

**Über Bären dienste
und erstaunte Bauklötze**
Idiome ohne freie Lesart in der HPSG

Jan-Philipp Soehn

Dissertation zur Erlangung des Grades ‚Doctor Philosophiae‘
eingereicht 2005 an der Philosophischen Fakultät
Friedrich-Schiller-Universität Jena

Für Laurin Maximilian
Zur Zeit der Anfertigung dieser Arbeit lernte er seine Muttersprache.

Inhaltsverzeichnis

Vorbemerkungen	xiii
1 Einführung	1
1.1 Motivierung der Thematik	1
1.2 Gliederung der Dissertationsschrift	4
2 Klassifikationskriterien	9
2.1 Auf den ersten Blick	9
2.2 Morphologische und lexikalische Kriterien	10
2.2.1 Eingefrorene Eigenschaften	10
2.2.2 Morphologische Anomalien	11
2.2.3 Unikale Elemente	11
2.2.4 Bedeutungsvolle vs. bedeutungsschwache Verben	15
2.3 Syntaktische Kriterien	16
2.3.1 Anomalien	16
2.3.2 Valenz	17
2.3.2.1 Interne vs. externe Valenz	17
2.3.2.2 Valenzänderung	17
2.3.3 Kategorieller Status der Argumente	19
2.3.4 Syntaktische Stabilität	20
2.3.4.1 Passivierbarkeit	21
2.3.4.2 Imperativbildung	23
2.3.4.3 Nominalisierbarkeit	24
2.3.4.4 Negation mit <i>nicht</i>	26
2.3.4.5 Negation mit <i>kein</i>	27
2.3.4.6 Referenzialität	29
2.3.4.7 Fragesatzbildung	29
2.3.4.8 Modifizierbarkeit	30
2.3.4.9 Quantifizierung	32
2.3.4.10 Andere Variationen	33
2.4 Semantische Kriterien	33
2.4.1 Motivierung und Durchsichtigkeit	33

2.4.2	Bildfelder und Tropen	36
2.4.2.1	Metaphern	37
2.4.2.2	Metonymie	39
2.4.3	Kompositionalität	39
2.5	Soziolinguistische Gesichtspunkte	43
2.5.1	Dialektale Varianten	46
2.5.2	Konnotationen	46
2.6	Kognitiver Aspekt: Kreative Modifikation	48
2.7	Zusammenfassung	53
3	Der Datenbereich	55
3.1	Definition	55
3.2	Kennzeichen	56
3.2.1	Unikale Elemente	57
3.2.2	Morphologische Anomalien	57
3.2.3	Syntaktische Hinweise	57
3.2.4	Selektionsrestriktionen	58
3.2.5	Motivierung und Dekomposition	59
3.2.6	Phraseoschablonen?	60
3.2.7	Seltene sprachliche Bilder?	61
3.3	Zusammenfassung	62
4	Herangehensweisen in der HPSG	65
4.1	Grundbausteine der HPSG	65
4.1.1	Der Formalismus	66
4.1.1.1	Die Objektebene: Perlen und Schnüre	66
4.1.1.2	Die Beschreibungsebene: Etiketten und Matrizen	67
4.1.2	Signatur und Theorie	68
4.1.3	Wichtige Attribute, Sorten und Prinzipien	69
4.2	Interne Selektion	71
4.3	Externe Selektion	75
4.3.1	Selektion von Lexemen	78
4.3.2	Analyse mit XSEL	80
4.4	Konstruktion	86
4.5	Kollokation	92
4.6	Zusammenfassung	96

5	Anknüpfungs- und Ausgangspunkte	97
5.1	Anknüpfen an die Herangehensweisen	97
5.1.1	Konstruktion	97
5.1.2	Externe Selektion	98
5.1.2.1	Lexemselektion → Listemselektion	98
5.1.2.2	Weitere Kritik am XSEL-Ansatz	101
5.1.3	Kollokation	104
5.2	Ein neuer alter COLL-Ansatz	104
5.2.1	Syntaktische Barrieren	105
5.2.2	Technische Voraussetzungen	108
5.3	Ausgangspunkte für die Analyse	113
5.3.1	Verben, Klammern und Komplexe	113
5.3.1.1	Das Subjekt	113
5.3.1.2	Der Verbalkomplex	115
5.3.1.3	Komplexere Prädikate	124
5.3.1.4	Passivierung	127
5.3.1.5	Nominalisierung	130
5.3.2	Ressourcen, Restriktionen und Referenzen	135
5.3.2.1	Semantik: LRS	135
5.3.2.2	Selektionsrestriktionen	143
5.3.2.3	Referenzialität	156
5.4	Zusammenfassung	159
6	COLLage der Analysen	161
6.1	Dekomponierbare IVPs	161
6.1.1	Analyse	161
6.1.2	Syntaktische Operationen	172
6.1.3	Mehrere Elemente auf der COLL-Liste	175
6.2	Nicht dekomponierbare IVPs	178
6.2.1	Analyse	178
6.2.2	Syntaktische Operationen	186
6.2.3	Unikale Elemente in PLEs	195
6.2.4	Idiomatische Adjunkte	196
6.2.5	Syntaktisch anomale IVPs	197
6.2.6	Externe Modifikation	200
6.3	Zusammenfassung und Diskussion	202
6.4	Noch einmal zum Vergleich: XSEL vs. COLL	208

7	Abschließendes Fazit und ausstehende Forschung	211
7.1	Zusammenfassung und Ausweitung	211
7.1.1	Ein lexikalistischer Ansatz	211
7.1.2	Kennzeichen im Visier	214
7.1.3	Weitere Phraseologismen	215
7.2	Horizonte	216
7.2.1	Phraseologisierung	216
7.2.2	Generalisierung unserer Herangehensweise	220
Literatur		223
A	IVP-Korpus	241
B	Zur HPSG-Signatur	245

Abbildungsverzeichnis

2.1	Syntaktische Operationen – Übersicht	34
2.2	Prisma von <i>Dampf ablassen</i>	42
2.3	Prisma von <i>Haare zu Berge stehen</i>	42
4.1	Eine einfache AVM	70
4.2	SUBCAT-Liste von <i>kommen</i> bei Krenn/Erbach	72
4.3	Interne Selektion in <i>jdm. den Garaus machen</i>	73
4.4	Skizze des LOCAL-Wertes von <i>machen</i>	74
4.5	XSEL: Änderungen in der Signatur	77
4.6	XSEL: Herkömmlicher und geänderter INDEX	79
4.7	Skizze des LEs von <i>Anhieb</i> mit XSEL	81
4.8	Skizze des LEs von <i>Brustton</i> mit XSEL	82
4.9	Beispiel für externe Selektion mit XSEL	82
4.10	Skizze des LEs von <i>Potte</i> mit XSEL	82
4.11	<i>zu Potte kommen</i> als PP mit festem Verb	83
4.12	Skizze des LEs von <i>fackeln</i> mit XSEL	85
4.13	Skizze der <i>spill_beans_idiom_phrase</i>	89
4.14	COLL-PRINZIP	93
4.15	Skizze des LEs von <i>Daffke</i> mit COLL	94
4.16	<i>dass Peter aus Daffke kommt</i> mit COLL	95
5.1	Neue Attribute von <i>head</i>	103
5.2	COLL-Liste in Richter/Sailer (1999a)	104
5.3	Neue Angemessenheitsbedingungen	108
5.4	Sortenhierarchie für <i>barrier</i>	109
5.5	Relationen für <i>barrier</i> -Untersorten	110
5.6	Beispiel für die Wirkungsweise des LIP	111
5.7	LE des Auxiliars <i>wird</i>	116
5.8	LE der Spur (vereinfacht)	117
5.9	LE der Verbspur (vorläufige Version)	118
5.10	DR für Verberststellung (V1-DR, vorläufige Version)	119
5.11	Analyse von <i>Widmen wird er die Dissertation seinem Neffen.</i>	120
5.12	MVF-DR (vorläufige Version)	122

5.13	Analyse von <i>Zum zweiten Mal die Weltmeisterschaft errang Clark</i>	123
5.14	Strukturbaum des erweiterten Prädikatskomplexes	126
5.15	Passiv-DR, die <i>haben</i> ausschließt	129
5.16	LRS Analyse von <i>jeder Student ein Buch liest</i>	142
5.17	Merkmalsgeometrie von <i>content</i>	143
5.18	Lokale und nicht lokale Semantik	143
5.19	Skizze des LEs von <i>liebt</i>	144
5.20	Merkmalsgeometrie unter CONTEXT	148
5.21	LE eines Namens nach Pollard/Sag (1994)	149
5.22	Instanziierung eines Objekts nach Pollard/Sag (1994)	149
5.23	Skizze des LEs von <i>essen</i> mit Selektionsrestriktion	151
5.24	Skizze des LEs von <i>Apfel</i> mit SATISFIES-Wert	152
5.25	Skizze einer Phrase mit Selektionsrestriktionen	153
5.26	Sortenhierarchie unter <i>phi</i>	158
6.1	Skizze des LEs von <i>staunen</i> in <i>Bauklötze staunen</i>	163
6.2	Analyse von <i>Bei Daimler-Chrysler staunt man Bauklötze</i>	164
6.3	LE der Verbspur (vorerst endgültige Version)	165
6.4	Semantik von (<i>dass</i>) <i>Peter Bauklötze staunt</i>	167
6.5	Skizze des LEs von <i>Bauklötze</i> in <i>Bauklötze staunen</i>	168
6.6	Skizze des LEs von <i>ausstellen</i>	171
6.7	Analyse von <i>Den Garaus macht den Seglern...</i>	173
6.8	Nominalisierungs-DR	175
6.9	COLL-Wert von <i>Potte</i> in <i>zu Potte kommen</i>	176
6.10	COLL-Wert von <i>Garaus</i>	177
6.11	Skizze des PLEs von <i>jdn. übers Ohr hauen</i>	183
6.12	Struktur von (<i>dass</i>) <i>Jens Uli übers Ohr haut</i>	185
6.13	Ausschnitt des PLEs von <i>Schindluder treiben</i>	187
6.14	DR für die Verbspur	190
6.15	V1-DR (endgültige Version)	191
6.16	MVF-DR (endgültige Version)	192
6.17	Nominalisierungs-DR für PLEs	194
6.18	Skizze des LEs von <i>Maulaffen</i> in <i>Maulaffen feilhalten</i>	196
6.19	Skizze des PLEs von <i>wie ein Schneekönig</i>	198
6.20	Skizze des PLEs von <i>gut Kirschen essen</i>	200
6.21	Skizze des PLEs von <i>das Handtuch werfen</i>	203

Vorbemerkungen

Danke

Allen voran möchte ich meiner Betreuerin Christine Römer danken, die viel Geschick bei der Auswahl des Themas bewies, mich stets sehr gut betreute und unsere Zusammenarbeit während der Entstehung dieser Dissertationsschrift sehr angenehm gestaltete.

Weiterhin darf mein Dank an Frank Richter nicht fehlen, der die Funktion des Zweitgutachters trotz aller anderweitigen Belastung bereitwillig übernommen hat und mir viele hilfreiche Ratschläge gab.

Natürlich danke ich auch Herrn Dekan Walter Ameling, der mir die Promotion an der Philosophischen Fakultät der Friedrich-Schiller-Universität Jena ermöglichte.

Meine Promotion entstand parallel zum DFG-Projekt „Syntax und Semantik idiomatisierter verbaler Phraseologismen in einem abgestuft kompositionellen Grammatikmodell“. Der Antrag für dieses Projekt, der auch der Ausgangspunkt für die vorliegende Arbeit war, wurde von Stefan Müller und Christine Römer ausgearbeitet. Ich danke der Deutschen Forschungsgemeinschaft für die Finanzierung des Projektes. Außerdem gilt mein Dank Lukas Scholz und Christian Schröter, die als studentische Hilfskräfte bei der Erstellung des Idiom-Korpus im Anhang mitgewirkt haben. In diesem Zusammenhang muss auch das Institut für Deutsche Sprache in Mannheim dankbar erwähnt werden, welches das Korpusrecherche- und -analysesystem COSMAS II zur Verfügung stellte. Viele der hier angeführten Belege wurden mit Hilfe dieses Systems gefunden. Ein Teil der Druckkosten für diese Arbeit konnte dank eines großzügigen Zuschusses von der Europäischen Gesellschaft für Phraseologie gedeckt werden.

Mein besonderer Dank gebührt Manfred Sailer, der oft alles stehen und liegen ließ, um meine Gedanken rund um die HPSG in geordnete Bahnen zu lenken. Darüberhinaus danke ich auch Stefan Müller, Tilman Höhle, Peter Gallmann, Markus Egg, Detmar Meurers und Doris Penka für interessante Diskussionen, wertvolle Hinweise und Korrekturen.

Teile dieser Dissertationsschrift wurden auf folgenden Konferenzen vorge-

stellt: Collocations 2003, Formal Grammar 2003 und 2004, Europhras 2004, HPSG Conference 2004 und 2005, außerdem auf dem Workshop „Empirie und Theorie idiosynkratischer Verbindungen“ 2004 in Tübingen. Ich danke den anonymen Gutachtern der Konferenzpapiere sowie den Teilnehmern aller Tagungen für ihre konstruktiven Kommentare und Fragen.

Ohne die Unterstützung meiner Familie wären mein Studium und auch meine Promotion wahrscheinlich nicht möglich gewesen. Ich danke meinen Eltern dafür, dass ich nie *am Hungertuch nagen* musste und sie mich stets mit ihrer Anteilnahme begleiteten, auch wenn für sie der Inhalt meines Studiums oder der vorliegenden Arbeit nicht völlig *klar wie Kloßbrühe* waren. Zuletzt gilt mein Dank meiner Verlobten Anne Glaß für ihre Unterstützung beim Korrekturlesen, ihre Geduld, wenn ich mal *Blut und Wasser schwitzte*, und ihre Liebe.

Zum Textsatz

Zum Textsatz dieser Arbeit ist zu sagen, dass natürlichsprachliche Beispiele, d.h. Phraseologismen oder Teile daraus als fortlaufend nummerierte Absätze erscheinen. Dort sowie im Fließtext werden die Beispiele *kursiv* gedruckt. Zur Hervorhebung etwa eines unikalenen Elementes oder eines Phraseologismus mit idiomatischer Lesart wird zusätzlich noch unterstrichen. Ein weiteres Mittel zur Hervorhebung ist der **Fettdruck**.

Fremdsprachliche Ausdrücke sowie Fachausdrücke, die sich nicht ins Deutsche übersetzen lassen, werden *schräg* gedruckt.

Ungrammatische Wendungen werden, wie üblich, mit einem Asterisk (*) gekennzeichnet. Bei Fällen, deren Grammatikalität zumindest für einige Sprecher fragwürdig ist, steht ein Fragezeichen voran. Idiomatische Wendungen, bei welchen die idiomatische Lesart wegfällt, die also nur noch wörtlich zu verstehen sind, werden durch ein vorausgehendes Ausrufezeichen markiert.

Wenn im Text über die Merkmalsgeometrie der HPSG gesprochen wird, so erscheinen Attribute bzw. Merkmale in KAPITÄLCHEN. Sorten, die ihrerseits auch Attributwerte sein können, werden klein und *kursiv* geschrieben. Relationen sind durch Courier-Schrift gekennzeichnet. Prädikate in der Semantik erscheinen ohne Serifen und erhalten einen Apostroph (prädikat'). Eine wichtige Rolle in der HPSG spielen auch sogenannte *tags*, die als kleine Kästchen mit Nummern (□) dargestellt werden.

Zur Orientierung für den Leser sei an dieser Stelle noch erwähnt, dass, um eine bessere Lesbarkeit zu gewährleisten, diese Arbeit im generischen Maskulinum verfasst wurde und Begriffe wie „Leser“ oder „Sprecher“ stets beide Geschlechter einschließen.

Die wichtigsten Abkürzungen

Abkürzungen im Text:

AP	Adjektivphrase
AVM	<i>attribute-value-matrix</i> , Merkmalsmatrix
DP	Determinansphrase
DR	Derivationsregel
EContP	External-Content-Prinzip
FVG	Funktionsverbgefüge
GB	Theorie des <i>Government & Binding</i>
GPSG	<i>Generalized Phrase Structure Grammar</i>
GPSR	Gültigkeitsprinzip der Selektionsrestriktionen
HPSG	<i>Head-Driven Phrase Structure Grammar</i>
IContP	Internal-Content-Prinzip
IVP	Idiomatisierter verbaler Phraseologismus
jmd.	jemand
jdn.	jemanden
jdm.	jemandem
jds.	jemandes
LE	Lexikoneintrag
LFG	<i>Lexical Functional Grammar</i>
LIP	Lizensierungsprinzip
LR	Lexikalische Regel
LRS	<i>Lexical Resource Semantics</i>
MRS	<i>Minimal Recursion Semantics</i>
NP	Nominalphrase
österr.	österreichisch
pers. Mitt.	persönliche Mitteilung
PLE	phrasaler Lexikoneintrag
PP	Präpositionalphrase
psoa	<i>parameterized-state-of-affairs</i>
PXS	Prinzip der Externen Selektion
RSRL	<i>Relational Speciate Re-entrant Language</i>
schweiz.	schweizerdeutsch
SP	Semantik-Prinzip
u. E.	unseres Erachtens
VP	Verbalphrase

Es folgt nun eine Liste mit den wichtigsten Attributen und Sorten der HPSG, inklusive der in dieser Arbeit neu eingeführten, und deren Abkürzungen. Wo Sorten gleichlautend mit Attributen sind, werden sie nicht extra aufgeführt. Zu deren hierarchischen Beziehungen siehe Appendix B.

ADJUNCT-DAUGHTER	ADJ	INITIAL	INIT
ARGUMENT	ARG	INHERITED	INHER, INH
ARGUMENT-STRUCTURE	ARG-ST	INPUT	IN
AUXILIARY	AUX	<i>internal</i>	<i>int</i>
BACKGROUND	BGR	INTERNAL-CONTENT	INCONT, INC
<i>background-condition</i>	<i>bgr-cond</i>	INVERTED	INV
CATEGORY	CAT, CT	LEXICAL	LEX
CLUSTER-DAUGHTER	CL	<i>lexical-resource-semantics</i>	<i>lrs</i>
COMPLEMENTS	COMPS	LF-LICENSER	LF-LIC
COMP-DAUGHTERS	C-DTRS	LISTEME	LSTM
<i>complete-clause</i>	<i>comp-cl</i>	LOCAL	LOC, LC
CONTENT	CONT	LOCAL-LICENSER	LOC-LIC
CONTEXT	CTXT	LOGICAL-FORM	LF
CONTEXT-OF-LEXICAL-LICENSING	COLL	MAIN	
CONTEXTUAL-INDICES	C-INDS	<i>meaningful-expression</i>	<i>me</i>
DAUGHTERS	DTRS	MODIFIED	MOD (obsolet)
DEFINITE	DEF	MUST-SATISFY	M-STF
DEPENDENCE-STRUCTURE	DEPS	<i>non-empty-list</i>	<i>nelist</i>
DOUBLE-SLASH	DSL	NON-LOCAL	NLOC
<i>embedded-sign</i>		<i>non-referential</i>	<i>non-ref</i>
<i>empty-list</i>	<i>elist</i>	<i>no-phi</i>	
<i>eventuality</i>		NUMBER	NUM
<i>extended-head-cluster-structure</i>	<i>ext-head-cl-struc</i>	OUTPUT	OUT
<i>external</i>	<i>ext</i>	PARTS	
EXTERNAL-CONTENT	EXCONT, EXC	PERSON	PER
EXTERNAL-SELECTION	XSEL	PHI	
GENDER	GEN	PHONOLOGICAL-LICENSER	PHON-LIC
HEAD	HD	PHONOLOGY	PHON
<i>head-cluster-structure</i>	<i>head-cl-struc</i>	<i>referential</i>	<i>ref</i>
HEAD-DAUGHTER	H-DTR	SATISFIES	STF
IDIOMATIC-ARGUMENT-STATUS	I-ARG-STAT	<i>selection-sort</i>	
INDEX	IDX	<i>selectional-restriction-imposed</i>	<i>sel-restr-imp</i>
		<i>selectional-restriction-satisfied</i>	<i>sel-restr-stf</i>
		SLASH	SL
		SPECIFIED	SPEC (obsolet)
		STORE	
		SUBCAT	SC

SUBJECT
SYNTAX-SEMANTICS

SUBJ
SYNSEM, SS

TO-BIND

TB

*unembedded-sign
utterance*

VALENCE
VARIABLE

VAL
VAR

1 Einführung

Damit unsere Sprache nicht noch mehr militarisiert wird,
müssen wir aufpassen wie die Schießhunde.

