte Trennung der einzelnen Link-Gruppen in separaten Style-Sheets durchzuführen. Die in Abbildung 7 von Kapitel 15 eingeführte Zwischenschicht könnte somit eventuell noch weiter ausdifferenziert werden: Neben der eigentlichen strukturellen³ Information des Links im HTML-Code könnte dort der Verweis auf diverse Style-Sheet-Klassen vermerkt sein, die ihrerseits in verschiedenen Style-Sheet-Dateien abgelegt sind. Ein Link könnte dadurch auf einfache Weise mit einem Typen – aber auch mit mehreren! – verbunden werden.

Folgendes vereinfachte Schaubild soll die Verhältnisse illustrieren:

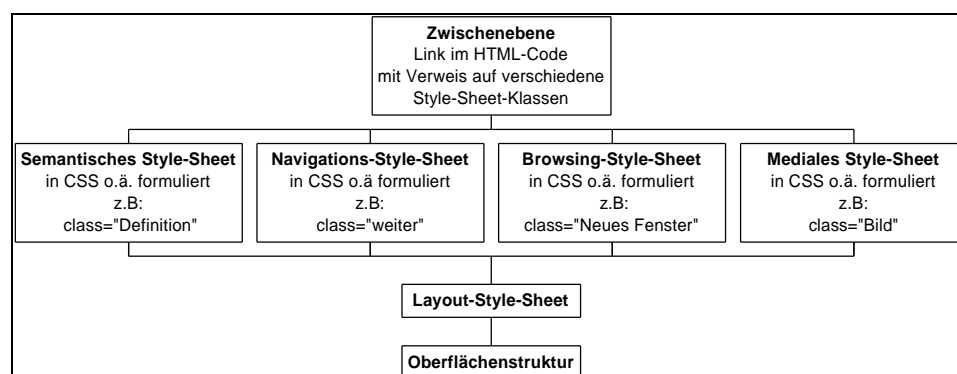


Abbildung 31: Style-Sheets der Link-Typenklassen

¹ Der Typ *Text* wird als Standard angenommen. Dieser braucht daher wohl nicht explizit kenntlich gemacht zu werden. Es wird davon ausgegangen, daß die meisten Hypertexte auch heute noch primär textuell ausgerichtet sind. In multimedialen Knoten haben die sekundären semiotischen Zeichen, wie etwa Bilder, oft (aber nicht immer, siehe hierzu auch Lobin 1999 S. 162) nur stützende Funktion für den textuellen Inhalt. Mediale Modifizierer sollten meiner Meinung nach daher nur eingesetzt werden, wenn gesamte Ziel-Knoten in einem sekundären semiotischen System vorliegen.

² Als Standardverhalten des Browsers wird hierbei die Anzeige des Ziel-Knotens im Fenster des aktuellen Knotens angenommen.

³ Nicht zu übersehen ist, daß auch die strukturelle Anordnung selbst und – jedenfalls in XML – die Benennungen der strukturellen Elemente bereits Träger von (impliziter) Semantik sind!

Fazit: Die vorgestellten Link-Typen und Modifizierer sind allgemeiner Art, daher offen für Hypertexte unterschiedlichster Funktion und Struktur. Welche Typen und Modifizierer aber tatsächlich für konkrete Hypertexte benutzt werden, läßt sich nur aus der individuellen Disposition des jeweiligen Hypertextes ableiten: So machen etwa mediale Modifizierer selbstredend nur in einem multimedialen Hypertext Sinn.

Angelehnt an die Analysekatogorien des TAH wird also eine Unterteilung in thematisch-funktionale, navigationsspezifische, browser-spezifische und mediale Link-Typen vorgeschlagen. Ein tatsächlicher Link kann hierbei mehrfach typisiert werden! Ausgegangen werden sollte von thematisch-funktionalen und/oder navigationsspezifischen Verhältnissen, die durch mediale und/oder browser-spezifische Aussagen modifiziert werden können. Durch getrennte Datenhaltung der Gruppen ließen sich hierbei komplexe Vorgänge bei mehrfach typisierten Links automatisieren: Ein Mapping zwischen semantischem Style-Sheet und Browsing-Style-Sheet könnte beispielsweise bewirken, daß ein Link vom semantischen Typ *Definition* hypertext-weit mit dem Browser-Verhalten vom Typ *Neues Fenster* verbunden wird. Ein allgemeines Style-Sheet könnte darüber hinaus noch die Aspekte des Layouts regeln – verschiedenen Link-Typen unterschiedliche Farben, Piktogramme oder Mouse-Over-Texte zuordnen.

Es sollte insgesamt jedoch darauf geachtet werden, daß das gesamte Typisierungssystem nicht zu komplex wird. Ansonsten würde der Leser nicht unterstützt, sondern wohl eher verwirrt werden. Genaue Aussagen zur Anzahl und paratextuellen Auszeichnung von Links bleiben zukünftiger empirischer Forschung überlassen.

Epilog: Einsatzmöglichkeiten des TAH

TAH ist ein Modell zur deskriptiven Analyse der Funktion und Struktur von Kommunikation, die durch Hypertexte stattfindet. Es ist darauf hin ausgerichtet, sämtliche medienspezifische Aspekte dieser Kommunikation zu beschreiben.

TAH ist kein präskriptives Modell, bietet also auch keine Rhetorik für die Produktion von Hypertexten an. Zur Kommunikationsqualität wird per se erst einmal nichts gesagt, aus den Beschreibungen der hypertextuellen Verhältnisse können jedoch Schlüsse hierfür gezogen werden: So wirkt sich etwa die An- bzw. Abwesenheit globaler und lokaler Kohärenz direkt auf die Qualität der Kommunikation aus. Diese Qualität zu messen und sie mit der anderer medialer Realisierungen von Informationsvermittlung zu vergleichen, sei jedoch empirischen Untersuchungen überlassen.

TAH soll ein Gespür für die vorhandenen Ebenen, Beschreibungsaspekte und Analysekatoren vermitteln, die bei der Untersuchung von Hypertexten eine Rolle spielen.¹ Insofern soll hiermit auch ein Input für folgende empirische Untersuchungen von Hypertexten aller Art gegeben werden: Durch eine vorab durchgeführte textlinguistische Betrachtung von Corpus-Texten mit TAH könnte gewährleistet werden, daß beim Aufstellen von Hypothesen und Definieren von Variablen für die eigentliche empirische Analyse nur ‚Gleiches mit Gleichem verglichen wird‘ – also nicht etwa funktionale mit strukturellen oder paratextuellen Aspekten gleichgesetzt werden. Stellen sich dann bei einer etwaigen zukünftigen empirischen Untersuchung von Lesepfaden leserseitige Präferenzen für bestimmte Link-Typen heraus, so könnten diese Unterschiede durch die Ebenen, Aspekte und Kategorien von TAH benannt, textlinguistisch interpretiert und damit ‚greifbar‘ gemacht werden.

In Kapitel 9.2.2 wurde ein schemabasiertes, konstruktivistisches Textverstehensmodell vorgestellt, von dem die Arbeit implizit ausgeht. Die kognitive Verarbeitung eines Textes wird hierbei, wie gesagt, von drei Faktoren bestimmt: durch den Text, durch die individuelle Disposition jedes Rezipienten sowie durch situative Einflüsse. TAH ist textzentriert, beschränkt sich in seiner

¹ Daß die medialen Bedingungen bei Untersuchungen der Kommunikationsqualität noch zu wenig berücksichtigt werden, bemängelt zu Recht Bucher 2001 S. 46.

Beschreibung also auf die Ebene des eigentlichen Textes sowie (mit Abstrichen) auf die der situativen Einflüsse. Die individuelle Disposition eines Lesers (bzw. verschiedener Lesergruppen) sowie die dynamische Interaktion von textinternen und textexternen Faktoren – der eigentliche Textverstehensprozeß also – kann von TAH nicht erfaßt werden.

