

LEXICOGRAPHICA Series
Maior

LEXICOGRAPHICA

Series Maior

Supplementary Volumes to the International Annual for Lexicography
Suppléments à la Revue Internationale de Lexicographie
Supplementbände zum Internationalen Jahrbuch für Lexikographie

Edited by

Sture Allén, Pierre Corbin, Reinhard R. K. Hartmann,
Franz Josef Hausmann, Ulrich Heid, Oskar Reichmann,
Ladislav Zgusta

104

Published in cooperation with the Dictionary Society of North America
(DSNA) and the European Association for Lexicography (EURALEX)

Matthias Kammerer

Lemmazeichentypen für deutsche Verben

Eine lexikologische und
metalexikographische Untersuchung

Max Niemeyer Verlag
Tübingen 2000



Für Bertel und Hartmut

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

[*Lexicographica / Series maior*]

Lexicographica : supplementary volumes to the International annual for lexicography / publ. in cooperation with the Dictionary Society of North America (DSNA) and the European Association for Lexicography (EURALEX). Series maior. – Tübingen : Niemeyer.

Früher Schriftenreihe

Reihe Series maior zu: Lexicographica

104. Kammerer, Matthias: Lemmazeichentypen für deutsche Verben. – 2000

Kammerer, Matthias: Lemmazeichentypen für deutsche Verben : eine lexikologische und metalexikographische Untersuchung / Matthias Kammerer. – Tübingen : Niemeyer, 2000

(Lexicographica : Series maior ; 104)

Zugl.: Heidelberg, Univ., Diss., 1999

ISBN 3-484-39104-9 ISSN 0175-9264

© Max Niemeyer Verlag GmbH, Tübingen 2000

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Printed in Germany.

Gedruckt auf alterungsbeständigem Papier.

Druckvorlage: Matthias Kammerer, Karlsruhe

Druck: Weihert-Druck GmbH, Darmstadt

Einband: Industriebuchbinderei Nädele, Nehren

Inhalt

Danksagung	VII
1 Vorbemerkungen	I
2 Was sind Lemmazeichentypen?	5
2.1 Die Konzeption des Wörterbuchartikeltyps	6
2.2 Die Konzeption des Lemmazeichentyps	15
2.3 Korrelation zwischen Lemmazeichentyp und Wörterbuchartikeltyp und weitere terminologische Unterscheidungen beim Lemmazeichentyp	20
2.4 Ein abschließender Blick auf ausgewählte metalexikographische Eigenschaftsausprägungen.	30
3 Sind Bedeutungen im Wörterbuch?	39
3.1 Bedeutung und Gebrauch	39
3.2 Die hiesige Konzeption des sprachlichen Zeichens	45
3.2.1 Prolegomena zum sprachlichen Zeichen	46
3.2.2 Festlegung des sprachlichen Zeichens	67
4 Konfliktbedingte Situationstypen von Wörterbuchkonsultationen bei Verben für allgemeine einsprachige Wörterbücher.	73
4.1 Typologie von konfliktbedingten Situationstypen von Wörterbuchkonsultationen.	74
4.2 Interdependenz zwischen Situationstypen von Wörterbuchkonsultationen, Fragehandlungen und Angabeklassen	99
5 Verbklassifikationen und exemplarische Teilkonstruktion von Verb-Lemmazeichentypen für WB-10	113
5.1 Einführender Überblick	113
5.2 Verbklassifikationen in ausgewählten Grammatiken	121
5.2.1 Morphologische Klassifizierungen.	122
5.2.2 Syntaktische Klassifizierungen	135
5.2.3 Semantische Klassifizierungen	159
5.3 Der Verthesaurus von BALLMER/BRENNENSTUHL	190
5.3.1 Einführende Darstellung	190
5.3.2 Lexikalsemantische Relationen	197
5.3.3 Festlegung der Akionalität für Verben des Deutschen	213
5.4 Semantische Verbklassifikationen und -analysen in ausgewählter Literatur	215
5.4.1 Die Verbklassifikation bei ENDRES (1973)	216
5.4.2 Die Verbklassifikation bei LEISI (1975)	219
5.4.3 Semantische Analyse der Handlungsverben bei SZLEK (1980)	224
5.4.4 Semantische Analyse der inchoativen Verben bei STORCH (1978)	226

5.4.5	Semantische Analyse der Aufforderungsverben bei VAN DER ELST (1982)	232
5.4.6	Semantische Analyse von Gebrauchsverben bei DITTMER (1985)	234
5.4.7	Semantische Analyse von Verben des Verfügungs- und Besitzwechsels bei BITTER (1991)	236
5.5	Semantische Kategorisierungen in sog. onomasiologischen Wörterbüchern.	239
5.6	Überprüfung der eruierten linguistischen Eigenschaftsausprägungen	250
6	Zur Bedeutungseruierung und -beschreibung in lexikographischen Prozessen	259
6.1	Zu Semantisierungsbemühungen und zur Bedeutungseruierung von thematischen Exzerptsegmenten.	259
6.1.1	Die Methode der Semantisierungsbemühungen	259
6.1.2	Die Methode der lexikographischen Bedeutungseruierung.	263
6.1.2.1	Beispiel für die Methode der lexikographischen Bedeutungseruierung	271
6.2	Bemerkungen zu Bedeutungsangaben in Lernerwörterbüchern	287
6.2.1	Enzyklopädisches und semantisches Wissen	287
6.2.2	Die Matrixframes EREIGNIS, HANDLUNG und ZUSTAND und ihre Funktion bei der Eruiierung des enzyklopädischen gegenstands- konstitutiven Bedeutungswissens.	294
7	Und zum Schluß: Zwei Wörterbuchartikel zu <i>wohnen</i>	325
7.1	Die linguistischen Eigenschaftsausprägungen von <i>wohnen</i>	325
7.2	Zwei Vorschläge zu einem Wörterbuchartikel zu <i>wohnen</i>	328
8	Anhang.	341
8.1	Verzeichnis der Siglen und Abkürzungen	341
8.2	Verzeichnis der Abbildungen.	349
8.3	Verzeichnis der Tabellen.	351
8.4	Verzeichnis der Wörterbuchartikel und -einträge.	352
8.5	Verzeichnis der Definitionen	353
8.6	Verzeichnis der Zitate.	354
9	Literatur und Corpora	357
9.1	Corpora	357
9.2	Bibliographien	358
9.3	Wörterbücher	359
9.4	Sonstige Literatur	369
	Namenindex.	395
	Sachindex	403
	Abstract.	421
	Résumé.	423

Danksagung

Zum Gelingen dieser Arbeit haben viele beigetragen, bei denen ich mich an dieser Stelle ganz herzlich bedanken möchte, denn ohne ihre wohlwollende Unterstützung, ihre Hinweise, Kritiken, Vorschläge und praktischen Tips wäre diese Arbeit nicht so weit gediehen, wie sie es jetzt ist.

Zuerst möchte ich mich bei meinen Eltern, Bertel und Hartmut, denen diese Arbeit auch gewidmet ist, dafür bedanken, daß sie mir die Voraussetzungen geschaffen haben, in Ruhe und mit Konzentration arbeiten zu können. Unauffällig haben sie wesentlich zum Gelingen der Arbeit beigetragen.

Darüber hinaus sind insbesondere zu nennen (in alphabetischer Reihenfolge): Dr. KLAUS-PETER KONERDING, der nicht nur ausgewählte Teile dieses Buches kritisch gelesen, sondern mir auch in vielen Gesprächen zu mehr Klarheit verholfen und mir wichtige Verbesserungsvorschläge unterbreitet hat; Prof. Dr. JOACHIM KORNELIUS, der sich dankenswerterweise sofort bereit erklärte, die Aufgabe des Coreferenten zu übernehmen; Prof. Dr. KLAUS MUDERSBACH, der sich viel Zeit genommen hat, den Abschnitt 6.1 genau zu studieren und mit mir zu besprechen, so daß ich noch manche wesentliche Anregung übernehmen und einarbeiten konnte; Dr. ANGELIKA STORRER, die freundlicherweise einige ausgewählte Ausschnitte prüfend gelesen hat, und schließlich Prof. Dr. Dr. h.c. HERBERT ERNST WIEGAND, der meine Studien nicht nur kritisch begleitet, sondern auch sehr intensiv betreut hat und mir in zahlreichen Besprechungen helfend und korrigierend zur Seite stand.

Schließlich möchte ich mich noch bei den Teilnehmerinnen und Teilnehmern des Doktorandenkolloquiums am Germanistischen Seminar der Universität Heidelberg für die konstruktiven Diskussionen bedanken und besonders herzlich bei Frau LARISSA BRÜTT (M.A.) für das genaue und schnelle Korrekturlesen des Typoskripts (wenn dennoch Fehler stehengeblieben sind, so ist das allein auf mich zurückzuführen).

Nicht zuletzt gilt mein Dank den Herausgebern von Lexicographica, Series Maior, und dem Max Niemeyer Verlag dafür, daß diese Arbeit in dieser wichtigen Reihe erscheinen konnte. Hier sind insbesondere Prof. Dr. FRANZ JOSEF HAUSMANN, Prof. Dr. OSKAR REICHMANN und Prof. Dr. LADISLAV ZGUSTA zu nennen.

Ihnen allen sei hier mein verbindlichster Dank ausgesprochen.

Karlsruhe, im Januar 2000

M. K.

I Vorbemerkungen

[...] die Spur der Sprache [...] – eine verwischte, aber verräterische Spur!

ROBERT MUSIL: Der Mann ohne Eigenschaften. Reinbek bei Hamburg 1986, p. 559

In dieser Arbeit geht es um die lexikologische und lexikographische Beschreibung der (Voll-) Verben der deutschen Standardsprache der Gegenwart im Hinblick auf allgemeine einsprachige Wörterbücher. Dazu sollen die Lemmazeichentypen, denen die Verben des Deutschen zugeordnet werden können, konstruiert werden.

Um dies durchführen zu können, wird in Kapitel 2 geklärt, was hier unter einem Lemmazeichentyp verstanden wird und wie sich die Korrelationen mit Wörterbuchartikeltypen gestalten. Demgemäß werden Lemmazeichentypen einerseits durch *linguistische* und andererseits durch *metalexikographische Eigenschaftsausprägungen* bestimmt. In dieser Arbeit wird nur auf erstere, die linguistischen Eigenschaftsausprägungen, näher eingegangen; für letztere wird nur ein cursorischer Überblick gegeben.

Eine linguistische Eigenschaftsausprägung ist immer eine Eigenschaftsausprägung *von* bzw. *zu etwas*. Insofern wird eine linguistische Eigenschaftsausprägung als ein 2-Tupel begriffen, das aus einem Attribut und einem Wert besteht, wobei beide theorieabhängig sind. Das drückt sich darin aus, daß die Menge der linguistischen Eigenschaftsausprägungen immer hinsichtlich einer *linguistischen Überdachung* (cf. Definition D 2-6, p. 17; dies ist z.B. eine Liste linguistischer Literatur oder Grammatiken und den darin ex- oder implizit enthaltenen Aussagen über sprachliche Phänomene) konstruiert wird und von dieser abhängig ist. Das hat zur Folge, daß es keine linguistischen Eigenschaftsausprägungen *per se* gibt, sondern daß diese grundsätzlich durch die linguistische Überdachung bestimmt werden.

Den linguistischen Eigenschaftsausprägungen können bestimmte *Angabe-* und *Kennzeichnungsklassen* korreliert werden. *Kennzeichnungsklassen* sind Klassen von Textsegmenten, die nur funktional, aber nicht positional segmentiert werden können. *Angabeklassen* sind Mengen von (konkreten) Angaben mit allgemeinem genuinen Zweck, die funktional-positional segmentiert werden können und daher immer eine Angabeform aufweisen. Auf Grund dieser Eigenschaft gehören zu den konkreten Angaben auch die typographischen Strukturanzeiger. Bei den Angabe- sowie bei den Kennzeichnungsklassen hingegen spielen die typographischen Strukturanzeiger keine Rolle.

Diese Korrelation von linguistischen Eigenschaftsausprägungen mit Angabe- und Kennzeichnungsklassen ist deshalb von größter Wichtigkeit, da sie den Anknüpfungspunkt zu den Wörterbuchartikeltypen darstellt, denn Wörterbuchartikeltypen werden hier als eine Menge von *n* abstrakten hierarchischen Mikrostrukturen begriffen. Auf diese Weise können den Lemmazeichentypen – bei denen verschiedene Arten unterschieden werden können – Wörterbuchartikeltypen zugeordnet werden.

In Kapitel 3 wird die aus forschungsstrategischen Gesichtspunkten notwendige Frage nach der Bedeutung gestellt, da sie eine zentrale Rolle bei der lexikographischen Beschreibung von Verben in allgemeinen einsprachigen Wörterbüchern spielt. Es wird dabei von der neuesten Konzeption von WIEGAND (1999a) ausgegangen und diese genau beschrieben. Im Anschluß daran wird dargelegt, was unter einem Lexem (i.S.v. sprachliches Zeichen) in dieser Arbeit verstanden werden soll (*cf.* Def. D 3-2, p. 67). Danach weist ein Lexem sechs »Aspekte« auf, die jedoch nicht als *Teile* eines Lexems zu verstehen sind; es sind dies: (1) die Wortart, (2) das morpho-semantiche Formenparadigma, (3) die Bedeutung(en), (4) die diasystematischen Markierungen, (5) das syntaktische Verhalten und (6) syntagmatische Einheiten, die mit dem Lexem korreliert sind. Es ist klar, daß eine solche Konzeption insbesondere im Hinblick auf Verben erstellt wurde und nicht beispielsweise im Hinblick auf Partikel.

Bevor nun die Lemmazeichentypen für allgemeine einsprachige Wörterbücher eruiert werden können, wird im Kapitel 4 dargelegt, welche Angabeklassen – bezogen auf Verblexeme – in einem allgemeinen einsprachigen Wörterbuch der deutschen Standardsprache enthalten sein müssen, wenn dieses Wörterbuch in konfliktbedingten Konsultationssituationen erfolgreich benutzbar sein soll. Dazu werden konfliktbedingte Situationstypen von Wörterbuchkonsultationen bei Verben für allgemeine einsprachige Wörterbücher im Anschluß an WIEGAND (1998) wiedergegeben, wobei Exhaustivität lediglich angestrebt wird, jedoch nicht erreichbar ist (dies drückt sich in den graphischen Darstellungen durch ein Dreieck „ Δ “ aus, das für einen nicht ausgeführten Teilbaum steht). Der Menge der konfliktbedingten Situationstypen von Wörterbuchkonsultationen bei Verben können dann Mengen von Fragen gegenübergestellt werden, auf die Wörterbuchangaben (und Kennzeichnungen) eine Antwort geben, sofern sie gemäß ihrem genuinen Zweck interpretiert werden. Diese Angabe- und Kennzeichnungsklassen seien dann konstitutiv für allgemeine einsprachige Wörterbücher der deutschen Standardsprache. Damit ist die Verbindung zwischen Situationstypen von Wörterbuchkonsultationen und Angabe- bzw. Kennzeichnungsklassen in Wörterbüchern gefunden und die Menge der Angabe- bzw. Kennzeichnungsklassen, die konstitutiv zu einem allgemeinen einsprachigen Wörterbuch gehört, festgelegt.

Im 5. Kapitel kann dann – immer bezogen auf die Ergebnisse aus Kapitel 3 (Lexemkonzeption) und 4 (konstitutive Angaben in allgemeinen einsprachigen Wörterbüchern) – ein Überblick über Klassifikationen und spezifische Charakteristika bei Verben gegeben werden, wobei diese Darstellung notgedrungen stark selektiv und interessegeleitet ist. Der rote Faden, der durch einen Großteil dieses Kapitels führt, ist die Grammatik von HELBIG/BUSCHA. An Hand dieser Grammatik werden nacheinander morphologische, syntaktische und semantische Klassifizierungen und Eigenschaften von Verben vorgestellt, mit anderen Darstellungen verglichen und diskutiert.

Ein weiterer Schwerpunkt dieses Kapitels bildet der Verbthesaurus von BALLMER/BRENNENSTUHL zusammen mit der Einleitung zu diesem Wörterbuch. Die Explikationen der Autoren werden ausführlich referiert, da gerade sie viele lexikalsemantische Relationen bei Verben beobachtet haben. Darüber hinaus werden weitere ausgewählte Verbklassifikationen kritisch vorgestellt, die m.E. in einer bestimmten Hinsicht exemplarisch sind. Jeder Abschnitt endet damit, daß der Lemmazeichentyp spezifiziert wird, und zwar immer bezogen auf die vorangegangenen Darlegungen.

Am Ende dieses Kapitels erfolgt eine systematische Verifikation und gegebenenfalls Ergänzung des konstruierten Lemmazeichentyps, so daß dieser schließlich exhaustiv präsentiert werden kann.

Im nachfolgenden Kapitel 6 wird explizit auf die Bedeutungseruierung und -beschreibung in lexikographischen Prozessen – bezogen auf einen bestimmten Situationstyp – eingegangen, wobei einerseits auf WIEGAND (1998 u. 1998a) und andererseits auf KONERDING (1993) Bezug genommen wird. Bei ersterem werden die Möglichkeiten der Bedeutungseruierung dargestellt und in einer *Methode der lexikographischen Bedeutungseruierung* systematisiert; zusätzlich wird eine phasisch gegliederte *Methode der Semantisierungsbemühungen* vorgestellt. Bei letzterem wird auf dessen Frame-Konzeption rekurriert und die Bedeutungsframes für Nomine für die hiesigen Bedürfnisse modifiziert aus der Beobachtung heraus, daß jedes Verb nominalisiert werden kann. Am Ende wird für ausgewählte Verben exemplarisch sowohl das gegenstandskonstitutive Bedeutungswissen als auch das enzyklopädische Bedeutungswissen an Hand einer framegestützten Analyse durchgeführt und die Ergebnisse in (provisorischen) framebasierten Wörterbuchartikeln präsentiert, die insbesondere hinsichtlich ihrer Form nicht als endgültig zu betrachten sind, sondern lediglich prinzipiellen Charakter haben.

Im letzten Kapitel 7 wird ein Vorschlag für einen Wörterbuchartikel zu *wohnen* unterbreitet, der erläutert und in Ansätzen diskutiert wird. Dieser Wörterbuchartikel beruht auf den Ergebnissen der Konstruktion der Lemmazeichentypen für Verben der deutschen Standardsprache, wie sie in dieser Arbeit im Hinblick auf allgemeine einsprachige Wörterbücher der deutschen Standardsprache der Gegenwart entwickelt wurden. Da dieser Wörterbuchartikel sehr umfangreich ist, wird abschließend noch eine gekürzte Version vorgestellt.