Friedrich Nowotny

1.1 Motivierung der Thematik

Automatische Übersetzungssysteme, *Workbenches* für Übersetzer, *Content Managing*, Auskunft- und Dialogsysteme, Informationsextraktion, Programme zur Zusammenfassung von Texten, *Information Retrieval*, Sprachsynthese – die automatische Sprachverarbeitung ist wieder im Aufwind.

Im Laufe des letzten Jahrzehnts hat sich besonders die korpusbasierte, statistische Sprachverarbeitung als zuverlässiger Ansatz etabliert, um robuste, anwendungsorientierte Systeme zu entwickeln.

Eine große Herausforderung für diese computerlinguistischen Anwendungen ist die Erkennung und Verarbeitung von Phraseologismen oder *idioms*. Bei verbalen Phraseologismen, dem Gegenstand dieser Dissertationsschrift, handelt es sich um Ausdrücke bestehend aus einem Verb und mindestens einem Komplement oder Adjunkt mit einer festen, oft metaphorischen Bedeutung. Diese kann nicht aus der Bedeutung ihrer Teile erschlossen werden.¹ Volk (1998) erläutert in seinem Artikel “*The Automatic Translation of Idioms*” einige Probleme, die bei der automatischen oder computerunterstützten Übersetzung von Phraseologismen auftauchen. Zuerst einmal besteht die Schwierigkeit darin, ein Idiom als solches zu erkennen. Für die Identifizierung der Wendungen, mit denen wir uns in dieser Arbeit beschäftigen wollen, gibt es einige Anhaltspunkte. Schwieriger zu identifizieren sind jedoch Ausdrücke wie:

¹In dieser Arbeit werden die Begriffe Phraseologismus, IVP, Idiom, engl. *idiom* und idiomatische Wendung synonym zueinander verwendet. Wir verstehen den Begriff des Phraseologismus im engeren Sinne: zur Polylexikalität (aus mehreren Wörtern bestehend) und Festigkeit (in Sprachgemeinschaft ähnlich wie ein Wort gebräuchlich) kommt als Eigenschaft die Idiomatizität hinzu (siehe Burger, 2003, S. 14f).

(1) *mit jdm. Pferde stehlen können*

(2) *gute Karten haben*

Wendungen dieser Art können auch wörtlich gemeint sein, d. h. ihre Bedeutung setzt sich dann aus der Bedeutung ihrer Teile zusammen.

Volk stellt zwei in ihrem Ansatz grundlegend verschiedene Systeme zur maschinellen Übersetzung vor, nämlich reine Übersetzungsprogramme (*machine translation systems* – MT), wie “T1 Professional” von Langenscheidt, und Übersetzungsspeichersysteme (*translation memory systems* – TrMem), wie den “Translation Manager” von IBM. Erstere Anwendungen leisten die eigentliche Übersetzungsarbeit, wobei TrMem-Systeme als Speicher für bereits korrekt übersetzte Phrasen dienen. Es wird in dem Aufsatz auf Vor- und Nachteile beider Gattungen eingegangen, wobei Volk zu dem Schluss kommt, dass sich beide Arten von Software ergänzen und man sie deshalb miteinander kombinieren muss. Es wird auch deutlich, dass die Verarbeitung von Ambiguitäten, welche in der natürlichen Sprache auch bzw. gerade in Redewendungen vorkommen, den MTs Probleme bereitet. Volk testet zwei Programme auf ihre Art, Idiome zu verarbeiten. Er schlägt in seinem Fazit vor, Archive für Phraseologismen einzubinden, so dass ein Programm durch einen Such- und Vergleichsvorgang ein Idiom erkennt oder zumindest warnt, dass ein Satz auch als Idiom verstanden werden könnte. Als Beispiele solcher Idiom-Datenbanken beschreibt er den “*Phrase Manager*” von Pedrazzini (1994) und das System „Phraseo-Lex“ von Keil (1997). Während “*Phrase Manager*” den Formalismus für eine systematische Repräsentation von Idiomen liefert, wurde „Phraseo-Lex“ sprachspezifisch für idiomatisierte verbale Phraseologismen (IVPs) im Deutschen entwickelt und kann detailliertere Unterscheidungen vornehmen.²

Diese Ausführungen sollen als Ausblick auf eine mögliche Anwendung der in der vorliegenden Arbeit vorgestellten Idiom-Analyse dienen. Gleichzeitig ist dies eine Motivation, sich eingehend und formal mit Phraseologismen auseinanderzusetzen. In dieser Arbeit geht es jedoch nicht darum, einen Weg zu

²Keil (1997) gibt neben der Beschreibung von “*Phrase Manager*” und der ausführlichen Vorstellung von „Phraseo-Lex“ einen Überblick über weitere computerlinguistische Forschung im Bereich der Phraseologismen. Sie weist darauf hin, dass außersprachliches Wissen – Weltwissen und Wissen, das auf Erfahrung und Erinnerung beruht – für die Interpretation von Phraseologismen wichtig ist, seine Implementierung aber bisher der KI-Forschung und der Computerlinguistik große Schwierigkeiten bereitet (für Forschungsarbeit in diese Richtung, insbesondere die Auflösung von Metonymien, siehe beispielsweise Markert, 1999).

einer Computer-Implementation von Phraseologismen aufzuzeigen, sondern sprachtheoretische Grundlagen zu diskutieren.

Im Allgemeinen werden bei der automatischen Sprachverarbeitung solche Systeme bevorzugt entwickelt, die eine akzeptable Laufzeit aufweisen, ein Maximum von Eingabesätzen korrekt erfassen und dabei die Vorzüge des Computers als Rechenmaschine ausnutzen. Statistische Systeme sind deshalb gerade das Nonplusultra. Leider wird man dabei aber kaum der Sprache als System mit ihren Gesetzmäßigkeiten und ihrer inneren Struktur gerecht. Äußerungen eines Sprechers sind nicht geleitet von Wahrscheinlichkeiten, sondern von weitgehend unbewusst erworbenen Gesetzmäßigkeiten der Phonologie, Morphologie, Syntax, Semantik und Pragmatik.³ In der vorliegenden Arbeit soll aus diesem Grund eine grammatische, linguistisch adäquate Analyse von Idiomen vorangetrieben werden.

Wir konkretisieren unsere Analyse im Rahmen der *Head-Driven Phrase Structure Grammar* (HPSG). Dieses Rahmenwerk legt einerseits größten Wert auf (zumindest) deskriptive Adäquatheit. Es eignet sich andererseits aber auf Grund seiner formalen Struktur auch dazu, ein Fragment, z. B. des Deutschen, in ein sprachverarbeitendes System zu implementieren. Einen Rahmen für eine Implementierung bieten die Systeme LKB (*Linguistic Knowledge Building*, siehe Copestake, 2002) oder TRALE (siehe Meurers/Penn/Richter, 2002). Konkrete Anwendungen sind beispielsweise das Babel-System von Stefan Müller, beschrieben in Müller (1996), und die linguistische Tiefenanalyse im „Verbmobil“, vgl. Müller/Kasper (2000).

Eine Beschreibung in der formalen HPSG bietet die Vorteile, den sprachlichen Phänomenen Rechnung zu tragen und geeignete Werkzeuge für eine konkrete Analyse bereitzustellen. Damit wird die Grundlage dafür geschaffen, dass ein System zur Sprachverarbeitung, das sich des HPSG-Formalismus bedient, auch Idiome erfassen kann.

Martina Keil (1997) bietet auf den ersten Seiten ihres Buches eine umfangreiche Einführung in die Eigenschaften, Besonderheiten und Klassifizierungen von Phraseologismen. Da wir das Rad nicht neu erfinden wollen, wird hier auf eine solche Einführung in die Phraseologie verzichtet. Dennoch sollen stichpunktartig die wichtigsten Eigenschaften genannt werden:

Phraseologismen fallen den Sprechern für gewöhnlich nicht auf, sie gehen mit ihnen selbstverständlich um. Weiterhin stellen Phraseologismen „vorgeformte Sprachkonstrukte“ dar, auf die in der Regel „als Ganzes“ (Keil, 1997, S. 2) zugegriffen wird. Dabei haben solche Wendungen eine besondere Bedeutung, die

³Vgl. den Überblick zur Spracherwerbsforschung von Rothweiler (2002)

sich nicht aus den Bedeutungen ihrer Lexeme zusammensetzt. Meist ist diese Bedeutung metaphorisch und kann mehrdeutig ausgelegt werden. Darüber hinaus zeichnen sich Phraseologismen durch eine gewisse, näher zu untersuchende Festigkeit aus, man kann also nur manche von ihnen variieren oder modifizieren. Wichtig für unsere Arbeit ist die Eigenschaft, dass in Phraseologismen zuweilen ungewöhnliche Wörter, gebundene Formative, auftreten.

Keil geht weiterhin auf Klassifizierungsarten bei Fleischer und Burger ein. Im Zentrum-Peripherie-Modell von Fleischer (1982) sind die verbalen Phraseologismen, oder auch „Phraseolexeme“, im Zentrum angeordnet und werden von adjektivischen (*frisch gebacken*), adverbialen (*durch die Bank*) und nominalen (*bessere Hälfte*) Phraseolexemen unterschieden. Zur Einteilung nach Burger siehe das folgende Kapitel.

1.2 Gliederung der Dissertationsschrift

Die vorliegende Arbeit gliedert sich weiterhin wie folgt.

Kapitel 2: Klassifikationskriterien

Nach der Hinführung zum Thema im ersten Kapitel werden eine genauere Einteilung der IVPs und einige ihrer spezifischen Merkmale vorgestellt. Zuerst werden morphologische und lexikalische Kriterien beleuchtet, wobei besonders auf unikale Elemente eingegangen wird. Es folgt eine Darstellung von syntaktischen Charakteristika wie Valenzeigenschaften und syntaktischer Stabilität bei Passivierung, Negation, Adjunktion und anderen Variationen. Weiterhin werden semantische Kriterien wie Motivation, Metaphorik und Kompositionalität berücksichtigt. Der wichtige Unterschied zwischen dekomponierbaren und nicht dekomponierbaren Idiomen wird herausgestellt und anhand eines graphischen Bedeutungsmodells veranschaulicht. Anschließend werden zwei wesentliche soziolinguistische Gesichtspunkte – dialektale Varianten und Konnotationen – erörtert. Da Idiome im individuellen Sprachgebrauch auf kreative Weise verändert werden können, bildet eine Diskussion der Idiom-Variabilität aus Sicht der kognitivbasierten Phraseologieforschung den Abschluss.

Ziel dieses Kapitels ist eine gründliche und für die weitere Diskussion grundlegende Untersuchung der beschriebenen Eigenschaften von Phraseologismen. Man muss dabei zu dem Schluss kommen, dass Phraseologismen keine der empirischen Kriterien einheitlich erfüllen, sondern dass innerhalb idiomatischer Wendungen ein hohes Maß an Idiosynkrasie vorliegt.

Kapitel 3: Der Datenbereich

Um den Umfang einer Dissertationsschrift in sinnvoller Weise einzuhalten, muss der Bereich der zu untersuchenden Daten eingegrenzt werden. Wir wollen uns dabei auf Idiome beschränken, die über keine freie Lesart verfügen.

Für ein Vorliegen einer idiomatischen Lesart und das gleichzeitige Fehlen der freien Bedeutung werden fünf formale Kennzeichen aufgedeckt: Das Vorkommen von unikalenen Elementen, morphologische Anomalien eines Idiom-Teils, syntaktische Hinweise (Anomalien innerhalb einer IVP und Änderung in der Valenzstruktur), Verletzung von Selektionsrestriktionen und eine nicht dekomponierbare Semantik.

Es wird gezeigt, dass das Vorkommen einer Phraseoschablone oder das Vorhandensein eines ungewöhnlichen semantischen Bildes nicht als formales Kennzeichen für eine ausschließlich gebundene Lesart gelten kann.

Kapitel 4: Herangehensweisen in der HPSG

Da dieser Arbeit die HPSG als Rahmentheorie zu Grunde liegt, werden bisherige Ansätze zur Idiom-Analyse in dieser Grammatiktheorie vorgestellt. Um dem mit der HPSG noch nicht vertrauten Leser ein Verständnis der Diskussion zu ermöglichen, werden Grundbausteine der HPSG und deren Formalismus kurz eingeführt.

Die kritische Untersuchung schließt die Herangehensweise über interne Selektion, einen Ansatz zur externen Selektion, eine Konstruktionsanalyse und die Analyse mittels eines Kollokationsmoduls ein.

Alle vier Herangehensweisen können entweder die Daten nicht adäquat beschreiben oder haben gewichtige Schwachpunkte, so dass man keinen Analysevorschlag unverändert für unseren Datenbereich übernehmen kann.

Kapitel 5: Anknüpfungs- und Ausgangspunkte

Die vier untersuchten Ansätze weisen Vorzüge auf, welche wir uns für die Idiom-Analyse zu Nutze machen wollen. Es werden daher einige Anknüpfungspunkte diskutiert, wie beispielsweise eine Selektion bestimmter Wörter. Es stellt sich heraus, dass eine frühere Version des in Kapitel 4 vorgestellten Kollokationsmoduls (Richter/Sailer, 1999a) besser zur Analyse geeignet ist, woraufhin dieser Ansatz konkretisiert und erweitert wird.

Anschließend werden weitere Grundlagen besprochen, wie beispielsweise die Behandlung der Verbstellung im Deutschen, Passivierung und Nominalisierung. Obwohl in dieser Arbeit der Schwerpunkt eindeutig auf syntaktischen

Gesichtspunkten liegt, wird ein in den letzten Jahren entwickeltes Semantik-Modul (LRS) für die HPSG vorgestellt, mit dessen Hilfe die konkreten Lexikoneinträge der Idiome skizziert werden. Im Zusammenhang mit der Semantik steht auch die Diskussion einer Integration von Selektionsrestriktionen und einer Analysemöglichkeit für nicht referenzielle Substantive.

Kapitel 6: COLLage der Analysen

In diesem Kapitel analysieren wir einige konkrete Beispiele von IVPs mit dem in Kapitel 5 entwickelten Kollokationsmodul. Dabei wird zwischen Beschreibungen der dekomponierbaren und der nicht dekomponierbaren Idiome differenziert, für die jeweils unterschiedliche Mittel verwendet werden müssen. Diese lassen sich jedoch in einer einheitlichen Gesamtanalyse verbinden. Weiterhin wird im Zusammenhang mit nicht dekomponierbaren IVPs die Frage nach idiomatischen Adjunkten und externer Modifikation diskutiert. Zusammenfassend werden die Vor- und Nachteile des neuen Kollokationsmoduls denjenigen der externen Selektion gegenübergestellt und die jeweiligen Anwendungsgebiete beider Mechanismen umrissen.

Kapitel 7: Abschließendes Fazit und ausstehende Forschung

Der in Kapitel 6 beschriebene Analysevorschlagn wird noch einmal zusammengefasst und dabei die Einbeziehung der Kennzeichen aus Kapitel 3 aufgegriffen. Der Ansatz wird daraufhin auf einen größeren Datenbereich skizzenhaft ausgedehnt.

Abschließend wird der Prozess der Phraseologisierung behandelt und die hierarchische Abstufung der sich in diesem Prozess befindlichen Wendungen betont. Eine mögliche Generalisierung unseres Ansatzes auf im phraseologischen Sinn weniger feste Wendungen erweist sich als interessanter und diskussionswürdiger Gegenstand weiterer Forschung.

Appendix A: Korpus und Belege

Um zu zeigen, dass wir hier nicht wie ein Blinder von der Farbe reden, wurde ein Idiom-Korpus (als XML-Datei auf beiliegender CD-ROM) zusammengestellt. Diese Sammlung, die aus 190 deutschen IVPs besteht, bildet die empirische Grundlage und damit einen zentralen Teil unserer Untersuchungen. Die dort aufgelisteten Wendungen stammen vor allem aus Dobrovol'skij (1988), Dobrovol'skij/Piirainen (1994), Fleischer (1997), Keil (1997), dem Duden 11

(2002), Burger (2003) und der eigenen Sprachkompetenz. Natürlich kann hier kein Anspruch auf Vollständigkeit erhoben werden, es handelt sich bei der Sammlung nur um Beispiele, anhand derer man das Gesagte illustrieren kann. Die Idiome wurden auf spezielle Merkmale hin ausgewählt und untersucht (vgl. Kapitel 2) und zu jedem dieser Punkte wurde ein Beleg aus Cosmas II (*Corpus Storage and Maintenance System*) des Instituts für Deutsche Sprache (IDS) in Mannheim oder mit Hilfe der Internet-Suchmaschine GOOGLE angeführt.

Appendix B: Zur HPSG-Signatur

Zuletzt werden Ausschnitte aus der HPSG-Signatur dargestellt, welche die in dieser Arbeit eingeführten neuen Sorten und Merkmale und deren Hierarchie zeigen. Die Signatur von Pollard/Sag (1994) wird zu Grunde gelegt und wir geben hauptsächlich die Abweichungen bzw. Neuerungen an. Darüberhinaus werden der Übersicht halber alle für die Analyse relevanten Prinzipien, Schemata und Relationen aufgeführt.

2 Klassifikationskriterien

Äußere Ordnung ist oft nur der verzweifelte Versuch,
mit einer großen inneren Unordnung fertig zu werden.

Albert Camus

2.1 Auf den ersten Blick

Zunächst kann man IVPs allgemein in zwei Hauptklassen einteilen. **Teil-idiomatische Wendungen** sind solche, bei denen Teile noch ihre wörtliche Bedeutung beibehalten, wie das Adjektiv in

(3) *dumm wie Bohnenstroh* „sehr dumm“

Bei **vollidiomatischen Wendungen** hat jede Komponente die wörtliche Lesart verloren, z. B. in folgendem IVP:

(4) *ins Gras beißen* „sterben“

Mit dieser Einteilung folgen wir Burger (2003, S. 38). Wir wollen hier Burgers weitaus detailliertere Klassifikation nicht wiedergeben, da sie viele Phraseologismen berücksichtigt, die nicht im Fokus dieser Arbeit liegen. Was IVPs ohne freie Lesart betrifft, werden sie von ihm als satzgliedwertige, referenzielle Phraseologismen bezeichnet, da sie keine ganzen Sätze bilden (im Gegensatz zu beispielsweise Sprichwörtern) und sich auf ein Objekt, einen Vorgang oder einen Sachverhalt in der Wirklichkeit beziehen (Burger, 2003, S. 36–41).

Untersucht man nun einzelne IVPs näher, stellt man fest, dass sie einige idiosynkratische Besonderheiten besitzen, was ihren inneren Aufbau und ihre Stabilität bzw. Variabilität betrifft. Keil (1997) widmet solchen Stabilitätsphänomenen einen eigenen Abschnitt (2.2). Wir wollen in diesem Kapitel einige Kriterien morphologischer, syntaktischer, semantischer und soziolinguistischer Art betrachten und auch den kognitiven Aspekt nicht außer Acht lassen. Dabei kann die Erfüllung mancher Kriterien nicht darüber hinwegtäuschen, dass innerhalb der IVPs eine große Varianz besteht und man in der

Regel Beispiele findet, die das untersuchte Kriterium gerade nicht erfüllen. So kann man z. B. nicht sagen, dass alle IVPs mit freier Lesart passivierbar sind. Diese Regularitäten zu finden ist noch weitgehend ein Forschungsdesiderat.

2.2 Morphologische und lexikalische Kriterien

Sehen wir uns zunächst einmal die Morphologie der einzelnen Bestandteile eines IVP an.