Der Rückzug auf die textinterne Ebene sollte jedoch nicht als Flucht vor einer ungleich komplexeren kognitiven Wirklichkeit der Textverarbeitung verstanden werden. Vielmehr geschieht dies aus der, bereits eingangs erwähnten Überzeugung heraus, daß ein inter- oder transdisziplinäres Wissenschaftsfeld stets von der Qualität des Inputs der integrierten (bzw. integrierenden) Einzeldisziplinen abhängt. Im Laufe der Arbeit sollte überdeutlich geworden sein, daß von seiten der Textlinguistik noch längst nicht alle Verhältnisse geklärt sind.

Auch die junge Hypertext-Theorie hat verständlicherweise noch in vielen Bereich Nachholbedarf. Insofern erschien es sinnvoll (und maßvoll), von einem funktionalen Textbegriff ausgehend mit TAH ein relativ beschreibungsmächtiges, rein textlinguistisches Instrumentarium zu entwickeln. Auf die Integration in ein übergeordnetes kognitives Modell der menschlichen Textverarbeitung wurde daher bewußt verzichtet, TAH soll als rein fachwissenschaftlicher Vorschlag zur textlinguistischen Beschreibung hypertextueller Ebenen gesehen werden. Ein integratives Modell der Textverarbeitung, welches nicht nur beschreibungs- sondern auch erklärungsadäquat ist, steht noch aus – auch wenn bereits einige erfolgversprechende Ansätze existieren (etwa Kintsch u. van Dijk 1978, Kintsch u. van Dijk 1983 sowie Bransford 1979).

Dank der rein deskriptiven Ausrichtung des TAH dürfte es jedoch zu keinen Unverträglichkeiten mit unterschiedlichen kognitiven Modellen der Textverarbeitung kommen: Auf die Entwicklung des TAH als primär textinternes Beschreibungsmodell hatte das der Arbeit inhärente, konstruktivistische Verständnis der Textverarbeitung wenig Einfluß, insofern dürfte TAH auch als Werkzeug für Untersuchungen zu benutzen sein, die mit anderen kognitiven Ansätzen arbeiten.

Anhang: Literatur und Verzeichnisse

A Glossar

Anmerkung: Sämtliche Termini der untenstehenden Definitionen, die einen Glossareintrag haben, sind im Fließtext in Kapitälchen gesetzt. Die große Mehrheit der Begriffe sind allgemein bekannte und eingeführte Fachbegriffe. Wurden Termini in dieser Arbeit geprägt oder modifiziert verwendet, so ist dies in den Definitionen explizit erwähnt.

| BEGRIFF | DEFINITION |
|---------------------|---|
| AUSGANGS-KNOTEN | Aktuell angezeigter Knoten im BROWSER, der den LINK-TEXT eines LINKS enthält, welcher auf einen ZIEL-KNOTEN verweist. |
| BROWSER (-SOFTWARE) | PROGRAMM für das Lesen von bzw. für die NAVIGATION in HTML- oder XML-HYPertexten. |
| BROWSER-VERHALTEN | siehe TRAVERSAL. |
| CLIENT | Arbeitsstation bzw. PROGRAMM im Netzwerk (z.B. ein BROWSER). Ein Client beansprucht Dienste irgendeiner Art von einem SERVERS. |
| CONTAINER | siehe ELEMENT |
| CSS | CASCADING STYLE SHEETS Einfache STYLE-SHEET-SPRACHE zur Formatierung von HTML-DOKUMENTEN, die sehr unterschiedlich von den aktuellen BROWSERN unterstützt wird. CSS2 ist im Gegensatz zu CSS1 auch für XML-DOKUMENTE zu verwenden. |
| DATEI | siehe KNOTEN. |
| DC | siehe DUBLIN CORE. |
| DOKUMENT | siehe KNOTEN. |
| DTD | DOCUMENT TYPE DEFINITION Interne oder externe Grammatik für SGML- und XML-DOKUMENTE. Die momentan bekanntesten DTDs sind die zu HTML – diese sind in den aktuellen Browsern implementiert. Grammatiken für XML-DOKUMENTE können auch durch SCHEMATA formuliert werden. |
| DUBLIN CORE | System von Metaangaben der <i>Dublin Core Metadata Initiative</i> zur semantischen Auszeichnung von HTML- oder XML-DOKUMENTEN |
| ELEMENT | Hauptbestandteil eines HTML- oder XML-DOKUMENTES. Liegt stets (außer bei leeren Elementen) als Kombination von öffnendem E. und schließendem Elemente vor: <element>inhalt</element>. Elemente können weitere Elemente enthalten – also eine verschachtelte Struktur besitzen. |
| E-TEXT | Elektronisch realisierter, traditioneller linearer Text. |
| FRAME | siehe FRAMESET. |
| FRAMESET | Technik in HTML, mittels der auf einer Bildschirmseite mehrere Knoten in jeweils eigenen Bereichen dargestellt werden können. |
| FTP | FILE TRANSFER PROTOCOL INTERNETPROTOKOLL, das auf dem TCP/IP aufsetzt. FTP dient dem Austausch von DATEIEN, die zumeist mittels eines FTP-CLIENTS von FTP-SERVERN heruntergeladen werden. |
| GUIDED TOUR | Geführte Unterweisung: Vom Autor angelegter linearer Pfad durch einen HYPertext. |
| HTML | HYPertext MARKUP LANGUAGE HYPertext-Dokumenttyp, der aus normalem Text und geschachtelten ELEMENTEN besteht. Die Elemente sind für die Struktur, das Layout und die LINKS zuständig. Die Trennung von Struktur und Layout ist nicht sauber vollzogen, kann aber durch Verwendung von CSS erreicht werden. HTML ist mit der Metasprache SGML formuliert, die neuere Version XHTML hingegen mit XML. |
| HTTP | HYPertext TRANSFER PROTOCOL Auf TCP/IP aufbauendes höheres PROTOKOLL für den Austausch von Ressourcen im WWW. |
| HYPER-LINK | siehe LINK. |