2 Was sind Lemmazeichentypen?

»Ich mußte ihnen alle möglichen Namen
geben, weißt du, damit ich in dem Buch
über sie schreiben konnte. – [...]«
KAREL ČAPEK: Der Krieg mit den Mol-
chen. Berlin, Weimar 1992, p. 48

Ein *Lemmazeichentyp* kann – zuerst einmal informell und auf rein intuitiver Basis formuliert – als eine Menge von Eigenschaften, die auf bestimmte Klassen von Lemmazeichen zutreffen, angesehen werden. Auf diese Weise können verschiedene Klassen von Lemmazeichen unterschieden werden. Jede Klasse von Lemmazeichen zeichnet sich gegenüber jeder anderen Klasse von Lemmazeichen dadurch aus, daß sie bestimmte und im Vergleich zu allen anderen Klassen von Lemmazeichen andere Angabeklassen in einem Wörterbuchartikel und/oder andere Eigenschaften eines Wörterbuchartikels (im weitesten Sinne) verlangt. Beispielsweise verlangt ein Substantiv in einem allgemeinen einsprachigen Wörterbuch des Deutschen u.a. eine Singularbildungsangabe – sofern es kein Pluraletantum ist –, während bei einem Verb eine grundsätzlich andere Morphologieangabe zu machen ist. Darüber hinaus gibt es auch Gemeinsames, z.B. die Lemmazeichengestaltungsangabe.

Eigenschaften eines Wörterbuchartikels sind beispielsweise seine Länge, der Standardisierungs- und Textverdichtungsgrad oder seine mediostrukturelle Vernetzung mit Abbildungen oder der Wörterbuchgrammatik im Vor- oder Nachspann. So kann beispielsweise die Länge des Wörterbuchartikels vom Lemmazeichentyp abhängig gemacht werden.

Ein Lemmazeichentyp kann demnach auch auf Grund der Menge der geforderten Angabeklassen und Wörterbuchartikeleigenschaften definiert werden. Er bestimmt auf diese Weise zusammen mit dem Wörterbuchprogramm, das im Instruktionbuch festgeschrieben ist, u.a. die Mikrostruktur der Wörterbuchartikel.

In Rekurs auf und teilweise in Modifikation von KAMMERER (1995, 93–96) soll nun der Begriff des *Lemmazeichentyps* genauer eingeführt und konkretisiert werden. Dazu muß in einem ersten Schritt auf den Begriff des *Wörterbuchartikeltyps* eingegangen werden.¹

1. Es wäre sehr wünschenswert gewesen, auch einen historischen Abriss darüber zu geben, welche Lemmazeichentypen von Lexikographen – bewußt oder unbewußt – unterschieden und angesetzt wurden; dies konnte aber hier nicht geleistet werden. Dennoch sei an dieser Stelle auf M. KRAMER und dessen Vorbericht zu seinem TEUTSCH-ITALIÄNISCHEN DICTIONARIUM (= KRAMER "TEUTITALDIC") hingewiesen, worin er entsprechend seiner Unterteilung der Stammwörter in *Substantive, Adjektive, Pronomen, Verben, Adverben, Interjektionen* und *Konjunktionen* zwischen sieben Lemmazeichentypen – allerdings unter einer anderen Terminologie – unterscheidet und diese detailliert expliziert (cf. KRAMER 1700[1982], o.Pag.). (Den Hinweis auf KRAMERS Vorwort verdanke ich H. E. WIEGAND; M. K.)

2.1 Die Konzeption des Wörterbuchartikeltyps

Ein gedrucktes Wörterbuch² besteht aus diversen funktionalen Komponenten, auch *Bauteile* genannt. Diese Bauteile sind *Texte* oder *Teiltexte*. Unter diesem Aspekt ist ein Printwörterbuch ein Textsortenträger, denn ein Printwörterbuch ist kein Text einer bestimmten Textsorte, sondern – aus drucktechnischen und buchbinderischen Gründen – eine Linearisierung von mehreren Texten, die verschiedenen Textsorten angehören. Ein Printwörterbuch ist deshalb ein *lexikographischer Textsortenträger* (LexTST). Unter einem *nichtfunktionalen* Gesichtspunkt weist ein Printwörterbuch als lexikographischer Textsortenträger obligatorisch entweder ein *zentrales Wörterverzeichnis* (z.WV) oder eine *Wörterverzeichnisreihe* (WVR) oder *Wörterverzeichnisblöcke* (WVB) sowie einen *Vorspann* (VorSp) auf; gegebenenfalls kann auf das zentrale Wörterverzeichnis bzw. auf die Wörterverzeichnisreihe noch ein *Nachspann* (NachSp) folgen. Dann bilden der Vorspann zusammen mit dem Nachspann relativ zum zentralen Wörterverzeichnis bzw. zur Wörterverzeichnisreihe einen *textuellen Rahmen*. Das zentrale Wörterverzeichnis bzw. die Wörterverzeichnisreihe bilden – bezogen auf den Vor- und den Nachspann – den *Binnentext*. Vorspann, Nachspann und das zentrale Wörterverzeichnis bzw. die Wörterverzeichnisreihe zusammen bilden einen *Textverbund* (oder: *Großtext*). Der Vorspann und der Nachspann gehören – zusammen mit den Einschüben ins zentrale Wörterverzeichnis – zu den *Umtexten* (oder auch: *Außertexten*). Die Umtexte, nämlich der Vor- und Nachspann, besitzen – bezogen auf das Wörterbuch – *keine* genuine Funktion, sondern werden lediglich durch ihre Position relativ zum zentralen Wörterverzeichnis bzw. zur Wörterverzeichnisreihe bestimmt. Deshalb sind sie auch keine *funktional-positional* segmentierbare Konstituenten eines Wörterbuches, sondern sie werden allein durch die Position von Texten und Teiltexten in der linearen Sukzession der (Teil-) Textanordnung relativ zum zentralen Wörterverzeichnis bestimmt (oder anders ausgedrückt: alles, was »vor« dem zentralen Wörterverzeichnis bzw. der Wörterverzeichnisreihe steht, gehört zum Vorspann, und alles, was »danach« steht, zum Nachspann). Zur Veranschaulichung sei das folgende Architekturbild wiedergegeben.

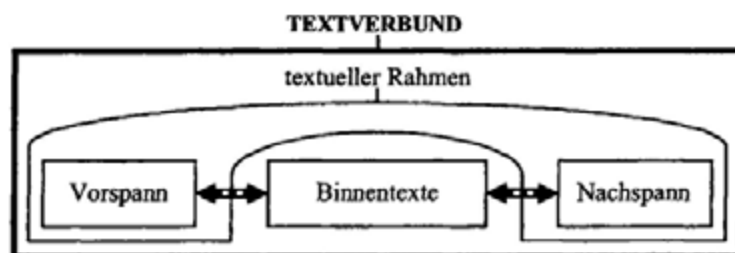


Abb. 2-1: Kommentiertes allgemeines Architekturbild zur Textverbundarchitektur und zum textuellen Rahmen (cf. KAMMERER/WIEGAND 1998[99], 226); *Darstellungskonventionen*: „x ↔ y“ bedeutet soviel wie *x steht als Ganzes oberhalb von y* und damit *y steht als Ganzes unterhalb von x*; die Umrandungszeichen dienen der kommentierenden Verdeutlichung

2. Cf. für eine Analyse von ¹LGWDAF: WIEGAND (1995); genaueres hierzu findet man bei KAMMERER/WIEGAND (1998[99]) und BERGENHOLTZ/TARP/WIEGAND (1999).

Die Strukturen, die dem Vor- und Nachspann zugeordnet werden können, sind *präzedentiv*. Das bedeutet, daß die zugehörigen Texte und Teiltex-te nicht in einer hierarchischen Beziehung zueinander stehen, sondern ausschließlich linear-präzedentiv organisiert sind. Die dem Vorspann zugeordnete Struktur ist die *präzedentive Vorspannstruktur*, die des Nachspanns ist die *präzedentive Nachspannstruktur*. Die Vorspannstruktur und die Nachspannstruktur bilden zusammen die *zweigeteilte präzedentive textuelle Rahmenstruktur*.

Neben der präzedentiven textuellen Rahmenstruktur, die nicht hierarchisch ist, gibt es noch die *partitive Textverbundstruktur*,³ zu deren Trägermenge u.a. der Vor- und der Nachspann sowie der lexikographische Textsortenträger, nicht aber das zentrale Wörterverzeichnis bzw. die Wörterverzeichnisreihe gehören. Es handelt sich hierbei um eine reflexive partielle Ordnung, wobei für die geordneten Paare der Textverbundstruktur der Relationsterm *x ist ein nichtfunktionaler textueller Teil von y* gilt. In der nachfolgenden Abbildung wird dafür beispielhaft eine Strukturdarstellung gegeben.

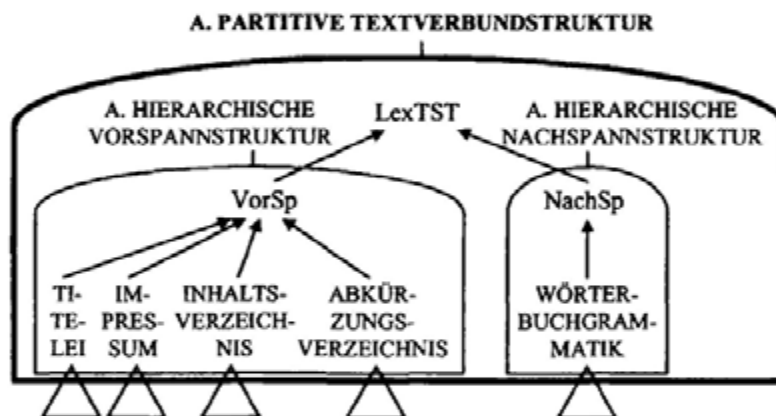


Abb. 2-2: Nicht vollständig ausgeführter kommentierter Strukturgraph zur abstrakten hierarchischen Textverbundstruktur; *Abkürzungen:* LexTST = lexikographischer Textsortenträger; VorSp = Vorspann; NachSp = Nachspann; A. = ABSTRAKT; *Darstellungskonventionen:* „ $x \rightarrow y$ “ bedeutet soviel wie *x ist ein nichtfunktionaler textueller Teil von y*; „ $\triangle$ “ steht für einen nicht ausgeführten Teilbaum

Neben der partitiven Textverbundstruktur und der präzedentiven textuellen Rahmenstruktur kann noch die *hierarchische Textverbundkonstituentenstruktur* unterschieden werden, zu deren Trägermenge der Vor- und Nachspann *nicht* gehören, sondern beispielsweise das geordnete Paar: (Textverbund; Titlei) oder das Paar: (Textverbund; Wörterbuchgrammatik). Die Relation mit dem Relationsterm *x ist ein funktionaler Teil von y* ist ein Element dieser Struktur. In Weiterführung des obigen Beispiels kann die nachfolgende Abbildung als Beispiel für eine Strukturdarstellung dienen:

3. Die Textverbundstruktur nannte WIEGAND (1995 u. 1996g) noch Textsortenträgerstruktur. Der dort verfolgte Ansatz barg jedoch einige Probleme, die in KAMMERER/WIEGAND (1998[99]) behoben wurden.

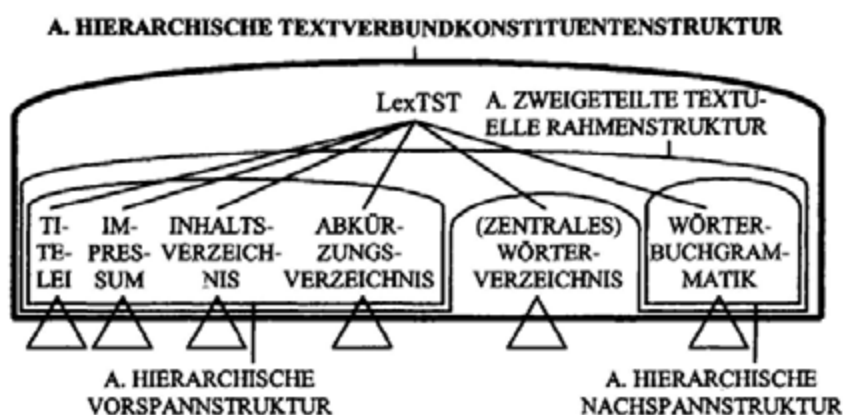


Abb. 2-3: Nicht vollständig ausgeführter kommentierter Strukturgraph zur abstrakten hierarchischen Textverbundkonstituentenstruktur; *Darstellungskonvention:* „x — y“ bedeutet (von unten nach oben gelesen) soviel wie x ist ein funktionaler Teil von y

Beide Strukturen können dann auch zusammengeführt werden. Auf diese Weise erhält man die *hierarchische vollständige Textverbundstruktur* (mit dem Vollständigkeitsgrad 1) (cf. z.B. KAMMERER/WIEGAND 1998[99], 232 seq.).

Das zentrale Wörterverzeichnis ist in der Regel nicht identisch mit einer geordneten Menge von Wörterbuchartikeln, sondern weist als unmittelbare Konstituenten einerseits *Zugriffsstrukturanzeiger* (ZSA), wie z.B. „A.“ im ¹DWB (I, 1/2), die ihrerseits durch (artikelexterne) Verweisangaben rechtserweitert sein können, wie z.B. „D, Dh, Th, p S. auch T“ in SCHÜTZEL ³AHDWB (24), und andererseits *Artikelstrecken* (AtS) auf. Daß dies aber nicht so sein muß, zeigt z.B. ²WB-UNTERFRANKEN: Dort gibt es keine Zugriffsstrukturanzeiger: Die Lemmastrecken werden dort ohne jegliche Einschübe oder nichttypographische Zugriffsstrukturanzeiger aneinandergereiht.⁴ Im PFEIFER ²ETYMBW sind jedoch, entgegen dem ersten Anschein, durchaus Zugriffsstrukturanzeiger zu beobachten, nämlich der größere Durchschuß vor einer neuen Artikelteilstrecke.⁵

Die hierarchische Konstituentenstruktur der Artikelstrecken heißt *hierarchische Artikelstreckenstruktur*. Die Artikelstreckenstruktur kann ihrerseits durch *Binnentexte* (bt) erweitert sein, wie dies beispielsweise in ¹LGWDAF der Fall ist; dann handelt es sich um eine *erweiterte hierarchische Artikelstreckenstruktur*.

Auf den Leitelementen des zentralen Wörterverzeichnisses ist mindestens eine *Zugriffsstruktur* definiert, mit deren Hilfe ein Benutzer-in-actu auf die Wörterbuchartikel des zentralen Wörterverzeichnisses gezielt zugreifen kann, falls er diese kennt. Bei Wörterbüchern mit initialalphabetischer Makrostruktur⁶ besteht die Hauptzugriffsstruktur in einer Zugriffsrelation, die i.allg. eine Erweiterung der »alphabetischen« Anordnung der Lemmazeichen ist. Wie

4. Dies ist ein Beispiel dafür, daß die Aussage: „Das Wörterverzeichnis (WV) besteht aus *Artikelstrecken* (AtS). Diese sind durch artikelexterne *nichttypographische Zugriffsstrukturanzeiger* (ZSA) voneinander getrennt“ (BERGENHOLTZ/TARP/WIEGAND 1999, 1766) nicht für alle Wörterbücher zutrifft.

5. An dieser Stelle wäre noch zu klären, auf welche Weise laufende Kolumnen, die ebenfalls Zugriffsstrukturanzeiger darstellen, zu analysieren sind. (Cf. hierzu neuerdings die Bemerkung in BERGENHOLTZ/TARP/WIEGAND 1999, 1764, nach der die lebende Kolumne als ein funktionales Element der Trägermenge der äußeren Schnellzugriffsstruktur interpretiert wird.)

6. Die Makrostruktur muß nicht immer alphabetisch sein, wie KESSELRING ¹DICHRON zeigt: Hier ist die Zugriffsstruktur beispielsweise chronologisch.

diese Anordnung genau aussieht, ist derzeit alles andere als klar. Klar ist nur, daß hier größtenteils völlige Unkenntnis herrscht. Zum Problembereich der *Erweiterung des Alphabets* seien lediglich folgende Anmerkungen gemacht: Die Erweiterung der *alphabetischen* Anordnung besteht u.a. in der Berücksichtigung von Buchstaben, die zwar Teil desjenigen Schriftsystems sind, in der die Sprache, die lexikographisch beschrieben werden soll, verschriftlicht wird, die aber nicht durch das in der Sprache geltende Alphabet geordnet werden. Im Deutschen sind die alphabetexternen, aber schriftsysteminternen Buchstaben sämtliche Umlaute sowie das β . Kompliziert wird es, wenn eine Sprache in zwei (oder mehreren) verschiedenen Schriftsystemen präsentiert werden kann, wobei sich dann die Frage stellt, welches Schriftsystem zur Präsentation der Lemmazeichen gewählt und inwieweit das alternative (bzw. die alternativen) Schriftsystem(c) im Wörterbuch berücksichtigt werden soll(en) (cf. z.B. für das Türkmenische WIEGAND 1994[95] und für das Usbekische WIEGAND 1996d). In der Regel wird darüber hinaus nicht zwischen Majuskeln und Minuskeln unterschieden (wie in einem solchen Fall die Zugriffsrelation auszusehen hat, wird in KAMMERER 1995, 7 dargestellt). Soll jedoch die Groß- und Kleinschreibung in der Zugriffsrelation berücksichtigt werden, so kann beispielsweise nicht festgelegt werden, daß alle Majuskeln vor den Minuskeln eingeordnet werden, denn dies hätte zur Folge, daß sämtliche Lemmazeichen, deren erster Buchstabe eine Versalie ist, vor allen Lemmazeichen stünden, deren erster Buchstabe eine Gemeine ist. Für das Deutsche würde das beispielsweise bedeuten, daß in jeder Lemmastrecke alle Substantive und Namen vor den übrigen Wortarten stünden. Schließlich blieben alle außeralphabetischen, schriftsystemexternen Buchstaben, wie z.B. ζ , φ , δ , ρ oder α (beispielsweise in α -Strahlen), sowie die nichtalphabetischen Buchstaben, wie Punkt, Klammer, Ausrufezeichen, Leerzeichen (!) etc., unberücksichtigt, die beispielsweise in Abkürzungen, Bindestrich-Ausdrücken, festen Wendungen (wie: *durch und durch*) oder bei Interjektionen auftreten. Schließlich muß noch auf jene nichtalphabetischen Zeichen hingewiesen werden, wie sie u.a. in der Mathematik auftreten, beispielsweise „!“ für „Fakultät“, „≡“ oder das Summenzeichen „ Σ “, oder die Logogramme, wie „§“, sowie die Währungszeichen „\$“, „£“ oder „€“ uvm.⁷

Jede Artikelstrecke besitzt als unmittelbare Konstituenten n *Wörterbuchartikel* (wa) (mit $n \geq 1$). Auf diesen Wörterbuchartikeln sind diverse Strukturen definiert; u.a. sind dies (in Weiterführung von BERGENHOLTZ/TARP/WIEGAND 1999, 1767):

- die Artikelkonstituentenstruktur (kurz: Artikelstruktur);
- die Mikrostruktur;
- die Makrostruktur;
- die Datendistributionsstruktur;
- die Positionsstruktur;
- die Angabestruktur;
- die Adressierungsstruktur;
- die Kohäsionsstruktur;
- die Thema-Rhema-Struktur;
- die Kohärenzstruktur;
- die innere Zugriffsstruktur;
- die innere Schnelzugriffsstruktur;
- die artikelinterne Verweis- oder Mediostruktur;
- die Mikroarchitektur;

7. Zu einigen Problemen von initialalphabetischen Makrostrukturen mit exhaustiver Alphabetisierung cf. BERGENHOLTZ/TARP/WIEGAND (1999, 1812–1814).