2.2.1 Eingefrorene Eigenschaften

Oft ist der **Numerus** des nominalen Komplements eingefroren. An Beispielen wie

(5) *einen Drehwurm/*Drehwürmer bekommen* oder

(6) *ein blaues Wunder/(**zwei blaue Wunder*) erleben*

sieht man, dass der Plural in manchen Wendungen nicht möglich ist. Manchmal muss auch der Plural stehen:

(7) *jds. Hoffnungen sind/(**Hoffnung ist*) geplatzt* oder

(8) *Bauklötze/(**einen Bauklotz*) staunen*

Auch das **Tempus** des Verbs kann eingefroren sein. In dem Idiom

(9) *etw. ist gehüpft wie gesprungen*

kommt die Kopula nur im Präsens vor, da die Aussage Allgemeingültigkeit ausdrückt. Ungrammatisch wäre demzufolge „*etw. wird gehüpft wie gesprungen sein*“.¹

Darüberhinaus kann auch der **Modus** des Verbs einfrieren. So gibt es die Wendung

(10) *mit allen Wassern gewaschen sein*

nur im Zustandspassiv und niemand kann sich beispielsweise mal eben **,mit allen Wassern waschen*“.²

¹Beachte: Das Tempus kann manchmal auch relativ zu einer Äußerung sein. Bei „Sie sagte, alles war gehüpft wie gesprungen“ unterliegt das Tempus des Idioms einer Zeitverschiebung auf Grund der indirekten Rede, zum Zeitpunkt der Äußerung wurde das Idiom dennoch im Präsens verwendet.

²Uns ist hier und generell klar, dass es in speziellen literarischen oder expressiven Kontexten zu einer solchen Verwendung kommen könnte.

2.2.2 Morphologische Anomalien

Neben ganz regulären Wortbildungen stößt man auch auf Anomalien, beispielsweise eine archaische Form des Dativ Plural³ wie in (10). In der Wendung

(11) *nicht viel Federlesens machen*⁴

ist das „s“ eine Genitivendung („*nicht viel des Federlesens*“), die in der Wortform eingefroren wurde, obwohl die Nominalphrase (NP) in der Wendung nun im Akkusativ steht. Dazu analog verhält sich auch das obligatorische Dativ-e in:

(12) *etw. schlägt zu Buche*

Eine weitere Anomalie ist die fehlende Kongruenz-Endung in (13).

(13) *sich bei jdm. lieb Kind machen*

Diese Anomalien zeigen, dass IVPs nicht unbedingt aus regulär vorhandenen Wörtern der deutschen Sprache im synchronen Sinn aufgebaut sein müssen, sondern idiosynkratische Eigenschaften haben, die man gesondert berücksichtigen muss.

2.2.3 Unikale Elemente

In IVPs können auch sogenannte unikale Elemente auftreten. Unikale Elemente sind Wörter mit extremer distributioneller Beschränktheit, deren syntaktische, semantische und selektionale Eigenschaften nicht ausreichend vorschreiben, in welchen Kontexten sie auftreten können. Beispiele dafür sind:

(14) *jdn. an den Bettelstab bringen*

(15) *Maulaffen feilhalten*

(16) *Fersengeld geben*

(17) *etw. auf dem Kerbholz haben*

Wie man leicht nachvollziehen kann, sind die unterstrichenen Wörter in der hier verwendeten Bedeutung nur in diesen (phraseologischen) Kontexten grammatisch und können nicht in anderen Umgebungen auftreten.

³Normal wäre *Gewässern*

⁴*Federlesen[s]* ist ein unikales Element. Da aber morphologische Anomalien auch bei freien Wörtern wie *lieb* und *Wasser* auftreten, ist ein eigener Abschnitt dafür gerechtfertigt.

Exkurs: Unikale Elemente erklärt

Abgrenzung von unikalenen Elementen

Bei den unikalenen Elementen oder auch unikalenen Komponenten handelt es sich um lexikalische Einheiten, die im Laufe der Sprachentwicklung „erstarrt“ sind. Früher einmal im normal gebräuchlichen Wortschatz vorhanden, kommen sie heute nur noch in ganz bestimmten Kontexten vor. Für sich alleine nehmen sie nicht mehr an produktiven Sprachschöpfungsprozessen teil. Es handelt sich um tradierte Sprachrelikte (vgl. Dobrovol'skij, 1988, S. 87).

In der englischen Literatur findet sich für diese Wörter neben *unique elements* auch der Begriff *cranberry words*, der von Aronoff (1976) geprägt wurde. Dieser Ausdruck leitet sich davon ab, dass Morphe wie *cran-* oder im Deutschen *Him-* in *Himbeere* nur in einer bestimmten Umgebung auftreten.

Weitere Beispiele⁵ für unikale Elemente sind:

(18) *Daimler-Benz-Aktionäre reden Tacheles.*

(19) *Ex-Generalsekretär Kanther mußte klein begeben.*

(20) *USA machen europäisches Fleisch madig.*

Man kann leicht feststellen, dass die unikalenen Elemente (unterstrichen) in dieser Bedeutung nur zusammen mit den jeweils fett gedruckten Wörtern vorkommen. Vergleiche dazu den Appendix in Soehn (2003), wo diese Fragestellung zu Grunde lag. Es handelt sich also um „phraseologisch gebundene Formative“, vgl. Fleischer (1997). Im ersten Beispiel kann es also nicht heißen „...sprechen Tacheles“, obwohl *reden* und *sprechen* eine sehr ähnliche Semantik haben. Genauso wenig kann man *groß begeben* oder etwas *madig reden*.

Die Abgrenzung von unikalenen Elementen ist umstritten. Wir folgen hier der Auffassung Dobrovol'skijs (1988, S. 93–96) und nehmen an, dass es unikale Komponenten mit frei vorkommenden Homonymen⁶ gibt. Das Wort *madig* in (20) bedeutet sicherlich etwas anderes als „voller Maden“ in *ein madiger Apfel*, genauso wie beispielweise die Elemente in *den Ausschlag geben* oder *jd. im Stich lassen*. Es handelt sich hier auch um unikale Elemente, weil *Ausschlag* und *Stich* in ihrer idiomatischen Bedeutung nur in diesen Kontexten auftreten, sie haben jedoch freie Homonyme, welche „Hautausschlag“ bzw.

⁵Beispiele hier aus COSMAS I des Instituts für Deutsche Sprache in Mannheim

⁶Homonyme sind Wörter mit gleichem Lautbild, aber unterschiedlicher Bedeutung und Etymologie.

einen anderen Stich (den im Kartenspiel, von Insekten oder einer Stichwaffe) bezeichnen.

Zu anderen Problemen der Abgrenzung siehe Fleischer (1997, S. 40–42).

Unikale Elemente können sowohl Nomen, Verben als auch Adjektive sein. Es handelt sich hierbei keineswegs um eine kleine, zu vernachlässigende Wortgruppe. Im Projekt A5 des SFB 441 der Universität Tübingen wurde eine Sammlung mit ungefähr 400 unikalenen Elementen im Deutschen erstellt. Diese Liste besteht größtenteils aus den Übersichten bei Dobrovol'skij (1988, S. 87ff), Dobrovol'skij/Piirainen (1994, S. 458ff) und Fleischer (1997, S. 37ff).

Klassifikation von unikalenen Elementen

Die Klassifikation unikalener Komponenten stellt sich in der Literatur als uneinheitlich dar. Da die Kriterien für eine bestimmte Klasse oft sehr unklar formuliert sind und sich manche Phraseologismen nicht eindeutig zuordnen lassen, gestaltete sich eine Einordnung aller im erwähnten Projekt gefundenen Beispiele stellenweise etwas willkürlich.

Grobklassifikation nach Nunberg/Sag/Wasow Zunächst kann man eine semantische Grobklassifizierung in **metaphorisch**, **dekomponierbar** und **nicht-dekomponierbar** vornehmen (siehe auch Kap. 2.4.3 und vergleiche Nunberg/Sag/Wasow, 1994). Dekomponierbare Phraseologismen mit unikalener Komponente sind z.B. *keinen Hehl aus etw. machen* oder *in Verruf geraten*, da man hier den einzelnen Komponenten eine entsprechende Bedeutung zuweisen kann, die dann zusammengesetzt die Bedeutung der Phrase bildet. Dies lässt sich anhand der Paraphrasierung zeigen: Die Bedeutung der genannten Wendungen ist „kein Geheimnis aus etw. machen“ und „einen schlechten Ruf bekommen“. Nicht dekomponierbar dagegen sind *Fersengeld geben*, *den Garaus machen* oder *Schindluder treiben*. Bei letzterer Phrase ist z.B. nicht klar, welche Bedeutung *Schindluder* bzw. *treiben* besitzen, um daraus die Gesamtbedeutung „jemanden/etwas schlecht behandeln“ abzuleiten. Als metaphorisch gelten Ausdrücke wie *aufpassen wie ein Schießhund* oder *sein Licht unter den Scheffel stellen*, wobei hier Scheffel zur unikalenen Komponente geworden ist (früher bekanntes Hohlmaß). Zur Metaphorik siehe auch Abschnitt 2.4.2.1.

Klassifikation nach Fleischer Wolfgang Fleischer (1997) nimmt eine andere Klassifikation vor, indem er sich auf die Typologie von Dobrovol'skij (1988) beruft und diese vereinfacht. Dobrovol'skij berücksichtigt nicht die semanti-

schen Eigenschaften, sondern die Wortstruktur der unikaligen Komponente und den Charakter bzw. den Grad der Determination durch die anderen Komponenten des Phraseologismus. Er unterscheidet zwei große Kategorien, nämlich Phraseologismen mit gebundenen Homonymen und solche mit gebundenen Formativen, wobei er die zweite Kategorie in vier Unterklassen teilt. Hier eine Übersicht der insgesamt 5 Kategorien mit je einem Beispiel (vgl. Dobrovol'skij, 1988, S. 93–107):

- Phraseologismen mit gebundenen Homonymen⁷: *Schmiere stehen* – ‚Schmiere‘ ist ein Homonym zur ‚fettenden Substanz‘.
- Phraseologismen mit gebundenen Grundmorphemen der Konstituenten: *sein Scherflein zu etw. beitragen* – ‚Scherflein‘ nur in dieser Umgebung, das Grundmorphem ist ‚Scherf‘
- Phraseologismen mit gebundenen Wortbildungskonstruktionen: *nicht viel Federlesens machen* – Bildung aus dem Syntagma ‚Federn lesen‘
- Phraseologismen mit Wortformanomalien: *gehupft wie gesprungen* – ‚gehupft‘ statt ‚gehüpft‘
- Phraseologismen mit sich wechselseitig determinierenden formal gebundenen Konstituenten: *Zeter und Mordio schreien* – beide Nomina treten nie einzeln auf.

Klassifikation nach Feyaerts Kurt Feyaerts (1994) unternimmt eine Kategorisierung der unikaligen Elemente im Niederländischen. Er berücksichtigt sieben Kriterien, darunter das ausschließliche Vorkommen der Form in festen Verbindungen, das Fehlen einer wendungsexternen Bedeutung und Wortformanomalien. Feyaerts stellt ein komplexes Typologisierungsschema auf und zeigt damit anschaulich die Heterogenität dieser Wörter „in semantischer, morphologischer, lexikalischer und syntaktischer Hinsicht“ (S. 159). Gerade deshalb gibt es wohl auch so unterschiedliche Ansätze zur Klassifizierung, wovon wir drei nun erwähnt haben.

Nach diesem Exkurs wenden wir uns einem weiteren lexikalischen Kriterium zu.

⁷Der Homonym-Begriff wird hier gemäß Dobrovol'skij (1988) verwendet.

2.2.4 Bedeutungsvolle vs. bedeutungsschwache Verben

Häufig treten in IVPs bedeutungsschwache Verben wie *machen*, *haben* und *sein* auf:

(21) *jdn. den Garaus machen*

(22) *Fracksausen haben*

(23) *dumm wie Bohnenstroh sein*

Mindestens genauso häufig jedoch sind Fälle mit bedeutungsstarken Verben, wie z. B. :

(24) *Bauklötze staunen*

(25) *sich etw. aus den Fingern saugen*

(26) *jdm. etw. in die Schuhe schieben*

Oft kommen auch bedeutungsvolle „Allerweltsverben“ wie *geben* oder *nehmen* vor.

Unerlässlich für die Eingrenzung des zu diskutierenden Datenbereichs, aber keinesfalls einfach, gestaltet sich die Differenzierung zwischen verbalen Phraseologismen und Funktionsverbgefügen (FVG). Während wir es bei den IVPs mit Wendungen mit idiomatischer Lesart zu tun haben, besitzen FVGs noch immer die freie Lesart. Lediglich hat das Verb die lexikalische Bedeutung als Vollverb verloren, siehe die Beispiele in (27) und (28).

(27) *etw. zur Aufführung bringen*

(28) *bei jdm. Billigung finden*

Manchmal stehen FVG-artige Wendungen an der Schwelle zum IVP, beispielsweise:

(29) *in Betracht ziehen*

(30) *zu Wege bringen*

Hier ist die wörtliche Bedeutung bereits sehr verblasst. Solche Wendungen sowie normale FVGs werden in dieser Arbeit nicht berücksichtigt, sondern sollen Gegenstand der weiteren Forschung sein (siehe auch Ende Kapitel 7).

2.3 Syntaktische Kriterien

Nach den morphologischen Merkmalen wenden wir uns nun den syntaktischen Eigenschaften der Bestandteile eines Idioms zu.

2.3.1 Anomalien

Auf der morphosyntaktischen Ebene spielen sogenannte Argumentmodifizierungen eine Rolle. Im Zuge der Idiomatisierung kann sich der Kasus eines Argumentes ändern. Vergleiche:

(31) *etw. kostet jdn. einen Geldbetrag* (nicht idiomatisch)

(32) *etw. kostet jdm. das Leben* (idiomatisch)

Der Akkusativ ist zum Dativ geworden. Geht man von einer freien Verwendung des Verbs aus, liegt ein anomaler Kasus vor.

Unter den IVPs finden sich auch eine Reihe syntaktisch-anomaler Konstruktionen. Diese Oberflächenanomalien weisen schon auf den ersten Blick auf die Idiomaticität einer Wendung hin. Hierzu zählen Rektionsanomalien (33, 34) und Anomalien im Gebrauch von Präpositionen (35), Pronomen (36) und Artikeln (37).

(33) *mit jdm. ist nicht gut Kirschen essen* – anomale Konstruktion aus *sein* + Verbalphrase (VP) + Adverb

(34) *jdn. Lügen strafen* – anderes Subkategorisierungsmuster als bei ‚strafen‘ in freier Verwendung

(35) *etw. ist nicht ganz ohne* – Präposition ohne Komplement

(36) *einen an der Waffel haben* – Pronomen ohne Antezedens

(37) *jmd. riecht Lunte* – fehlender Artikel bei zählbarem Nomen

Zwei Beispiele aus anderen Sprachen:

(38) *to follow suit* (jdm. etw. gleichtun) – fehlender Artikel bei einem *count noun*, im Englischen normalerweise ungrammatisch

(39) *avoir pignon sur rue* („Giebel zur Straße hin haben“ – gut situiert sein) – fehlender Artikel, im Französischen normalerweise ungrammatisch

Diese Wendungen folgen nicht den normalen Regeln der Syntax. Laut Keil (1997, S. 21) handelt es sich hierbei jedoch um eine recht kleine Gruppe von IVPs, während die meisten syntaktisch wohlgeformt sind.

2.3.2 Valenz

2.3.2.1 Interne vs. externe Valenz

Bezüglich der Valenz der Verben in einem IVP kann man, in Anlehnung an Keil (1997) und Burger (2003), zwischen wendungsinternen und wendungsexternen Argumenten unterscheiden. Ein internes Argument ist fester Bestandteil der Wendung und kann nicht ausgetauscht werden. Bei einer Änderung oder Auslassung würde die Wendung ihre Idiomaticität verlieren. Externe Argumente hingegen sind solche, die vom Verb subkategorisiert werden, aber je nach Kontext unterschiedlich sein können. Ein Beispiel:

(40) *Er habe wohl den Kopf verloren, erklärte seine Verteidigerin.*⁸

Hier ist *den Kopf* internes Argument von *verlieren*, das Subjekt *Er* ist extern und kann sogar ein Pronomen sein, das sich auf einen Referenten außerhalb der Phrase bezieht.

Die Begriffe ‚intern‘ und ‚extern‘ in der Phraseologie sind zu unterscheiden von der gleichlautenden Dichotomie in manchen Syntax-Theorien. Dort sind interne Argumente diejenigen, die innerhalb der maximalen Projektion des Verbs (VP) auftreten müssen, das externe Argument (das Subjekt) wird dagegen in vielen Arbeiten außerhalb der VP generiert.

2.3.2.2 Valenzänderung

Die Valenzstruktur einiger IVPs entspricht derjenigen, die auch einer freien Verbverwendung zu Grunde liegt. Im IVP

(41) *Ihm fällt ein Stein vom Herzen.*

hat das Verb die gleiche Valenz wie in *Ihm fällt ein Glas vom Tablett* mit dem Schema: (*etw.*) *fällt* (*jdm.*) (*von etw.*); allerdings nur, wenn man nicht zwischen externen und internen Argumenten unterscheidet. Tut man dies doch, liegt eine Änderung der Valenzstruktur vor: Das direkte Objekt und die Präpositionalphrase (PP) in (41) sind interne Argumente, nur das Subjekt ist extern. Ein weiteres Beispiel ist der IVP

(42) *Er hat die Gabe, Ruhe in die Sache zu bringen.*

Dort ist auch das direkte Objekt und die PP zum internen Komplement geworden, extern ist nur noch das Subjekt. Im Vergleich dazu: *Er hat den Auftrag,*

⁸Zürcher Tagesanzeiger, 14.06.1997, S. 15

die Bücher in die Bibliothek zu bringen. – (jmd.) bringt (etw.) (wohin). Es hat bei dem Verb des IVP also eine Änderung der syntaktischen Funktion zweier Argumente und ein Thetarollenwechsel der PP stattgefunden.

Man kann hier von einer Valenzreduzierung sprechen, wenn man nur die externen Argumente zählt. So hat *bringen* in diesem Fall zwei externe Argumente verloren. Beispiele für eine echte Valenzreduzierung, wobei insgesamt weniger Argumente vorhanden sind als in der freien Verwendung, stellen folgende Wendungen dar, geht man von der normalen Valenzstruktur (jmd.) gibt (jdm.) (etw.) aus.

(43) *jmd./etw. gibt den Ausschlag*

(44) *jmd. gibt Fersengeld*

(45) *jmd. gibt Obacht*

Ein weiteres Beispiel ist *das Handtuch werfen* im Gegensatz zu *das Handtuch in die Ecke werfen*. Idiomatisch kann nur der erste Ausdruck verstanden werden, eine PP, welche die Richtung des Werfens angibt, darf nicht vorkommen. Es gibt auch echte Valenz erhöhungen. Zum Beispiel eine zweistellige in der Wendung (jmd.) (jdm.) (mit etw.) in den Ohren liegen im Gegensatz zur freien Verwendung von *liegen* in (jmd./etw.) (wo) liegt.⁹ Auch kam es zu einem Thetarollenwechsel, da *in den Ohren* nicht mehr die Rolle *Location* innehat.

Torzova (1983) unternimmt empirische Erhebungen, indem sie die Rate der Valenzänderung der Idiome aus „Werken der schöngestigen Literatur des XX. Jahrhunderts“ (ebd., S. 283) bestimmt. Sie findet eine externe Valenzreduzierung um ein Argument bei 62% und um zwei Argumente bei 6% der untersuchten Wendungen. Bei 10% kommt es zu einer Valenzhöhung. Nur bei 22% stimmen die Valenzverhältnisse überein.

In unserem Korpus (im Appendix A) finden wir folgende Verhältnisse: Eine Valenzreduzierung um ein Argument beobachtet man bei 76,8% von 190 IVPs, um zwei Argumente bei 6,8%. Eine Erhöhung der Valenzstellen stellt man bei 22,6% um ein Argument fest und bei 4,7% sogar um zwei Argumente.

⁹Valenzänderungen ganz allgemein sowie auch die Argumentmodifizierungen (s. o.) kann man mittels lexikalischer Redundanzregeln erfassen. Siehe u. a. Jackendoff (1975) und für die HPSG Meurers (1995) und Müller (2000) zur lexikalischen Passiv-Regel. Normalerweise sind solche Regeln dazu da, systematische Beziehungen zwischen lexikalischen Einheiten zu erfassen. Da man bei idiomatischen Valenzänderungen nicht von systematisch sprechen kann, bleibt zu prüfen, ob diese Herangehensweise adäquat ist.