| BEGRIFF | DEFINITION |
|----------------------|---|
| HYPERTEXT | Im elektronischen Medium realisierte, strukturell nicht-lineare und potentiell multimedial ausgerichtete Texte. Durch das Lesen eines Hypertextes erschafft der Leser per Navigation einen individuellen, linearen Lesepfade. |
| HYPERTEXT-LINK | siehe LINK. |
| HYPERTEXT-NETZ | Sammlung von HYPERTEXTEN, die zumeist inhaltlich-thematisch verwandt sind und sich häufig durch gemeinsames Layout und ähnlichen Stil auszeichnen. Häufig befindet sich ein HYPERTEXT-Netz auch auf einer gemeinsamen SITE. |
| HYPERTEXT-SYSTEM | System zur Speicherung und Ansicht (evtl. auch Manipulation) von HYPERTEXTEN. In der vorliegenden Arbeit werden folgende Subsysteme von Hypertext-Systemen angesprochen: Datenhaltung (SEMANTISCHE TIEFENSTRUKTUR), Ebene des QUELLCODES (ZWISCHENSCHICHT) und Darstellung (OBERFLÄCHENSTRUKTUR). |
| INTERNET | Weltweiter Zusammenschluß (heterogener) Computernetze. Die Kommunikation basiert auf dem TCP/IP-PROTOKOLL. Darauf bauen zahlreiche höhere PROTOKOLLE auf, wie etwa auch das HTTP für das HYPERTEXT-SYSTEM WWW. |
| INTERNET EXPLORER | BROWSER-SOFTWARE der Firma <i>Microsoft</i> . |
| INTERNET-PROTOKOLL | siehe TCP/IP. |
| KNOTEN | auch DATEI o. DOKUMENT; Grundlegende Informationseinheit von HYPERTEXTEN, die bzgl. Inhalt, Größe und Form nicht exakt festgelegt ist. Neben LINKS sind KNOTEN die Hauptkomponenten von HYPERTEXTEN. (siehe auch RESSOURCE) |
| LEERES ELEMENT | Element ohne schließende Klammer und damit auch ohne Elementinhalt. Leere Elemente können durch Attribute näher spezifiziert werden: z.B. in HTML |
| LINK | Eine direkt ausführbare Relation zwischen KNOTEN, Teilen von verschiedenen Knoten oder desselben Knotens, die mittels eines LINK-TEXTES (bzw. LINK-ANKERS) im Ausgangs-Knoten visualisiert wird. Per Mausklick auf den Link-Text wird der ZIEL-KNOTEN abhängig vom TRAVERSAL-Verhalten angezeigt. |
| LINK-ANKER | In irgendeiner Form kenntlich gemachtes Objekt innerhalb eines KNOTENS, das angeklickt werden kann und damit als Ausgangspunkt eines Links dient. Ist der Link-Anker rein textuell realisiert, so wird auch von LINK-TEXT gesprochen. |
| LINK-TEXT | siehe LINK-ANKER. |
| LINK-URSPRUNG | siehe AUSGANGS-KNOTEN. |
| LINK-ZIEL | siehe ZIEL-KNOTEN. |
| LYNX | Textbasierte BROWSER-SOFTWARE. |
| MARKUP-SPRACHE | Sprache zur Auszeichnung von Texten durch wohldefinierte ELEMENTE. Markup kann sich sowohl auf Druckanweisungen, Bildschirmlayout oder aber auch auf Struktur, Semantik und/oder Logik von Texten beziehen. |
| MOUSE-OVER-TEXT | Text der erscheint, wenn der Mauszeiger sich über eine sensitive Fläche auf dem Bildschirm, etwa einen HYPERTEXT-LINK, bewegt |
| NAMESPACE | auch NAMENSRAUM; Eine Deklaration innerhalb eines XML-DOKUMENTES verweist auf einen Namespace, der dann unter einem Kürzel die Namen dieses Namensraumes im aktuellen Dokument verfügbar macht – ohne daß dabei Namenskonflikte mit den Elementen des aktuellen Dokumentes entstehen können. |
| NAVIGATION | Lesen eines HYPERTEXTES; entspricht der Konstruktion eines Lesepfades entlang der vom Autor gesetzten Links |
| NETSCAPE NAVIGATOR | BROWSER-SOFTWARE der Firma <i>Netscape</i> . |
| OBERFLÄCHEN-STRUKTUR | In dieser Arbeit verwendeter Terminus für die für den Leser sichtbare Ebene eines HYPERTEXTES. Auf dieser Ebene setzen die verschiedenen Beschreibungsaspekte der TAH-Analyse an. Die Ansicht der Oberflächenstruktur ist abhängig von der Codierung der ZWISCHENSCHICHT. Dort liegt der HTML bzw. XML-Code sowie eventuelle STYLE-SHEETS. Die Zwischenschicht wiederum kann je nach Implementierung aus Datenbanken o.ä. abgeleitet sein (= TIEFENSTRUKTUR). |
| OPERA | BROWSER-SOFTWARE der Firma <i>Opera Software AS</i> . |
| POP-UP-FENSTER | Ein TRAVERSAL-Verhalten: Das Ziel eines LINKS wird in einem neuen Fenster dargestellt. |
| PROGRAMM | Syntaktisch korrekt geformtes Produkt einer PROGRAMMIERSPRACHE zur Ausführung von Prozeduren. |
| PROGRAMMIER-SPRACHE | Syntaktisch exakt festgelegte Regeln zur Manipulation und Verknüpfung von Datenbereichen. Höhere Programmiersprachen streben hierfür eine immer abstraktere Formulierung des Lösungswegs an. Beispiele für solche imperativen |

| BEGRIFF | DEFINITION |
|---------------------|---|
| | Programmiersprachen sind Pascal, Perl und C. Programmiersprachen, die für einen bestimmten Bereich geschrieben wurden und deshalb tendenziell einfacher gehalten sind, werden oft als SKRIPT-SPRACHE bezeichnet (etwa Perl, PHP o. JavaScript). Die Übergänge sind jedoch fließend (z.B. die äußerst komplexe ‚Skriptsprache‘ Perl). |
| PROTOKOLL | Gesamtheit der Regeln, die den Informationsfluß zwischen dem Computer und seiner Peripherie bzw. zwischen mehreren vernetzten Computern regelt. Der Informationsfluß im INTERNET basiert auf verschiedenen geschichteten Protokollen (siehe auch TCP/IP u. HTTP). |
| PULL-DOWN-MENÜ | Menü, das nach unten aufgezogen wird. Sichtbar ist zuerst lediglich ein Feld mit einem Begriff. Per Mausclick rollt das zugehörige Menü nach unten und zeigt alle Einträge an. |
| QUELLCODE | Allgemein: In einer höheren PROGRAMMIERSPRACHE geschriebenes PROGRAMM, welches dann noch von einem Compiler in eine Form gebracht werden muß, die von der Computerzentraleinheit (CPU) verstanden wird. HTML- und XML-Dokumente liegen als (semi)strukturierter (und evtl. semantisch angereicherter) Quelltext vor, der aus reinem Text und definierten MARKUP-Zeichen besteht und für Menschen lesbar ist. Der Quellcode wird in dieser Arbeit auf der Ebene der (semantischen) ZWISCHENSCHICHT positioniert, da er einerseits aus der Datenhaltung des HYPERTEXT-SYSTEMS erst erzeugt sein kann und andererseits vom BROWSER für den Leser erst zur Ansicht gebracht werden muß. |
| RDF | RESOURCE DESCRIPTION FRAMEWORK XML-Anwendung für den Austausch von Metaangaben in HTML- oder XML-DOKUMENTEN. RDF ist eine mögliche Realisierungsform des SEMANTIC WEB. |
| RELATION | Beziehung zwischen Textteilen bzw. KNOTEN. Eine Relation kann als ausführbarer LINK realisiert sein, muß dies aber nicht. XLINK ist ein Modell zur Beschreibung solcher Relationen zwischen RESSOURCEN. |
| RESSOURCE | Jedes mittels einer URI adressierbare Objekt. Eine Ressource kann ein HYPERTEXT-KNOTEN, ein Teil-KNOTEN, ein Bild, Teil eines Bildes o.ä. sein. Bezieht man Ressource auf KNOTEN, so sind diese Termini synonym zu verwenden. |
| SCHEMA | Grammatik für XML-DOKUMENTE, die im Gegensatz zu DTDs selbst in XML formuliert ist und eine weitergehende Typenbeschreibung erlaubt. |
| SEMANTIC WEB | Semantische Repräsentation der im WWW vorhandenen RESSOURCEN, die beispielsweise durch RDF und/oder TOPIC-MAPS realisiert sein kann. |
| SERVER | Rechner oder Dienst, der alle CLIENTS im Netzwerk ‚bedient‘ (z.B. diesen RESSOURCEN von HYPERTEXTEN zur Verfügung stellt) |
| SGML | STANDARD GENERALIZED MARKUP LANGUAGE Metasprache zur Beschreibung von MARKUP-SPRACHEN. SGML ist der Vorläufer zu XML. |
| SITE | Mediale Realisierung eines HYPERTEXTES im WWW. Eine Site umfaßt in der Regel einen Hypertext oder ein HYPERTEXT-NETZ. |
| SKRIPT | siehe PROGRAMM. |
| SKRIPTSPRACHE | siehe PROGRAMMIERSPRACHE. |
| STRETCHTEXT | Ein BROWSER-VERHALTEN; siehe TRAVERSAL |
| STYLE-SHEET | Vorlage mit Formatierungsanweisungen, die DOKUMENTEN zugeordnet und für Bildschirmlayout, Druckbild, Sprachausgabe o.ä. verantwortlich ist. In dieser Arbeit wird die Aufteilung der Information für Semantik, Layout und BROWSER-VERHALTEN von Links in semantische Style-Sheets (semantische Typen), Navigations-Style-Sheets (navigationsspezifische Typen), mediale Style-Sheets (mediale Typen) und Browsing-Style-Sheets (TRAVERSAL des Browsers) vorgeschlagen. |
| STYLE-SHEET-SPRACHE | Sprache zur Formulierung von STYLE-SHEETS |
| TAG | siehe ELEMENT. |
| TCP/IP | TRANSMISSION CONTROL PROTOCOL/INTERNET PROTOCOL Protokolle der Transport- und Adressierungsschicht des INTERNET auf die höhere Protokolle wie HTTP aufbauen. |
| TIEFENSTRUKTUR | In dieser Arbeit verwendeter Terminus für Datenhaltung aller Art eines HYPERTEXT-SYSTEMS. Hieraus wird bei Anfragen der Quellcode der ZWISCHENSCHICHT generiert. |
| TOPIC-MAP | Abstrakte Form der Wissensrepräsentation, die konkreten RESSOURCEN etwa aus dem WWW zugeordnet sein kann. In dieser Arbeit werden Repräsentationen wie Topic-Maps auf Ebene der TIEFENSTRUKTUR eines HYPERTEXT-SYSTEMS angesiedelt. Topic-Maps sind eine mögliche Realisierungsform des SEMANTIC WEB. |