- die Suchbereichsstruktur;
- die Suchbereichsarchitektur.

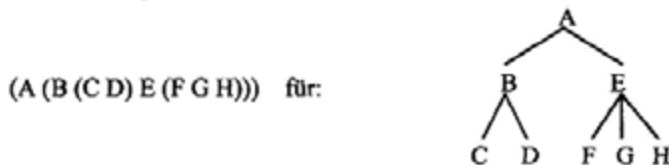
Jedem in einem Wörterbuch gegebenen Wörterbuchartikel kann auf Grund einer Analyse eine abstrakte und eine konkrete hierarchische Mikrostruktur (MiS_h^a bzw. MiS_h^k) zugeordnet werden, aus der die partitive als auch die präzedentive abstrakte bzw. konkrete Mikrostruktur hervorgeht. Auch umgekehrt können auf der Grundlage von Mikrostrukturen, die im Rahmen eines Instruktionbuches vorgegeben sind, Wörterbuchartikel erstellt werden, die nach dem Abfassen der Wörterbuchartikel jeweils genau einer abstrakten hierarchischen Mikrostruktur entsprechen (cf. hierzu exemplarisch KAMMERER 1995, insbes. 123–160). – Unterscheiden sich zwei Wörterbuchartikel eines bestimmten Wörterbuches nicht hinsichtlich der ihnen zugewiesenen *abstrakten* hierarchischen Mikrostruktur, so kann man sie demselben *Wörterbuchartikeltyp* (*wat*) zuordnen. Sind die beiden zugehörigen abstrakten hierarchischen Mikrostrukturen jedoch grundsätzlich verschieden, dann gehören beide Wörterbuchartikel verschiedenen Wörterbuchartikeltypen an.⁸ Wörterbuchartikeltypen sind immer auf ein bestimmtes Wörterbuch (*WB*) oder einen Wörterbuchtyp (*WB-x*) bezogen, weshalb sie mit dem Namen des betreffenden Wörterbuches bzw. Wörterbuchtyps indiziert sind (also z.B.: wat^{WB} oder: wat^{WB-x}). (Im folgenden wird die Indizierung mit dem Namen des Wörterbuch(typ)s weggelassen, da sie sich von selbst versteht.) Damit kann definiert werden:

[D 2-1: Wörterbuchartikeltyp]

Ein *Wörterbuchartikeltyp* (*wat*) eines Wörterbuches (wat^{WB}) oder Wörterbuchtyps (wat^{WB-x}) ist eine Menge von *n* abstrakten hierarchischen Mikrostrukturen (mit $n \in \mathbb{N}^*$).

Anmerkung 1: Ein Wörterbuchartikeltyp ist immer wörterbuch- (bzw. wörterbuchtyp-) spezifisch, weshalb er mit einem Index für das Wörterbuch oder den Wörterbuchtyp versehen wird (z.B.: wat^{WB-24}).

Anmerkung 2: Die Notation der abstrakten hierarchischen Mikrostrukturen, die die Elemente der Wörterbuchartikeltypen sind, muß nicht durch eine graphische Darstellung erfolgen, sie kann beispielsweise auch in der Kammerschreibweise durchgeführt werden; dadurch würde ein maschineller Vergleich erleichtert werden. Beispiel:



Je nach dem, ob $n > 1$ ist oder ob $n = 1$ ist, kann zwischen einem *engen* und einem *weiten* *Wörterbuchartikeltyp* unterschieden werden; es wird deshalb definiert:

[D 2-2: enger Wörterbuchartikeltyp]

Ein *enger Wörterbuchartikeltyp* (${}^e\text{wat}$) eines Wörterbuches oder Wörterbuchtyps ist ein Wörterbuchartikeltyp, für den gilt, daß die Kardinalität *n* der Menge von abstrakten hierarchischen Mikrostrukturen, durch die er festgelegt wird, gleich 1 ist (mithin: $[{}^e\text{wat} = \text{wat}] \leftrightarrow [|\text{wat}| = 1]$).

Anmerkung: Besteht die Menge aller Wörterbuchartikeltypen (*WAT*) ausschließlich aus engen Wörterbuchartikeltypen, so heißt auch die Menge aller Wörterbuchartikeltypen *eng* (${}^e\text{WAT}$).

(Es gilt also dann: $[WAT = {}^eWAT] \leftrightarrow [\forall \text{wat}_i \in WAT: |\text{wat}_i| = 1]$).

8. Wie die Formulierung „grundsätzlich verschieden“ zu verstehen ist, wird an späterer Stelle geklärt.

[D 2-3: weiter Wörterbuchartikeltyp]

Ein *weiter Wörterbuchartikeltyp* (${}^w\text{wat}$) eines Wörterbuches oder Wörterbuchtyps ist ein Wörterbuchartikeltyp, für den gilt, daß die Kardinalität n der Menge von abstrakten hierarchischen Mikrostrukturen, durch die er festgelegt wird, echt größer als 1 ist (mithin: $\{\text{wat} = {}^w\text{wat}\} \leftrightarrow [|\text{wat}| > 1]$).

Anmerkung 1: Besitzt die Menge aller Wörterbuchartikeltypen (WAT) mindestens einen weiten Wörterbuchartikeltyp ${}^w\text{wat}$, als Element, so heißt auch die Menge aller Wörterbuchartikeltypen *weit* (${}^w\text{WAT}$).

(Es gilt also dann: $[\text{WAT} = {}^w\text{WAT}] \leftrightarrow [\exists \text{wat}_i \in \text{WAT}: |\text{wat}_i| > 1]$).

Anmerkung 2: Für alle Wörterbuchartikeltypen wat_i , die Elemente einer weiten Menge von Wörterbuchartikeltypen ${}^w\text{WAT}$ sind, gelte einschränkend, daß sie paarweise disjunkt sind.

(Es gilt also dann: $\forall \text{wat}_i \in {}^w\text{WAT} \forall \text{wat}_j \in {}^w\text{WAT}: [\text{wat}_i \cap \text{wat}_j = \emptyset]$.)

Fürs erste wird – der Einfachheit wegen – von engen Wörterbuchartikeltypen ausgegangen. – Alle Wörterbuchartikeltypen eines Wörterbuches können in der *Menge aller Wörterbuchartikeltypen* (WAT) dieses Wörterbuches (WAT_{WB}) bzw. eines Wörterbuchtyps ($\text{WAT}_{\text{WB-X}}$) zusammengefaßt werden.

Für enge Wörterbuchartikeltypen gelten zusammenfassend folgende Beziehungen (wobei der Ausdruck „ $\text{MiS}_n^a(\text{wa}_i)$ “ zu lesen ist als *die abstrakte hierarchische Mikrostruktur vom Wörterbuchartikel i*):

$\text{WAT}_{\text{WB}} = \{\text{wat}_1^{\text{WB}}, \text{wat}_2^{\text{WB}}, \dots, \text{wat}_n^{\text{WB}}\}$ mit $1 \leq n \leq \text{Anzahl der Wörterbuchartikel des Wörterbuches WB.}$

$\text{WAT}_{\text{WB-X}} = \{\text{wat}_1^{\text{WB-X}}, \text{wat}_2^{\text{WB-X}}, \dots, \text{wat}_i^{\text{WB-X}}, \dots\}$ mit $i \in \mathbb{N}^*$

$\text{wat}_k = \{\text{MiS}_n^a(\text{wa}_i)\}$

$[\text{MiS}_n^a(\text{wa}_i) = \text{MiS}_n^a(\text{wa}_j)] \rightarrow [(\text{MiS}_n^a(\text{wa}_i) \in \text{wat}_k) \wedge (\text{MiS}_n^a(\text{wa}_j) \in \text{wat}_k)].$

Hierzu nun ein Beispiel. Im DUDEN-²⁰1 finden sich die folgenden beiden Kurzartikel:

[E 2-1: DUDEN-²⁰1, 400^b]

Knj|t|ter, der; -s, -

[E 2-2: DUDEN-²⁰1, 394^a]

Kj|t|tel, der; -s, -

Die abstrakte hierarchische Mikrostruktur zu diesen beiden Wörterbuchartikeln, zusammen mit einem Ausschnitt aus den dazu k-isomorphen konkreten hierarchischen Mikrostrukturen, wird in der nachstehenden Abbildung wiedergegeben.

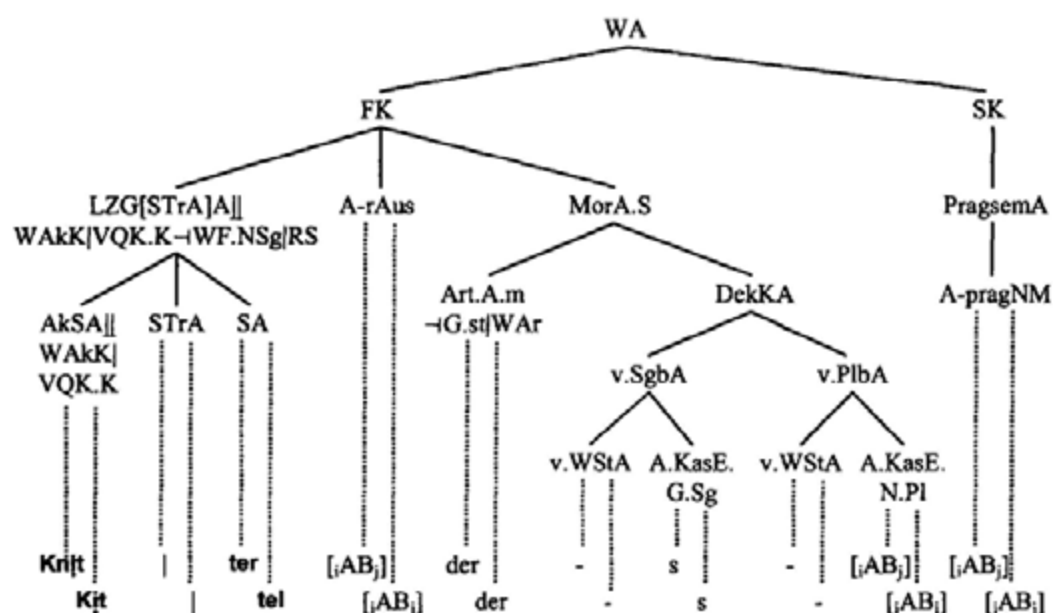


Abb. 2-4: Strukturgraph⁹ zur (abstrakten und dazu isomorphen konkreten) einfachen hierarchischen Mikrostruktur, die zu E 2-1 und E 2-2 gehören; *Abkürzungen*: WA = Wörterbuchartikel; FK = Formkommentar; LZG[STrA]A||WAKK|VQK.K-|WF.NSg|RS = Lemmazeichengestaltangabe, die um eine Silbentrennungsangabe (STrA) binnenerweitert und um eine Wortakzentkennzeichnung unten erweitert ist, die gleichzeitig eine Vokalquantitätskennzeichnung für die Kürze (WAKK|VQK.K) ist, und aus der die Wortform für den Nominativ Singular (WF.NSg) und die Rechtschreibung (RS) erschließbar sind; „|“ = zugleich; „||“ = unten erweitert; „-“ = daraus ist erschließbar; A-rAus = Angabe zur regelmäßigen Aussprache; MorA.S = Morphologieangabe beim Substantiv; AkSA||WAKK|VQK.K = Akzentsilbenangabe (AkSA), die unten durch eine Wortakzent- und Vokalquantitätskennzeichnung für die Kürze erweitert ist; Art.A.m-|G.st|WAr = Artikelangabe maskulinum, aus der das stabile Genus (G.st) sowie die Wortart (WAr) erschließbar sind; DekKA = Deklinationsklassenangabe; v.SgbA = verdichtete Singularbildungsangabe; v.PlbA = verdichtete Pluralbildungsangabe; v.WStA = verdichtete Wortstammangabe; A.KasE.G.Sg = Angabe der Kasusendung (A.KasE) für den Genitiv Singular (G.Sg); A.KasE.N.Pl = Angabe der Kasusendung für den Nominativ Plural (N.Pl); [iABj] = Angabeblank (als Angabeform!) mit *i* und *j* als Variablen für den unmittelbaren linken bzw. rechten Kontext (= terminale Angabe); SK = semantischer Kommentar; PragsemA = pragmatisch-semantische Angabe; A-pragNM = Angabe zur pragmatischen Nullmarkierung; *Darstellungskonventionen*: „x — y“ bedeutet (von oben nach unten gelesen) soviel wie *y* ist unmittelbares Konstitut zu *x* bzw. (von unten nach oben gelesen) *x* ist unmittelbare Konstituente von *y*; „x y“ bedeutet (von unten nach oben gelesen) soviel wie *x* ist ein Element von *y*

Ein Wörterbuchartikel, dessen abstrakte hierarchische Mikrostruktur nicht mit der von E 2-1 und E 2-2 identisch ist, ist beispielsweise:

[E 2-3: DUDEN-¹⁰1, 401⁶]

Kog|ge, dic; -, -n (dickbauchiges Hanseschiff)

Hier sieht die abstrakte einfache hierarchische Mikrostruktur, zusammen mit einem Ausschnitt aus der dazu isomorphen konkreten einfachen Mikrostruktur, wie folgt aus:

9. Zum Problem der nichtverzweigenden Knoten cf. jetzt MEYER/WIEGAND (1999), wo statt „— SK — PragsemA — A-pragNM“ analysiert wird: „— SK=PragsemA=A-pragNM“ mit „x=y“ bedeutet soviel wie *x* ist identisch mit *y* (cf. auch WIEGAND 1999c, 308, Abb. 3). Auf diese Weise erhält man einen einzigen komplex etikettierten Knoten. – In den folgenden Strukturgraphen wird weiterhin von nichtverzweigenden Knoten ausgegangen; diese können jedoch jederzeit problemlos in einen einzigen verzweigenden nichtterminalen bzw. in einen einzigen terminalen Knoten überführt werden.

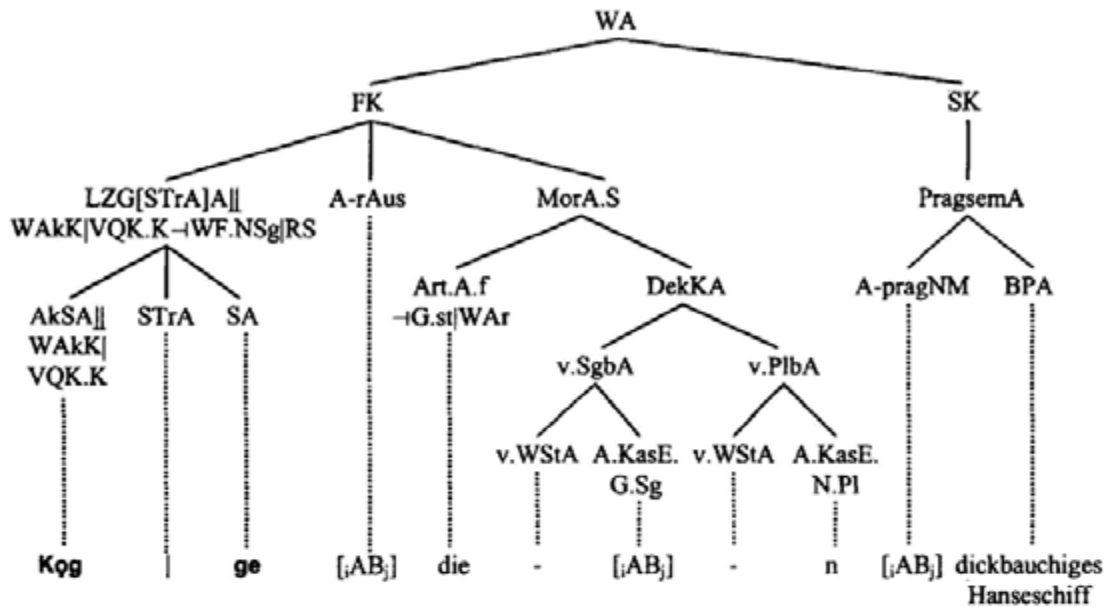


Abb. 2-5: Strukturgraph zur (abstrakten und dazu isomorphen konkreten) einfachen hierarchischen Mikrostruktur, die zu E 2-3 gehören; *Abkürzungen:* Art.A.f-|G.st|WAr = Artikelangabe femininum, aus der das stabile Genus sowie die Wortart erschließbar sind; BPA = Bedeutungsparaphrasenangabe

Daraus folgt, daß, wenn die Konzeption eines engen Wörterbuchartikeltyps zugrunde gelegt wird, die Wörterbuchartikel zu *Knitter* und *Kittel* im DUDEN-²⁰1 *per definitionem* zum selben Wörterbuchartikeltyp gehören, während der Wörterbuchartikel zu *Kogge* zu einem anderen Wörterbuchartikeltyp gehört, denn es gilt: Die abstrakten hierarchischen Mikrostrukturen zu den Wörterbuchartikeln von *Knitter* und *Kittel* sind identisch, während die zum Wörterbuchartikel von *Kogge* eine andere ist. Zur Verdeutlichung kann auch kurz geschrieben werden (indem für den betreffenden Wörterbuchartikel das zugehörige Lemma in Fettschrift angegeben wird):

$$\begin{aligned} \text{MiS}_h^*(\text{Knitter}) &= \text{MiS}_h^*(\text{Kittel}) \Rightarrow [(\text{MiS}_h^*(\text{Knitter}) \in \text{'wat}_i) \wedge (\text{MiS}_h^*(\text{Kittel}) \in \text{'wat}_i)] \\ \text{MiS}_h^*(\text{Kogge}) &\neq \text{MiS}_h^*(\text{Knitter}) \Rightarrow [(\text{MiS}_h^*(\text{Kogge}) \notin \text{'wat}_i) \wedge (\text{MiS}_h^*(\text{Kogge}) \in \text{'wat}_i)]. \end{aligned}$$

Unter gewissen Bedingungen kann es aber ratsam sein, einen *weiten* Wörterbuchartikeltyp zugrunde zu legen; hierzu ein Beispiel. Gegeben sei der Wörterbuchartikel zu *Rinde* aus dem DUDEN-²⁰1:

[**R** 2-4: DUDEN-²⁰1, 603^c]
Rjn|de, die; -, -n

Die abstrakte hierarchische Mikrostruktur zu diesem Wörterbuchartikel sieht wie folgt aus:

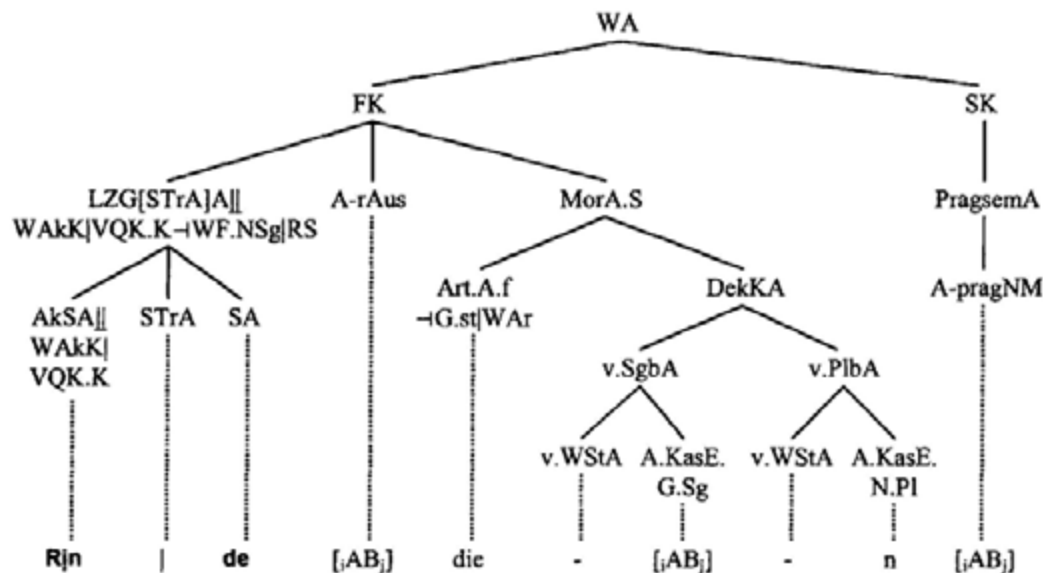


Abb. 2-6: Strukturgraph zur (abstrakten und dazu isomorphen konkreten) einfachen hierarchischen Mikrostruktur, die zu E 2-4 gehören

Die abstrakte hierarchische Mikrostruktur von E 2-4 unterscheidet sich von der zu E 2-1 und E 2-2 nur darin, daß die Morphologieangabe beim Substantiv nicht aus einer Artikelangabe *maskulinum* und einer Deklinationsklassenangabe besteht, sondern aus einer Artikelangabe *femininum* und einer Deklinationsklassenangabe. Bei einer engen Konzeption der Wörterbuchartikeltypen sind die Wörterbuchartikeltypen, zu denen die Wörterbuchartikel zu *Rinde* und zu *Kittel* (bzw. auch *Kogge*) gehören, verschieden, da sie sich in der Artikelangabe unterscheiden. Unter bestimmten Umständen kann es aber ratsam sein, weite Wörterbuchartikeltypen anzusetzen und dafür zu sorgen, daß z.B. die Wörterbuchartikel zu *Rinde* und zu *Kittel* zu demselben Wörterbuchartikeltyp gehören. Dies ist beispielsweise dann der Fall, wenn die Wörterbuchartikel bildschirmorientiert abgefaßt werden und für die lexikographische Bearbeitung der lexikalischen Einheiten *Rinde*, *Kittel* und *Kogge* jeweils dieselbe Eingabemaske vorgesehen ist, da sie (thorieintern) zum selben Lemmazeichentyp und damit Wörterbuchartikeltyp gehören (cf. KAMMERER 1995). Die Diversifizierung erfolgt schließlich durch die Dateneingabe, in diesem Fall durch „der“ bzw. durch „die“. (Dabei ist darauf zu achten, daß „der“ bzw. „die“ keine Angaben, sondern *Vorangaben* sind; cf. auch p. 270 und insbes. WIEGAND 1998, 217.) Ist die Dateneingabe zu einem Wörterbuchartikel abgeschlossen und hat darauf MSG (der *Mikrostrukturgenerator*; cf. WIEGAND 1998, 215) operiert, dann können (müssen aber nicht) enge Wörterbuchartikeltypen unterschieden werden. Zusammenfassend gelten nun folgende Beziehungen:

$$\begin{aligned}
 \text{MiS}_i^j(\text{Knitter}) &\in {}^w\text{wat}_i \\
 \text{MiS}_i^j(\text{Kittel}) &\in {}^w\text{wat}_i \\
 \text{MiS}_i^j(\text{Rinde}) &\in {}^w\text{wat}_i \\
 \text{MiS}_i^j(\text{Kogge}) &\notin {}^w\text{wat}_i \\
 \text{MiS}_i^j(\text{Kogge}) &\in {}^w\text{wat}_j \\
 \text{MiS}_i^j(\text{Rinde}) &\neq \text{MiS}_i^j(\text{Kittel}) \quad \text{und:} \quad \text{MiS}_i^j(\text{Rinde}) \neq \text{MiS}_i^j(\text{Knitter}).
 \end{aligned}
 \quad \text{mit } i, j \in \mathbb{N}^*$$

Da jedem Wörterbuchartikel eines Wörterbuches genau eine abstrakte hierarchische Mikrostruktur zugewiesen werden kann, ist jeder Wörterbuchartikel zusammen mit genau einem Wörterbuchartikeltyp Element einer Relation $\mathbb{R}$. Der zweistellige Relationsterm von $\mathbb{R}$ lautet: ... *gehört zum Wörterbuchartikeltyp* ... Da $\mathbb{R}$ linkstotal und rechtseindeutig ist, handelt es sich genau genommen um eine Funktion. Es gilt also:

$$\forall wa \in WA \exists! wat \in WAT [(wa, wat) \in \mathbb{R} \wedge \forall wat' \in WAT ((wa, wat') \in \mathbb{R} \rightarrow (wat = wat'))]$$

oder einfacher ausgedrückt:

$$\forall wa \in WA \exists! wat \in WAT ((wa, wat) \in \mathbb{R}).$$

Die Anzahl der Wörterbuchartikeltypen in einem gegebenen Wörterbuch kann maximal so groß sein wie die Anzahl der Wörterbuchartikel dieses Wörterbuches. In der Regel wird die Anzahl der Wörterbuchartikeltypen eines Wörterbuches deutlich unter der Anzahl der Wörterbuchartikel des Wörterbuches liegen. Die Anzahl der Wörterbuchartikeltypen hängt mit dem Grad der Standardisierung sowie mit dem lexikographischen Datenangebot zusammen; es seien diesbezüglich folgende Hypothesen formuliert, die empirisch verifiziert werden müßten:

Hypothesen:

- Je höher der Grad der Standardisierung der Wörterbuchartikel eines Wörterbuches ist, um so niedriger ist die Anzahl der Wörterbuchartikeltypen dieses Wörterbuches und umgekehrt.
- Je niedriger (bei gleichem Standardisierungsgrad) das lexikographische Datenangebot eines Wörterbuches ist, um so niedriger ist auch die Anzahl der Wörterbuchartikeltypen dieses Wörterbuches und umgekehrt.

Das heißt: In einem Wörterbuch mit stark standardisierten Mikrostrukturen gibt es – wenn die erste Hypothese korrekt ist – genuin weniger Wörterbuchartikeltypen als in einem Wörterbuch, in dem die Wörterbuchartikel nur schwach standardisiert sind.

2.2 Die Konzeption des Lemmazeichentyps

Jede lexikalische Einheit (LE) einer natürlichen Sprache kann genau einem Lemmazeichentyp (lzt) zugeordnet werden, mit anderen Worten: Alle lexikalischen Einheiten sind zusammen mit genau einem Lemmazeichentyp Element der Zuordnungsrelation $\mathbb{S}$, deren Relationsterm lautet: ... *gehört zum Lemmazeichentyp* ... Diese Zuordnungsrelation $\mathbb{S}$ ist linkstotal und rechtseindeutig, so daß es sich wiederum um eine Funktion handelt, denn es gilt:

$$\forall LE \in LE \exists! lzt \in LZT ((LE, lzt) \in \mathbb{S}).$$

WIEGAND legt einen Lemmazeichentyp informell wie folgt fest:

[Z 2-1: WIEGAND 1998, 214 seq.]

Ein Lemmazeichentyp [= LZT; M. K.] ist eine Menge von linguistisch und metalexikographisch bestimmten Eigenschaften wie z. B.:

$LZT_x = \{\text{Substantiv, Simplex, normale Realisierung, Bezeichnung für natürliche Art, n-fach polysem; Verweisverfahren V3, Artikelgrößenklasse } G_x\}$

Jede Menge von LZT_i spezifiziert eine Menge von Eigenschaften des zu schreibenden Wörterbuchartikels. Die linguistisch bestimmten Eigenschaften legen fest, welche Angabeklassen (mit Angaben, die sich auf den Wörterbuchgegenstand beziehen) zu berücksichtigen sind; zu jedem Lemmazeichentyp gehört daher eine Menge von Angabeklassen, auf der bestimmte Ordnungsrelationen definiert sind, so daß zu jedem Lemmazeichentyp eine *abstrakte hierarchische Mikrostruktur* gehört [...].

Ein Lemmazeichentyp wird demnach allgemein als eine Menge aufgefaßt, in der Eigenschaftsausprägungen zweier verschiedener Kategorien zusammengefaßt sind. Hier ist allerdings einzuwenden, daß einer gegebenen lexikalischen Einheit genuin keine metalexikographischen Eigenschaften zukommen; diese erhält die lexikalische Einheit erst in Bezug auf einen spezifischen Wörterbuchtyp bzw. in Bezug auf ein konkretes Wörterbuch. So variiert beispielsweise die Artikelgrößenklasse mit dem Wörterbuchtyp: Eine Interjektion, die in einem allgemeinen zweisprachigen Wörterbuch lexikographisch behandelt wird, wird u.U. einer kleineren Artikelgrößenklasse angehören, als wenn sie in einem einsprachigen Spezialwörterbuch für Partikel lexikographisch beschrieben wird. Was also im obigen Zitat beschrieben wird, ist kein *wörterbuchtypunspezifizierter* Lemmazeichentyp, sondern ein *wörterbuchspezifizierter*, nämlich im Hinblick auf ein bestimmtes Wörterbuch (oder: im Hinblick auf einen bestimmten Wörterbuchtyp). Der im Zitat angeführte Lemmazeichentyp muß demnach mit dem Wörterbuch(typ) indiziert werden, auf das/den er bezogen ist (also: lzt_x^{WB} bzw. $\text{lzt}_x^{\text{WB-}y}$). Ein *wörterbuchtypunspezifizierter* Lemmazeichentyp muß dann – in einer ersten, später zu präzisierenden Annäherung – wie folgt gefaßt werden:

[D 2-4: wörterbuchtypunspezifizierter Lemmazeichentyp]

Ein *wörterbuchtypunspezifizierter* Lemmazeichentyp (lzt_x) ist eine Menge von ausschließlich linguistischen Eigenschaftsausprägungen (IE). Kurz:

$$\text{lzt}_x = \{\text{IE}_1, \text{IE}_2, \dots, \text{IE}_n\} \quad \text{mit } n, x \in \mathbb{N}^+.$$

Anmerkung: *Expressis verbis* sei darauf hingewiesen, daß, wenn $n = 1$ ist, die Kardinalität der Menge der linguistischen Eigenschaftsausprägungen 1 ist – dieser Fall wird in der Definition damit ausdrücklich berücksichtigt.

Die metalexikographischen Eigenschaftsausprägungen (mE), wie sie im obigen Zitat angesprochen wurden, werden erst durch das Wörterbuch oder den Wörterbuchtyp bestimmt, auf das/den der Lemmazeichentyp bezogen ist. Ein solcher *wörterbuchspezifizierter* Lemmazeichentyp ist dann ein 2-Tupel, bestehend aus einer Menge von linguistischen und aus einer Menge von metalexikographischen Eigenschaftsausprägungen. Es gilt demnach:

[D 2-5: wörterbuchspezifizierter Lemmazeichentyp]

Ein *wörterbuchspezifizierter* Lemmazeichentyp (lzt_x^{WB} bzw. $\text{lzt}_x^{\text{WB-}y}$) ist ein 2-Tupel, bestehend aus einer Menge von linguistischen Eigenschaftsausprägungen (IE) und einer Menge von metalexikographischen Eigenschaftsausprägungen (mE), der in Bezug auf ein Wörterbuch (WB) oder einen Wörterbuchtyp (WB- y) konzipiert ist. Kurz:

$$\text{lzt}_x^{\text{WB}} \text{ bzw. } \text{lzt}_x^{\text{WB-}y} = \langle \{\text{IE}_1, \text{IE}_2, \dots, \text{IE}_m\}, \{\text{mE}_1, \text{mE}_2, \dots, \text{mE}_n\} \rangle \quad \text{mit } m, n, x, y \in \mathbb{N}^+.$$

Aber es ist noch eine weitere Präzisierung notwendig. Dazu betrachte man die Wortartangabe „Substantiv“ im obigen Zitat Z 2-1, p. 15. Die linguistische Eigenschaftsausprägung, ein Substantiv zu sein, dürfte an sich wenig problematisch sein. Schwieriger wird es allerdings bereits bei Verben: *müssen* ist sicherlich ein Verb, aber es stellt sich die Frage: soll es – außer daß es ein Modalverb ist – auch als Hilfsverb gelten? ENGEL (¹1996, 391) setzt auf Grund der verwirrenden Terminologie in den verschiedenen Grammatiken erst gar keine Hilfsverben an, sondern nennt diese *Auxiliarverben* und grenzt diese ex- wie intensional von den Modalverben

ab, zu denen *müssen* gehört. Bei HELBIG/BUSCHA (1986, 33 u. 123) zählt *müssen* hingegen zu den Hilfsverben, genauer: zu den modalen Hilfsverben, die – auch Modalverben genannt – eine Unterkategorie der Hilfsverben bilden. Im Duden-³⁴ (92 seq.) werden die Modalverben als eigene Kategorie neben den Hilfsverben angesetzt, Auxiliärverben gibt es hier im Gegensatz zu ENGEL (³1996) nicht. Aber selbst die allgemeine Aussage, *müssen* sei ein Hilfsverb, wurde von manchen Forschern bestritten, und *müssen* wurde zu den Vollverben gezählt (cf. hierzu: HELBIG 1995, 208).

Nochmals komplizierter wird es, wenn man Wörter wie z.B. *ja* oder *oh!* betrachtet: Für HELBIG/BUSCHA (1986, 477 u. 529) ist *ja* eine Partikel bzw. ein Satzäquivalent, Duden-³⁴ (371) kennt *ja* nur als Partikel, genauer als Modalpartikel, und ENGEL (³1996, 231, 743 u. 772) rechnet *ja* den Partikeln zu, unterscheidet aber drei verschiedene Unterklassen: (1) intensivierender Konjunkt, (2) Antwortpartikel als Satzäquivalent¹⁰ und (3) Abtönungspartikel.

Aus diesen Ausführungen sollte deutlich geworden sein, daß die linguistischen Eigenschaftsausprägungen nie theorieelos (im weitesten Sinne) vorgegeben sind, sondern immer im Hinblick auf bestimmte Auffassungen konzipiert werden. Diese, hier salopp so bezeichneten *Auffassungen*, können in ausgearbeiteten linguistischen Theorien greifbar sein, sie sind aber auch in Aussagen zu finden, wie man sie beispielsweise in einer Schulgrammatik antrifft:

[Z 2-2: HERINGER 1995, 16]

Verben bilden den Kern von Verbalphrasen.

Solche Aussagen sind keine Aussagen im philosophischen Sinne, insofern es hier nicht auf ihren Wahrheitswert ankommt. Ihnen kann zweifellos ein Wahrheitswert zugeordnet werden, wenn sie z.B. auf eine konkrete Grammatik bezogen werden (cf. z.B. die vorangegangene Fußnote), aber darum geht es hier nicht. Vielmehr dienen die Aussagen dazu, bestimmte Feststellungen zu treffen, die von einem Lexikographen akzeptiert werden können oder nicht. Solche Aussagen findet man nicht nur in linguistischen Theorien, sondern auch in Wörterbüchern, beispielsweise hinsichtlich der diatopischen oder diastratischen Markierung eines Lexems, in Grammatiken, in Einführungsbüchern in die Linguistik etc.

Diese explizit gegebenen Aussagen kontrastieren mit impliziten Aussagen über linguistische Phänomene. So können sämtliche Textinstanzen¹¹ als extraindividuelle, externalisierte und insofern objektive Manifestationen von impliziten Aussagen über die Sprache angesehen werden, in der sie formuliert wurden, sofern sie von einem Individuum nach der Textinstanzwahrnehmung und Textbildung als intraindividueller Textinhalt für kognitive Verarbeitungen und sprachbezogene Fragestellungen zur Verfügung stehen. Schließlich beruht die erfolgreiche und sinnvolle Verwendung von Corpora in der Linguistik auf dieser Prämisse.

Die Menge der linguistischen Aussagen, in deren Rahmen der Lemmazeichentyp konstruiert wird, heiße *linguistische Überdachung*¹² (LÜ). Diese kann wie folgt definiert werden:

[D 2-6: linguistische Überdachung]

Die *linguistische Überdachung* (LÜ) dient zur Festlegung der Eigenschaftsausprägungen eines Lemmazeichentyps und ist eine echte Teilmenge der Menge an expliziten und impliziten Aussagen über sprachliche Phänomene (wie man sie z.B. in linguistischen Theorien und in Corpora findet).

10. Dies wäre bei HELBIG/BUSCHA (1986) eine *Contradictio in adjecto*!

11. Zur Text-Terminologie cf. WIEGAND (1993[94], 245 seq.)

12. Der Terminus *linguistische Überdachung* darf nicht verwechselt werden mit dem Terminus *lexikographische Überdachung* (der Dialekte) bei WIEGAND (1986c, 206).

Anmerkung: Diese expliziten und impliziten Aussagen über sprachliche Phänomene werden z.B. in einem Ausschnitt der linguistischen Literatur greifbar und determinieren auf dieser Grundlage die Lemmazeichentypen hinsichtlich ihrer Eigenschaftsausprägungen.

Für die Konzeption der Lemmazeichentypen bedeutet dies, daß immer n (explizite oder implizite) Aussagen (A) (mit $n \geq 1$) in den Lemmazeichentyp einfließen. Diese linguistischen Aussagen determinieren in ihrer Gesamtheit das So-Sein der Lemmazeichentypen, indem sie festlegen, wie sich die einzelnen linguistischen Eigenschaftsausprägungen gestalten.

Dies hat jetzt den Anschein, als werde ein Lemmazeichentyp allein durch die linguistische Überdachung bestimmt und sei somit in einem gewissen Sinne objektivierbar – dies ist allerdings nicht ganz korrekt und muß relativiert werden. Die linguistische Überdachung liegt zwar als eine Menge von Daten vor, diese müssen aber von Subjekten (i.e. den Lexikographen) zu Informationen verarbeitet werden. Insofern fließt immer das interpretierende Subjekt in die linguistische Überdachung mit ein. Denn nur so ist es erklärlich, daß dieselben Daten (z.B. die Verwendungsweisen eines Wortes in einem Corpus) von verschiedenen Subjekten u.U. verschieden interpretiert werden, beispielsweise hinsichtlich des Polysemiegrades.

Wird eine linguistische Aussage A_i durch eine andere linguistische Aussage $A_{i'}$ ersetzt, kann sich das So-Sein der Lemmazeichentypen ändern, indem sich eine linguistische Eigenschaftsausprägung ändert. Damit kann außerdem einhergehen, daß sich die Anzahl der Lemmazeichentypen vergrößert oder verkleinert. Unter bestimmten Bedingungen kann eine linguistische Aussage A_i nicht durch eine linguistische Aussage $A_{i'}$ ersetzt werden, und zwar genau dann nicht, wenn die Ersetzung zu Inkonsistenzen oder Inkompatibilitäten führt. Ist dies der Fall, muß entweder die Ersetzung rückgängig gemacht werden oder die Aussage(n), denen gegenüber die Inkonsistenzen auftreten, müssen ebenfalls durch geeignete andere ersetzt werden. Diese sind dann ihrerseits daraufhin zu überprüfen, ob sie nicht zu Inkonsistenzen und Inkompatibilitäten führen.