Die Valenzstruktur stimmte nur bei 2,6% mit der freien Verwendung überein. Bei 12,6% der Fälle kam es gleichzeitig zu einer Valenzreduzierung und -erhöhung, beispielsweise bei *jdm. das Herz ausschütten*. Hier ist das direkte Objekt zum internen Argument geworden und das indirekte Objekt kam hinzu. Da Torzova auch Redewendungen mit freier Lesart berücksichtigt und ihr Korpus wesentlich größer zu sein scheint, weichen ihre Ergebnisse von den unsrigen ab. Überdies sind einige ihrer Urteile nicht nachvollziehbar. Beispielsweise analysiert sie die Valenz der Wendung *jdm. in die Arme fallen* als übereinstimmend mit derjenigen in der freien Verwendung (*etw. fällt wohin*), wo doch eine Internalisierung der Lokativ-PP und eine Erhöhung um ein obligatorisches Dativobjekt stattfand.

Generell kann man beobachten, dass bei vielen Fällen ein schon vorhandenes Argument „internalisiert“ wird, nämlich wenn es zu einer einstelligen Valenzreduktion kommt.

2.3.3 Kategorieller Status der Argumente

Informationen über den kategoriellen Status der Argumente umfassen die Wortart (oder *Part-of-Speech*) und die Information, ob es sich beim Argument um ein Wort oder eine Phrase handelt. Die Wortarten, wie beispielsweise Nomen, Pronomen und Adjektiv, haben dabei spezifische grammatikalische Eigenschaften. So werden Nomina im Deutschen dekliniert und können einen Artikel bei sich tragen, Adjektive können kompariert werden etc. Diese Informationen muss man bei einzelnen Lexemen nicht mehr in ihren Lexikoneintrag schreiben, sondern sie leiten sich einfach von der Wortart des Lexems ab.

Interne Argumente von Phraseologismen haben dabei grundsätzlich die Gestalt von Argumenten oder Adjunkten in freien Wendungen. Im Korpus (Appendix A) finden wir folgende Arten regelhafter interner Argumente vor:

- Nominalphrasen, realisiert als direktes Objekt (*den Vogel abschießen*) oder Subjekt (*jds. Thron wackelt*)
- Präpositionalphrasen (*auf die schiefe Bahn geraten*)
- Adjektive (*jdn. kalt erwischen*)
- Partikel (*immer auf dem Sprung sein*)

Jedoch gibt es bei Phraseologismen auch Fälle, in welchen die normalen Eigenschaften der Wortarten gewissermaßen außer Kraft gesetzt worden sind. So

hat das nominale Argument bei *Lunte riechen* keinen Artikel. Solche Irregularitäten (vgl. auch Abschnitt 2.3.1) müssen dann explizit im Lexikon angegeben werden, weil sie sonst im Widerspruch zu den normalen Annahmen im Zusammenhang mit der Wortart stehen. Folgende interne Argumente gehorchen nicht den allgemeinen grammatischen Regeln, sondern müssen einzeln im Lexikon spezifiziert werden:

- Nominale Paarformeln (*auf Stein und Bein schwören*)
- Präpositionalphrasen – in vielen Fällen kann das nominale Komplement nicht modifiziert werden (*einen Streit vom Zaun brechen*)
- Komplemente, die aus Adjektiv und Präpositionalphrasen bestehen – die sogenannten stehenden Vergleiche (z. B. *dumm wie Brot sein*, siehe Agricola, 1992, S. 29)

In Kapitel 6 werden wir auf viele der genannten Fälle noch zurückkommen und zeigen, wie man idiosynkratische Information im Lexikoneintrag im Rahmen einer HPSG-Grammatik aufnehmen kann.

2.3.4 Syntaktische Stabilität

Bei IVPs gibt es syntaktische Restriktionsbeschränkungen, d. h. IVPs weisen eine gewisse Inflexibilität gegenüber frei formulierten Phrasen auf. Wenn diese Beschränkungen verletzt werden, geht oft die idiomatische Lesart verloren. Manchmal resultiert eine bestimmte syntaktische Operation¹⁰ auch in einem ungrammatischen Satz. Keil (1997) führt im Rahmen ihrer Darstellung des Systems „Phraseo-Lex“ einige syntaktische Operationen an, die Hinweise auf die Stabilität eines Idioms geben. Es handelt sich um folgende Transformationen: Passivierung, Bildung des Imperativs, Nominalisierung, Negation mit *nicht* und *kein*, Relativsatzbildung, Fragesatzbildung, Modifizierung durch Adjunkte und Quantifizierung. Es sei angemerkt, dass die Referenzialität von IVP-Konstituenten die Voraussetzung ist, Relativsätze zu bilden und auch den Fragetest bei den internen Argumenten anzuwenden. Im Folgenden sollen unsere Daten daraufhin überprüft werden, ob diese genannten und wichtigen Transformationen dort möglich sind. Im Anhang A sind dabei für die einzelnen IVPs nur die Kriterien Referenzialität, Passivierung, Nominalisierung, die Expansion durch Adjunkte und die Negationsmöglichkeiten dargestellt.

¹⁰Die Begriffe „Operation“ und später „Transformation“ werden hier nicht im technischen Sinn einer Transformationsgrammatik gebraucht, sondern dienen als metaphorische Beschreibung für jede Art von syntaktischer Abweichung vom einfachen, aktivischen Deklarativsatz.

2.3.4.1 Passivierbarkeit

Hinsichtlich der Möglichkeit, IVPs zu passivieren, muss man unterscheiden zwischen der syntaktischen und der semantischen Sicht. Oft kann man IVPs passivieren und erhält durchaus grammatische Sätze.

(46) *jdm. wird ein Bär aufgebunden*

(47) *Der Stein wurde ins Rollen gebracht.*

Bei vielen Wendungen geht dabei aber die idiomatische Lesart verloren:

(48) *!Verloren wurde der Kopf.*

(49) *!Dampf wurde abgelassen.¹¹*

Manche Idiome kann man allerdings aus grammatikalischen Gründen nicht passivieren. Dies trifft in erster Linie auf diejenigen IVPs zu, bei welchen das Verb auch in der freien Verwendung nicht passivierbar ist, wie *haben* und *müssen*. Es gibt bei der Passivierbarkeit allerdings subtile Abstufungen, also auch Verben, die „nicht gerne“ im Passiv auftreten, wie z. B. *riechen* oder *nagen*. Für Idiome, die nur eine gebundene Lesart haben, gilt ebenfalls, dass manche wegen des Verbs nicht passiv-fähig sind, wie folgende Wendung:

(50) *die Spendierhosen anhaben*

Diejenigen, die passivierbar sind, verlieren ihre idiomatische Lesart nicht, vergleiche:

(51) *Ein Scherflein wurde (von jdm.) beigesteuert.*

(52) *Ihm wurde eine Abfuhr erteilt.*

Wenn eine Passivform ungebräuchlich ist, wie in den beiden folgenden Beispielen, liegt das meist an den Eigenschaften des Verbs.

¹¹Die Wendungen (46) bis (49) haben, obwohl sie vornehmlich als idiomatisch erkannt werden, auch eine freie Lesart. Wir vertreten in dieser Arbeit ein Konzept der Grammatik, die in erster Linie den einzelnen Satz berücksichtigt, und betrachten die Idiome hier ohne kontextuelle Information (zur Unterscheidung Satz-/Textgrammatik siehe Fleischer/Helbig/Lerchner, 2001, Kap. 5). Sicherlich könnte man einen Kontext finden, in dem Idiome wie (49) auch im Passiv noch idiomatisch sind. Jedoch spielen dabei pragmatische Gesichtspunkte eine große Rolle: Der Kontext erlaubt oft die verschiedensten Deutungen, vgl. hierzu auch die Abschnitte 2.3.4.8 und 2.6.

(53) *?Ein blaues Wunder wird erlebt. Aber besser:
Was hier erlebt wurde, war wirklich ein blaues Wunder.*

(54) *??Es wurde auf dem aufsteigenden Ast gesessen.*

Jedoch gibt es auch Fälle, in denen das Verb passivierbar ist, das Idiom aber nicht oder höchst selten im Passiv auftritt.

(55) *Ein Armutszeugnis wird ihm (von jdm.) ausgestellt.*

In der Regel scheint es, als ob IVPs ohne freie Lesart dann passivierbar sind, wenn auch das Verb in freier Verwendung passivierbar ist. Wie gezeigt, bleibt die idiomatische Interpretation bestehen.

Bei IVPs, welche sowohl eine wörtliche als auch eine idiomatische Lesart aufweisen, können passivierte Formen oft nicht mehr idiomatisch interpretiert werden.

Das Argument Fleischers (1997, S. 50), dass Passivierung ausgeschlossen wird, um die Adjazenz der festen Komponenten des Idioms beizubehalten, verliert an Boden, wenn man folgende Daten berücksichtigt:

(56) *Kunst darf einfach nicht über den Kamm der Ö-Norm, der Rentabilitätsberechnungen, der Aufwandsgerechtigkeit geschoren werden.*¹²

(57) *Den Garaus macht den Seglern die Langleinenfischerei.*¹³

Dobrovol'skij (1999) leistet einen Beitrag zu dieser Fragestellung, indem er nach allgemeinen Regeln für die Passivierung deutscher Idiome sucht. Sein Fazit (Dobrovol'skij, 1999, S. 37) sei hier kurz zusammengefasst. Es gibt vier Bedingungen für eine Idiom-Passivierung:

1. Semantische Voraussetzung: Das Idiom muss als eine agentiv-transitive VP interpretierbar sein.¹⁴
2. Grammatische Voraussetzung: Es muss eine Akkusativ-NP vorhanden sein, die bei der Passivierung als Subjekt fungiert.

¹²Salzburger Nachrichten, 25.05.1992

¹³Salzburger Nachrichten, 15.11.2000

¹⁴Gemeint ist hier nicht nur die wörtlich verstandene VP, sondern auch die Interpretation bzw. Paraphrase. Einen nicht dekomponierbaren IVP wie *das Zeitliche segnen* kann man nicht passivieren, da die Bedeutung „sterben“ intransitiv ist (siehe auch Newmeyer, 1974, S. 330).

3. Diese Akkusativ-NP muss eine relativ selbstständige Bedeutung haben, sonst wäre die Passivierung aus kommunikativer Sicht sinnlos.

4. Ist dies nicht der Fall und eine Passivierung findet trotzdem statt, so muss im Aktantenpotenzial der Wendung eine offene Stelle für eine NP oder PP vorhanden sein, welche die Funktion des Themas übernimmt.

Eine Nichteinhaltung dieser Regeln schließt nach Dobrovol'skij eine Passivierung aus. Er geht hier allerdings lediglich vom Standard-Passiv im Deutschen aus, wobei ein Akkusativobjekt zum Subjekt wird. Das unpersönliche Passiv lässt er aus nicht nachvollziehbaren Gründen außer Acht, obwohl er selbst ein Beispiel dafür (S. 25, ebd.) angibt:

(58) *jdm. wird auf die Sprünge geholfen*

Ein Korpusbeleg in unserem Appendix ist folgender:

(59) *Auch da wird gefressen wie ein Scheunendrescher.*¹⁵

Dobrovol'skijs Voraussetzungen gelten somit nur für die „normale“ Passivtransformation von transitiven Verben. Festzuhalten ist, dass Idiome im Wesentlichen dann passivierbar sind, wenn auch das Verb in seiner freien Verwendung passivierbar ist. Siehe Abschnitt 5.3.1.4 für eine Diskussion innerhalb der HPSG und die Behandlung idiosynkratischer Passivierbarkeit.

2.3.4.2 Imperativbildung

Generell kann man IVPs in den Modus des Imperativs setzen, wenn die Paraphrase im Imperativ auch grammatisch ist. Bei den allermeisten IVPs in unserem Korpus ist dies der Fall, hier nur zwei Beispiele:

(60) *Zeige dich erkenntlich!*

(61) *Drehe ihr doch den Geldhahn zu!*

Allerdings gibt es auch Fälle, in denen eine nichtidiomatische Wendung im Imperativ zulässig ist („*Komm, sei doch spendabel!*“), eine Form der idiomatischen Entsprechung aber nicht:

(62) **Komm, habe die Spendierhosen an!* vergleiche aber:

(63) *Komm, zieh die Spendierhosen an!*

¹⁵<http://forum.yellopet.de/beitrag-10208.htm> (26.07.04);

Dies liegt in diesem Beispiel daran, dass man von *haben* meistens keinen (positiven) Imperativ bilden kann, außer in „*Hab Acht!*“. Manchmal ist ein positiver Imperativ nicht sinnvoll, wie bei folgenden Wendungen.

(64) *?Geräte in Verruf!*

(65) *?Lege jedes Wort auf die Goldwaage! Aber:
Lege nicht jedes Wort auf die Goldwaage!*

Gerade dieses Beispiel zeigt, dass auch pragmatische Gründe einer Imperativbildung im Wege stehen. So kann nur die negative Variante des Imperativs sinnvoll (als Ermahnung) interpretiert werden. Ebenso verhält es sich bei der Entsprechung von „*Habe Angst!*“ in (66). Wenn man auf die Umstände keinen Einfluss hat, macht eine Äußerung wie (67) ebenfalls wenig Sinn.

(66) **Habe Fracksausen! Aber: Nun habe mal kein Fracksausen!*

(67) **Habe es schwer! Aber: Nimms nicht so schwer!*

Bei einigen IVPs jedoch ist eine Imperativbildung höchst ungebräuchlich bzw. die Grammatikalität ist abhängig vom Sprecher, obwohl man den Imperativ der Paraphrase problemlos bilden kann.

(68) *?Gewinne die Oberhand!* vs. „Sei erfolgreicher als andere!“

(69) *?Schneide ihr den Lebensfaden ab!* vs. „Töte sie!“

Keil (1997, S. 26) führt darüberhinaus noch Wendungen an, die den Imperativ wegen einer Anomalie oder einer funktionalen Beschränktheit nicht zulassen.

(70) **Friss einen Narren an ihm!* (Valenzabweichung)

(71) **Friss einen Besen!* (Funktional auf 1. Person Singular beschränkt)

2.3.4.3 Nominalisierbarkeit

Nominalisierung ist normalerweise eine Veränderung der Kategorie auf morphologischer Ebene. Darunter fällt die Derivation, wobei morphologische Prozesse wie Affigierung von Derivationsmarkern (*laufen* \rightarrow *Läufer*) oder die Rückbildung (*laufen* \rightarrow *Lauf*) eine Rolle spielen. Die zweite Art der Nominalisierung ist die Konversion oder auch Infinitivnominalisierung. Hier ändert

sich die Wortart, ohne dass eine morphologische Veränderung sichtbar ist (*laufen* \rightarrow *Laufen*). Wir wollen auf die morphologische Nominalisierung und die komplexen semantischen Prozesse, die dabei stattfinden, nicht eingehen und verweisen auf Hall (1980), der diese Prozesse diskutiert und eine transformationsgrammatische Analyse der Nominalisierung vorschlägt. Für eine Nominalisierungsanalyse im HPSG-Paradigma sei auf Reinhard (2001) verwiesen, die eine vollständige Morphologie-Komponente für die HPSG entwickelt und vor allem Nominalisierungen auf *-ung* untersucht.

Konversion kann auch auf syntaktischer Ebene stattfinden. Gallmann (1990) nennt dieses Phänomen „Nominalisierung beliebiger Sprachauschnitte“ (ebd. S. 86ff). Beispiele aus Gallmann (1985) und (1990) sind:

(72) sein dauerndes *Andern-auf-die-Zehen-Treten*

(73) die Sitte des *In-den-April-Schickens*

(74) das *So-Tun-als-ob*

(75) Das ist doch zum *An-die-Decke-Gehen!*

(76) Er ist ein rechter *Guckindieluft*.

Gallmann (1990, S. 86) schlägt folgende Regel zu deren Bildung vor:

(77) [beliebiger Sprachauschnitt] \rightarrow N

Gallmann nimmt dabei an, dass ein Ausschnitt aus einem Satz mit beliebiger Komplexität noch einmal zum Input der Morphologie-Komponente wird, um den nominalen Charakter zu erhalten. Er argumentiert, dass (i) eine Genuszuweisung stattfindet (meist Neutrum), (ii) es zu morphologischen Anpassungen kommt, (iii) das verbartige Element keine Personalendung trägt und (iv) die nominalen Elemente Teile von Komposita sein können (z. B. *Sauregurkenzeit*).

Man sieht bereits an den oben erwähnten Beispielen, dass die neu entstandenen Nomen aus Verben und deren Komplementen bestehen können. Auch bei IVPs ist solch eine Nominalisierung möglich:

(78) *Kann man das eigentlich lernen, das Sich-aus-den-Fingern-saugen oder erfordert das jahrelange Übung?*¹⁶

¹⁶Google (23.10.03): <http://www.nensch.de/comments/2003/8/29/143842/065/15>

Meibauer (2003) sammelte etwa 300 Beispiele solcher Nominalisierungen (oder nach Meibauer „Phrasenkomposita“), gibt einen Überblick über die formalen Eigenschaften und skizziert ein Analysemodell, welches eine nicht lineare Interaktion zwischen Morphologie und Syntax beinhaltet. Meibauer kritisiert Gallmanns Ansatz insofern, als er die unzureichende Spezifikation des Inputs in (77) und die daraus folgende Übergenerierung aufdeckt. In Meibauers Artikel geht es jedoch ausschließlich um die möglichen Erstglieder von Phrasenkomposita. Die mögliche Konversion des Kopfes von Verb zu Nomen lässt er unberücksichtigt. Erstglieder der Phrasenkomposita sind immer ganze Phrasen (z. B. NPs oder PPs); Artikel sind dabei nur in lexikalisierten Phrasen möglich (**der Die-grauen-Schlafen-Effekt* aber *die Das-verflichte-siebte-Jahr-Problematik*). Genitivattribute bzw. Appositionen und Relativsätze können nicht vorkommen (**der [Graue-Schlafen, die so cool aussehen]-Effekt*). So erklärt sich auch, warum folgendes Beispiel ungrammatisch ist.

(79) **In der Geisterbahn war es ein einziges Kalt-den-Rücken-Hinunterlaufen.*

In diesem Satz wurde versucht, alle idiomatischen Bestandteile in die Nominalisierung mit aufzunehmen. *Kalt* ist jedoch ein Adverb zu *hinunterlaufen* und bildet mit *den Rücken* keine Konstituente.

Für eine Betrachtung dieses Phänomens im HPSG-Rahmen siehe Abschnitt 5.3.1.5.

2.3.4.4 Negation mit *nicht*

Bei IVPs ist eine Negation im Normalfall möglich, d. h. es gelten hierfür die normalen Regeln, die auch auf nicht idiomatische Sätze zutreffen. Folgende Beispiele weisen eine Negation mit *nicht* auf:

(80) *Es ist deshalb unbedingt ratsam, sich zuerst zu überlegen, wie man sein Anliegen formulieren will, damit man nicht ins Fettnäpfchen tritt.*¹⁷

(81) *Bei der letzten Erhöhung meckerte selbst die FDP, die ansonsten kostendeckenden Gebühren nicht abhold ist.*¹⁸

(82) *Wir lassen uns nicht über den Löffel balbieren.*¹⁹

¹⁷St. Galler Tagblatt, 22.06.1999, Ressort: TB-STE;

¹⁸Züricher Tagesanzeiger, 24.06.1998, S. 23, Ressort: Stadt Winterthur;

¹⁹Mannheimer Morgen, 30.08.1995, Wirtschaft;

IVPs, welche bereits eine Negation enthalten, lassen sich nicht noch einmal negieren, ohne dass eine doppelte Verneinung oder eine ungrammatische Äußerung daraus resultiert:

(83) ?*Er sieht nicht den Wald vor lauter Bäumen nicht.*

(84) ?*Er könnte ihr niemals kein Haar krümmen.*²⁰

Der Bedeutung folgender Wendung ist bereits eine Negation inhärent, weshalb eine Negierung mit *nicht* unzulässig ist.

(85) **Einen Teufel werde ich nicht tun!*

Idiome mit einem unbestimmten Artikel verhalten sich bei Negation mit *nicht* auf erwartbare Weise. Entweder *nicht* wird fokussierend gebraucht:

(86) *Er hat bei ihm nicht **einen** Stein im Brett.*

Hier ist man versucht, *einen* zu betonen, und damit der Wendung die Nuance „nicht einmal einen einzigen“ zu geben. Oder *nicht* steht selbst im Fokus und es liegt eine sogenannte Echo-Negation vor, in der im Skopus von *nicht* die positive Form des Idioms wiederholt oder zitiert wird.