| BEGRIFF | DEFINITION |
|-------------------------------|--|
| TRANSMISSION CONTROL PROTOCOL | siehe TCP/IP. |
| TRAVERSAL | Verhaltensweisen der BROWSER-SOFTWARE bei Klick auf einen LINK. Das Link-Ziel kann in heutigen Implementierungen den aktuellen Knoten ersetzen, im aktuellen Fenster zusätzlich zum AUSGANGS-KNOTEN (etwa als STRETCHTEXT) angezeigt oder in einem neuen Fenster geöffnet werden. Traversal-Verhalten könnte via BROWSING-STYLE-SHEETS auf der Ebene der ZWISCHENSCHICHT codiert sein. |
| URI | UNIVERSAL RESOURCE IDENTIFIER Sammelname für URL oder URN. |
| URL | UNIVERSAL RESOURCE LOCATOR Eindeutige Adresse einer RESSOURCE im INTERNET in der Form <i>protokoll://server/pfad</i> |
| URN | UNIFORM RESOURCE NAME Eindeutige Bezeichnung einer RESSOURCE im INTERNET. |
| W3C | WORLD WIDE WEB CONSORTIUM Zusammenschluß verschiedener Firmen und Wissenschaftler zur Spezifizierung von INTERNET-Standards wie etwa HTTP, HTML, XML u.v.m. |
| WEBSITE | siehe SITE. |
| WWW | WORLD WIDE WEB Alle RESSOURCEN im Computernetz INTERNET, die über das PROTOKOLL HTTP zu einem weltweiten HYPERTEXT verknüpft sind. Das WWW ist das momentan populärste HYPERTEXT-SYSTEM. |
| XHTML | EXTENSIBLE HYPERTEXT MARKUP LANGUAGE Neuformulierung von HTML 4.0 mittels XML. |
| XLINK | XML LINKING LANGUAGE Relationsmodell für XML-Dokumente. Es kann uni- u. multidirektionale Hyperlinks beschreiben. Xlinks können außerhalb der am LINK beteiligten RESSOURCEN gespeichert werden. |
| XML | EXTENSIBLE MARKUP LanGuage Metasprache wie – und echte Teilmenge von – SGML zur Formulierung von MARKUP-SPRACHEN. In XML ist die neue Version von HTML (= XHTML) formuliert. |
| XML-TOPIC-MAP | siehe TOPIC-MAP |
| XPATH | XML PATH LANGUAGE Adressierungssprache für Teile von XML-Dokumenten, die sowohl von der STYLE-SHEET-SPRACHE XSLT als auch von der Fragmentidentifizierersprache XPOINTER verwendet wird. |
| XPOINTER | Fragmentidentifizierersprache zum Adressieren in XLINK-Ausdrücken |
| XSL(T) | EXTENSIBLE STYLESHEET LANGUAGE (TRANSFORMATIONS) Style-Sheet-Sprache für XML. XSLT vermag in XML-DOKUMENTEN komplexe Transformationen vorzunehmen und ist damit CSS weit überlegen. |
| ZIEL-KNOTEN | KNOTEN, der das Ziel eines LINKS darstellt. |
| ZWISCHENSCHICHT | In dieser Arbeit verwendeter Terminus für die Ebene des HTML- oder XML-QUELLCODES sowie möglicher STYLE-SHEETS für Semantik, Darstellung und BROWSER-VERHALTEN. |

Tabelle 41: Glossar

B Fragebogen der Pilotstudie

Fragen nach allgemeiner Erfahrung mit dem WWW:

1. Seit wie vielen Jahren nutzen Sie das WWW?
2. Wie gut sind Ihre Fähigkeiten in HTML?
3. Wie gut sind Ihre Fähigkeiten in CSS?
4. Wie gut sind Ihre Fähigkeiten in XML?
5. Wie gut sind Ihre Fähigkeiten in RDF?

Fragen nach Erfahrung mit typisierten Links:

1. Wissen Sie, was ein Hypertext-Link ist?
2. Wissen Sie, was ein typisierter Link ist?
3. Nennen Sie einen Hypertext im WWW, der typisierte Links verwendet.
4. Wie waren diese Links typisiert (Mouse-Over-Text, Piktogramm, Farbe, Mouse-Over-Text und Piktogramm etc.)
5. Haben Ihnen die Link-Typen bei der Entscheidung geholfen, den entsprechenden Link zu verfolgen? Wenn ja, inwiefern?
6. Empfanden Sie die typisierten Links in irgendeiner Weise als unangenehm? Wenn ja, warum?
7. Was ist Ihnen persönlich wichtiger: Struktur, Semantik oder Layout einer Seite?

Fragen nach der Einschätzung von typisierten Links:

1. Welche Anzahl von Link-Typen würden Sie als sinnvoll erachten?
2. Was könnte Ihrer Meinung nach das Problem bei zu vielen Link-Typen sein?
3. Wie würden Sie das Layout der typisierten Links gestalten (Mouse-Over-Text, Piktogramm, Farbe)
4. Was unterscheidet Ihrer Meinung nach den Link-Typ *Beispiel* vom Typ *Pop-Up-Fenster*?
5. Was unterscheidet Ihrer Meinung nach den Link-Typ *Visualisierung* vom Typ *Flash-Animation*?
6. Was unterscheidet Ihrer Meinung nach den Link-Typ *Nächstes Kapitel* vom Typ *Definition*?

C Corpusverzeichnis

[GRAMMIS]

Grammis. Das grammatische Informationssystem des ids

[Online im WWW unter: <http://hypermedia.ids-mannheim.de/grammis/> (zuletzt aufgerufen am 03.01.2002)]

[SELF]

SELFHTML. Die Energie des Verstehens: HTML-Dateien selbst erstellen. Version 7

[Online im WWW unter: <http://selfaktuell.teamone.de/archiv/> (zuletzt aufgerufen am 23.02.2002)]

[SELF8]

SELFHTML. Die Energie des Verstehens: HTML-Dateien selbst erstellen. Version 8

[Online im WWW unter: <http://selfhtml.teamone.de/> (zuletzt aufgerufen am 23.02.2002)]

[SPORT1]

Sport 1. Sportportal

[Online im WWW unter: <http://www.sport1.de/> (zuletzt aufgerufen am 03.02.2002)]

[W3C]

W3C. World Wide Web Consortium

[Online im WWW unter: <http://www.w3c.org/> (zuletzt aufgerufen am 21.02.2002)]

D Literaturverzeichnis

[Agosti 1996]

Agosti, M.: Information retrieval and Hypertext. Boston 1996

[Akscyn 1989]

Akscyn, R.: Hypertext '89. New York 1989

[Ambron u. Hooper 1988]

Ambron, S. u. Hooper, K. (Hrsg.): Interactive multimedia. o.O. 1988

[Andersen 1990]

Andersen, P. B.: A Theory of Computer Semiotics. Semiotic approaches to construction and assessment of computer systems. Cambridge 1990

[Andersen 1992]

Andersen, P. B.: Towards an Aesthetics of Hypertext Systems: A Semiotic Approach.
In: Streitz, N., Risk, A. u. André, J. (Hrsg.): Hypertext: Concepts, Systems and Applications Cambridge 1992

[Ansel Suter 1995]

Ansel Suter, B.: Hyperlinguistics. Hypertext-Lernumgebungen im Akademischen Kontext: Eine Fallstudie. Zürich 1995

[Assfalg 1993]

Assfalg, R.: The hypertext internet connection. Konstanz 1993

[Astleitner 1997]

Astleitner, H.: Lernen in Informationsnetzen. Theoretische Aspekte und empirische Analysen des Umgangs mit neuen Informationstechnologien aus erziehungswissenschaftlicher Perspektive. Berlin u.a. 1997

[Atkinson u. Knee 1990]

Atkinson, S.D. u. Knee, M. (Hrsg.): Hypertext/Hypermedia. An annotated Bibliography. N.Y. u.a. 1990

[Austin 1962]

Austin, P.: How to do things with words. Oxford 1962

[Ballstaedt u.a. 1981]

Ballstaedt, S.P. , Mandl, H. , Schnotz, W. u. Tergan, S.O.: Texte verstehen, Texte gestalten. München, Wien u. Baltimore 1981

[Barrett 1989]

Barrett, E.: Text, ConText and HyperText. Cambridge u.a. 1989

[Barrett 1991]

Barrett, E.: The Society of Text. Cambridge u.a. 1991

[Bearman 1997]

Bearman, D.: Relation Element Working Draft.