Die linguistische Überdachung kann im einzelnen wie folgt konzipiert werden: Gegeben sei eine Bibliographie linguistischer Literatur, Grammatiken, Wörterbücher u.ä. sowie gegebenenfalls diverser Corpora. Die bibliographischen Angaben in der Bibliographie werden mittels der Sigle B_i durchgezählt ($i \in \mathbb{N}^*$). Die linguistische Überdachung wird nun durch Nennung der Siglen angegeben, die auf diejenige Literatur- und Corporaangaben in der Bibliographie verweisen, in denen die betreffende linguistische Aussage zu finden ist. Diese Menge von Siglen wird der linguistischen Eigenschaftsausprägung angehängt, so daß man z.B. erhält:

Substantiv $_{B_{14}B_{17}B_{29}}$

Auf die Reihenfolge der Siglen kommt es nicht an. So kann ein wörterbuchtypenspezifizierter Lemmazeichentyp wie folgt angegeben werden:

$lzt_{4712} = \{IE_{B_7}, IE_{B_{39}B_{14}}, IE_{B_{32}B_{36}B_{39}}, IE_{B_{21}}\}$

Wenn klar ist, was als linguistische Überdachung für die jeweilige linguistische Eigenschaftsausprägung gilt, kann die Liste mit den Siglen für die bibliographischen Angaben weggelassen werden.

Die linguistische Überdachung, die für ein gegebenes Wörterbuch gilt, ist in einem Wörterbuch grundsätzlich anzugeben, und zwar entweder im Vorspann oder im Nachspann – eine Forderung, die nur in den seltensten Fällen praktiziert wird. Eine Ausnahme bildet diesbezüg-

lich das FWB (cf. dessen „Lexikographische Einleitung“, das „Verzeichnis der Quellen“ sowie das „Verzeichnis der Sekundärliteratur“). Der Übersichtlichkeit wegen sollte zwischen Literatur, aus der explizite Aussagen gewonnen wurden (z.B. Grammatiken und andere Wörterbücher), von Literatur, aus der implizite Aussagen gewonnen wurden (wozu i.allg. die Corpora zu zählen sind), unterschieden werden. Auf diese Weise kann sich ein Wörterbuchbenutzer ein Bild darüber machen, auf welcher Grundlage Entscheidungen über den lexikographisch behandelten Phänomenbereich im Wörterbuch getroffen wurden.

Die Zuordnungsrelation $\S$ wurde als eine Relation festgelegt, die als Elemente geordnete Paare von gegebenen lexikalischen Einheiten und Lemmazeichentypen besitzt; die Zuordnungsrelation $\S$ ordnet also lexikalische Einheiten einem bestimmten Lemmazeichentyp zu. Der Lemmazeichentyp wurde – im Falle des wörterbuchtypunspezifisierten Lemmazeichentyps – als eine Menge von linguistischen Eigenschaftsausprägungen bzw. – im Falle des wörterbuchspezifizierten Lemmazeichentyps – als ein 2-Tupel, bestehend aus einer Menge von linguistischen und aus einer Menge von metalexikographischen Eigenschaftsausprägungen, eingeführt. Die bisherige Annäherung an den Begriff des Lemmazeichentyps kann so als eine *intensionale* Bestimmung des fraglichen Terminus charakterisiert werden. Wird die *intensionale* Bestimmung *extensionalisiert*, erhält man eine Menge von Lemmazeichen, nämlich genau diejenige Klasse von Lemmazeichen, auf die die metalexikographischen und/oder linguistischen Eigenschaftsausprägungen zutreffen.

Auf Grund dieser Beobachtung ist es sinnvoll, zwischen dem *intensionalen* und dem *extensionalen* Lemmazeichentyp zu unterscheiden; es sei definiert:

[D 2-7: intensionaler Lemmazeichentyp]

Ein *intensionaler Lemmazeichentyp* (lzt^{int}) ist ein wörterbuchtypunspezifizierter oder ein wörterbuchspezifizierter Lemmazeichentyp.

[D 2-8: extensionaler Lemmazeichentyp]

Ein *extensionaler Lemmazeichentyp* (lzt^{ex}) ist eine Menge von lexikalischen Einheiten (LE), die die metalexikographischen und/oder linguistischen Eigenschaftsausprägungen, wie sie im *intensionalen* Lemmazeichentyp gegeben sind, aufweisen.

Die Zuordnung von *intensionalem* und *extensionalem* Lemmazeichentyp zu Mengen von lexikalischen Einheiten geschieht mittels der bereits eingeführten Relation $\S$, die hier lediglich genauer charakterisiert wurde.

Es gilt zusammenfassend:

$$lzt_m^{int} = \{IE_1, IE_2, \dots, IE_n\}$$

$$lzt_m^{int} \mapsto lzt_m^{ex}$$

$$lzt_m^{ex} = \{LE_1, LE_2, \dots, LE_l\}$$

$$\langle \{LE_1, LE_2, \dots, LE_l\}, lzt_m^{int} \rangle \in \S$$

$$\langle \{LE_1, LE_2, \dots, LE_l\}, lzt_m^{ex} \rangle \in \S$$

mit $l \in \mathbb{N}$ und $m, n \in \mathbb{N}^*$

„ $x \mapsto y$ “ bedeutet soviel wie x ist die *Intension* von y bzw. y ist die *Extension* von x

bzw.

$$lzt_m^{int\omega} = \langle \{IE_1, IE_2, \dots, IE_n\}, \{mE_1, mE_2, \dots, mE_o\} \rangle$$

$$lzt_m^{int\omega} \mapsto lzt_m^{ex\omega}$$

$$lzt_m^{ex\omega} = \{LE_1, LE_2, \dots, LE_l\}$$

$$\langle \{LE_1, LE_2, \dots, LE_l\}, lzt_m^{int\omega} \rangle \in \S$$

$$\langle \{LE_1, LE_2, \dots, LE_l\}, lzt_m^{ex\omega} \rangle \in \S$$

$$\langle \{IE_1, IE_2, \dots, IE_n\}, \{mE_1, mE_2, \dots, mE_o\} \rangle = lzt_m^\omega = \{LE_1, LE_2, \dots, LE_l\}$$

mit $\omega = „WB“$ oder „WB- x “,

$l \in \mathbb{N}$ und $m, n, o, x \in \mathbb{N}^*$

Allgemein kann nun der Lemmazeichentyp wie folgt definiert werden:

[Ü 2-9: Lemmazeichentyp]

Ein *Lemmazeichentyp* ist entweder ein intensionaler Lemmazeichentyp oder ein extensionaler Lemmazeichentyp.

2.3 Korrelation zwischen Lemmazeichentyp und Wörterbuchartikeltyp und weitere terminologische Unterscheidungen beim Lemmazeichentyp

Jetzt stellt sich die Frage, wie sich die Verbindung zwischen dem Lemmazeichentyp und dem Wörterbuchartikeltyp gestaltet. Die Lösung dieses Problems soll an Hand der kritischen Darstellung eines kleinen Beispiels geschehen, wobei gleichzeitig die Definition des Lemmazeichentyps weiter präzisiert wird. Im Zentrum der Betrachtung stehe das Verb *grinsen* als eine lexikalische Einheit des Deutschen. Vorerst sei folgender Lemmazeichentyp angesetzt, wobei der Lemmazeichentyp nicht vollständig wiedergegeben wird, was durch die drei Punkte „...“ angezeigt wird; die linguistische Überdachung wird durch die Literatur, die in den bibliographischen Angaben (= B) genannt wird, abgesteckt:¹³

- LÜ: B₁ = ULRICH ENGEL: *Deutsche Grammatik*. 3., korrigierte Aufl. Heidelberg: Julius Groos 1996.
 B₂ = SIEBS. *DEUTSCHE AUSSPRACHE*. Reine und gemäßigte Hochlautung mit Aussprachewörterbuch. Hrsg. v. HELMUT DE BOOR, HUGO MOSER u. CHRISTIAN WINKLER. 19., umgearbeitete Aufl. Berlin: Walter de Gruyter 1969.
 B₃ = DUDEN. *RECHTSCHREIBUNG DER DEUTSCHEN SPRACHE*. 20., völlig neu bearb. u. erw. Aufl. Hrsg. v. der Dudenredaktion. Auf der Grundlage der amtlichen Rechtschreibregeln. [= Der Duden in 10 Bänden. Das Standardwerk zur deutschen Sprache. Bd. 1.] Mannheim, Leipzig, Wien, Zürich: Dudenverlag 1991.
- lzt(*grinsen*) =
 lzt₄₇₁₂ = {Verb, zweisilbig, Initialakzent, kurzer betonter Vokal, (regelgemäße) reine Hochlautung des Infinitivs: ['grɪnzən], (regelgemäße) gemäßigte Hochlautung des Infinitivs: ['grɪnz̥n], orthographische Realisierung des Infinitivs: <grinsen>, schwaches Verb, Perfektbildung mit *haben*, Stammformen: *grinsen* – *grinste* – *gegrinst*, nicht passivfähig, monosem, Tätigkeitsverb, ...}.

Dieser Lemmazeichentyp soll in einem Wörterbuch lexikographisch bearbeitet werden, das zu dem Typ der allgemeinen einsprachigen Wörterbücher gehört (= WB-10; cf. zur Bezeichnungskonvention: WIEGAND 1998, 310); das bedeutet für den Lemmazeichentyp mit der Nummer 4712:

- lzt₄₇₁₂^{WB-10} = ({Verb, zweisilbig, Initialakzent, kurzer betonter Vokal, (regelgemäße) reine Hochlautung des Infinitivs: ['grɪnzən], (regelgemäße) gemäßigte Hochlautung des Infinitivs: ['grɪnz̥n], orthographische Realisierung des Infinitivs: <grinsen>, schwaches Verb, Perfektbildung mit *haben*, Stammformen: *grinsen* – *grinste* – *gegrinst*, nicht passivfähig, monosem, Tätigkeitsverb, ...}, {mE₁, mE₂, ..., mE_n}).

13. Da jeweils klar ist, ob es sich um einen intensionalen oder extensionalen Lemmazeichentyp handelt, wird im Folgenden auf die Indexierung mit „int“ bzw. „ex“ der besseren Lesbarkeit wegen verzichtet.

Spätestens hier stellt sich jedoch ein Problem, genauer: ein Adäquatheitsproblem.¹⁴ Wenn der Lemmazeichentyp lzt_{4712}^{WB-10} so gegeben ist, wie er jetzt exemplarisch eingeführt wurde, so gibt es keine weitere lexikalische Einheit des Deutschen, die dem Lemmazeichentyp lzt_{4712}^{WB-10} zugeordnet werden könnte, denn diejenige lexikalische Einheit des Deutschen, die neben dem Verb *grinsen* zum Lemmazeichentyp lzt_{4712}^{WB-10} zu rechnen wäre, müßte u.a. folgende Eigenschaft besitzen: Die (regelgemäße) reine Hochlautung des Infinitivs müßte ['grɪnzən] sein. Dies ist aber eine Eigenschaftsausprägung, die ausschließlich dem dt. Lexem *grinsen* zukommt und sonst keinem anderen Wort des Deutschen. Das hat zur unerwünschten Folge, daß die Anzahl der Lemmazeichentypen etwa gleich der Anzahl der Lexeme einer Sprache ist. Dies ist aber nicht sinnvoll, da der Lemmazeichentyp, wie in der eingangs formulierten informellen Umschreibung festgestellt wurde, gerade mehrere Lexeme – desselben Typs – in einer Menge zusammenfassen soll.

Eine Lösung für dieses Problem kann durch die Unterscheidung zwischen *idiohabituellen* und *habituellen linguistischen Eigenschaftsausprägungen* einer lexikalischen Einheit herbeigeführt werden. So ist die reine Hochlautung des Infinitivs (nämlich: ['grɪnzən] bei dem Verb *grinsen*) eine idiohabituelle linguistische Eigenschaftsausprägung, da sie in der Regel nur bei einem einzigen Lexem auftritt, während die Eigenschaft, ein Verb zu sein, eine habituelle linguistische Eigenschaftsausprägung ist, die auf mehr als ein Lexem zutrifft. Es muß demnach vereinbart werden, daß idiohabituelle Eigenschaftsausprägungen nichts zur Konstruktion des Lemmazeichentyps beitragen, mithin also nicht Element des Lemmazeichentyps sein können. Auf diese Weise wird garantiert, daß ein Lemmazeichentyp nicht nur auf ein einziges Lexem einer Sprache zutrifft, sondern prinzipiell auf mehrere Lexeme.

Eine idiohabituelle sowie eine habituelle linguistische Eigenschaftsausprägung kann nun wie folgt definiert werden:

[D 2-10: idiohabituelle linguistische Eigenschaftsausprägung]

Eine *idiohabituelle linguistische Eigenschaftsausprägung* ist eine linguistische Eigenschaftsausprägung, die immer nur einer einzigen lexikalischen Einheit (einer Sprache L) zukommt.

[D 2-11: habituelle linguistische Eigenschaftsausprägung]

Eine *habituelle linguistische Eigenschaftsausprägung* ist eine linguistische Eigenschaftsausprägung, die prinzipiell mehreren lexikalischen Einheiten (einer Sprache L) zukommt.

Anmerkung 1: Wichtig zu beachten ist hier, daß eine linguistische Eigenschaftsausprägung *prinzipiell* mehreren lexikalischen Einheiten zukommt; dies bedeutet, daß in einer Sprache L eine linguistische Eigenschaftsausprägung als eine habituelle gelten kann, obwohl sie faktisch nur an einer einzigen lexikalischen Einheit LE beobachtet wird. Ausschlaggebend ist also, ob eine linguistische Eigenschaftsausprägung *genuin* nur auf eine einzige lexikalische Einheit LE zutreffen kann oder ob sie lediglich *akzidentiell* nur bei einer einzigen lexikalischen Einheit auftritt.

Anmerkung 2: Die (idio)habituelle Eigenschaftsausprägung ist von einem engen bzw. weiten Wörterbuchartikeltyp strikt zu unterscheiden.

Diese Konzeption hat nun zur Folge, daß beispielsweise die idiohabituelle linguistische Eigenschaftsausprägung „orthographische Realisierung des Infinitivs: <grinsen>“ aus der Menge

14. Gemäß PRODAINFORMAD (dieses Akronym steht für: *Problem- resp. Fragestellung, Daten, intuitiver resp. informaler Lösungsansatz, formale Lösung, Adäquatheitsprüfung*; cf. MUDERSBACH 1997, 202 seq.) muß die Adäquatheit der formalen Lösung an Hand der empirischen Daten – die in diesem Fall nicht vorliegen können, da Lemmazeichentypen etwas sind, die konstruiert werden – und der informalen Lösung geprüft werden. Da bei der informalen Lösung von *Klassen* die Rede ist, ist es wenig sinnvoll, bei der formalen Lösung ausschließlich Mengen der Kardinalität 1 zu unterscheiden.

der linguistischen Eigenschaftsausprägungen beim Lemmazeichentyp $\text{lzt}_{4712}^{\text{WB-10}}$ getilgt werden muß. Die Frage ist dann allerdings, wie die orthographische Realisierung im Lemmazeichentyp berücksichtigt werden kann, ohne daß dies eine idiohabituelle linguistische Eigenschaftsausprägung ist. Der hier verfolgte Lösungsvorschlag (der auch anders hätte ausfallen können) ist der, daß zwischen *Attributen* und ihren *Werten* differenziert wird (auch *Slot* und *Filler* usw. genannt). So gesehen, besteht jede linguistische Eigenschaftsausprägung IE genau genommen aus einem 2-Tupel, nämlich aus einem Attribut und einem Wert; es gilt:

[D 2-12: linguistische Eigenschaftsausprägung]

Eine *linguistische Eigenschaftsausprägung* ist ein 2-Tupel, bestehend aus einem Attribut und einem Wert.
Kurz: $\text{IE} = \langle \text{ATTRIBUT}, \text{WERT} \rangle$.

Die Menge der linguistischen Eigenschaftsausprägungen werden in der Menge E^L zusammengefaßt.

Kurz: $E^L = \{\text{IE}_1, \text{IE}_2, \dots, \text{IE}_n\}$ mit $n \in \mathbb{N}^*$.

Anmerkung 1: Für den Ausdruck „ $\langle \text{ATTRIBUT}, \text{WERT} \rangle$ “ wird im folgenden auch einfach notiert:

$\text{ATTRIBUT}:\underline{\text{WERT}}$.

bzw. auch:

$\text{ATTRIBUT}:[\text{WERT}]$.

Anmerkung 2: Linguistische Eigenschaftsausprägungen können auch geschachtelt werden; das bedeutet, daß statt:

$A:[B], B:[C]$.

auch kurz geschrieben werden kann:

$A:[B:[C]]$.

Anmerkung 3: Für die Wiedergabe der den Werten zugrundeliegenden Grundmenge gelten folgende Schreibkonventionen: Eine *exklusive Alternation* von Werten wird durch einen senkrechten Strich „|“ gekennzeichnet, z.B. „[ja | nein]“; ein *Bereich* von Werten wird durch die linke und rechte Grenze sowie durch zwei Punkte angezeigt, z.B. „[0%..100%]“; kann kein Bereich von einzusetzenden Werten spezifiziert werden (z.B. wenn es sich um die Orthoepie oder Orthographie von lexikalischen Einheiten handelt), so werden lediglich drei Punkte notiert: „[...]“; eine *inklusive Alternation* von Werten wird durch einen Schrägstrich „/“ markiert, z.B. „[ja: [0%..100%] / nein: [0%..100%]]“.

Anmerkung 4: Statt *Eigenschaftsausprägung* wird mitunter verkürzend nur *Eigenschaft* gesagt; wenn sich von selbst versteht, daß es sich um eine linguistische und nicht um eine metalexikographische Eigenschaftsausprägung handelt, dann kann das Adjektivattribut *linguistisch* weggelassen werden.

Damit ist ein wörterbuchtypenspezifizierter Lemmazeichentyp eine Menge von 2-Tupeln:

$\text{lzt}_x = \{\langle \text{ATTRIBUT}, \text{WERT} \rangle_1, \langle \text{ATTRIBUT}, \text{WERT} \rangle_2, \dots, \langle \text{ATTRIBUT}, \text{WERT} \rangle_n\}$.

Auf das obige Beispiel angewendet, bedeutet das, daß die linguistische Eigenschaftsausprägung *Verb* ein Wert ist, der zum Attribut *Wortart* gehört. Es muß also notiert werden:

Wortart:Verb

bzw.:

Wortart:[Verb].