(87) Sprecher A: *Der hat doch einen an der Waffel!*

Sprecher B: *Nein, er hat **nicht** einen an der Waffel.*

(88) *Er schießt **heute mal nicht** einen Bock.*

Solche Wendungen werden normalerweise mit *kein* negiert (s. u.), eine Eigenschaft, die auch auf nicht idiomatische Phrasen zutrifft.

2.3.4.5 Negation mit *kein*

Eine andere Form der Verneinung ist die Negation mit *kein*, einem Artikel, der Negation und Indefinitheit ausdrückt.

(89) *Abweichend von der Tradition halten wir euch mal keine Standpauke oder geben euch kluge Worte mit auf den Weg!...]*²¹

²⁰Im Gegensatz zu Keils Urteil (Keil, 1997, S. 216) halten wir eine Negation dieser Wendung mit *nicht* für unmöglich.

²¹Mannheimer Morgen, 19.06.2000, Lokales;

- (90) *Schüssel scherzte, der Finanzminister brauche sich nicht zu fürchten; er habe gar keine Spendierhosen an, sondern nur 'Billig-Jeans', und Klima strich die Geige dazu.*²²

Oftmals als Konstituentennegation bezeichnet, wird hier, bei weitem Skopus von *kein*, die ganze Proposition negiert.

Bei definiten nominalen Komplementen geht bei einer Negation mit *kein* entweder die idiomatische Lesart verloren (92) oder es kommt zur Ungrammatikalität (94).

- (91) *Sie hat nicht die Flinte ins Korn geworfen.* gegenüber
 (92) *!Sie hat keine Flinte ins Korn geworfen.*
 (93) *Er hat heute noch nicht das Zeitliche gesegnet.* gegenüber
 (94) **Er hat heute noch kein Zeitliches gesegnet.*

Ein weiterer Fall ist die Wendung *Bauklötze staunen*, bei der man das ebenfalls nicht mit *kein* negieren kann. Hingegen kann man bei dem Ausdruck *Bauklötze stapeln* (kein IVP) diese Art der Negation anwenden.

Der Artikel *kein* kann in der Regel den indefiniten nominalen Komplementen immer voranstehen. Man kann also *jemandem keinen Bären aufbinden* und man wird hier (**k*)*einen Teufel tun*, dies für ungrammatisch zu halten. Letztgenannte Wendung steht für wenige Fälle, bei denen kein negativer Artikel stehen darf.

Bei artikellosen Komplementen geht häufig die idiomatische Lesart verloren:

- (95) *!Er hat kein Blut und Wasser geschwitzt.*

Bei artikellosen Nomen, die von einer Präposition regiert werden, ist die Verneinung auch nicht möglich, wahrscheinlich, weil eine Definitheit angenommen werden muss.

- (96) *jdn. in (*keine) Mitleidenschaft ziehen*
 (97) *jdn. zu (*keinen) Paaren treiben*
 (98) *jdn. in (*keinen) Misskredit bringen*

²²Kleine Zeitung, 22.01.1999;

Darüber hinaus gibt es Sonderfälle. So muss die Wendung *jdm. kein Haar krümmen* bzw. *niemandem ein Haar krümmen* immer eine Negation enthalten, jedoch nicht unbedingt innerhalb einer Konstituente, vgl. *niemals jdm. ein Haar krümmen*.

Insgesamt kann keine allgemeingültige Regel gefunden werden, außer dass man den Artikel als Anhaltspunkt hat. Gerade bei unikalen Elementen muss man die Möglichkeit der Negation einzeln untersuchen.

2.3.4.6 Referenzialität

Das Kriterium der Referenzialität ist zwar eher semantischer Natur, spielt aber bei einigen syntaktischen Transformationen eine Rolle. So hängt es eng mit der Möglichkeit der Relativsatz-Transformation zusammen. Auf die wendungsinternen Argumente kann man sich in der Regel nicht durch ein Pronomen beziehen, d. h. die Argumente sind nicht referenziell. So kann man nicht sagen **Das Ohr; über das er mich gehauen hatte, blutete*. Manchmal ist ein Satz noch grammatisch, wie z. B. *!Die Nase, auf der er ihm herumtanzte, war schon ganz platt.*, aber die idiomatische Bedeutung ist hier verloren gegangen und man ist versucht, an eine Episode aus Gullivers Reisen zu denken. Dennoch gibt es ein paar IVPs, bei denen Referenzialität beibehalten wurde. So ist folgender Satz durchaus wohlgeformt und die idiomatische Lesart bleibt erhalten.

- (99) *Legendär ist bereits das Fettnäpfchen, in das die Bahn bei der Einführung ihres neuen Preissystems getreten ist: [...]*²³

2.3.4.7 Fragesatzbildung

Beim Fragetest kommt es darauf an, ob ein wendungsinternes Argument referenziell ist. Ist es das nicht, scheitert der Test:

- (100) **Über was haut er sie? – Übers Ohr.*
 (101) **Über welches Ohr haut er sie?*

Dies kann man mit allen Wendungen, die im Anhang A als „nicht referenziell“ spezifiziert sind, nachprüfen.

Allerdings kann man i. A. bei IVPs eine Entscheidungsfrage bilden, wobei die idiomatische Lesart erhalten bleibt.

- (102) *Haut er sie übers Ohr?*
 (103) *Tanzt sie ihm auf der Nase herum?*

²³Google (27.01.04): <http://www.aichberger.de/42D-Preise-mit-System.htm>

2.3.4.8 Modifizierbarkeit

Eine weitere Restriktionsbeschränkung stellt die Fähigkeit dar, modifizierende Adjunkte zuzulassen. Einige IVPs erlauben es nicht, ihre internen Argumente zu modifizieren:

(104) *Nun nagten sie am (*großen/*dünnen) Hungertuch.*

Es gibt jedoch genügend Phraseologismen, die eine Modifikation zulassen und keineswegs feste, unveränderliche Wortketten bilden. Die Modifizierbarkeit muss dabei für jeden Fall einzeln untersucht werden.

(105) *Der eine oder andere Gast trinke vielleicht ein Bierchen weniger, habe die kürzeren Spendierhosen an oder sich nach der „Völlerei“ über die Festtage Mässigkeit auferlegt, erzählen die Wirte.²⁴*

(106) *Welch erbärmliches Armutszeugnis stellt sich der Senat aus, wenn er mit ein paar Dutzend Schmierern und Vandalen nicht fertig wird!²⁵*

Manchmal kann man nur bestimmte Adjunkte benutzen. Zwar ist *jdm. einen großen Bären aufbinden* zulässig, aber nicht *einen braunen Bären*. Hier spielt offenbar auch die semantische Interpretation eine Rolle (vgl. parallel dazu *jdm. eine große/*braune Lüge erzählen*). Ein weiteres Beispiel ist das Idiom *ins Gras beißen*. Man kann zwar *schnell* oder *unvorhergesehen ins Gras beißen*, so wie man auch *schnell* oder *unvorhergesehen sterben* kann, jedoch kann man nicht das Bild als solches modifizieren (außer bei Wortspielen):

(107) *(*fest) ins Gras beißen; ins *grüne/*hohe Gras beißen*

Das Phänomen der Modifizierbarkeit ist weitgehend idiosynkratischer Natur. Dies wird z. B. klar, wenn man nachprüft, ob man Adjunkte aus einem semantischen Feld beliebig austauschen kann:

(108) *fröhliche/heitere/*ausgelassene/*glückliche Urstände feiern.*

Dabei kommt es natürlich darauf an, wie fein man ein semantisches Feld definiert. Letztendlich wird man aber sagen wollen, dass solche willkürlichen Modifikationen als Kollokationen zu erfassen sind, wobei manche Adjunkte einfach gebräuchlich sind und andere wiederum nicht. Die Tatsache allein, dass

²⁴St. Galler Tagblatt, 10.01.1998, Ressort: TB-SG

²⁵Berliner Morgenpost, 12.11.98, Ressort: MEINUNG, S. 4

kein Beleg für ein bestimmtes Adjunkt gefunden wurde (z. B. für *ausgelassene Urstände*), bedeutet zudem nicht zwingend, dass dieses Adjunkt unzulässig ist.

Im Zusammenhang mit der Modifikation, wie wir sie verstehen, d. h. Expansion durch Attribute, steht Burgers Unterscheidung zwischen Variabilität und Modifikation (siehe Burger, 2003, S. 25–28). Zu Variabilität zählt Burger usu-elle Erscheinungen wie *jdn. auf den Arm/auf die Schippe nehmen* oder *auf dem auf-/absteigenden Ast sitzen*, wo entweder zwei Idiome synonym sind, obwohl sie verschiedene interne Argumente haben, oder wo eine Antonymie vorliegt. Modifikation ist für Burger ein okkasionelles Phänomen, d. h. ein Idiom wird nur in bestimmten Äußerungen abgewandelt. Oft können diese Abwandlungen jedoch über das einmalige Vorkommen hinaus gebräuchlich werden, vgl. *schlechter Rat ist teuer* oder *zu schlechter Letzt*.

In dieser Arbeit sollen okkasionelle Modifikationen, die beispielsweise Burger (2003, Kap. 7.2.2), Wotjak (1992) und Sabban (1998) ausführlicher behandeln, außer Acht gelassen werden. Autoren nehmen sich oft eine große Freiheit, Idiome zu verändern. Meist merkt der Leser aber sofort, dass es sich dabei um eine ungewöhnliche Abwandlung eines Idioms handelt, und inferiert, dass diese Abwandlung sein besonderes Augenmerk verdient, siehe auch Abschnitt 2.6.

Wir interessieren uns hier für Modifikationen im System, d. h. welche Änderungen bestimmte Phraseologismen normalerweise zulassen, ohne dass damit ein besonderer Effekt erzielt werden soll. Burger nennt dies Variabilität, bei Dobrovol'skij sind dies phraseologische Varianten. Dobrovol'skij (1988, S. 161–174) unterscheidet fünf Modifikationstypen, die mit je einem Beispiel wiedergegeben werden sollen (vgl. Wotjak, 1992, S. 7).

1. Strukturelle Varianten (Variierungen, welche die innere Organisation des Idioms nicht verändern): *sich mit Hand und Fuß/mit Händen und Füßen gegen etw. sträuben*
2. Strukturelle Synonyme (strukturelle Kongruenz und teilweise Identität des lexikalischen Bestands): *jdm. den Fehdehandschuh ins Gesicht/vor die Füße werfen*
3. Strukturelle Antonyme (Korrelierende Phraseologismen mit identischer Struktur und teilweise Identität des lexikalischen Bestands): *auf dem aufsteigenden Ast/dem absteigenden Ast sitzen*
4. Phraseologische Serien (Variation eines „nichtidiomatischen“ Adjektivs): *ein armes/gutes/großes/hohes Tier*

5. Konversive (Variation eines „neutralen“ Verbs v. a. bzgl. der Aktionsart):
sich auf die faule Haut legen/auf der faulen Haut liegen

Wenn wir sagen, dass eine Modifikation für einen bestimmten Phraseologismus zulässig ist, ist gemeint, dass es viele Belege für eine solche Modifikation gibt. Okkasionelle Modifikationen werden, wie schon gesagt, nicht berücksichtigt.

2.3.4.9 Quantifizierung

Die letzte Operation, die hier untersucht werden soll, ist die der Quantifizierung von Idiom-Konstituenten. Es ist klar, dass hier die Semantik eine Rolle spielen muss. Wenn ein Objekt nur einmal auftritt, kann man es nicht mit *mehrere* oder *viele* modifizieren. Gleichzeitig muss man im Bild bleiben: Wenn man davon ausgeht, dass ein Mensch nur einen Lebensfaden besitzt, kann man ihm folglich nicht **einige Lebensfäden abschneiden*.

Allerdings ist die Möglichkeit der Quantifizierung nicht nur semantisch gesteuert. Dort, wo das Objekt im Plural steht, ist manchmal eine Quantifizierung dennoch unmöglich:

(109) *jdm. *viele/*mehrere/*wenige/die Leviten lesen*

(110) **viele/*mehrere/*wenige/keine Sperenzchen machen* (auch ohne Artikel)

Eine grammatische, aber nicht idiomatische Äußerung bekommt man bei anderen Wendungen, wie

(111) *!Er wirft die Handtücher.*

Bei einer Reihe von Idiomen ist eine Quantifizierung durchaus zulässig. Dazu folgende Beispiele:

(112) *jdm. einige Standpauken halten*

(113) *sich etw. aus allen zehn Fingern saugen*

Man muss also auch bei dieser Operation feststellen, dass es sich um eine idiosynkratische Fähigkeit handelt.

2.3.4.10 Andere Variationen

Riehemann (2001) problematisiert ebenfalls die Variationsmöglichkeiten, welche bei Idiomen zulässig sind. Zusätzlich zu den bereits erwähnten Transformationen führt sie unter anderen an:

- Anhebungsstrukturen:
Die Standpauke scheint bereits gehalten worden zu sein.
- Kontrollstrukturen:
**Das Tanzbein versuchte, geschwungen zu werden.*
- Clefting:
Es ist das gleiche Fettnäpfchen, in das er immer wieder tritt.

Riehemann zeigt daran, dass bestimmte Ansätze der Idiom-Analyse an den Variationen scheitern, und argumentiert damit zu Recht, dass man Idiome nicht als feste Konstituenten („complete, fixed constituents“, Riehemann, 2001, S. 63) betrachten kann.

Soweit die Diskussion der syntaktischen Kriterien bei Phraseologismen. Einen zusammenfassenden Überblick bietet die Tabelle 2.1.

2.4 Semantische Kriterien

Bei der semantischen Analyse von Idiomen steht natürlich die Idiomatizität im Vordergrund und speziell die Frage, wie sich die Bedeutung der Wendung zusammensetzen lässt. Wir wollen uns aber zuerst einmal die Bedeutungen selbst ansehen.

2.4.1 Motivierung und Durchsichtigkeit

In einschlägigen populärwissenschaftlichen Texten oder Fernsehsendungen werden häufig Redewendungen erklärt und es wird erläutert, woher die Ausdrücke kommen. Dies zeigt, dass bei vielen Idiomen heute nicht mehr klar ist, wie sie zum Zeitpunkt der Entstehung motiviert wurden. Auf Grund nicht mehr benutzter Wörter in den Bestandteilen wie z. B. *„feilhalten“* in *Maulaffen feilhalten* ist die Durchsichtigkeit ebenfalls nicht mehr gegeben.

Manche IVPs sind durchsichtig und motiviert, wie *„Dampf ablassen“*, *„Dampf draufhaben“* oder *„unter Dampf stehen“*, da uns heute Geräte, die mit Dampfdruck arbeiten – und sei es nur der Schnellkochtopf – geläufig sind. Bei vielen Redewendungen muss man jedoch nachschlagen, wie sie entstanden bzw. wie

Kriterium	Grammatikalität generell	Besonderheiten bei	
		IVPs mit auch freier Lesart	Nicht-dekomponierbaren IVPs
Passivierbarkeit	je nach Verbeigenschaften und sem. Interpretation	idiom. Lesart geht teilw. verloren	
Imperativbildung	je nach Verbeigenschaften und sem. Interpretation; unmöglich bei Anomalien		je nach Paraphrase
Nominalisierbarkeit	Konversion möglich, aber idiosynkr. Ausnahmen		
Negation mit <i>nicht</i>	Nicht bei: vorhandener Negation, unbest. Artikel; zudem idiosynkratische Ausnahmen		
Negation mit <i>kein</i>	Nicht bei: best. oder fehlendem best. Artikel; idiosynkr. Ausnahmen	idiom. Lesart geht teilw. verloren	
Referenzialität	idiosynkratisch		i. A. nicht, nur in Wortspielen
Fragesatzbildung	möglich, Frage nach einzelnen Bestandteilen setzt Referenzialität voraus	idiom. Lesart geht i. A. nicht verloren	
Modifizierbarkeit	idiosynkratisch; Adjunkte richten sich nach Interpretation		
Quantifizierbarkeit	je nach Bild oder Interpretation, meist idiosynkratisch	idiom. Lesart geht teilw. verloren	

Abbildung 2.1: Syntaktische Operationen – Übersicht

sie motiviert sind, und sie sind nicht mehr durchsichtig (*Maulaffen feilhalten; Fersengeld geben*).

Es gibt jedoch einen Unterschied zwischen Motivierung und Durchsichtigkeit. Motivierung bedeutet, dass ein IVP aus einem bestimmten Grund entstanden ist, z. B. ein Zitat aus einem berühmten Werk der Literatur. So kommt die Wendung *jdm. die Leviten lesen* aus dem Mönchswesen und leitet sich vom lateinischen Namen des dritten Buches Mose im Alten Testament ab.

Durchsichtigkeit bedeutet, dass man noch nachvollziehen kann, warum man einen IVP in einem bestimmten Kontext gebraucht. Es ist verständlich, dass man Emotionen meist mit Temperatur gleichsetzt (*vor Wut kochen; das Eis brechen*). Man kann aber nicht nachvollziehen, warum man über Gaffer sagt, sie würden „*Maulaffen feilhalten*“, oder warum jemand manchmal „*keinen Bock auf etw. hat*“.

Manchmal findet eine sogenannte volksetymologische Remotivierung statt (siehe Keil, 1997, S. 84). Opake bzw. unmotivierte Redewendungen bekommen eine andere Motivation zugewiesen. Ein Beispiel ist *Bauklötze staunen*. Dass die Wendung laut Duden 11 (2002) aus dem Bild *staunen* = „große Augen machen“ entstanden ist und aus den ‚großen Augen‘ = Glotzaugen \rightarrow Klotzaugen \rightarrow Klötze \rightarrow Bauklötze wurden, weiß heute niemand mehr. Stattdessen führen viele ein Bild an, in dem man wie ein Kind über hoch aufgetürmte Bauklötze staunt oder ähnliches. Hier fand also eine Remotivierung statt.

Harald Burger hat in einem Vortrag (Jena, 15.01.2003) den Gesichtspunkt der Lebendigkeit eines Idioms vorgestellt. Lebendig ist ein Idiom genau dann, wenn es morphosyntaktisch nicht stark restringiert ist, es nach einem Modell gebildet wurde oder selber als Modell fungiert, es leicht motivierbar ist, im Text eine auffällige Form darstellt, ein gewisses Spektrum an Variation auftritt, es als bekannt eingestuft wird und in Korpora mit einer gewissen Häufigkeit auftritt.

Ein Beispiel für ein Modell ist ‚A übertrifft B durch x‘ in „*Anton trinkt Bert unter den Tisch*“ oder „*Anton redet Bert in Grund und Boden*“. Eine neue Wendung nach diesem Muster ist dann „*Pavarotti singt seine Kollegen aus dem Saal*“.

Burger führt als Beispiele lebendiger Idiome unter anderen „*bittere Pillen schlucken*“ oder „*jdm. die rote Karte zeigen*“ an. Man beachte, dass diese Idiome neben der gebundenen auch eine wörtliche Lesart aufweisen.

2.4.2 Bildfelder und Tropen

Redewendungen sind oft ‚bildkräftig‘. Burger (2003) diskutiert den Begriff des sprachlichen Bildes und definiert ‚bildkräftig‘ in dem Sinne, dass eine bildhafte wörtliche Bedeutung vorliegt und „die Projektion der wörtlichen auf die phraseologische Bedeutung für Sprecher des Deutschen leicht nachvollziehbar ist“ (Burger, 2003, S. 94).

Wir beschäftigen uns hier jedoch gerade nicht mit Wendungen, die eine wörtliche Bedeutung besitzen. Burger zitiert Röhrich und Mieder (1977), wonach eine stärkere Bildhaftigkeit dann gegeben ist, wenn die Durchsichtigkeit schwächer ist. Er weist damit auf einen immanenten Widerspruch hin, der darin besteht, dass der höchste Grad an Bildlichkeit dem höchsten Grad an Idiomatizität entsprechen soll. Gerade solche IVPs, die keine freien Bedeutungen mehr haben, können aber nicht im gleichen Sinne bildhaft sein wie Wendungen, die sich auf wörtlich zu nehmende Vorgänge beziehen (*etw. vom Zaun brechen vs. jdm. einen Korb geben*).