[Online: <http://dublincore.org/documents/1997/12/19/relation-element/>] o.O. 1997

[Begeman u. Conklin 1988]

Begeman, M.L. u. Conklin, J.: The right tool for the job. In: Byte (Oct. 1988) S. 255-268. 1988

[Begeman u. Conklin 1989]

Begeman u. Conklin, M.L.: gIBIS: A tool for all reasons.

In: Journal of American Society for Information Science 40 (3) S. 200-213. 1989

[Berk u. Devlin 1991]

Berk, E. u. Devlin, J.: A hypertext timeline.

In: Berk, E. u. Devlin, J. (Hrsg.): Hypertext/Hypermedia handbook New York 1991

[Berners-Lee 1989]

Berners-Lee, T.: Information Management: A Proposal.

[Online: <http://www.w3.org/History/1989/proposal.html>] o.O. 1989

[Bernes-Lee 1999]

Berners-Lee, T.: Semantic Web Road map.

[Online: <http://www.w3.org/DesignIssues/Semantic.html>] o.O. 1999

[Berners-Lee 1999a]

Berners-Lee, T.: Der Web-Report. Der Schöpfer des World Wide Web über das grenzenlose Potential des Internets. München 1999

[Berners-Lee FAQ]

Berners-Lee, T.: Press FAQ.

[Online: <http://www.w3.org/People/Berners-Lee/FAQ.html>] o.O.

[Berners-Lee u. Lassila 2001]

Berners-Lee, T. u. Lassila, O.: The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In: Scientific American (May 2001)

[Online: <http://www.sciam.com/2001/0501issue/0501berners-lee.html>] o.O.

[Bernstein 1988]

Bernstein, M.: The bookmark and the compass: Orientation tools for hypertext users. In: SIGOIS Bulletin. Vol.9 No.4 o.O. 1988

[Bernstein 1999]

Bernstein, M.: Structural Patterns and Hypertext Rethoric.

[Online: http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/45.html] o.O. 1999

[Bieber 1999]

Bieber, M.: Hypermedia: A Design Philosophy.

[Online: http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/55.html] o.O. 1999

[Blum u. Bucher 1998]

Blum, J. u. Bucher, H.-J.: Die Zeitung: Ein Multimedia. Textdesign – ein Gestaltungskonzept für Text, Bild und Grafik. Konstanz 1998

[Bolter 1989]

Bolter, J. D.: Beyond Word Processing: The Computer an a New Writing Space.

In: Language & Communication 9 2/3. O.O. 1989

[Bolter 1991]

Bolter, J. D.: Writing Space. The Computer, Hypertext, and the History of Writing. Hillsdale 1991

[Brailsford 1999]

Brailsford, D. F.: Separable hyperstructure and delayed link binding.

[Online: http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/56.html] o.O. 1999

[Bransford 1979]

Bransford, J. D.: Human cognition : Learning, understanding and remembering. Belmont 1979

[Brinker 2001]

Brinker, K.: Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. Berlin 2001

[Bromme 2002]

Bromme, R.: Writing hypertext and learning : conceptual and empirical approaches. London 2002

[Brown 1991]

Brown, H.: Hypermedia, hypertext and object oriented databases. London 1991

[Bry u. Kraus 2001]

Bry, F. u. Kraus, M.: Advanced Modeling and Browsing of Technical Documents.

[Online: <http://www.pms.Informatik.uni-muenchen.de/publikationen/PMS-FB/PMS-FB-2001-11.pdf>] 2001

[Bucher 1999]

Bucher, H.-J.: Die Zeitung als Hypertext. Verstehensprobleme und Gestaltungsprinzipien für Online-Zeitungen. In: Lobin, Henning (Hrsg.): Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologie und Hypertext Engineering. S. 9-32. Konstanz 1999

[Bucher 2001]

Bucher, H. J.: Von der Verständlichkeit zur Usability. Rezeptionsbefunde zur Nutzung von Online-Medien. In: Osnabrücker Beiträge zur Sprachtheorie 63 S. 45-66. Osnabrück 2001

[Bush 1945]

Bush, V. (1945): As We May Think. In: Atlantic Monthly 176 (1). o.O. 1945

[Bußmann 1990]

Bußmann, H.: Lexikon der Sprachwissenschaft. Stuttgart 1990

[Cailliau 1999]

Cailliau, R.: Hypertext in the Web – a History.

[Online: http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/62.html] o.O. 1999

[Campel 1995]

Campel, K.: Coherence, continuity and cohesion. Hillsdale 1995

[Chaffin u. Herrman 1984]

Chaffin, R. u. Herrman, D.J.: The similarity and diversity of semantic relations.
In: Memory and Cognition 12 (2) S. 134.141. o.O. 1984

[Chaffin u. Herrman 1986]

Chaffin, R. u. Herrman, D.J.: Relation Element Theory: A new Account of the Representation and Processing of Semantic Relations. In: Gorgeim, D. u. Hoffmann, R. (Hrsg.): Learning and Memory: The Ebbinghaus Centennial Conference S. 221-245. Hillsday 1986

[Conklin 1987]

Conklin, J: Hypertext: An introduction and survey. In: Computer Magazin 20 o.O. 1987

[Cordes u. Streitz 1992]

Cordes, R. u. Streitz, N. (Hrsg.): Hypertext und Hypermedia. Konzepte und Anwendungen auf dem Weg in die Praxis. Berlin u.a. 1992

[Davis 1999]

Davis, H. D.: Hypertext Link Integrity.

[Online: http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/54.html] o.O. 1999

[De Beaugrande 2001]

De Beaugrande, R.-A.: Cognition and Technology.

[Online: <http://beaugrande.bizland.com/Technology.htm>] o.O. 2001

[De Beaugrande u. Dressler 1981]

De Beaugrande, R.-A. u. Dressler, W.U.: Einführung in die Textlinguistik (Konzepte der Sprach- und Literaturwissenschaft 28). Tübingen 1981

[DeRose 1999]

De Rose, S. J.: XML Linking.

[Online: http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/47.html] o.O. 1999

[Diepold u. Richter 1998]

Diepold, P. u. Richter, C.: Meta-Tags. Hilfe zum Auffinden von Web-Seiten.

[Online: <http://dbs.schule.de/bonsai/metatags.html>] o.O.

[Doland 1988]

Doland, V.M.: The hermeneutics of Hypertext. In: International Online Information Meeting, 12th. Oxford 1988

[Dölling 2001]

Dölling, E.: Multimediale Texte: Multimodalität und Multicodalität. In: Hess-Lüttich, E. W. B. (Hrsg.): Medien, Texte und Maschinen. Angewandte Mediensemiotik. Tübingen 1997

[Dorsdowski 1984]

Dorsdowski, G. (Hrsg.): Duden Grammatik der deutschen Gegenwartssprache (Der Duden in 10 Bänden; Band 4). Mannheim 1984

[Dressler 1978]

Dressler, W. (Hrsg.): Textlinguistik. Darmstadt 1978

[Duncan 1989]

Duncan, E. B.: Structuring Knowledge Bases for Designers of Learning Materials. In: Hypermedia 1 (1) London 1989

[Eisenberg 1994]

Eisenberg, P.: Grundriß der deutschen Grammatik. Stuttgart u. Weimar 1994

[Engelbart 1963]

Engelbart, D.C: A conceptual framework for the augmentation of man's intellect.