Auf Grund dieser Unterscheidung können nun auch idiohabituelle linguistische Eigenschaftsausprägungen im Lemmazeichentyp berücksichtigt werden, indem lediglich das Attribut genannt wird, während der Wert offen bleibt; zur Kennzeichnung der Wertleerstelle, die lemmazeichentypspezifisch gefüllt werden muß, diene ein Unterstrich: „___“ bzw. drei Punkte („[...]“). Auf diese Weise muß die Eigenschaft *orthographische Realisierung des Lemmazeichens* nicht im Lemmazeichentyp getilgt werden, sondern bleibt als Element erhalten. Auf der

Grundlage dieser Überlegungen und Schreibkonventionen muß der oben eingeführte Lemmazeichentyp wie folgt korrigiert werden:¹⁵

$lzt_{4712}^{WB-10} = \langle \{ \text{Wortart: } \underline{Verb}_B, \text{ Silbenzahl: } \underline{2}_B, \text{ Akzent: } \underline{initial}_B, \text{ Akzentquantität: } \underline{kurz}_B, \text{ RegelgemäßeReine-HochlautungDesInfinitivs: } \underline{\quad}_B, \text{ RegelgemäßeGemäßigteHochlautungDesInfinitivs: } \underline{\quad}_B, \text{ OrthographieDesInfinitivs: } \underline{\quad}_B, \text{ Konjugiertheit: } \underline{schwach}_B, \text{ PerfektbildungMit: } \underline{haben}_B, \text{ Stammformen: } \underline{\quad}_B, \text{ Passivfähig: } \underline{nein}_B, \text{ Polysemiegrad: } \underline{1}, \text{ Geschehensart: } \underline{Tätigkeitsverb}_B, \dots \rangle, \{ mE_1, mE_2, \dots, mE_n \} \rangle$.

Nun kann der wörterbuchspezifizierte Lemmazeichentyp neu definiert werden:

[D 2-5': wörterbuchspezifizierter Lemmazeichentyp]

Ein *wörterbuchspezifizierter Lemmazeichentyp* (lzt_x^{WB} bzw. lzt_x^{WB-y}) ist ein 2-Tupel, bestehend aus einer Menge von linguistischen Eigenschaftsausprägungen (IE) und einer Menge von metalexikographischen Eigenschaftsausprägungen (mE), der hinsichtlich eines Wörterbuches (WB) bzw. Wörterbuchtyps (WB-x) und hinsichtlich einer linguistischen Überdachung (LÜ) konzipiert ist. Kurz:

lzt_x^{WB} bzw. $lzt_x^{WB-y} = \langle \{ IE_1, IE_2, \dots, IE_m \}, \{ mE_1, mE_2, \dots, mE_n \} \rangle$
 $= \langle \{ (ATTRIBUT, WERT)_1, (ATTRIBUT, WERT)_2, \dots, (ATTRIBUT, WERT)_m \}, \{ mE_1, mE_2, \dots, mE_n \} \rangle$
 mit $m, n \in \mathbb{N}^*$.

Bislang wurde die Frage, wie der Lemmazeichentyp mit dem Wörterbuchartikeltyp relationiert ist, immer noch nicht beantwortet; doch dies ist jetzt auf der Grundlage des bisher Erarbeiteten relativ einfach möglich. Dazu soll das Beispiel für einen Lemmazeichentyp präzisiert werden.

Weiterhin stehe das dt. Verb *grinsen* im Zentrum der Betrachtung. Der Lemmazeichentyp von *grinsen* sei nun in Bezug auf den Typ des Rechtschreibwörterbuches (= WB-3) konzipiert, genauer: hinsichtlich des DUDEN-²⁰1. Der Lemmazeichentyp weise keine metalexikographischen Eigenschaftsausprägungen auf, weshalb diese Menge leer ist. Die linguistische Überdachung sei wie folgt gegeben:

LÜ: B₁ = *Duden. Aussprachewörterbuch. Wörterbuch der deutschen Standardaussprache*. 3., völlig neu bearb. u. erw. Aufl. Bearb. v. MAX MANGOLD in Zusammenarbeit mit der Dudenredaktion. [= Der Duden in 10 Bänden. Das Standardwerk zur deutschen Sprache. Bd. 6.] Mannheim, Wien, Zürich: Bibliographisches Institut; Dudenverlag 1990.

B₂ = ULRICH ENGEL: *Deutsche Grammatik*. 3., korrigierte Aufl. Heidelberg: Julius Groos 1996.

B₃ = Bonner Zeitungscorpus

B₄ = LIMAS-Corpus

Für die lexikographische Beschreibung von Verben in Wörterbüchern vom Typ *Rechtschreibwörterbuch* sei folgende Richtlinie genuin, wie sie u.a. im DUDEN-²⁰1 formuliert ist:

[Z 2-3: DUDEN-²⁰1, 10 seq.]

2. Verben (Tätigkeitswörter, Zeitwörter)

- a) Bei den schwachen Verben werden im allgemeinen keine Formen angegeben, da sie regelmäßig im Präteritum (erste Vergangenheit) auf -te und im Partizip II (2. Mittelwort) auf -t ausgehen.

[...]

15. Es ist zu beachten, daß im Gegensatz zu WIEGAND (1998, 215 seq.) hier die Festlegung des Polysemiegrades der Festlegung des Lemmazeichentyps vorgängig ist, während bei WIEGAND der Polysemiegrad erst während der Dateneingabe an Hand von systematischen Fragen und Handlungsaufforderungen festgelegt wird. Es ist jedoch nach dem bisher Erläuterten ein leichtes, den Lemmazeichentyp so zu modifizieren, daß auch der Polysemiegrad offen bleibt.

- b) Bei den Verben, deren Stamm mit einem S-Laut oder Zischlaut endet (s, ß, sch, z, tz), wird die 2. Person Singular im Indikativ des Präsens (Wirklichkeitsform der Gegenwart) angegeben, weil -e- oder -es- der Endung gewöhnlich ausfällt.

Beispiele: zischen; du zischst; lesen; du liest; sitzen; du sitzt[.]

Daraus folgt für den Lemmazeichentyp $lzt_{4712}^{DUDEN-20I}$:

$lzt_{4712}^{DUDEN-20I} = \langle \{ \text{Wortart: } \underline{Verb}_{B_2}, \text{ Silbenzahl: } \underline{2}_{B_3B_4}, \text{ Akzent: } \underline{initial}_{B_1}, \text{ Akzentquantität: } \underline{kurz}_{B_1}, \text{ Regelgemäße Orthoepie: } \underline{ja}_{B_1}, \text{ OrthographieDesInfinitivs: } \underline{\quad}_{B_3B_4}, \text{ Konjugiertheit: } \underline{schwach}_{B_3B_4}, \text{ StammendungZischlaut: } \underline{ja}_{B_3B_4}, \text{ Silbe1: } \underline{\quad}_{B_3B_4}, \text{ Silbe2: } \underline{\quad}_{B_3B_4}, \text{ PragmatischeMarkierung: } \underline{keine}_{B_3B_4} \}, \{\} \rangle$.

Dieser Lemmazeichentyp kann nun mit dem Wörterbuchartikeltyp aus dem DUDEN-²⁰I, zu dem der Wörterbuchartikel zu *grinsen* im DUDEN-²⁰I gehört, korreliert werden. Dazu muß zuerst der Wörterbuchartikeltyp angegeben werden, indem die abstrakte hierarchische Mikrostruktur (zusammen mit einem Ausschnitt aus der dazu k-isomorphen konkreten hierarchischen Mikrostruktur) wiedergegeben wird, die sich in der nächsten Abbildung findet.

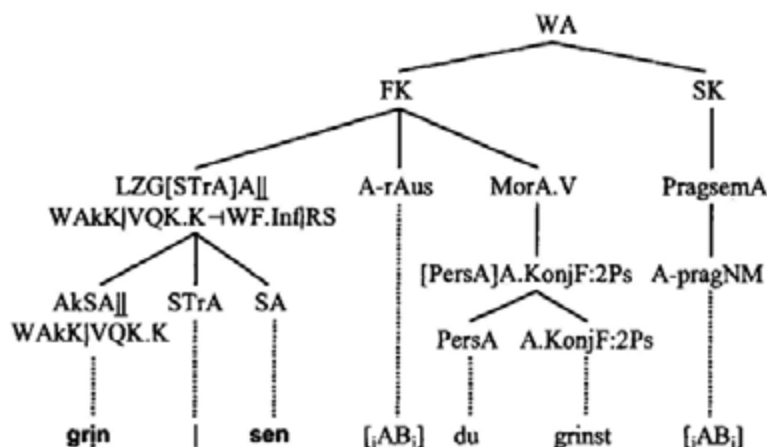


Abb. 2-7: Strukturgraph zur (abstrakten und dazu isomorphen konkreten) einfachen hierarchischen Mikrostruktur, die zum Wörterbuchartikel zu *grinsen* aus dem DUDEN-²⁰I (314^c) gehören; Abkürzungen: MorA.V = Morphologieangabe beim Verb; [PersA]A.KonjF:2Ps = eine um eine Personenangabe links erweiterte Angabe der Konjugationsform für die 2. Pers. Sg. (= 2) Präs. (= Ps); WF.Inf = Wortform des Infinitivs

Betrachtet man sowohl den Wörterbuchartikeltyp, der durch die abstrakte hierarchische Mikrostruktur gegeben ist, als auch den Lemmazeichentyp $lzt_{4712}^{DUDEN-20I}$, so fallen mehrere Parallelen auf: Die linguistische Eigenschaftsausprägung „Silbenzahl: $\underline{2}_{B_3B_4}$ “ beispielsweise verlangt neben zwei Silbenangaben eine Silbentrennungsangabe, da die Angabe der Silbenanzahl und der Silbengrenze im DUDEN-²⁰I und damit auch in dessen Wörterbuchprogramm vorgesehen ist; allgemeiner formuliert: Wenn die Silbenzahl n ist, dann sind n Silbenangaben und $n-1$ Silbentrennungsangaben nötig (mit $n \geq 1$). Die linguistische Eigenschaftsausprägung „StammendungZischlaut: $\underline{ja}_{B_3B_4}$ “ verlangt eine bestimmte Form der Morphologieangabe, nämlich eine Personenangabe zusammen mit einer Angabe der Konjugationsform für die 2. Pers. Sg. Präs., wie sie im obigen Zitat Z 2-3 expliziert wurde, etc. Insgesamt lassen sich folgende Abhängigkeiten beobachten:

linguistische Eigenschaftsausprägung	korrelierte Angabeklasse / Kennzeichnungsklasse	Angabe(n) / Kennzeichnung
Wortart: <i>Verb_{B1}</i>	MorA.V	?
Silbenzahl: <i>2_{B1B4}</i>	AkSA; SA; STra	grin; sen;
Silbe1: <i>____{B1B4}</i>	AkSA	grjn
Silbe2: <i>____{B1B4}</i>	SA	sen
Akzent: <i>initial_{B1}</i>	WakK	.
Akzentquantität: <i>kurz_{B1}</i>	VQK.K	.
RegelgemäßeOrthoepie: <i>ja_{B1}</i>	A-rAus	[iABj]
OrthographieDesInfinitivs: <i>____{B1B4}</i>	LZGA	grin sen
Konjugiertheit: <i>schwach_{B2}</i>	MorA.V	?
StammendungZischlaut: <i>ja_{B1B4}</i>	[PersA]A.KonjF:2Ps	du grinst
PragmatischeMarkierung: <i>keine_{B1B4}</i>	A-pragNM	[iABj]

Tab. 2-1: Korrelationen zwischen linguistischen Eigenschaftsausprägungen eines Lemmazeichentyps, den Angabeklassen der abstrakten hierarchischen Mikrostruktur eines Wörterbuchartikeltyps und den Angaben der präzedentiven konkreten Mikrostruktur des Wörterbuchartikels zu *grinsen*; *Darstellungskonvention:* „?“ bedeutet soviel wie *ist fraglich und kann nicht genau angegeben werden*

Diese Zusammenhänge können auch wie folgt dargestellt werden:

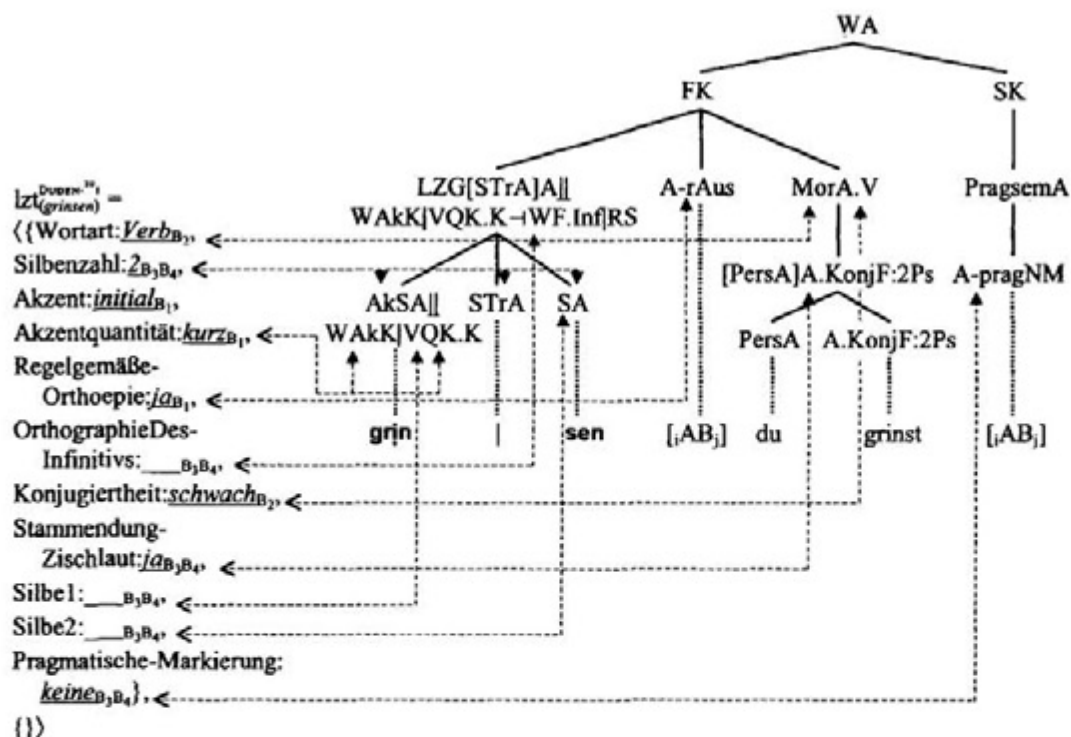


Abb. 2-8: Strukturgraph zur (abstrakten und dazu isomorphen konkreten) einfachen hierarchischen Mikrostruktur, die zum Wörterbuchartikel zu *grinsen* aus dem DUDEN-¹⁹1 (314⁵) gehören, wobei Angabeklassen bzw. Kennzeichnungen mit dem Lemmazeichentyp *lzt_{DUDEN-19}(grinsen)* korreliert sind; *Darstellungskonvention:* „x <----> y“ bedeutet soviel wie: *x ist mit der Angabeklasse/Kennzeichnung y korreliert*

In einem Rechtschreibwörterbuch, in dem Silbentrennungsangaben und Akzentangaben nicht vorgesehen sind, wie z.B. dem DUDEN (¹1890), sähe der zugehörige Lemmazeichentyp anders aus, beispielsweise wie folgt:

$$lzt_{4712}^{\text{DUDEN}^{(1890)}} = \langle \{ \text{Wortart: } \underline{\text{Verb}}, \text{ OrthographieDesInfinitivs: } \underline{\quad}, \text{ Konjugiertheit: } \underline{\text{schwach}}, \dots \}, \{ \} \rangle.$$

Hier wird deutlich, daß der Lemmazeichentyp sehr stark vom Wörterbuch(typ) abhängig ist, auf das/den er bezogen ist. So gilt beispielsweise:

$$lzt_{4712}^{\text{DUDEN}^{(1890)}} \neq lzt_{4712}^{\text{DUDEN}^{201}} \neq lzt_{4712}^{\text{WB-3}} \neq lzt_{4712}^{\text{DUDEN}^{(1890)}}.$$

Folgender Zusammenhang muß jedoch betont werden: Gegeben sei der Lemmazeichentyp Nr. 4712 für den Wörterbuchtyp *Rechtschreibwörterbuch* (= WB-3; also: $lzt_{4712}^{\text{WB-3}}$) sowie für die konkreten Wörterbücher DUDEN-²⁰1 (= $lzt_{4712}^{\text{DUDEN}^{201}}$) und DUDEN (¹1890) (= $lzt_{4712}^{\text{DUDEN}^{(1890)}}$), nämlich:

$$\begin{aligned} lzt_{4712}^{\text{WB-3}} &= \langle E_1, \{ \dots \} \rangle \\ lzt_{4712}^{\text{DUDEN}^{201}} &= \langle E_2, \{ \dots \} \rangle \\ lzt_{4712}^{\text{DUDEN}^{(1890)}} &= \langle E_3, \{ \dots \} \rangle. \end{aligned}$$

Nun sollten – falls DUDEN-²⁰1 und DUDEN (¹1890) tatsächlich zum Wörterbuchtyp *Rechtschreibwörterbuch* gehören – die folgenden Beziehungen gelten:

1. $E_1 \subseteq E_2$;
2. $E_1 \subseteq E_3$;
3. $E_1 \subseteq E_2 \cap E_3$;
4. $E_2 \cap E_3 \neq \emptyset$.

Dies ist aber nicht der Fall! Dazu beachte man: In einem Wörterbuch, das zum Typ *Rechtschreibwörterbuch* gehört, sind genuin Angaben zur Rechtschreibung zu verzeichnen, wozu u.a. die Angabe der korrekten Silbentrennung gehört. Es muß also für den Lemmazeichentyp von *grinsen* hinsichtlich des Wörterbuchtyps WB-3 u.a. gelten:

$$\begin{aligned} \text{OrthographieDesInfinitivs: } \underline{\quad} &\in E_1 \\ \text{Silbenzahl: } \underline{\quad} &\in E_1 \\ \text{Silbe1: } \underline{\quad} &\in E_1 \\ \text{Silbe2: } \underline{\quad} &\in E_1 \\ \text{Silbe3: } \underline{\quad} &\in E_1 \dots \end{aligned}$$

Die korrekte Silbentrennungsangabe wird aber im DUDEN (¹1890), der laut Titel zum Typ der Rechtschreibwörterbücher gehört, nicht angegeben. Deshalb gilt hier:

$$\begin{aligned} \text{OrthographieDesInfinitivs: } \underline{\quad} &\in E_3 \\ \text{Silbenzahl: } \underline{\quad} &\notin E_3 \\ \text{Silbe1: } \underline{\quad} &\notin E_3 \\ \text{Silbe2: } \underline{\quad} &\notin E_3 \\ \text{Silbe3: } \underline{\quad} &\notin E_3 \dots \end{aligned}$$

und damit gilt hier nicht $E_1 \subseteq E_3$, sondern $E_1 \not\subseteq E_3$. Auf Grund dieser Eigenschaften gehört der DUDEN (¹1890) nicht zum Typ der Rechtschreibwörterbücher. Um jedoch eine Redeweise, wie: »Der DUDEN (¹1890) ist ein defizientes Rechtschreibwörterbuch« o.ä., zu vermeiden, bietet es sich an, Formulierungen, wie »Der DUDEN (¹1890) entspricht nicht dem Idealtypus eines Rechtschreibwörterbuches« oder »Der DUDEN (¹1890) entspricht nicht dem Prototyp eines Rechtschreibwörterbuches«, zu wählen.