Dennoch beinhalten auch IVPs ohne freie Lesart bestimmte Bilder bzw. kommen aus verschiedenen thematisch-sachlichen Gebieten, von denen wir hier einige anführen wollen.

- Menschliches Handeln: *sich ein Armutszugnis ausstellen; das Kriegsbeil begraben*
- Menschlicher Körper: *jdm. mit etw. in den Ohren liegen; jdm. schwillt die Zornader an*
- Mechanik/Maschinen/Technik: *essen wie ein Scheunendrescher; auf Hochtouren laufen; Lunte riechen*
- Jargon aus Berufsgruppen (Handel, Jagdwesen, Seefahrt): *jedes Wort auf die Goldwaage legen, etw. auf dem Kerbholz haben; auf der Abschussliste stehen; Schlagseite haben; jdn. auf dem Kieker haben* (Fernrohr)
- Küche: *in Teufels Küche kommen; vor Wut kochen*
- Übersinnliches: *mit jdm. Engelsgeduld haben*
- Farbensymbolik: *sein blaues Wunder erleben* (blau ist Farbe der Täuschung, Lüge)
- Tierwelt: *aufpassen wie ein Schießhund; nur ein Katzensprung weit*
- Natur: *auf dem aufsteigenden Ast sitzen; auf dem Holzweg sein*

2.4.2.1 Metaphern

Typisch für Idiome ist ihre Metaphorizität. Christa Baldauf (1997) gibt einen Überblick über die Metapherndiskussion, kritisiert dabei Lakoff/Johnson (1980) und entwickelt das Metaphernverständnis weiter. Wichtig ist hierbei, wie auch Lakoff und Johnson betonen, dass Metaphorik ein allgegenwärtiger Bestandteil unserer Sprache ist und keineswegs als sprachlicher Sonderfall behandelt werden darf. Standardbeispiele sind *leere Worte* oder *Wahlkampf*, bei denen die zu Grunde liegenden Metaphern WÖRTER SIND BEHÄLTER²⁶ und POLITIK IST KRIEG dem Sprecher gar nicht mehr bewusst sind.

Baldauf modifiziert und präzisiert die Metaphertheorie von Lakoff/Johnson innerhalb der germanistischen kognitiven Semantik. Dabei ersetzt sie auch deren Klassifikation durch vier Klassen, die sich am Herkunftsbereich orientieren. Diese vier Klassen (Baldauf, 1997, S. 83f), zusammen mit Beispielen aus unserem Korpus, sind im Folgenden aufgeführt.²⁷

1. **Attributsmetaphern:** Darstellung von Eigenschaften physischer Wahrnehmung auf abstrakte Eigenschaften.
Beispiele: MANGEL AN EMOTIONEN IST KÄLTE → *das Eis brechen*; EMOTIONALITÄT IST WÄRME → *vor Wut kochen*
2. **Ontologische Metaphern:** Abstrakte Bereiche werden als Objekte dargestellt.
Beispiel: ABSTRAKTA SIND OBJEKTE → *jdm. den Lebensfaden abschneiden, das Kriegsbeil begraben*
3. **Bildschematische Metaphern:** Projektion gestalthafter Struktur in abstrakte Bereiche.
Beispiele: MENSCH IST BEHÄLTER FÜR EMOTIONEN → *Dampf ablassen*; LEBENSUMSTÄNDE SIND BEHÄLTER → *in Verruf geraten*; LEBENSWANDEL IST WEG → *auf die schiefe Bahn geraten*; GUTE STELLUNG IST OBEN → *die Oberhand gewinnen/Oberwasser haben*

²⁶Per Konvention werden Metaphern in Versalien geschrieben.

²⁷Baldauf stellt jedoch in ihrer abschließenden Diskussion fest, dass die verschiedenen Klassen nicht unabhängig nebeneinander stehen. Weiterhin macht sie auf das Problem aufmerksam, dass es gegenüber ihrer Klassifikation noch andere Metaphernkonzepte und -systeme gibt und dass es keinesfalls leicht ist, sprachliche Daten den Metaphernsystemen zuzuordnen. Bei einer theoretischen Festlegung von Metaphernkonzepten und der Zuordnung der Daten hat nicht nur die Art und Größe des ausgewählten Korpus großen Einfluss, sondern auch individuelle Ansichten zur Interpretation (vgl. Baldauf, 1997, Teil III).

4. **Konstellationsmetaphern:** Abbildung ganzer, gestalthafter Konstellationen in abstrakte Bereiche.

Beispiel: GELD IST WASSER → *jdm. den Geldhahn zudrehen*; ABSTRAKTA SIND TIERE → *den Drehwurm bekommen*; ABSTRAKTA SIND PERSONEN → *etw. feiert fröhliche Urstände*

Manchmal spielen in einer Redewendung auch mehrere Metaphern eine Rolle. Zur Bedeutung *jdm. einen Bären aufbinden* tragen die Metaphern GEDANKEN SIND OBJEKTE (zu 2.) und ABSTRAKTA SIND LEBEWESEN (zu 4.) bei. Das *aufbinden* entspricht hier: „jdn. etw. glauben manchen“.

Bei *jdm. geht ein Licht auf* schwingen die Metaphern POSITIV IST HELL (zu 1.), ABSTRAKTA SIND OBJEKTE (zu 2.) und ZUNAHME IST AUFWÄRTSBEWEGUNG (zu 3.) mit.

In diesem Zusammenhang spielt der Begriff des Bildfeldes eine Rolle. Ein Bildfeld entsteht beim Zusammenwirken von Bildspender und -empfänger durch einen sprachlichen Akt. Bei der Metapher GELD IST WASSER wäre der Bildspender das Wasser und der Empfänger das Geld. Bildfelder sind leicht erkennbar, da sie kulturell verankert sind und zum semantischen Bestand eines Sprachkreises gehören. Wolfgang Liebert (2002) geht detailliert auf den Begriff des Bildfeldes in synchroner Perspektive ein. Nach einer Darstellung der Etymologie des Begriffes und der klassischen Definition umreißt er den modernen Bildfeldbegriff. Dieser wurde gegenüber früheren Konzepten um kategorieübergreifende Bedeutungsbeziehungen erweitert. Liebert illustriert dies mit Wendungen aus der GELD IST WASSER Metaphorik in (114) oder (115).

(114) *Reichlich Geld fließt aus Brüssel.* (Bild ist im Verb enthalten.)

(115) *jdm. den Geldhahn zudrehen* (Bild ist im Nomen enthalten.)

Beide Ausdrücke stammen aus dem gleichen Denkmodell. Jedoch musste beim klassischen Bildfeldbegriff das Bild noch der gleichen Wortkategorie angehören und man hätte bei obigen Beispielen von zwei verschiedenen Wortfeldern gesprochen.

Liebert untersucht in seinem Artikel weiterhin Valenz, Morphosyntax und Wortbildungsphänomene (darunter auch Idiombildung), geht auf semantische Innovation durch Bildfelder und pragmatische Aspekte ein und gibt einen weiten Literaturüberblick zu Studien einzelner Bildfelder. Nicht zuletzt kommt er auf die praktische Anwendung von Bildfeldanalysen zu sprechen. So werden Metaphern gezielt bei der Erklärung und Popularisierung von Fachwissen eingesetzt, finden Anwendung in der Psychotherapie („*im Nebel umherwandern*“

als Bild für Orientierungslosigkeit eines Patienten) und erleichtern beispielsweise auch die Benutzung eines EDV-Systems durch graphische Benutzeroberflächen (z.B. Fenster-Metapher bei Windows™ oder Mac OS™).

Eine Unterart der Metapher ist die **Synästhesie**. Dabei werden bei Beschreibungen Eindrücke eines Sinnesorgans auf ein anderes übertragen. Beispiele sind *laute Farben* bzw. *klangvolle Bilder* und *süßer Ton*. Ein Beispiel für ein IVP aus unserem Korpus in Appendix A ist *Fäuste schmecken*: „*Der aufrechte Charakter hingegen zeigt dem Rüpel, wie Fäuste schmecken.*“²⁸

2.4.2.2 Metonymie

Ein anderer rhetorischer Tropus ist die Metonymie oder Teil-Ganzes-Beziehung, bei der ein Ausdruck durch eine sachlich verwandte Bezeichnung ersetzt wird: Ein „Wort bezieht sich nicht auf die in seiner wörtlichen Bedeutung vorgegebene Art von Entität, stattdessen referiert es auf eine andere Art von Entität, wobei es aber einen systematischen Zusammenhang zwischen beiden Arten von Entität geben muss.“ (Egg, 2004, S. 36)

Ein amüsantes Beispiel (in Anlehnung an ähnliche von Nunberg, 1979) ist die Ersetzung des Gastes durch das Gericht, welches er bestellt hat: „*Das verärgerte Schnitzel will zahlen*“. Auch bei IVPs tritt dieses Stilmittel auf. Folgende Wendungen aus Appendix A sind metonymisch, eine Erklärung folgt in Klammern.

(116) *Fersengeld geben* (der Mensch flieht mit dem ganzen Fuß)

(117) *Augenschondienst haben* (der ganze Mensch schont sich im Schlaf)

(118) *das Zeitliche segnen* (sich vom Irdischen verabschieden für ‚sterben‘)

(119) *das Gesicht verlieren* (Gesicht für ‚Ansehen‘)

(120) *den Kopf verlieren* (Kopf für ‚Gedanken/Kontrolle‘)

2.4.3 Kompositionalität

Die Bedeutung nicht idiomatischer Wendungen ist kompositional, d. h. die Bedeutung der gesamten Wendung lässt sich aus der Bedeutung ihrer wohlgeformten Teile rekonstruieren. Dies wird als Kompositionsprinzip bezeichnet, welches Gottlob Frege zugeschrieben wird.²⁹

²⁸Gefunden in Google am 08.10.03: <http://www.rasputin.de/NoSex/alltag.html>

²⁹Jedoch weist Janssen (1983, Kap. I.2) darauf hin, dass Frege selbst dieses Prinzip nie so formulierte und dass der Name „Freges Prinzip“ wohl eher als Ehrung seiner Leistungen als Semantiker zu verstehen ist (vgl. auch Löbner, 2003, S. 20).

Idiomatische Wendungen weisen oft keine Kompositionalität auf. Am Anfang des Kapitels wurde zwischen teildiomatischen und vollidiomatischen Wendungen unterschieden. Damit hängt auch die Frage nach der Komposition zusammen, jedoch muss man zwischen beiden Begrifflichkeiten differenzieren. Dass man bei idiomatischen Wendungen zwischen dekomponierbaren und nicht dekomponierbaren Wendungen unterscheidet, geht v. a. auf die Dekompositionshypothese von Gibbs und seinen Kollegen (z. B. Gibbs et al., 1989) zurück, welche in ähnlicher Weise schon von Rajchštejn (1980), Burger/Buhofer/Sialm (1982) und Dobrovol'skij (1988) postuliert wurde, siehe auch den Artikel von Nunberg/Sag/Wasow (1994).

Kompositionalität heißt in unserem Zusammenhang, dass man die Bedeutung des IVPs auf seine Bestandteile „verteilen“ kann. Auch wenn keine wörtliche Interpretation möglich ist, kann eine Wendung deshalb immer noch kompositional sein. An dieser Stelle sind ein paar Beispiele angebracht.

(121) *Fracksausen haben*

(122) *Dampf ablassen*

Diese Idiome sind kompositional bzw. dekomponierbar. Im Beispiel (121), das man mit „Angst haben“ paraphrasieren kann, hat *Fracksausen* die Bedeutung „Angst“ und *haben* behält seine wörtliche Bedeutung bei. Deshalb ist (121) teildiomatisch. Die Wendung in (122) dagegen ist vollidiomatisch: Eine Paraphrase wäre „seinem Ärger Luft machen“, wobei *Dampf* hier „Ärger“ bedeutet und *etw. ablassen* „etw. Luft machen“ entspricht. Dennoch ist sie dekomponierbar.

(123) *sich freuen wie ein Schneekönig*

(124) *jmd. geht der Arsch auf Grundeis*

Hier handelt es sich um zwei nicht-kompositionale Wendungen, da man nicht sieht, in welcher Weise die Bestandteile zur Gesamtbedeutung beitragen. Insbesondere die PP in (123), die so viel wie „sehr“ bedeutet, ist nicht dekomponierbar.³⁰ Das Beispiel (123) ist dabei teildiomatisch, da *freuen* seine lexikalische Bedeutung behält. (124) ist eine vollidiomatische, nicht dekomponierbare Wendung, die dieselbe Bedeutung wie (121) besitzt.

³⁰Die Urteile zur Dekomponierbarkeit bei den Beispielen im Anhang sind durchweg so begründet: nur wenn jeweils ein Teil der Bedeutung jedem einzelnen Wort zugewiesen werden kann (auch innerhalb PPs), liegt Dekomponierbarkeit vor.

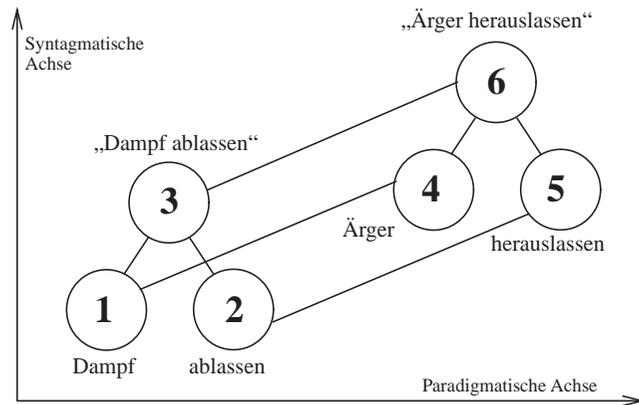
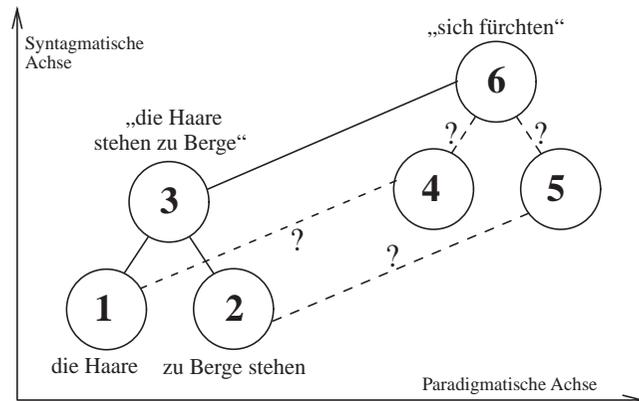
Ein Modell, das die Frage nach der Kompositionalität veranschaulicht und gleichzeitig noch semantische Eigenschaften wie Motiviertheit berücksichtigt, haben Geeraerts/Bakema (1993) entwickelt. Kurt Feyaerts beschreibt dieses **prismatische Bedeutungsmodell** detailliert und verständlich und bewertet es als einen „wertvollen Vorschlag“ (Feyaerts, 1994, S. 138).

Wir sind davon überzeugt, daß das prismatische Bedeutungsmodell als heuristisches Mittel zur Beschreibung der phraseologischen Semantik im allgemeinen eine gute Chance hat, weil es den verschiedenen semantischen Aspekten Rechnung trägt und sie miteinander verbindet. Motiviertheit und Isomorphie (paradigmatische und syntagmatische Transparenz) nehmen dabei eine zentrale Position ein. Aus dem Prismamodell wird ersichtlich, daß bei der Interpretation eines Idioms jede Komponente oder jedes Morphem, sei es auf wendungsexterner oder auf wendungsinterner Ebene, berücksichtigt werden muß. (Feyaerts, 1994, S. 160)

Das Prisma entsteht durch eine Kreuzklassifizierung zwischen der Syntagmatik und der Paradigmatik einerseits und zwischen der wendungsexternen (wörtlichen) und wendungsinternen Bedeutung andererseits. Mit Syntagmatik ist das Verhältnis zwischen Komponenten- und Gesamtbedeutung gemeint; Paradigmatik beschreibt das Verhältnis zwischen wendungsexterner und idiomatischer Bedeutung.

Wir wollen das Modell mit Hilfe des Beispiels *Dampf ablassen* illustrieren (Abb. 2.2). Die Knoten ① und ② stehen für die wörtliche Bedeutung der beiden Lexeme in der Wendung. Die Gesamtbedeutung „Dampf ablassen“ setzt sich aus ihnen zusammen und wird durch ③ repräsentiert. Es findet nun eine Ableitung der Bedeutung in ③ statt auf die wendungsinterne Bedeutung „Ärger herauslassen“ (in ⑥). Da diese idiomatische Bedeutung wiederum aus zwei Teilen besteht, also kompositional ist (Geeraerts und Bakema sprechen hier von Isomorphie), erfolgt eine Abbildung der beiden Bedeutungsbestandteile auf die Knoten ④ und ⑤. Man kann nun eine Rückinterpretation (entlang der beiden parallelen Projektionslinien) vornehmen und damit *Dampf* (①) die idiomatische Bedeutung „Ärger“ und *ablassen* (②) die Bedeutung „herauslassen“ zuweisen.

Wenn man mit diesem Modell nicht kompositionale Wendungen beschreiben möchte, werden die Projektionslinien an geeigneter Stelle unterbrochen. Beispielsweise kann man bei der Wendung *mir stehen die Haare zu Berge* („ich fürchte mich“) keine Projektion von der idiomatischen Bedeutung in ⑥ auf deren Teile vornehmen. Siehe dazu die Abbildung 2.3.

Abbildung 2.2: Prisma von *Dampf ablassen*Abbildung 2.3: Prisma von *Haare zu Berge stehen*

Mit dem prismatischen Modell wird auch dem Begriff der Motivierung Rechnung getragen, da man eine Verbindung zwischen der wörtlichen und der idiomatischen Bedeutung eines Lexems sowohl über die Gesamtbedeutung als auch auf Komponentenebene herstellt. Hinter „Ärger“ steht die Metapher DRUCK-IN-BEHÄLTER, welcher durch *Dampf* im physikalischen Sinn durchaus auch entstehen kann. So wird das Lexem *Dampf* auch in anderen Wendungen in der Bedeutung „Ärger“ oder „angestaute Emotion“ gebraucht. Beide Bedeutungen verhalten sich wie Polyseme zueinander. *Dampf ablassen* ist also schon durch die Projektion von ① nach ④ motiviert.

Zum Thema der „Nicht-Kompositionalität“ merkt Dobrovol'skij (1995) an, dass es kein hinreichendes Kriterium von Idiomatizität sein kann, da es – wie wir bereits gesehen haben – auch dekomponierbare Idiome gibt. Ob ein bestimmtes Idiom dekomponierbar ist, ist subjektiv und die Beurteilung richtet sich laut Dobrovol'skij nach der Organisation des verwendeten Lexikons und nach ökonomischen Kriterien.³¹ Oft werden den Teilen nicht dekomponierbarer Idiome auch eigene Bedeutungen zugewiesen und man verwendet sie als semantisch selbstständige Einheiten, wie z. B. *Fettnäpfchen*.

Die Möglichkeit, *Fettnäpfchen* auch ohne die Distribution *ins – treten* zu verstehen und die ganze Aussage sinnvoll zu interpretieren, legt nahe, daß man als Sprecher Idiome dieser Art zwar als Ganzes speichert, jedoch bereit ist, auch ihre Konstituenten als selbstständige Wörter mit spezifischen Bedeutungen aufzufassen. (Dobrovol'skij, 1995, S. 25)

Dabei sagt Dobrovol'skij keinesfalls, dass alle Idiome als dekomponierbar analysiert werden und gibt einige Beispiele von nicht dekomponierbaren Phraseologismen an.

Nach diesen Ausführungen wollen wir die soziolinguistischen Kriterien betrachten.

2.5 Soziolinguistische Gesichtspunkte

Jeder Mensch, also jeder Sprecher einer bestimmten Sprache, ist Teil einer Gruppe. Seine Identität wird von dieser Zugehörigkeit definiert. Werner Veith spezifiziert diesen Punkt folgendermaßen:

³¹Man kann dem Wort *Patsche* die Bedeutung „schlimme Situation“ zuweisen und hat damit die Idiome *in der Patsche sitzen*, *in die Patsche geraten* und *jdm. aus der Patsche helfen* einheitlich beschrieben.