In: Howerton, P.D. u. Weeks, D.C. (Hrsg.): Vistas in information handling. Vol.1. Washigton 1963

[Engelbart u. English 1968]

Engelbart, D.C. u. English, W.K: A reseach center for augmentation of human intellect.

In: AFIPS Proceedings, Fall Foint Computer Coference o.O. 1968

[Fickert 1992]

Fickert, T.: Multimediales Lernen. Grundlagen, Konzepte, Technologien. Wiesbaden 1992

[Freisler 1994]

Freisler, S: Hypertext : ein semantischer, struktureller und medialer Vergleich zwischen linearen Texten und Hypertexten. Erlangen-Nürnberg 1994

[Freytag 1990]

Freytag, J.: Betriebsanleitung war ein ziemlicher Wälzer. Hypertext und das Verhalten der Informationsaufnahme. In: Börsenblatt für den deutschen Buchhandel. Fachzeitschrift für Verlagswesen und Buchhandel (Erscheinen eingestellt) 157 (24) Leipzig 1990

[Friederici 1999]

Friederici, A. D.: Language Comprehension: A Biological Perspective. Berlin u.a. 1999

[Gaggi 1997]

Gaggi, S.: From text to hypertext. Decentering the subject in fiction, film, the visual arts and electronic media. Philadelphia 1997

[Genette 1989]

Genette, G.: Paratexte. Das Buch vom Beiwerk des Buches. Frankfurt / Main 1989

[Genette 1993]

Genette, G.: Palimpseste. Die Literatur auf zweiter Stufe. Frankfurt / Main 1993

[Gerdes 1997]

Gerdes, H.: Lernen mit Hypertext (Aktuelle psychologische Forschung 18). Berlin u.a. 1997

[Gloor u. Saxer 1990]

Gloor, P. A. u. Saxer K.H.: Navigation im Hyperraum: Fisheye Views in Hypercard. In: Gloor, P.A. u. Streitz, N.A. (Hrsg.): Hypertext und Hypermedia. Von theoretischen Konzeptionen zur praktischen Anwendung. Berlin 1990

[Gloor u. Streitz 1990]

Gloor, P. A. u. Streitz, N.A. (Hrsg.): Hypertext und Hypermedia. Von theoretischen Konzepten zur praktischen Anwendung. Berlin u.a. 1990

[Glowalla 1992]

Glowalla, U.: Hypertext und Multimedia. Berlin 1992

[Glück 2000]

Glück, H.: Metzler Lexikon Sprache (CD-ROM: Elektronische Bibliothek Bd. 34). Stuttgart 2000

[Gülich u. Raible]

Gülich, E. u. Raible, W.: Linguistische Textmodelle. Grundlagen und Möglichkeiten. München 1977

[Haake 1997]

Haake, J. M.: Autorensysteme für die kooperative Erstellung von Hyperdokumenten. München 1997

[Hammwöhner 1993]

Hammwöhner, R.: Kognitive Plausibilität: Vom Netz im (Hyper)Text zum Netz im Kopf. In: Nachrichten für Dokumentation. 44 o.O. 1993

[Hamp u. Feldweg 1997]

Hamp, B. u. Feldweg, H.: Germanet – a lexical-semantic net for german. In: Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resource for NLP Applications. Madrid 1997

[Hasebrook 1995]

Hasebrook, J.: Multimediapsychologie: eine neue Perspektive menschlicher Kommunikation. Heidelberg, Berlin u. Oxford 1995

[Helbig 1997]

Helbig, G.: Zu Problemen der linguistischen Beschreibung des Dialogs im Deutschen. In: Deutsch als Fremdsprache 12 o.O. 1997

[Holly u. Biere 1998]

Holly, W u. Biere, B. U. (Hrsg.): Medien im Wandel. Opladen u.a. 1998

[Horn 1989]

Horn, R. E.: Mapping Hypertext. Analysis, Linkage and Display of Knowledge for the Next Generation of On Line Text and Graphics. Lexington 1989

[Idensen 1993]

Idensen, H.: Hypertext als Utopie: Entwürfe postmoderner Schreibweisen und Kulturtechniken. In: Nachrichten für Dokumentation. 44 o.O. 1993

[Irlter u. Barbieri 1991]

Irlter, W. u. Barbieri, G.: Farbmarkierung im Hypertext als Orientierungs- und Lernhilfe. In: Maurer, H. (Hrsg.): Hypertext/Hypermedia '91 Berlin 1991

[Isenberg 1970]

Isenberg, H.: Der Begriff „Text“ in der Sprachtheorie.
In: Bericht 8 der Arbeitsgruppe Strukturelle Grammatik Berlin 1970

[Isenberg 1976]

Isenberg, H.: Einige Grundbegriffe für eine linguistische Texttheorie.
In: Danes, F. u. Viehweger, D. (Hrsg.): Probleme der Textgrammatik I Berlin 1976

[Jackendoff 1983]

Jackendoff, R.: Semantics and Cognition. Cambridge 1983

[Jonassen 1989]

Jonassen, D.H.: Hypertext/Hypermedia. New York 1989

[Jonassen 1990]

Jonassen, D.H.: Designing hypermedia for learning. Heidelberg 1990

[Jonassen 1993]

Jonassen, D.H.: Structural knowledge. Techniques for representation, conveying, and acquiring structural knowledge. Hillsdale 1993

[Kallmeyer 1974]

Kallmeyer, W. u.a. (Hrsg.): Lektürekolleg zur Textlinguistik (2 Bde.). Frankfurt 1974

[Kintsch u. van Dijk 1978]

Kintsch, W. u. van Dijk, T.: Toward a model of text comprehension: A construction-integration model. In: Psychological Review 85 S. 363-394. o.O. 1978

[Kintsch u. van Dijk 1983]

Kintsch, W. u. van Dijk, T.: Strategies of discourse comprehension. London 1983

[Kleiber 1998]

Kleiber, G.: Prototypensemantik. Eine Einführung. Tübingen 1998

[Klein u. Stutterheim 1987]

Klein, W. u. von Stutterheim, C.: Quaestio und referentielle Bewegung in Erzählungen.
In: Linguistische Bericht 109 o.O. 1987

[Klein u. Stutterheim 1991]

Klein, W. u. von Stutterheim, C.: Text structure and referential movement.
In: Sprache und Pragmatik 22 o.O. 1991

[Knott 1998]

Knott, A.: The Underlying theory: Rhetorical Structure Theory. In: Online Manual ILEX
[Online: <http://cirrus.dai.ed.ac.uk:8000/illex/Manual/ts/DiscStruct/discstruct/node1.html>] o.O. 1998

[Kuhlen 1991]

Kuhlen, R.: Hypertext. Ein nicht lineares Medium zwischen Buch und Wissensbank. Berlin 1991

[Kuhlen u.a. 1989]

Kuhlen, R., Böhlen, M., Diefenbach, M., Reck, W. u. Weber, H.: Hypertext – Grundlagen und Funktionen der Entlinearisierung von Text. Teil I: Modellierung und Realisierung einer Hypertextbasis in einem Ausbildungssystem. In: Nachrichten für Dokumentation, 40, 5 Darmstadt 1989

[Labarta Postigo 1996]

Labarta Postigo, M.: Linguistische Eigenschaften und Verständlichkeit von Software-Dokumentationen: eine anwendungsorientierte Untersuchung (Europäische Hochschulschriften: Reihe 21, Linguistik Bd. 180). Berlin u.a. 1996

[Lakoff 1987]

Lakoff, G.: Women, Fire and Dangerous Things. What Categories Reveal About the Mind. Chicago u. London 1987

[Landow 1992]