Die Ausspracheangabe, wie man sie im DUDEN-²⁰1 vorfindet, ist – im Gegensatz zur Silbentrennungsangabe – keine genuine Angabe eines Wörterbuches vom Typ *Rechtschreibwörterbuch*. Deshalb gehört die linguistische Eigenschaftsausprägung „RegelgemäßeOrthoepie:ja“ auch nicht notwendigerweise zu einem Lemmazeichentyp, der für Wörterbücher des Typs WB-3 entwickelt wurde, sondern zu Wörterbüchern, die zum Typ *Aussprachewörterbuch* (= WB-5) gehören. Es gilt demnach:

RegelgemäßeOrthoepie:ja $\notin \mathbb{E}$.

Der Lemmazeichentyp mit der Nummer 4712 für Aussprachewörterbücher (= $\text{lzt}_{4712}^{\text{WB-5}}$) könnte dann wie folgt aussehen:

$\text{lzt}_{4712}^{\text{WB-5}} = \{ \{ \text{Akzent:} \underline{\text{initial}}, \text{Akzentquantität:} \underline{\text{kurz}}, \text{Orthoepie: } \underline{\quad} \}, \{ \dots \} \}$

oder:

$\text{lzt}_{4712}^{\text{WB-5}} = \{ \{ \text{Akzent:} \underline{\text{initial}}, \text{Akzentquantität:} \underline{\text{kurz}}, \text{RegelgemäßeReineHochlautungDesInfinitivs: } \underline{\quad}, \text{RegelgemäßeGemäßigteHochlautungDesInfinitivs: } \underline{\quad} \}, \{ \dots \} \}$.

– Die in der obigen Tabelle 2-1, p. 25, aufgelisteten Abhängigkeiten zwischen den linguistischen Eigenschaftsausprägungen des Lemmazeichentyps und den Angabeklassen der abstrakten hierarchischen Mikrostruktur sowie den Angaben der konkreten präzedentiven Mikrostruktur lassen folgende Beobachtungen zu: Es gibt linguistische Eigenschaftsausprägungen, denen zwar eine Angabeklasse zugeordnet werden kann, nicht aber eine konkrete Angabe (verdeutlicht durch das kursive Fragezeichen). So verlangt beispielsweise die Eigenschaft „Wortart:Verb“ allein noch nicht die Angabe der Konjugationsform der 2. Pers. Sg. Präs. (dies wird erst durch die Eigenschaft „StammendungZischlaut:ja“ gemäß Zitat Z 2-3, p. 23, erforderlich). Dasselbe gilt für die Eigenschaft „Konjugiertheit:schwach“. Dennoch können beide Eigenschaften fraglos der Angabeklasse MorA.V zugeordnet werden, denn eines ist sicher: Die linguistische Eigenschaftsausprägung eines Lexems, ein Verb zu sein, erfordert eine grundsätzlich andere Morphologieangabe als die linguistische Eigenschaftsausprägung, ein Substantiv oder Adjektiv zu sein. Ebenso *kann* die Eigenschaft „Konjugiertheit:schwach“ weitere Angabeklassen fordern, die allerdings hier nicht realisiert sind. Daraus ist ersichtlich, daß die linguistischen Eigenschaftsausprägungen teilweise selbst zueinander in einem Abhängigkeitsverhältnis stehen: So ist es nicht möglich, daß, wenn Wortart:Verb $\in \mathbb{E}^1$ ist, auch z.B. Genus:femininum $\in \mathbb{E}^1$ ist; mit anderen Worten: Gewisse linguistische Eigenschaftsausprägungen schließen bestimmte andere Eigenschaftsausprägungen aus und eröffnen im Gegenzug die Möglichkeit, daß bestimmte weitere linguistische Eigenschaftsausprägungen Elemente von $\mathbb{E}^1$ sind.

Außerdem kann der obigen Tabelle 2-1 entnommen werden, daß einer linguistischen Eigenschaftsausprägung u.U. mehrere Angabeklassen korreliert sind, wie das bei „Silbenzahl:2“ der Fall ist. Diese Eigenschaft des Lemmazeichens hat zur Folge, daß zwei Silbenangaben gemacht werden müssen, wovon die eine eine Akzentsilbenangabe ist, die sich aus der Eigenschaftsausprägung „Akzent:initial“ ergibt, sowie eine Silbentrennungsangabe. Umgekehrt kann es aber auch sein, daß zwei verschiedene Eigenschaften genau eine einzige Angabe zur Folge haben, wie das bei den Eigenschaften „Akzent:initial“ und „Akzentquantität:kurz“ der Fall ist: Beide Eigenschaften werden durch den Unterpunkt „ “ realisiert, der gleichzeitig eine

Wortakzentkennzeichnung (WakK) und eine Vokalquantitätskennzeichnung der Kürze (VQK.K) ist.

Schließlich können bestimmte Eigenschaftsausprägungen durch andere ersetzt werden, ohne daß sich der Lemmazeichentyp ändert; ein Beispiel wäre die Ersetzung von „Akzent: *initial*“ durch „Initialakzent: *ja*“. Des weiteren könnten die Eigenschaften „Initialakzent: *ja*“ und „Akzentquantität: *kurz*“ durch „Initialakzent: *kurz*“ ersetzt werden; mit anderen Worten: Eine linguistische Eigenschaftsausprägung IE_1 kann durch eine andere linguistische Eigenschaftsausprägung IE_2 ersetzt werden, wenn sie identisch sind. Ebenso können die linguistischen Eigenschaftsausprägungen IE_3 und IE_4 durch IE_5 ersetzt werden, wenn der propositionale Gehalt von IE_5 dem von IE_3 & IE_4 entspricht (wobei das kaufmännische *und* als Konkatenator von propositionalen Gehalten zu interpretieren wäre).

Die in der Tabelle 2-1 aufgelisteten Korrelationen machen darüber hinaus noch folgendes deutlich: Ein Lemmazeichentyp mit seinen *linguistischen* Eigenschaftsausprägungen determiniert zwar einen großen Teil der in einem Wörterbuchartikel geforderten Angabeklassen und damit zum großen Teil die Trägermenge der zugehörigen Mikrostruktur, was sich wiederum auf die Selektion des Wörterbuchartikeltyps auswirkt, aber keineswegs die Mikrostruktur selbst! Die Mikrostruktur – verstanden als ein 2-Tupel, bestehend aus einer Trägermenge und n Relationen, die auf der Trägermenge definiert sind (mit $n \geq 1$), – wird durch die linguistischen Eigenschaftsausprägungen nicht vorgegeben. So könnte die Mikrostruktur von *grinsen* bei derselben angenommenen linguistischen Eigenschaftsausprägung auch wie folgt aussehen:

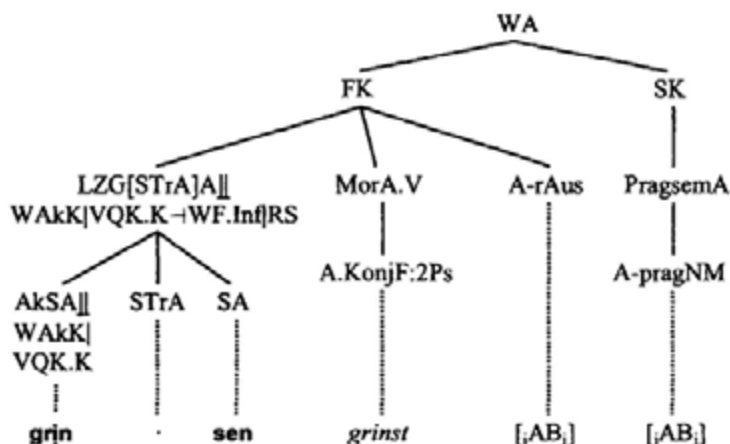


Abb. 2-9: Strukturgraph zur (abstrakten und dazu isomorphen konkreten) einfachen hierarchischen Mikrostruktur, die zu einem fiktiven Wörterbuchartikel zu *grinsen* gehören

Dies ist offensichtlich eine andere abstrakte hierarchische Mikrostruktur als diejenige im DUDEN-²⁰ I.

Aus dem Erläuterten kann nun gefolgert werden, daß zwar zwischen den linguistischen Eigenschaftsausprägungen eines gegebenen Lemmazeichentyps und den Angabeklassen, wie sie in der abstrakten hierarchischen Mikrostruktur eines Wörterbuchartikeltyps auftreten, eine Zuordnungsrelation $\mathbb{Z}$ besteht, diese ist aber keine 1:1-Beziehung, sondern eine $m:n$ -Beziehung (mit $m, n \geq 1$). Der Relationsterm von $\mathbb{Z}$ lautet: ... wird/werden zugeordnet der/den An-

gabeklasse(n) bzw. *Kennzeichnungsklasse(n)* ... Diese Relation $\mathbb{Z}$ ist zwar linkstotal, aber nicht rechtseindeutig. Damit handelt es sich nicht um den Spezialfall einer Funktion.

Die Zuordnungsrelation $\mathbb{Z}$ stellt damit die gesuchte Verbindung zwischen dem Wörterbuchartikeltyp und dem Lemmazeichentyp her. Es gilt (mit *a* als Abkürzung für *Angabe-/Kennzeichnungsklasse aus einer abstrakten hierarchischen Mikrostruktur eines Wörterbuchartikels*):

$$\forall IE \in \mathbb{E}^1 \exists a [(IE, a) \in \mathbb{Z}].$$

Betrachtet man nochmals das obige Beispiel mit *grinsen*, so sind folgende 2-Tupel Elemente der Relation $\mathbb{Z}$:

(Wortart: <i>Verb</i> _{B₂})	MorA.V)	∈ $\mathbb{Z}$
(Silbenzahl: <i>2</i> _{B₃B₄})	AkSA)	∈ $\mathbb{Z}$
(Silbenzahl: <i>2</i> _{B₃B₄})	SA)	∈ $\mathbb{Z}$
(Silbenzahl: <i>2</i> _{B₃B₄})	STrA)	∈ $\mathbb{Z}$
(Silbe1: <i>gr</i> _{B₃B₄})	AkSA)	∈ $\mathbb{Z}$
(Silbe2: <i>in</i> _{B₃B₄})	SA)	∈ $\mathbb{Z}$
(Akzent: <i>initial</i> _{B₃})	WakK)	∈ $\mathbb{Z}$
(Akzentquantität: <i>kurz</i> _{B₃})	VQK.K)	∈ $\mathbb{Z}$
(RegelgemäßeOrthoepie: <i>ja</i> _{B₃})	A-rAus)	∈ $\mathbb{Z}$
(OrthographieDesInfinitivs: <i>grin</i> _{B₃B₄})	LZGA)	∈ $\mathbb{Z}$
(Konjugiertheit: <i>schwach</i> _{B₃})	MorA.V)	∈ $\mathbb{Z}$
(StammendungZischlaut: <i>grin</i> _{B₃B₄})	[PersA]A.KonjF:2Ps)	∈ $\mathbb{Z}$
(PragmatischeMarkierung: <i>keinen</i> _{B₃B₄})	A-pragNM)	∈ $\mathbb{Z}$

Es stellt sich nun die Frage, wie $\mathbb{Z}$ festgelegt wird. Dies ist die Aufgabe desjenigen, der das Instruktionbuch erarbeitet, denn im Instruktionbuch ist nicht nur anzugeben, welche lexikographischen Datenangebote auf welche Weise präsentiert werden sollen, sondern vor allem zwei weitere Dinge: (1) welche Lemmazeichentypen im Hinblick auf das zu erarbeitende Wörterbuch unterschieden werden und (2) welchen Angaben die einzelnen linguistischen Eigenschaftsausprägungen korreliert sind, mit anderen Worten: welche Elemente $\mathbb{Z}$ enthält. Darüber hinaus wird über die Angabe des Datentypensortiments im Instruktionbuch, das auch das Wörterbuchprogramm beinhaltet, festgelegt, welche Angabeform die konkreten Angaben am Ende besitzen und welche typographischen Strukturanzeiger an welcher Stelle in der präzedentiven Artikelstruktur gesetzt werden. So wäre es denkbar, daß Lemmata nicht fett, sondern kursiv oder mit einer anderen Farbe, beispielsweise blau, ausgezeichnet werden (wie z.B. in LEX-¹INF) und daß die Bedeutungsparaphrasenangabe in einer anderen (beispielsweise serifenlosen) Schrift erscheint, wie das z.B. im ¹WMU der Fall ist. Aber diese Zuordnungen werden nicht über die Relation $\mathbb{Z}$ festgelegt.

Beim Wörterbuchartikeltyp wurde ein enger und ein weiter Wörterbuchartikeltyp unterschieden (cf. D 2-2, p. 10, und D 2-3, p. 11). Ganz entsprechend kann auch zwischen *engen* und *weiten Lemmazeichentypen* unterschieden werden. Diese lassen sich wie folgt definieren:

[D 2-13: enger Lemmazeichentyp]

Ein *enger Lemmazeichentyp* ist ein intensional gegebener Lemmazeichentyp, für den folgende Bedingungen gelten:

1. Die Menge der linguistischen Eigenschaftsausprägungen E^1 enthält Elemente der Form:
ATTRIBUT: WERT.
2. Wenn die Menge der linguistischen Eigenschaftsausprägungen E^1 auch Elemente der Form:
ATTRIBUT: ____

enthält, so handelt es sich immer um idiohabituelle linguistische Eigenschaftsausprägungen.

[D 2-14: weiter Lemmazeichentyp]

Ein *weiter Lemmazeichentyp* ist ein intensional gegebener Lemmazeichentyp, für den folgende Bedingung gilt:

1. Die Menge der linguistischen Eigenschaftsausprägungen E^L enthält mindestens 1 Element der Form:

ATTRIBUT: _____
das eine habituelle linguistische Eigenschaftsausprägung ist.

Das bedeutet, daß ein Lemmazeichentyp, wenn er *weit* ist, lexikalische Einheiten zusammenfaßt, die weiter differenziert werden *könnten*. Dies wird aber – aus entsprechenden Gründen – nicht getan, weil z.B. eine weitergehende Diversifizierung zu einer unökonomischen Arbeitsweise führen würde. So könnte in dem obigen Beispiel die habituelle linguistische Eigenschaftsausprägung „Polysemiegrad: 1“ durch „Polysemiegrad: _____“ ersetzt werden, wodurch aus dem engen Lemmazeichentyp lzt_{4712}^{WB-10} ein weiter Lemmazeichentyp lzt_{4712}^{WB-10} werden würde. Für den extensionalen Lemmazeichentyp bedeutet dies, daß alle als *n*-fach polysem interpretierten Lemmazeichen (mit $n \geq 2$) zusammen mit allen monosemen Lemmazeichen in einer Menge zusammengefaßt werden, sofern sie die übrigen genannten linguistischen Eigenschaften besitzen.

Auf diesem Hintergrund kann nun bei den extensionalen Lemmazeichentypen zwischen *homogenen* und *heterogenen* Lemmazeichentypen unterschieden werden:

[D 2-15: homogener Lemmazeichentyp]

Ein *homogener Lemmazeichentyp* ist ein extensionaler Lemmazeichentyp, der durch einen engen (intensionalen) Lemmazeichentyp gegeben ist.

[D 2-16: heterogener Lemmazeichentyp]

Ein *heterogener Lemmazeichentyp* ist ein extensionaler Lemmazeichentyp, der durch einen weiten (intensionalen) Lemmazeichentyp gegeben ist.

Ein Lemmazeichentyp ist *n*-fach *heterogen*, wenn in der Menge der linguistischen Eigenschaftsausprägungen E^L des zugehörigen weiten Lemmazeichentyps *n* Elemente der Form:

ATTRIBUT: _____
enthalten sind, die alle habituell sind.

2.4 Ein abschließender Blick auf ausgewählte metalexikographische Eigenschaftsausprägungen

Bei allem bisher Erläuterten wurde die Menge der metalexikographischen Eigenschaftsausprägungen noch nicht berücksichtigt. Auf diese soll nun näher eingegangen werden. Dabei ist darauf hinzuweisen, daß von den metalexikographischen Eigenschaftsausprägungen eines Lemmazeichentyps die metalexikographischen Eigenschaftsausprägungen des zentralen (bzw. des 1., 2., ...) Wörterverzeichnis eines Wörterbuches *etc.* zu unterscheiden sind. Beispielsweise werden bestimmte Aspekte der Makrostruktur eines Wörterverzeichnisses nicht durch den Lemmazeichentyp festgelegt, sondern durch das schriftlich fixierte Wörterbuchprogramm; d.h.: Entweder ist die Makrostruktur eines Wörterverzeichnisses initialalphabetisch oder finalalphabetisch (sofern Buchstabenketten als Leitelementträger fungieren), aber nicht abwechselnd einmal final- und einmal initialalphabetisch. (Das bedeutet nicht, daß in-

nerhalb eines Wörterbuches das eine Wörterverzeichnis eine initialalphabetische Makrostruktur aufweisen kann und ein anderes Wörterverzeichnis im selben Buchbinderband eine finalalphabetische.) Ebenso *kann* – neben anderen Wörterbuchstrukturen – für bestimmte Aspekte der Mikrostruktur sowie der Mikroarchitektur gelten (unten werden Beispiele genannt, wo dies nicht der Fall ist), daß diese für die einzelnen Wörterverzeichnisse pauschal im Wörterbuchprogramm festgelegt sind und nicht durch den jeweiligen Lemmazeichentyp bestimmt werden.

Somit ist zusammenfassend festzustellen: Alle Aspekte von Wörterbuchstrukturen, die für einzelne Wörterverzeichnisse, Einschübe, sonstige Umtexte *etc.* eines Wörterbuches als solche verbindlich sind, gehören nicht zu den metalexikographischen Eigenschaftsausprägungen eines Lemmazeichentyps und werden im Instruktionsbuch festgehalten. Nur diejenigen Aspekte von Wörterbuchstrukturen, die vom Lemmazeichentyp abhängig sind bzw. sein können, sind Inhalt der metalexikographischen Eigenschaftsausprägungen eines Lemmazeichentyps.