Unter soziologischen Gesichtspunkten sind für die Identitätsbildung entscheidend u. a. die Zugehörigkeit zu einem sozialen Gebilde (z. B. einer Schicht oder einer Gruppe), auch zu einer bestimmten Religion bzw. Ethnie in multi-ethnischen Gesellschaften [...]; ferner: die familiäre Lage, auch in materieller Hinsicht, Geschlecht und Alter, Erziehung und Beruf, die Nationalität, die politische Orientierung u. a. Die Sprache in der Umgebung des Individuums, die diesem über soziale und linguale Beziehungen vermittelt wird, führt zu dessen sprachlicher Identität. (Veith, 2002, S. 31)

Das bedeutet, dass die persönliche Sprachgestaltung mit den gesellschaftlichen Verhaltenserwartungen in Einklang gebracht wird – es findet eine sprachliche Identitätsbildung statt. Die Soziolinguistik untersucht Sprache unter diesem Aspekt, d. h. die linguale Interaktion und die linguale Variation in der Gesellschaft.

Innerhalb einer Gesellschaft findet man folgende Kommunikationstypen:

- **Gemeinsprache** als „Instrumentarium an sprachlichen Mitteln, über das alle Angehörigen einer Sprachgemeinschaft verfügen und das deshalb die sprachliche Verständigung zwischen ihnen möglich macht“ (zit. nach W. Schmidt in Veith, 2002, S. 25f)
- Sprachvarietäten, bzw. orts- und bevölkerungsspezifische **Dialekte**
- **Fachsprachen**, die der fachspezifischen Kommunikation dienen
- **Sondersprachen** wie Rotwelsch, die Gruppenmitglieder als solche identifizieren
- **Jargon** als situationsabhängiger Spezialfall von Sondersprachen (z. B. Schülerdeutsch)
- sprachliche **Register**, wie „baby talk“, Umgangssprache oder gehobene, förmliche Ausdrucksweise
- **Gesamtsprache**, welche alle vorhergehenden Punkte umfasst

Nunberg/Sag/Wasow (1994) betonen, dass Idiome konventionalisiert sind. Idiomatische Lesarten sind damit Teil des sprachlichen Wissens einer Gesellschaft. Ihre Bedeutung kann nicht allein aus den Idiom-Teilen rekonstruiert werden, sondern die idiomatische Bedeutung ist erlernt.

Idiome zählen häufig zum Register der Umgangssprache, innerhalb der verschiedenen Jargons oder Sondersprachen gibt es verschiedene Idiome. So stammt eine gewisse Anzahl an IVPs aus der Jugendsprache, z. B.:

(125) *keinen Bock auf etw. haben*

(126) *die Flatter machen*

(127) *zu Potte kommen* oder *in die Puschen kommen*

(128) *einen an der Waffel haben*

(129) *jdm. geht der Arsch auf Grundeis*

(130) *jmd. hat es voll drauf*

Veith gibt Folgendes als Grund für die Häufung der Phraseologismen in der Jugendsprache an:

Die Kommunikationspraxis der Jugendlichen verfolgt sowohl das Ziel der Abgrenzung gegenüber den Erwachsenen als auch das Spiel mit der Sprache und die Lust an der Originalität. (Veith, 2002, S. 67)

Hinzu kommt, dass eine gewisse Vagheit der Idiome dazu führt, dass sie in verschiedenen Kontexten einsetzbar sind. So kann die Wendung in (130) Ausdruck genereller Akzeptanz sein, oder sich auf eine bestimmte Fertigkeit beziehen.

Die gehäufte Präsenz von IVPs in bestimmten Subsystemen wie Jargon oder Fachsprache ist also dadurch erklärbar, dass man sich erstens durch den Gebrauch von Idiomen von anderen abgrenzen möchte (bewusst oder unbewusst) und zweitens einen Sachverhalt treffender beschreiben kann, als durch eine standardsprachliche Formulierung. Zum letzten Punkt trägt einerseits die Vagheit der Semantik mancher IVPs bei, andererseits die oft sehr reiche Metaphorik.

Bei diesen Ausführungen wird deutlich, dass die Phraseologie voller Gegensätzlichkeiten steckt: einerseits die Verwendung **konventionalisierter** Ausdrücke aus Lust an der **Originalität**, andererseits die Tatsache, dass die **Vagheit** der Bedeutung eines Idioms dazu dient, einen Sachverhalt **treffender** zu beschreiben. Dies macht die Phraseologie zu einem interessanten Forschungsgebiet, aber birgt gleichzeitig auch Schwierigkeiten in Beschreibung, Abgrenzung und v. a. formaler Erfassung.

2.5.1 Dialektale Varianten

Es gibt natürlich auch dialekt spezifische Phraseologismen, wie berlinisch *etw. aus Daffke tun* („etw. aus Spaß/zum Trotz machen“) oder schweizerdeutsch *einen Bart einfangen* („sich blamieren“). Diese Wendungen kommen so nicht in der Standardsprache vor, sind also eigenständige Bildungen. Oft gehen in die Phraseologismen auch regional spezifische Begriffe ein, z. B. (i) österreichisch *des ziaht sich wie ein Strudlteig* oder (ii) im Schweizerdeutschen *d Finke chlopfe*, das soviel bedeutet wie „weggehen“ (*Finke* sind „Hausschuhe“, die man klopft bzw. mit denen man Lärm macht). Manchmal sind sich die Idiome in den verschiedenen Dialekten jedoch erstaunlich ähnlich. Das kommt einerseits daher, dass Metaphern nicht auf den Dialekt beschränkt sind, sondern auf der sprachunabhängigen Konzeptebene nahe liegen. Vergleiche:

(131) *aufpassen wie ein Heftelmacher* (mitteldt.) gegenüber
aufpassen wie ein Haftelmacher (österr.)

Ein Haftel- oder Heftelmacher ist jemand, der Drahtspangen (Hafteln) für Mieder herstellt. Diese Arbeit erfordert eine hohe Aufmerksamkeit. Weiterhin:

(132) *etw. unter den Teppich kehren* (binnendt.) gegenüber
etw. unter den Teppich wischen (österr./schweiz.)

(133) *die Faust in der Tasche ballen* (binnendt.) gegenüber
die Faust im Sack machen (schweiz.)

Andererseits gibt es auch Fälle, in denen Phraseologismen aus der Standardsprache vollständig übernommen wurden. Ein Beleg dafür ist das Beispiel im Schweizerdeutschen *Haar uf de Zäh ha* von *Haare auf den Zähnen haben*. Burger (2003, S. 198) gibt an, dass die Hälfte seines Materials an schweizerdeutschen Phraseologismen nur Varianten der binnendeutschen Entsprechungen sind. Die andere Hälfte sind eigenständige Bildungen.

2.5.2 Konnotationen

Da es in diesem Abschnitt 2.5 darum geht, wie Sprache in der Interaktion mit anderen verwendet wird, wollen wir noch einen Blick auf die konnotativen Bedeutungselemente werfen, die mit der Benutzung verbaler Phraseologismen

einhergehen. Palm (1997) gibt uns eine Zusammenstellung von Konnotationen, die hier mit je einem Beispiel wiedergegeben werden.³²

- Emotionale Bedingungen
 - scherzhaft *jmd. kriegt sein Jackenfett [ab]*
 - ironisch *etw. passt wie die Faust aufs Auge*
 - verhüllend *jmd. schneidet jdm. den Lebensfaden ab*
 - abwertend *jmd. macht sich bei jdm. lieb Kind*
 - anerkennend *etw. ist nicht ohne*
- Kommunikative Ebene
 - umgangssprachlich *jmd. macht die Flatter*
 - salopp *jmd. staunt Bauklötze*
 - derb/vulgär *jdm. geht der Arsch auf Grundeis*
 - offiziell *etw. schlägt zu Buche*
 - feierlich/gehoben *jmd. segnet das Zeitliche*
- Funktionsbereiche
 - administrativ *jmd. wird bei jdm. vorstellig*
 - juristisch *zur Entscheidung reif sein**
 - Medizin *jmd. ist ans Bett gefesselt**
 - Militär *jmd. ist beim Bund**
 - Sport *jmd. schießt ein Eigentor**
 - [sonstige Fachsprachen *starke Verben** (ling.); *rebooten** (inf.)]
- Soziale Geltung
 - Jugendsprache *jmd. hat keinen Bock auf etw.*
 - Familie *jmd. macht ein Bäuerchen*
 - Bildungssprache *in medias res**

³²Wendungen mit einem Asterisken erscheinen nicht im Appendix, da sie entweder keine idiomatische Lesart besitzen oder nicht ausschließlich idiomatisch verstanden werden können.

Weiterhin geht Palm auf die Regionalität des Phrasemgebrauchs (siehe dazu Abschnitt 2.5.1) und auf die Zeitgebundenheit ein. Dazu zählt sie Archaismen wie *in Ermangelung eines besseren* auf und führt Zitate von Schriftstellern vergangener Jahrhunderte an. Sie bemerkt zu Recht, dass es bei einer Aufzählung nicht bleiben kann, und fragt nach einem Zusammenhang zwischen Idiomatizität und Konnotation. Um diesem auf die Spur zu kommen, stellt sie ein Modell zur Sem-Analyse von Phrasemen vor, womit sie die verbreitete Annahme, „der Anteil der konnotativen Seme im Phrasem steige mit dem Grade seiner Idiomatizität“ (Palm, 1997, S. 24), widerlegen will. In dem Modell wird eine Sem-Analyse in denotative und konnotative Seme, eine Transformationsanalyse der Phrasemkomponenten (von wendungsexterner zu wendungsinterner Bedeutung) und eine Semsynthese über die quantitative und qualitative Distribution beider Semarten vorgenommen.

Leider handelt es sich hier um kein formal ausformuliertes Modell. In den von ihr gegebenen Beispielen wirken die einzelnen Schritte recht willkürlich und man kann, außer dass Seme gezählt werden, nicht sehen, wie ein Schritt vom vorigen abgeleitet wird. Dennoch zeigen ihre Beispiele, dass eine starke konnotative Markierung nicht mit einem höheren Grad an Idiomatizität korreliert.

2.6 Kognitiver Aspekt: Kreative Modifikation

Zum Abschluss dieser Diskussion der Klassifikationskriterien seien noch einmal die okkasionellen Modifikationen aufgegriffen (vgl. Abschnitt 2.3.4.8), die einen Beweis für die Kreativität, mit der Sprecher mit Idiomen umgehen, darstellen. Beispiele für diese Kreativität bzw. das sogenannte *idiom breaking* findet man sehr häufig.

(134) *Der Vorwurf paßte wie die sprichwörtliche Faust aufs Auge.*³³

(135) *Alle Figuren, die da von Joana Maria Gorvin bis zu Ernst Stötzner und Dörte Lyssewski gespielt werden, kennt man aus dem Berliner Effeß.*³⁴

(136) *Zu Saisonbeginn wollte man bei steigender Arbeitslosigkeit eigentlich Franz Xaver Kroetz' „Furcht und Hoffnung“ spielen. Erlebte aber blaue (Forderungs-)Wunder des Autors und sonst noch einiges...*³⁵

³³Frankfurter Rundschau, 23.01.1997, S. 1;

³⁴Die Presse, 07.02.1992;

³⁵Kleine Zeitung, 22.10.1996;

(137) *Was zunächst als geheime Operation gedacht war, spricht sich recht rasch im Ort herum, und vor allem die Damen zeigen sich pikiert, dass die Herren der Schöpfung mit ihren Freiersfüßen ausgerechnet diesen ungewöhnlichen Weg beschreiten.*³⁶

(138) *Entweder wir bauen neue Stadien oder wir beißen nicht nur sportmäßig ins Gras, sondern weiterhin in kalte Wurstsemmeln.*³⁷

Gerade eine Einfügung von *sprichwörtlich* wie in (134) ist häufig zu beobachten. Im Englischen gibt es das analoge Phänomen, z.B. *kick the proverbial bucket*. Kommt dieses Adjunkt vor, dient es als Hinweis darauf, dass der Sprecher sich der Sprichwörtlichkeit seiner Äußerung bewusst ist. Er zitiert sozusagen das Idiom bzw. weist ihm eine Markiertheit zu. In (134) modifiziert *sprichwörtlich* weniger das Nomen *Faust* an sich, sondern macht deutlich, dass eine solche Faust nicht wirklich existiert. Eine Paraphrase von (134) wäre etwa: *Der Vorwurf passte, sprichwörtlich gesagt, wie die Faust aufs Auge*.

Dabei steht die Kreativität, mit Idiomen umzugehen, im Widerspruch zur klassischen Sicht, dass man Idiome nicht modifizieren kann. So wurden Idiome in der früheren Phraseologieforschung als feste Einheiten im Lexikon betrachtet und damit wurde ihnen die Eigenschaft, unveränderlich zu sein, zugeschrieben.

Jedoch können offenbar diese festen Lexikoneinheiten recht flexibel gehandhabt werden. Normalerweise ist *Effeß* nicht modifizierbar, was aber einen Sprecher nicht davon abhielt, (135) zu produzieren. Gerade diese Kreativität ist es, die die Sprache in hohem Maße bereichert. Oftmals kommt es zu Sprachspielereien, in denen idiomatische Wörter wie *beißen* in (138) im wörtlichen Sinn benutzt werden, in diesem Beispiel sogar elliptisch. Zu untersuchen, welche Wirkung Sprachspiele mit Idiomen haben, fällt in den Bereich der Pragmatik.

Dobrovolskij beobachtet folgende Arten von Modifikation (1997, S. 72):

- Morphosyntaktische Veränderung des Idioms (nach Numerus, Artikelgebrauch, Präpositionen, Diminutivbildung etc.)
jdm. etw. in die Hand/in die Hände spielen
- Lexikalische Substitution (auch: Kontamination zweier Idiome)
jdn. über den Löffel balbieren/barbieren

³⁶Mannheimer Morgen, 19.04.2001, Ressort: Kino;

³⁷Kleine Zeitung, 08.06.1997;

- Lexikalisch-syntaktische Transformationen wie Eliminierung von Konstituenten (*in der Klemme [sitzen]*) oder auch Passivierung, Aktionsart-Transformationen etc.
- Textsemantische Modifikationen
Dem Fett (statt jdm.) geht es an den Kragen.

Als Gründe für diese Variierungen gibt Dobrovol'skij neben einfachen Versprechern die Einbettung des Idioms in den Kontext (im Rahmen des Usus) und kreative Sprachspiele zum Zweck eines besonderen perlokutiven Effekts an. Dennoch gibt es Evidenz für das Vorhandensein einer kanonischen Form der Phraseologismen (vgl. Riehemann, 2001, S. 32ff). Zudem müssen bestimmte Konstituenten eines Idioms immer beibehalten werden, ohne die das Idiom nicht mehr zu erkennen ist. Diese sogenannten Schlüsselkonstituenten (*keys*) wurden erstmals von Cacciari/Tabossi (1988) im Rahmen ihrer Konfigurationshypothese postuliert. Auch wenn die linguistische Natur der *keys* ungeklärt ist, gibt es experimentell nachweisbare Hinweise darauf, dass solche Konstituenten helfen, ein Idiom bei der Verarbeitung zu identifizieren. Strittig ist jedoch, ob es in jedem Idiom nur einen *key* gibt und welcher das jeweils ist. Siehe hierzu auch die Diskussion in Dobrovol'skij (1997, S. 42).

Dobrovol'skij (ebd.) gibt einen Überblick über Theorien und Ansätze in der kognitivbasierten Phraseologieforschung. Dort geht es v.a. um Speicherung und Verarbeitung idiomatischer Ausdrücke. Die Sichtweise, Idiome seien starre Einheiten im Lexikon, trifft nur auf eine kleine Gruppe von Phraseologismen zu, beispielsweise *gang und gäbe* oder *in Bausch und Bogen*. Diese Ausdrücke lassen sich nicht modifizieren und verhalten sich in mancherlei Hinsicht wie Simplizia. Cacciari/Glucksberg (1991) weisen experimentell nach, dass solche Einheiten kognitiv anders verarbeitet werden als transparente Idiome (vgl. dazu auch Dobrovol'skij, 1995, S.41ff).

In dieser Arbeit folgen wir der Dekompositionshypothese und es wurde schon deutlich, dass innerhalb der Idiome eine Unterscheidung bezüglich deren Teilbarkeit bzw. Kompositionalität getroffen wird. Der Sprecher ist sich über die Autonomie der Idiom-Konstituenten bewusst, weshalb er einzelne Konstituenten modifizieren kann. Dabei kann es sogar vorkommen, dass Teile von nicht dekomponierbaren Idiomen modifiziert werden, was durch folgendes Beispiel (Dobrovol'skij, 1997, S. 26) belegt ist.

- (139) Ned: *By and large*, people are well off these days.
Mark: *By and not-so-large!* Have you seen the figures on homelessness in Amerika[sic]?

Es scheint also, dass sogar Konstituenten solcher unteilbaren Idiome für manche Sprecher immer noch einen gewissen Grad an Autonomie besitzen. Es kann dabei nicht Aufgabe einer Idiomanalyse sein, solche Fälle zu berücksichtigen. Ebenso wenig kann eine Grammatik alle Fälle dieser Art erfassen oder beschreiben. Ein Sprecher ist sich zwar der kanonischen Form eines Idioms bewusst, weiß also, wie es normalerweise verwandt wird, setzt sich aber oft absichtlich darüber hinweg, um seiner Äußerung die gewünschte Wirkung zu verleihen.

In diesem Zusammenhang spielt die Verarbeitung von Idiomen eine nicht zu vernachlässigende Rolle. Dobrovol'skij (1997) stellt drei Hypothesen zur kognitiven Idiom-Verarbeitung vor, welche die Wechselwirkung von wörtlicher und gebundener Lesart modellieren. Erstens gibt es die Hypothese ‚literal vor figurativ‘, die davon ausgeht, dass grundsätzlich alle Äußerungen zuerst wörtlich interpretiert werden. Vertreter der Hypothese ‚figurativ vor literal‘ dagegen meinen, dass bei Idiomen immer erst auf die gebundene Lesart zugegriffen wird. Drittens gibt es die Hypothese der simultanen Verarbeitung, nach der beide Lesarten abgerufen werden und es sozusagen zu einem Wettlauf zwischen ihnen kommt. Unter Berücksichtigung der Semantik des Kontextes gewinnt dann eine der beiden Interpretationen.

Alle drei Hypothesen lassen sich experimentell bestätigen. Diese Widersprüchlichkeit führt Dobrovol'skij (1997, S. 35ff) auf die Heterogenität der Phraseologismen zurück. Demnach stimmen die Hypothesen nur bei der Verarbeitung von bestimmten Idiomtypen. Begegnet man einem (fast) unbekanntem Idiom, wird man es sicherlich erst einmal wörtlich interpretieren. Umgekehrt greift man bei einem sehr geläufigen Idiom zuerst auf die figurative Lesart zurück. Bei Idiomen, die auf dieser Geläufigkeitsskala zwischen den Extremen liegen, ist die Hypothese der simultanen Verarbeitung am plausibelsten. Dass es so viele Sprachspielereien mit Idiomen gibt, liegt daran, dass bei den meisten Idiomen die wörtliche Lesart auch zur Verfügung steht.

Das Phänomen des *idiom breaking* bzw. der Modifikation von Idiomen ist also produktiv und somit kann man in der Beschreibung niemals sämtliche Fälle berücksichtigen. Es bleibt der weiteren Forschung überlassen, bei dieser Kreativität bestimmte Muster herauszufinden und beispielsweise Änderungen, die nie an Idiomen vorgenommen werden, zu erkennen und auszuschließen.

Dobrovol'skij (1997) bietet in dieser Hinsicht einige Anhaltspunkte. Zusätzlich zur Unterscheidung dekomponierbar vs. nicht dekomponierbar teilt er die dekomponierbaren Idiome in (i) **potenziell teilbare** und (ii) **aktuell teilba-**

re **Wendungen** auf.³⁸ Potenziell teilbare Idiome bestehen aus Konstituenten, die zwar semantisch selbstständig sind, aber nicht als autonome Lexeme vorkommen, z. B. *Bock* in der Bedeutung „Lust“ in *keinen Bock auf etw. haben*. Aktuell teilbare Wendungen hingegen beinhalten Konstituenten, die auch außerhalb des Idioms in der idiomatischen Bedeutung vorkommen, wie *Trumpf* („Vorteil“) in *einen Trumpf ausspielen*.