Landow, G. P.: Hypertext. The Convergence of Contemporary Critical Theory and Technology. Baltimore u. London 1992

[Landow 1994]

Landow, G.P.: Hypertext theory. Baltimore u.a. 1994

[Lennon 1997]

Lennon, J. A.: Hypermedia Systems and Applications: World Wide Web and Beyond. Berlin u.a. 1997

[Linke u.a. 1991]

Linke, A., Nussbaumer, M. u. Portmann, P.R.: Studienbuch Linguistik (Reihe Germanistische Linguistik 121). Tübingen 1991

[Lobin 1999]

Lobin, H.: Intelligente Dokumente. Linguistische Repräsentation komplexer Inhalte für die hypermediale Wissensvermittlung. In: Lobin, Henning (Hrsg.): Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologie und Hypertext Engineering. S. 155-178. Opladen/Wiesbaden 1999

[Lobin 1999a]

Lobin, H. (Hrsg.): Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologie und Hypertext Engineering. Opladen/Wiesbaden 1999

[Lucarell u. Zanzi 1993]

Lucarella, D. u. Zanzi, A.: Information Retrieval from Hypertext: An approach using plausible inference. In: Information Processing & Management Vol. 29, No. 3 New York 1993

[Mackeldey 1987]

Mackeldey, R.: Alltagssprachliche Dialoge. Leipzig 1987

[Mann 1999]

RST Homepage.

[Online: <http://www.sil.org/linguistics/rst/>] o.O. 1999

[Mann 1999a]

Mann, B.: An Introduction to Rhetorical Structure Theory (RST).
[Online: <http://www.sil.org/linguistics/rst/rintro99.htm>] o.O. 1999

[Mann 1999b]

Mann, B.: RST Relation Definitions. A Tool for Analyzing Text.
[Online: <http://www.sil.org/linguistics/rst/reldefs.htm>] o.O.

[Mann u. Thompson 1987]

Mann, W.C. u. Thompson, S.A.: Rhetorical Structure Theory: Description and Construction of Text Structures. In: Kempen, G. (Hrsg.): Natural Language Generation. New Results in Artificial Intelligence, Psychology and Linguistics S. 84-95. Dordrecht 1987

[Marchionini u. Shneiderman 1988]

Marchionini, G.; Shneiderman, B.: Finding facts versus browsing knowledge in
Finding facts versus browsing knowledge in hypertext systems. In: IEEE Computer 1988/1 o.O. 1988

[Marshall u. Irish 1989]

Marshall, C.C. u. Irish, P.M.: Guided tours and on-line presentations: How authors make existing hypertext intelligible for readers. In: ACM-Hypertext 1989. New York 1989

[McKnight u.a.1993]

McKnight, C., Dillon, A. u. Richardson, J. (Hrsg.): Hypertext. A psychological perspective. New York u.a. 1993

[Meibauer 2001]

Meibauer, J.: Pragmatik: Eine Einführung. Tübingen 2001

[Meyerhoff 1994]

Meyerhoff, D. B.: Hypertext und tutorielle Lernumgebungen: ein Ansatz zur Integration. München u.a. 1994

[Moulthrop 1991]

Moulthrop, S.: Beyond the Electronic Book: A Critique of Hypertext Rhetoric. In: Proceedings of the ACM Conference on Hypertext. San Antonio 1991

[Münz 1997a]

Münz, S.: Hypertext. [Online: <http://www.teamone.de/hypertext/>] o.O. 1997

[Münz 1997b]

Münz, S.: SELFHTML. Die Energie des Verstehens: HTML-Dateien selbst erstellen. Version 7
[Online: <http://www.teamone.de/selfhtml/>] o.O. 1997

[Münz u. Nefzger 1999]

Münz, S. u. Nefzger, W.: HTML 4.0 Handbuch. HTML, JavaScript, DHTML, Perl. München 1999

[Münz 2001]

Münz, S.: SELFHTML. Die Energie des Verstehens: HTML-Dateien selbst erstellen. Version 8
[Online: <http://selfhtml.teamone.de/>] o.O. 2001

[Münz 2002]

Münz, S.: HTML & Web-Publishing Handbuch (Gesamtausgabe in 2 Bd.). München 2002

[Negroponte 1995]

Negroponte, N.: Being digital. London 1995

[Nelson 1967]

Nelson, T.: Hypertext Note 1: Brief Words on Hypertext. o.O. 1967

[Nelson 1974]

Nelson, T.: Dream Machines: new freedoms through computer screens. A minority Report. Computer Lib: You can and must understand computers now. Chicago 1974

[Nelson 1987]

Nelson, T.: Literary machines. o.O. 1987

[Neubert u. Schmidt 1994]

Neubert, S. u. Schmidt, G.: Hypertext und künstliche Intelligenz. Aktuelle Trends 1993 (Universität Karlsruhe Forschungsberichte 289). Karlsruhe 1994

[Nielsen 1990]

Nielsen, J.: Hypertext and Hypermedia. Boston u.a. 1990

[Nielsen 2000]

Nielsen, J.: Designing web usability: the practice of simplicity. Indianapolis 2000

[Nix u. Spiro 1990]

Nix, D. u. Spiro, R. (Hrsg.): Cognition, Education and Multimedia. Exploring Ideas in High Technology. Hillsdale 1990

[Nyce u. Kahn 1991]

Nyce, J. / Kahn, P. (Hrsg.): From Memex to Hypertext: Vannevar Bush and the Mind's Machine. London 1991

[Ott 1996]

Ott, S.: Technische Dokumentation im Unternehmen: Grundlagen und Fallbeispiele. (Beiträge zur Betriebslinguistik 1). Paderborn 1996

[Pörings u. Schmitz 1999]

Pörings, R. u. Schmitz, U.: Sprache und Sprachwissenschaft. Eine kognitiv orientierte Einführung. Tübingen 1999

[Quasthoff 1980]

Quasthoff, U. M.: Erzählen in Gesprächen. Tübingen 1980

[Rada 1991]

Rada, R.: Hypertext. Form Text to Experttext. London 1991

[Rada 1995]

Rada, R.: Interactive Media. New York 1995

[Ramm 1994]

Ramm, M.: Interactive Multimedia and Narrative Art. o.O. 1994

[Rapp u. Zumbach 2001]

Rapp, A. u. Zumbach, J.: Wissenserwerb mit Hypermedien. Eine kognitionswissenschaftliche Betrachtung. In: Osnabrücker Beiträge zur Sprachtheorie 63 S. 45-66. Osnabrück 2001

[Reichenberger u.a. 1995]

Reichenberger, K., Rondhuis, K. J. , Kleinz, J. u. Bateman, J.: Effective Presentation of Information Through Page Layout: a Linguistically-Based Approach. In: Electronic Proceedings of the ACM Workshop on Effective Abstractions in Multimedia
[Online: <http://www.eecs.tufts.edu/~isabel/reichen/page-layout.html>] o.O. 1995

[Reimer 1991]

Reimer, U.: Einführung in die Wissensrepräsentation. Netzartige und schemabasierte Repräsentationsformate. Stuttgart 1991

[Retterer 1991]

Retterer, O.J.: Learning from hypertext: The effect of reading interactive text containing non-sequential, associate linkages on comprehension. Toledo 1991

[Rickheit u. Bock 1983]

Rickheit, G. u. Bock, M. (Hrsg.): Psycholinguistic studies in language processing (Research in text theory 7). Berlin u.a. 1983

[Rickheit u. Strohner 1993]

Rickheit, G. u. Strohner, H.: Grundlagen der kognitiven Sprachverarbeitung. Modelle, Methoden, Ergebnisse. Stuttgart 1993

[Riggert 1998]

Riggert, W.: Betriebliche Informationskonzepte: von Hypertext zu Groupware. Braunschweig u. Wiesbaden 1998

[Rösner u. Stede 1993]

Rösner, D. u. Stede, M.: Untersuchungen zur Struktur von Texten: RST-Analysen deutscher Texte. In: KI – Künstliche Intelligenz: Forschung, Entwicklung, Erfahrungen Baden-Baden 1993

[Rouet 1996]

Rouet, J.-F.: Hypertext and cognition. New York 1996.