Beispielsweise wird durch den Lemmazeichentyp – wie bereits erwähnt – nicht festgelegt, ob die Makrostruktur in einem gegebenen Wörterverzeichnis, in dem das Lemmazeichen lexikographisch behandelt werden soll, initialalphabetisch oder finalalphabetisch sein soll – dies wird vielmehr durch das Wörterbuchprogramm festgelegt. Im Gegensatz dazu bestimmt die metalexikographische Eigenschaftsausprägung eines Lemmazeichentyps auf Grund der formalen Präsentation durch Buchstaben, wo das Lemma in der Zugriffsrelation ($\mathbb{Z}\mathbb{R}$) auftritt (genauer ausgedrückt: mit welchen anderen Lemmata es als 2-Tupel Element der Zugriffsrelation ist). Angenommen, in einem Wörterbuch müßten die Lemmazeichen *keifen* und *keilen* lexikographisch behandelt werden. An Hand der metalexikographischen Eigenschaftsausprägungen der beiden Lemmazeichentypen kann nicht festgestellt werden, ob die Makrostruktur, zu deren Trägermenge sie gehören, finalalphabetisch oder initialalphabetisch ist, um hier nur die einfachsten Formen der Makrostruktur zu betrachten (für weitere Unterscheidungen cf. insbes. WIEGAND 1998b). Wenn aber durch das Wörterbuchprogramm vorgegeben ist, daß sie initialalphabetisch sein soll, dann gilt für die metalexikographischen Eigenschaftsausprägungen von *keifen* und *keilen*:

keifen < keilen < keimen

(mit „ $x < y$ “ bedeutet soviel wie *x geht y in der Makrostruktur voraus*);

(mithin gilt: $\langle \text{keifen}, \text{keilen} \rangle \in \mathbb{Z}\mathbb{R}$; $\langle \text{keifen}, \text{keimen} \rangle \in \mathbb{Z}\mathbb{R}$; $\langle \text{keilen}, \text{keimen} \rangle \in \mathbb{Z}\mathbb{R}$;

und: $\langle \text{keimen}, \text{keilen} \rangle \notin \mathbb{Z}\mathbb{R}$; $\langle \text{keimen}, \text{keifen} \rangle \notin \mathbb{Z}\mathbb{R}$; $\langle \text{keilen}, \text{keifen} \rangle \notin \mathbb{Z}\mathbb{R}$).

Wenn darüber hinaus feststeht, daß es für das Wörterbuch kein weiteres lexikographisch zu behandelndes Lemmazeichen *lz* gibt, für das gilt:

keifen < lz < keilen,

weil es sich z.B. um ein Wörterbuch der Verbvalenz handelt und weil es im Gegenstandsreich kein weiteres Verb gibt, das initialalphabetisch nach *keifen* und gleichzeitig vor *keilen* eingereiht werden müßte, dann gilt außerdem:

keifen << keilen

(mit „ $x << y$ “ bedeutet soviel wie *x geht y in der Makrostruktur unmittelbar voraus*).

Dies bedeutet, daß – hinsichtlich der Makrostruktur – in der metalexikographischen Eigenschaftsausprägung eines jeden Lemmazeichentyps angegeben ist, in welcher Ordnungsrelation

das durch das Lemma repräsentierte Lemmazeichen zu allen anderen Lemmata des Wörterverzeichnisses steht.

Ebenso kann die metalexikographische Eigenschaftsausprägung eines Lemmazeichentyps weitere Strukturen mitbestimmen, z.B. die Suchbereichsarchitektur oder die innere Schnellzugriffsstruktur. Im Folgenden werden nur noch diejenigen metalexikographischen Eigenschaftsausprägungen genauer betrachtet, die Wörterbuchartikel bestimmen; dies kann unter drei Aspekten geschehen:

- hinsichtlich der *Form*;
- hinsichtlich des *Inhalts* und
- hinsichtlich der Art und Weise, wie er erstellt wird (*Wörterbuchartikelmanagement*, oder kurz: *Artikelmanagement*).

Zu denjenigen metalexikographischen Eigenschaftsausprägungen, die die *Form* bestimmen, gehören u.a.:

- der Wörterbuchumfang;
- die Wörterbuchartikelgrößenklasse (oder: Artikelgrößenklasse) und
- die Wörterbuchartikelform (oder: Artikelform)

sowie u.a. die folgenden Strukturen:

- die Artikelkonstituentenstruktur;
- die Mikrostruktur;
- die Mikroarchitektur;
- die innere (Schnell-) Zugriffsstruktur;
- die Suchbereichsstruktur und
- die Suchbereichsarchitektur.

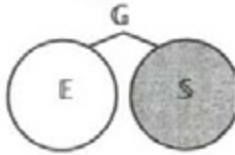
Metalexikographische Eigenschaftsausprägungen, die den *Inhalt* eines Wörterbuchartikels bestimmen, sind u.a.:

- der Grad der Buchungswürdigkeit eines Lemmazeichens;
- die Datendistributionsstruktur und
- die Mediosstruktur.

Bei der Datendistributionsstruktur gibt es grundsätzlich nur sechs Möglichkeiten der Datenverteilung, die im nachfolgenden Exkurs kurz erläutert werden. Die Wahl eines Verteilungsmusters legt damit eine metalexikographische Eigenschaftsausprägung des Lemmazeichentyps fest.

Exkurs

Die sechs möglichen Verteilungsmuster, wie sie in BERGENHOLTZ/TARP/WIEGAND (1999, 1782seq.) im Zusammenhang mit der Datendistributionsstruktur informell expliziert werden, können formal (d.h. mittels VENN-Diagrammen und Aussagen über die Mengenbeziehungen) leicht wie folgt gefaßt werden (es bedeuten: G = Grundmenge der (fach-) enzyklopädischen Dateneinheiten eines Wörterbuches; E = Menge der (fach-) enzyklopädischen Dateneinheiten in synopseorientierten Einzelartikeln; S = Menge der (fach-) enzyklopädischen Dateneinheiten in Synopseartikeln; Q = irgendeine ausgezeichnete Menge von (fach-) enzyklopädischen Dateneinheiten; $\emptyset$ = leere Menge):

– Verteilungsmuster 1:¹⁶

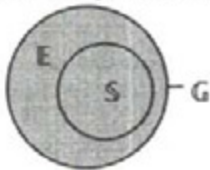
$$\begin{aligned}\text{Es gilt: } & G \neq \emptyset \\ & E = \emptyset \\ & E \cup S = G \\ & S = G\end{aligned}$$

– Verteilungsmuster 2:



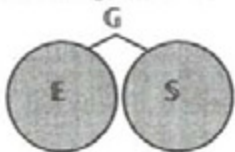
$$\begin{aligned}\text{Es gilt: } & G \neq \emptyset \\ & E = S = G\end{aligned}$$

– Verteilungsmuster 3:



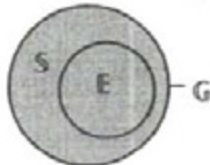
$$\begin{aligned}\text{Es gilt: } & G \neq \emptyset \\ & E = G \\ & S \subset G\end{aligned}$$

– Verteilungsmuster 4:



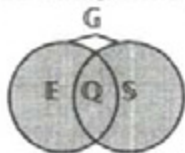
$$\begin{aligned}\text{Es gilt: } & G \neq \emptyset \\ & E \cap S = \emptyset \\ & E \cup S = G\end{aligned}$$

– Verteilungsmuster 5:



$$\begin{aligned}\text{Es gilt: } & G \neq \emptyset \\ & S = G \\ & E \subset G\end{aligned}$$

– Verteilungsmuster 6:



$$\begin{aligned}\text{Es gilt: } & G \neq \emptyset \\ & E \cup S = G \\ & E \cap S = Q \\ & Q \neq \emptyset\end{aligned}$$

$$\begin{aligned}E \cup Q &= E \\ S \cup Q &= S \\ E \setminus S &\neq \emptyset \\ S \setminus E &\neq \emptyset\end{aligned}$$

Obwohl sich die hiesigen Ausführungen lediglich auf (fach-) enzyklopädische Daten und ihre Verteilung auf Einzel- und Synopseartikel beziehen (genauer: Es werden hier nur Formen der *einfachen monoalphabetischen artikelheterogenen Datendistributionsstruktur* betrachtet; cf. BERGENHOLTZ/TARP/WIEGAND 1999, 1778–1780), können daraus dennoch alle grundsätzlich möglichen Distributionsmöglichkeiten abgelesen werden, die leicht auch auf bestimmte andere Datendistributionsstrukturen übertragen werden können.

Ende des Exkurses. –

16. Für die nun folgende Darstellung der VENN-Diagramme, bei denen symbolisch die Elemente einer Menge durch eine geschlossene Kurve in der Ebene umrandet werden, gelte, daß die einzelnen Elemente nicht explizit aufgeführt, sondern durch eine Graufäche angedeutet werden. Dies bedeutet, daß die Visualisierung der leeren Menge mit Hilfe eines VENN-Diagramms dadurch geschieht, daß die von der Kurve umschlossene Fläche nicht grau eingefärbt wird. – Die Mengenbezeichner (wie z.B. G oder E) werden den VENN-Diagrammen entweder einbeschrieben oder mittels eines Striches zugeordnet.

Metalexikographische Eigenschaftsausprägungen, die das *Management* eines Wörterbuchartikels bestimmen, sind u.a.:

- der Bearbeiter;
- die Dateneingabemaske und
- die Zeitvorgabe für die Bearbeitung.

Nun einige erläuternde Worte dazu. – Eine metalexikographische Eigenschaftsausprägung einer lexikalischen Einheit kann deren *Buchungswürdigkeit* betreffen und die drei Werte *buchungswürdig*, *bedingt buchungswürdig* und *nicht buchungswürdig* annehmen. Diese Eigenschaft kann durch die Häufigkeit des Auftretens in den zugrunde gelegten Corpora bestimmt werden oder dadurch, daß das Lemmazeichen in irgendeiner Weise markiert ist, so daß es zum Lemmabestand des zu erarbeitenden Wörterbuches zu zählen ist oder nicht. Beispielsweise ist der medizinische Ausdruck *Arteriosklerose* im Hinblick auf ein linguistisches Fachwörterbuch oder auf ein zweisprachiges Wörterbuch des deutschen Grundwortschatzes nicht buchungswürdig, nicht jedoch in Bezug auf ein medizinisches Fachwörterbuch oder ein allgemeines einsprachiges Wörterbuch der deutschen Gegenwartssprache. Insofern wird die metalexikographische Eigenschaftsausprägung nicht allein durch das Lemmazeichen bestimmt, sondern auch wieder durch das Wörterbuch, das erarbeitet werden soll, und auch durch das Wörterbuchprogramm.

Mit der Eigenschaftsausprägung der Buchungswürdigkeit hängt die Eigenschaftsausprägung der *Wörterbuchartikelgrößenklasse* (oder kurz: *Artikelgrößenklasse*) eng zusammen. Je nach Genauigkeitsgrad lassen sich mehrere Größenklassen unterscheiden, wie z.B.:

- Einzeilenartikel;
- Zweizeilenartikel etc.;
- Kurzartikel;
- Standardartikel;
- Großartikel;
- Magnumartikel;
- Spaltenartikel;
- Seitenartikel etc.

Die Artikelgrößenklasse hat – neben Eigenschaften der Mikrostruktur – direkten Einfluß auf die Benutzerfreundlichkeit und die Benutzungsstrategien bezüglich eines Wörterbuches. So gilt in der Regel: Je kürzer ein Wörterbuchartikel ist (und damit: je näher die Angaben beim Lemma stehen) um so schneller werden die gesuchten Angaben von einem Benutzer-in-actu gefunden. Und umgekehrt: Je umfangreicher ein Wörterbuchartikel ist, um so benutzerfreundlicher ist er (wenn nicht durch die Wahl geeigneter Mikrostrukturen gegengesteuert wird, wie z.B. durch (einfach) architektonisch ausgebaute, einfache gemischt-semiintegrierte Mikrostrukturen bei zweisprachigen Wörterbüchern).

Neben den Verweisartikeln, die genuin Einzeilenartikel sind, wirkt sich die Buchungswürdigkeit in der Länge der Wörterbuchartikel aus. Lemmazeichen, die nur peripher zum lexikographisch zu behandelnden Wortschatz gehören, werden eher in Kurzartikeln abgehandelt, Lemmazeichen hingegen, die für das zu erarbeitende Wörterbuch zentral sind, werden eher in einem Standardartikel oder gar Großartikel behandelt. Auch der Polysemiegrad wirkt sich hinsichtlich der Artikelgrößenklasse aus. Wiederum wird die Eigenschaftsausprägung der Artikelgrößenklasse nicht vom Lemmazeichen allein determiniert, sondern auch durch das Wörterbuch, das geschrieben werden soll.

Hierzu gehört auch, daß die Anzahl der Angaben eines bestimmten Typs in einem Wörterbuchartikel durch das Lemmazeichen bestimmt werden können. So kann es sein, daß die Anzahl der Beispiel- und Belegangaben je nach Lemmazeichen und Wörterbuchtyp variieren: In einem Lernerwörterbuch wird man auch zu lexikalischen Einheiten, die eher dem sog. Grundwortschatz angehören, mehr Beispielangaben machen als in einem einsprachigen allgemeinen Wörterbuch. Damit würde die Anzahl der Beispielangaben in einem Wörterbuchartikel zur metalexikographischen Eigenschaftsausprägung gehören.

Der Wörterbuchumfang wird einerseits durch das Wörterbuchprogramm (soll es ein Einbänder oder ein Mehrbänder werden, ein Taschenwörterbuch oder ein Großwörterbuch) und andererseits durch ökonomische Überlegungen bestimmt. Dabei gilt nicht in jedem Falle, daß eine durch den Verlag auferlegte Umfangsbegrenzung negative Konsequenzen für die Qualität des Wörterbuches haben muß; denn diesbezüglich schreibt W. DULTZ im Vorwort seines Wörterbuches:

[Z 2-4: 'DBG-FWL, Vorwort (o. Pag.)]

Die vom Verlag aus Gründen der Kostenersparnis gewünschte Beschränkung auf denselben Umfang [wie in der ersten Auflage; M.K.] hatte ein sehr wertvolles Nebenergebnis zur Folge: nämlich die Straffung des Fremdwörtergutes. Ballast, der heute noch ganze Spalten der gängigen Fremdwörterbücher füllt, konnte über Bord geworfen werden. Altbekannte Fremdwörter, über deren Bedeutung sich wahrscheinlich kein Benutzer eines Lexikons mehr informiert, wie Apparat, Barbier, Coca-Cola, Depesche, Etappe, finanzieren, harmonisch, jonglieren, Kaffee, Lawine, Mandarine, Minister, Nonne, Oper, Parlament, Rezept, Semester, Tanker, Universität, Vase, Zelluloid u. v. a., konnten eliminiert und durch neue, teilweise bisher überhaupt noch nicht erläuterte Fremdwörter eingetauscht werden.

Unter dem Gesichtspunkt der *Wörterbuchartikelform* (oder kurz: *Artikelform*) können u.a. unterschieden werden:

- (reiner und erweiterter) Verweisartikel;
- synopseorientierter Verweisartikel;
- Einzelartikel;
- synopseorientierter Einzelartikel und
- Synopseartikel (oder: Übersichtsartikel).

Die Ausprägungen der Wörterbuchartikelform können durch die linguistischen Eigenschaftsausprägungen des Lemmazeichentyps bestimmt werden. So werden beispielsweise im FWB systematisch Verweisartikel zu Varianten der schriftlichen Realisierung sowie bei Wortbildungen geschrieben. Darüber hinaus kann die Wörterbuchartikelform auch durch andere Konstanten determiniert werden, die mit den linguistischen Eigenschaftsausprägungen nichts zu tun haben. Ein Beispiel hierfür wäre, daß bei den Zahl-Lemmazeichen *null* bis *zwölf* auf Grund des Wörterbuchprogramms zwölf Verweisartikel zu schreiben sind (nämlich zu *null* bis *sieben* sowie zu *neun* bis *zwölf*), in denen alle eine Verweisangabe auf den Wörterbuchartikel zu *acht* gemacht werden muß, da dieser vor allen anderen Zahlen in der (initialalphabetischen) Makrostruktur steht, denn es gilt:

acht < drei < eins < elf < fünf < neun < null < sechs < sieben < vier < zehn < zwei < zwölf
(mit „x < y“ bedeutet soviel wie x geht y in der Makrostruktur voraus).

Hier wäre die Tatsache, daß beim Lemmazeichen *zwei* ein Verweisartikel zu schreiben ist, keine metalexikographische Eigenschaftsausprägung, die durch das Lemmazeichen bestimmt werden würde, sondern durch das Wörterbuchprogramm.

Die *mediostrukturellen Eigenschaften* eines Wörterbuchartikels gehören ebenfalls zu den metalexikographischen Eigenschaftsausprägungen, bestimmen aber nicht die Form eines Wörterbuchartikels, sondern dessen Inhalt(e). Dabei wird die Mediostruktur einerseits durch das Lemmazeichen und andererseits durch weitere Faktoren festgelegt. Beispielsweise kann es in einem Wörterbuch vorgesehen sein, bei allen unregelmäßig konjugierten Verben eine Verweisangabe auf die jeweilige Konjugationstabelle im Nachspann zu machen. Insofern wird die Mediostruktur durch eine linguistische Eigenschaftsausprägung des Lemmazeichens festgelegt. Denkbar wäre auch, daß auf Grund des Wörterbuchprogramms Verweisangaben auf Bildtafeln zu machen sind, wenn es sich um nennlexikalische Ausdrücke von Artefakten oder natürlichen Arten handelt, zu denen eine Bildtafel existiert. In diesem Fall würde die Mediostruktur nicht durch die metalexikographische Eigenschaftsausprägung des Lemmazeichens bestimmt werden.

Nach WIEGAND (1996c) kann zwischen der *wörterbuchinternen* und der *intertextuellen* Mediostruktur unterschieden werden. Während die wörterbuchinterne Mediostruktur nur jene expliziten und impliziten Verweisangaben und -kennzeichnungen als Trägermenge besitzt, mit denen ein Lexikograph auf einen anderen Teiltext desselben Wörterbuches verweist, gehören zur intertextuellen Mediostruktur die *wörterbuchvernetzende*, die *quellenbezogene* und die *literaturbezogene* Mediostruktur. Als wörterbuchvernetzende Mediostruktur gelten Verweisungen auf andere Wörterbücher, als quellenbezogene Mediostruktur diejenigen Verweisungen, die das Wörterbuch virtuell mit Quellen verbinden, und als literaturbezogene Mediostruktur die Menge derjenigen Verweisangaben, aus denen ein Benutzer-in-actu einen Verweis auf wissenschaftliche und sonstige Literatur erschließen kann.

Innerhalb der wörterbuchinternen Mediostruktur kann unterschieden werden, wo die Verweisung steht sowie wohin mit ihr innerhalb des Wörterbuches verwiesen wird. Daraus ergeben sich 9 Möglichkeiten der Verweisung, die in der nachstehenden Tabelle wiedergegeben werden.

Textueller Ort der Verweisangabe	Verweisziel: in den/das X hinein		
	X = Vorspann	X = Wörterverzeichnis	X = Nachspann
Vorspann	×	×	×
Wörterverzeichnis	×	×	×
Nachspann	×	×	×

Tab. 2-2: Typen von wörterbuchinternen Verweisangaben bzw. -kennzeichnungen

Insbesondere die intertextuelle Mediostruktur wird *nicht* durch das Lemmazeichen und damit durch dessen metalexikographische Eigenschaftsausprägungen bestimmt, sondern durch andere Faktoren, wie z.B. durch das Wörterbuchprogramm und die Lage der wissenschaftlichen Literatur zu einem Lemmazeichen.

Das Wörterbuchartikelmanagement und die Konsistenzüberprüfung von Verweisangaben kann sehr aufwendig sein. Wird jedoch ein Computer beim Erstellen der Wörterbuchartikel eingesetzt, kann dieser relativ viel Arbeit und Verwaltungsaufwand übernehmen (cf. hierzu KAMMERER 1995, bes. 123–160).