[Runkehl u.a. 1998]

Runkehl, J., Schlobinski, P. u. Siever, T.: Sprache und Kommunikation im Internet. Überblick und Analysen. Opladen u.a. 1998

[Sandig 1973]

Sandig, B.: Beispiele pragmalinguistischer Textanalyse. In: Der Deutschunterricht 25, Heft 1 S. 5-23. o.O. 1973

[Sarre 1991]

Sarre, F.: Zur Generierung von Hypertextstrukturen aus linearen Texten durch Partitionierung und Strukturanreicherung. München 1991

[Schmidt 1973]

Schmidt, S.J.: Texttheorie. Probleme einer Linguistik der sprachlichen Kommunikation. München 1973

[Schmitz 1995]

Schmitz, U.: Neue Medien und Gegenwartssprache. Lagebericht und Problemskizze. [Online: http://www.linse.uni-essen.de/obst/obst_inh.htm] o.O. 1995

[Schmitz 1996]

Schmitz, U.: Zur Sprache im Internet. Skizze einiger Eigenschaften und Probleme. [Online: http://www.linse.uni-essen.de/papers/sprache_internet.htm] o.O. 1996

[Schmitz 1997]

Schmitz, U.: Schriftliche Texte in multimedialen Kontexten. In: Weingarten, J. (Hrsg.): Sprachwandel durch Computer. S. 131-158. Opladen 1997

[Schnupp 1992]

Schnupp, P.: Hypertext. München u.a. 1992

[Schuler 1990]

Schuler, W.: Author's Argumentation Assistant (AAA): A Hypertext-based authoring tool for argumentative texts. In: Streitz, N. Rizk, A. u. Andre, J. (Hrsg.): Hypertext. Concepts, Systems and Applications. Proceedings of the first European Conference on Hypertext S. 137-151. Cambridge 1990

[Schwarz 1992]

Schwarz, M.: Einführung in die Kognitive Linguistik. Tübingen 1992

[Schweikle 1990]

Schweikle, G. u. Schweikle, I. (Hrsg.): Metzler-Literatur-Lexikon: Begriffe und Definitionen. Stuttgart 1990

[Searle 1969]

Searle, J. R.: Speech acts. An essay in the philosophy of language. Cambridge 1969

[Shneiderman u.a. 1991]

Shneiderman, B., Kreitzberg, C. u. Berk, E.: Editing to structure a readers experience. In: Berg, E. u. Devlin, J.: Hypertext/hypermedia handbook S. 55-64. McGraw Hill, New York 1991

[Sowarka u.a. 1983]

Sowarka, B., Abel, U. u. Michel, J.: Menschliche Textverarbeitung und propositionale Analyse: eine Anleitung zur propositionalen Darstellung von Texten. Berlin 1983

[Sowinski 1983]

Sowinski, B.: Textlinguistik. Eine Einführung. Stuttgart 1983

[Stevick 1970]

Stevick, Ph: The chapter in fiction. Theories of narrative division. New York 1970

[Storrer 1999]

Storrer, A.: Kohärenz in Text und Hypertext.

In: Lobin, Henning (Hrsg.): Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologie und Hypertext Engineering S. 33-66. Opladen u. Wiesbaden 1999

[Storrer 2000]

Storrer, Angelika: Was ist "hyper" am Hypertext.

[Online: <http://www.ids-mannheim.de/grammis/storrer/hyper.pdf>] o.O. 2000

[Storrer 2001a]

Storrer, Angelika: Digitale Wörterbücher als Hypertexte: Zur Nutzung des Hypertextkonzepts in der Lexikographie.

[Online: <http://www.ids-mannheim.de/grammis/storrer/wobu.pdf>] o.O. 2001

[Storrer 2001b]

Storrer, Angelika: Schreiben, um besucht zu werden: Textgestaltung fürs World Wide Web.

[Online: <http://www.ids-mannheim.de/grammis/storrer/trier.pdf>] o.O. 2001

[Storrer 2001c]

Storrer, Angelika: Neue Medien – neue Stilfragen: Das World Wide Web unter stilistischer Perspektive. [Online: <http://www.ids-mannheim.de/grammis/storrer/stil.pdf>] o.O. 2001

[Strohner 1990]

Strohner, H.: Textverstehen. Kognitive und kommunikative Grundlagen der Sprachverarbeitung. Wiesbaden 1990

[Sucharowski 1996]

Sucharowski, W.: Sprache und Kognition. Neuere Perspektiven in der Sprachwissenschaft. Wiesbaden 1996

[Tergan 1993]

Tergan, S.O.: Zum Aufbau von Wissensstrukturen mit Texten und Hypertexten. In: Nachrichten für Dokumentation. 44. o.O. 1993

[Toulmin 1975]

Toulmin, St.: Der Gebrauch von Argumenten. Kronberg 1975

[Unz 2000]

Unz, D.: Lernen mit Hypertext. Informationssuche und Navigation. Münster 2000

[Utting u. Yankelovich 1989]

Utting, K. u. Yankelovich, N.: Context and orientation in hypermedia networks. In: ACM Transactions on Information Systems 7 o.O. 1989

[van Dijk 1980a]

van Dijk, T. A.v.: Textwissenschaft. Eine interdisziplinäre Einführung. Tübingen 1980

[van Dijk 1980b]

van Dijk, T. A.v.: Macrostructures. Hillsdale 1980

[Vater 1994]

Vater, H.: Einführung in die Textlinguistik. München 1994

[W3C 1990]

W3C: Design Issues.

[Online: <http://www.w3.org/DesignIssues/>] o.O. 1990

[W3C 1998]

W3C: HTML 4.0 Specification. W3C Recommendation, revised on 24-Apr-1998 W3C Recommendation, revised on 24-Apr-1998.

[Online: <http://www.w3.org/TR/1998/REC-html40-19980424>] o.O. 1998

[W3C 2001]

W3C: Resource Description Framework (RDF) Schema Specification 1.0. W3C Candidate Recommendation 27 March 2000.

[Online: <http://www.w3.org/TR/2000/CR-rdf-schema-20000327>] o.O. 2001

[Wagner 1998]

Wagner, F.: Sind Printmedien im Internet Online-Medien?. In: Pfammatter, R. (Hrsg.): Multi-Media-Mania. Reflexionen zu Aspekten neuer Medien. S. 191-211. Konstanz 1998

[Wawrzyniak 1980]

Wawrzyniak, Z.: Einführung in die Textwissenschaft. Probleme der Textbildung im Deutschen. Warschau 1980

[Wedeles 1965]

Wedeles, L.: Prof. Nelson Talk Analyzes P.R.I.D.E. In: VASSAR MISCELLANY NEWS February 3, 1965
[zitiert nach der Online-Version: http://iberia.vassar.edu/~mijoyce/MiscNews_Feb65.html] o.O. 1965

[Weingarten 1997]

Weingarten, R.: Textstrukturen in neuen Medien: Clusterung und Aggregation.
In: Weingarten, J. (Hrsg.): Sprachwandel durch Computer S. 215-238. Opladen 1997

[Weingarten 1997a]

Weingarten, R. (Hrsg.): Sprachwandel durch Computer. S. 215-238. Opladen 1997

[Wenz 2001]

Wenz, K.: Zeichen lesen: Hypertext revisited. In: Hess-Lüttich, E. W. B. (Hrsg.): Medien, Texte und Maschinen. Angewandte Mediensemiotik. Wiesbaden 2001

[Wenz 2001a]

Wenz, K.: Cybertextspace.
[Online: <http://65.107.211.206/cpace/ht/wenz/contents.html>] o.O. 2001

[Wingert 1993]

Wingert, B.: Äußerer und innerer Hypertext: Eine notwendige Differenzierung, verdeutlicht am FlusserHypertext. In: Nachrichten für Dokumentation. 44 o.O. 1993

[Woodhead 1991]

Woodhead, N.: Hypertext and hypermedia. Wilmslow u.a. 1991

[Zimmer 1992]

Zimmer, D. E.: Die Elektrifizierung der Sprache. Zürich 1992