Zusätzlich kann man nicht dekomponierbare Phraseologismen nach deren semantischer Wertigkeit kategorisieren. Dobrovol'skij (1997) bezieht sich auf Rajchštejn (1980), der die **Nullwertigkeit** von der **indirekten** und der **direkten Wertigkeit** unterscheidet. Nullwertig sind nicht teilbare Idiome dann, wenn gemeinsame semantische Komponenten in der Idiom-Bedeutung und im semantischen Potenzial der Einzelkonstituenten völlig fehlen. Beispielsweise gibt es in der semantischen Struktur von *ins Gras beißen* keinerlei gemeinsame Komponenten mit der Semantik von „sterben“.

Unter der indirekten Wertigkeit versteht Rajchštejn eine Korrelation zwischen der Idiom-Bedeutung und der Semantik seiner Teile im wendungsexternen Gebrauch. So stehen die *Ohren in jdm. mit etw. in den Ohren liegen* für das Gehör bzw. die Aufmerksamkeit des Adressaten (vgl. auch *die Ohren spitzen* oder *die Ohren auf Durchzug stellen*). Das Idiom ist nicht dekomponierbar, jedoch besitzt die Konstituente *in den Ohren* eine gewisse semantische Autonomie.

Die direkte Wertigkeit kommt bei allen nicht idiomatischen Phraseologismen vor. Deren Konstituenten tragen eine relativ selbstständige Bedeutung im Rahmen der Wendung.

Diese feineren Unterscheidungen können herangezogen werden, um idiomatische Wendungen noch genauer nach ihrer Möglichkeit zur Modifikation zu untersuchen. Dobrovol'skij hält fest:

Die Analyse des empirischen Materials läßt annehmen, daß zwischen dem Grad der semantischen Autonomie der Idiom-Konstituenten und der Flexibilität der Idiom-Struktur eine tendenzielle Korrelation besteht. (1997, S.101).

Er geht davon aus, dass es sich bei der Kompositionalität sowie bei der semantischen Autonomie der Konstituenten um ein graduelles Phänomen handelt (vgl. auch Geeraerts/Bakema, 1993). „Der Grad der Autonomie ist im ganzen mit dem Grad der Idiom-Variabilität isomorph.“ (Dobrovol'skij, ebd.). Dabei ist die Kompositionalität nicht der einzige Faktor der Variabilität. Es spielen

³⁸Eine andere Terminologie, die benutzt wird, ist paradigmatische bzw. paradigmatisch-syntagmatische Teilbarkeit.

außerdem sicherlich das Weltwissen des Sprechers sowie textsemantische Faktoren eine größere Rolle.

Um dieser Frage weiter nachzugehen, müsste man auf das menschliche System der Sprachproduktion im Allgemeinen zu sprechen kommen. Bierwisch/Schreuder (1992), die das Sprachproduktionsmodell von Levelt (1989) ergänzen, betonen, dass die Komponenten der Sprachproduktion, die bei der Konzeptualisierung einer Äußerung eine Rolle spielen, in hoch komplexer Weise zusammenwirken.

2.7 Zusammenfassung

In diesem Kapitel wurden zahlreiche Eigenschaften von Idiomen vorgestellt, die als Kriterien zu deren Klassifikation dienen können.

Zuerst unterschieden wir zwischen vollidiomatischen und teildiomatischen Wendungen.

Daraufhin wurden morphologische und lexikalische Kriterien näher betrachtet, darunter bestimmte „eingefrorene“ Eigenschaften, morphologische Anomalien, unikale Elemente und die Art der Verben.

Der dritte Punkt war die Syntax, wobei wir syntaktische Anomalien und Valenzänderungen ausmachten und die Kategorie der Argumente bestimmten. In diesem Zusammenhang wurde die Fähigkeit von Idiomen, sich bestimmten syntaktischen Operationen zu unterziehen, untersucht.

Viertens haben wir semantische Kriterien wie Motivierung und Durchsichtigkeit, Metaphorik, wobei wir auf die Klassifizierung von Christa Baldauf eingegangen, und Kompositionalität beleuchtet. Diese wurde mit dem prismatischen Bedeutungsmodell von Geeraerts und Bakema veranschaulicht. Auf die Dichotomie von dekomponierbaren und nicht dekomponierbaren IVPs werden wir in Kapitel 6 zurückkommen, da sie bei der dort diskutierten Analyse eine wichtige Unterscheidung darstellt.

Weiterhin sind wir auf soziolinguistische Gesichtspunkte eingegangen und haben uns mit der Frage nach dialektalen Varianten und den Konnotationen von Phraseologismen auseinandergesetzt, wobei die Aufstellung von Christine Palm hilfreich war.

Zuletzt wurde ein Ausblick auf die kognitive Phraseologieforschung gegeben und in diesem Zusammenhang haben wir besonders die Modifikation der Idiome diskutiert, ein Phänomen, das innerhalb der Grammatik nicht ausschöpfend modelliert werden kann.

Nach dieser Darstellung von Klassifikationskriterien ist deutlich geworden, dass keine der vorgestellten Eigenschaften allein auf IVPs ohne freie Lesart zutrifft. Man kann genauso wenig sagen, dass IVPs ohne freie Lesart alle dieses oder jenes Kriterium gemeinsam haben. Leider scheint es so, als ob sich IVPs ohne freie Lesart nicht derart klassifizieren lassen, dass bei allen Wendungen innerhalb einer Klasse die gleichen syntaktischen Operationen möglich sind. Man kann zwar beispielsweise alle IVPs, die nominalisiert werden können, zusammenfassen, das heißt aber nicht, dass alle IVPs dieser Gruppe dann automatisch auch bestimmten anderen Operationen unterzogen werden können.

In der Literatur werden Idiome oft in bestimmte Kategorien eingeteilt und man gibt vorschnell pauschale Urteile über die so klassifizierten Idiome ab. Es sollte in diesem Kapitel klar werden, dass man es in der Idiomatik mit einem sehr idiosynkratischen Verhalten einzelner Idiome zu tun hat und dass es bei sämtlichen Klassifizierungsversuchen stets Ausnahmen gibt.

Wir wollen im folgenden Kapitel versuchen, anhand der vorgestellten Kriterien einige Kennzeichen für eine idiomatische Lesart zu finden, um den Datenbereich für die vorliegende Arbeit sinnvoll einzugrenzen.

3 Der Datenbereich

Man muss vieles übersehen, um schauen zu können.

Emanuel von Bodman

3.1 Definition

Im vorhergehenden Kapitel wurde deutlich, dass es sich bei der Phraseologie um ein umfangreiches Gebiet handelt. Verschiedene Forschungsstrategien, eine Vielzahl an Ergebnissen, Übersichten, Wörterbüchern und Klassifikationsarten wurden veröffentlicht. Daher muss bei einer Arbeit wie der vorliegenden eine Einschränkung des Datenbereichs und der Analysekriterien getroffen werden.

Hier sollen diejenigen Idiome untersucht werden, die keine freie Lesart besitzen, also diejenigen, bei denen nur eine gebundene bzw. idiomatische Lesart verfügbar ist. Zur Illustration sei das Beispiel *jdm. einen Korb geben* herangezogen: Die freie Lesart wäre, dass tatsächlich ein Korb den Besitzer wechselt, während die gebundene Lesart eine Absage an jemanden beinhaltet. Im Gegensatz dazu gibt es den IVP *jdm. eine Abfuhr erteilen*. Bei dieser Wendung liegt keine freie Lesart vor. Nur die idiomatische Bedeutung „jdn. zurückweisen“ kann dem Satz zugewiesen werden.

So einfach, wie es den Anschein erwecken mag, ist eine Unterscheidung zwischen freier und gebundener Lesart jedoch nicht. Nehmen wir den IVP *jdm. einen Bären aufbinden*. Hier ist eigentlich keine wörtliche Lesart vorhanden, da man normalerweise niemandem ein Exemplar dieser Tiergattung auf den Rücken bindet. Die übertragene Bedeutung ist „jdm. eine Lüge glaubhaft machen“. Allerdings kann man sich schon gewisse Kontexte vorstellen, in denen wirklich jemand einem anderen einen Bären auf den Rücken bindet, die Wendung also doch wörtlich verstanden werden will. Bei IVPs dieser Art handelt es sich deshalb um Grenzfälle. Manchmal sind jedoch die verwendeten Bilder so abstrus, dass eine wörtliche Lesart nur schwer vorstellbar ist. Beim genannten IVP ist das sicherlich der Fall, noch klarer wird es bei *Schmetterlinge im Bauch haben*: Hier stellt man sich ein Bild von lebenden Faltern vor, die jemandem ein kribbliches Gefühl im Bauch verschaffen – jedoch hat norma-

lerweise niemand je lebende Schmetterlinge im Bauch. Im Folgenden sollen Kennzeichen aufgelistet werden, die möglichst eindeutig auf eine idiomatische Lesart als die einzig vorhandene hinweisen.

In unserem Datenbereich wollen wir außer den IVPs, die neben ihrer gebundenen Lesart auch eine freie aufweisen, sämtliche Phraseologismen, die keine idiomatische Lesart besitzen, unberücksichtigt lassen. Burger (2003) zählt zu diesen Phraseologismen „im weiteren Sinn“ auch die Funktionsverbgefüge (FVG). Beispiele für FVGs (oder *support verb constructions*) sind *etw. in Betrieb nehmen* oder auch *etw. zur Aufführung bringen*. In diesen Konstruktionen ist die Bedeutung der Verben verblasst, ihr semantischer Beitrag besteht zumeist nur in Aspekt und Aktionsart (so bei *Einfluss haben/nehmen/erhalten*, siehe Gallmann, 1999).

Manchmal gibt es in einem FVG ein unikales Element, wie in folgenden Wendungen:

(140) *in Betracht ziehen/kommen*

(141) *außer Acht lassen*

Bei diesen Wendungen handelt es sich eindeutig um Nominalisierungen von *betrachten* und (*be*)*achten*, deren Bedeutung erhalten geblieben ist. Unikal sind die Nomen deshalb, weil sie nicht mehr frei verwendet werden. Jedoch gibt es FVGs, die eine übertragene Bedeutung haben, die also IVPs sehr ähnlich sind (siehe Abschnitt 2.2.4):

(142) *sich in Acht nehmen*

Hier liegt zwar auch eine Nominalisierung von (*be*)*achten* vor, die Wendung bedeutet aber so viel wie „auf der Hut sein/vorsichtig sein“. Diese und andere FVG-ähnliche Wendungen wie die beiden folgenden sollen in unseren Datenbereich aufgenommen werden, da keine freie Lesart vorhanden ist.

(143) *jdm. etw. madig machen*

(144) *sich mausig machen*

3.2 Kennzeichen

Kennzeichen oder *marker* für Idiome ohne freie Lesart erlauben, diese als solche sofort zu erkennen. Diese Kennzeichen sollten so beschaffen sein, dass

sowohl der Mensch als auch ein System für automatische Sprachverarbeitung erkennt, dass hier keine wörtliche Lesart vorliegt. Was könnten dies für Kennzeichen sein?

3.2.1 Unikale Elemente

Wie im Abschnitt 2.2.3 ausführlich behandelt, sind unikale Elemente auf bestimmte Kontexte begrenzt. Idealerweise sieht man diese Distribution schon an ihrem Lexikoneintrag (siehe folgende Kapitel). Deshalb ist das Vorkommen eines unikaligen Elementes ein zuverlässiges Zeichen dafür, dass es sich bei einem zu verstehenden bzw. zu verarbeitenden Satz um einen Phraseologismus (im weiteren Sinn) handeln muss. In den allermeisten Fällen weist dieser Phraseologismus nur eine idiomatische Lesart auf, da die Kontexte der unikaligen Elemente idiomatischer Natur sind. Außer die Beispiele in (140) und (141) sind uns kaum unikale Elemente bekannt, die eine freie Lesart besitzen.

Einen Sonderfall stellen gebundene Homonyme dar, wie beispielweise *den Ausschlag geben* oder *jdn. im Stich lassen*, vgl. Kapitel 2.2.3. Ein System zur maschinellen Sprachverarbeitung wird auch den Lexikoneintrag für den „Hautausschlag“ bzw. für den „Insektenstich“ finden. Beim Wort *Levit* in *jdm. die Leviten lesen* handelt es sich auch um ein gebundenes Homonym, das zwar aus „Levitikus“ entstanden ist, nun aber gleich lautet mit der Bezeichnung der Mitglieder des israelitischen Stammes Levi. Somit tritt eine Ambiguität auf. Wir haben uns daher entschlossen, in der Übersicht im Appendix diese Wörter zwar als unikale Elemente zu bezeichnen, jedoch deren Vorkommen im Idiom nicht als Marker zu werten.

3.2.2 Morphologische Anomalien

Auch morphologische Anomalien kommen nicht in freier Verwendung vor und sind eben nur noch in gewissen Kontexten - den Phraseologismen - erhalten geblieben. Begegnet man nun Phrasen wie *lieb Kind* oder *mit allen Wassern*, ist das ebenfalls ein deutliches Kriterium für ein Idiom.

3.2.3 Syntaktische Hinweise

Phraseologismen weisen manchmal eine abweichende syntaktische Struktur auf. Die Verben besitzen dann eine andere Valenzstruktur als in der normalen Verwendung. Die Wendung *jdm. die Leviten lesen* illustriert diesen Sach-

verhält. Man kann im Standarddeutschen normalerweise niemandem „etwas lesen“, sondern nur etwas **vor**lesen. Das Verb *lesen* hat in der genannten Wendung folgende Valenzstruktur: wendungsintern ein direktes Objekt, das auf ein Lexem fixiert ist (*die Leviten*), wendungsextern ein Subjekt und ein Dativ-Objekt. Genau dieses Dativ-Objekt fehlt im freien Gebrauch von *lesen*: *jmd. liest etw.*

Ein weiteres Beispiel sind wieder unsere Bauklötze. Die Valenzstruktur *jmd. staunt etw. über etw.* liegt nur im idiomatischen *staunen* vor. Bei der frei verwendeten Entsprechung fehlt das Akkusativ-Objekt. Diese Liste ließe sich noch fortführen. Folglich ist die Valenzänderung (und die damit häufig einhergehenden Thetarollenwechsel) ein Kennzeichen für eine gebundene Lesart. Einen anderen syntaktischen Hinweis bilden die Argumentmodifizierungen und die kategoriellen Anomalien (Abschnitt 2.3.1).

3.2.4 Selektionsrestriktionen

Manchmal sind in idiomatischen Ausdrücken auch die Selektionsrestriktionen der Verben verletzt. Selektionsrestriktionen sind Beschränkungen auf die möglichen Komplemente. So muss das direkte Objekt von *verschlingen* etwas Essbares sein. Ist es das nicht, so wie in „*Er hat das Buch in einer Nacht verschlungen*“, weiß der Leser, dass es sich um eine metaphorische bzw. übertragene Bedeutung handeln muss. In unseren Beispielen im Anhang gibt es dazu auch einige Fälle, unter anderen:

(145) *jdm. sein Herz ausschütten*

(146) *jdm. fällt ein Stein vom Herzen*

(147) *einen Teufel tun*

Verletzt wurden hier folgende Selektionsrestriktionen: (i) *Ausschütten* kann man normalerweise nur einen Behälter. Man hat also hier ein Beispiel für eine HERZ-IST-BEHÄLTER-Metapher. (ii) Ein Gegenstand kann nur von einem anderen Objekt herunterfallen, das sich in erhöhter Position befindet und als „Stellfläche“ für Gegenstände dienen kann. Dies kann man von einem inneren Organ wie dem Herzen nicht behaupten. (iii) Das direkte Objekt von *tun* kann nur eine Aktion sein. Die Deutung der Wendung ist nicht sicher.

Es gibt in der Literatur eine Diskussion darüber, ob Selektionsbeschränkungen Teil der Semantik sind oder von der Pragmatik behandelt werden sollten. So argumentiert beispielsweise Jackendoff (1993), eine Selektionsrestriktion sei

Teil der Verbbedeutung und sollte gänzlich in die Argumentstruktur des Verbs integriert sein (ebd., S. 53). Androutsopoulos/Dale (2000) legen hingegen dar, dass Selektionsrestriktionen Tatsachen in der Welt erfassen, die zwar im Allgemeinen, aber nicht notwendigerweise immer wahr sind (ebd., S. 1). Weil die Verletzung einer Selektionsrestriktion stark vom jeweiligen Kontext abhängig ist und manchmal auch bewusst übergangen wird (s.o.), sehen sie darin eher ein pragmatisches Phänomen.

Eine weitere Diskussion betrifft die Unterscheidung zwischen Selektionsbeschränkungen und sortalen Beschränkungen (vgl. Dölling, 1992 und Lang, 1994). Erstere sind nach den genannten Autoren Teil der sprachspezifischen und somit lexikalischen Information, während letztere universell gelten und der konzeptuellen Struktur zugeordnet werden. Beispiele für Sorten sind **Ding** („Die Münze ist geschmolzen.“) gegenüber der **Dingart** („Münzen werden geprägt.“), weiterhin **Konfiguration** („Der Haufen wird sortiert.“) und **Gruppe** („Das Team liegt in Führung.“). Dabei bilden sortale Beschränkungen eine echte Teilmenge der Selektionsbeschränkungen und letztere sind feiner abgestuft. Auch innerhalb der HPSG gibt es verschiedene Ansätze, mit Selektionsbeschränkungen umzugehen, siehe dazu v. a. Androutsopoulos/Dale (2000). In Kapitel 5.3.2.2 werden wir noch darauf zurückkommen. Festzuhalten bleibt, dass verletzte Selektionsbeschränkungen ein viertes Kennzeichen für eine gebundene Lesart darstellen.¹

3.2.5 Motivierung und Dekomposition

Wie in Abschnitt 2.4.3 dargestellt, gibt es Idiome mit dekomponierbarer und nicht dekomponierbarer Bedeutung. Man kann dies auch mit den Termen durchsichtige bzw. nicht durchsichtige Bedeutung ausdrücken. Gereon Müller (2000) nimmt eine feinere Unterscheidung vor und trennt zwischen opaken, semi-opaken, semi-transparenten und transparenten Idiomen.²

¹Man beachte, dass im Skopus einer Negation verletzte Selektionsrestriktionen wieder grammatisch sind (*Peter aß kein Motorrad*). Wir nehmen mit Sailer (2004a, S. 202) an, dass hier zwar eine Verletzung der Selektionsrestriktion vorliegt, die Negation als unabhängiger semantischer Operator die Ungrammatikalität aber wieder aufhebt. Es gibt noch weitere Kontexte, in welchen verletzte Selektionsrestriktionen nicht mehr ins Gewicht fallen, siehe dazu die Diskussion in Abschnitt 5.3.2.2.

²Anhand seiner Beispiele bleibt unklar, worin sich semi-opake und semi-transparente Idiome unterscheiden. Diese Einteilung macht er an syntaktischer Variabilität von nur vier Tests fest: Alle vier Klassen können **topikalisiert** werden, **passivieren** lassen sich semi-opake, semi-transparente und transparente Idiome, eine **Wh-Bewegung** tritt nicht bei opaken und semi-opaken Idiomen auf (bei semi-transparenten und transparenten Wendungen ist dies möglich) und eine **Links-Dislokation** ist lediglich bei transparenten Idiomen möglich.

Wenn Idiome nicht mehr dekomponierbar und nicht mehr durchsichtig sind, haben die Bestandteile endgültig ihren Status als Einheiten der Syntax verloren. Es kommt zu einem Widerspruch zwischen der polylexikalischen Form und der holistischen Bedeutung. Diese Eigenschaft bezeichnet man auch als anti-ikonisch. Dies kann auch als ein Idiomkennzeichen angesehen werden.

Damit haben wir nun eindeutige Hinweise an der Hand, die es uns erlauben, fast schon auf den ersten Blick zu sehen, wann es sich bei einer Äußerung um ein Idiom handelt. Ist dabei aber stets eine freie Lesart ausgeschlossen? Die Antwort darauf muss positiv ausfallen, da IVPs mit freier Lesart nur aus solchen Komponenten aufgebaut sein können, die auch frei vorkommen. Bei den ersten drei Kennzeichen ist dies nicht der Fall (außer bei einer geringen Anzahl unikalischer Elemente). Ferner dürfen grammatische Sätze mit freier Lesart die Selektionsrestriktionen nicht verletzen. Tun sie es doch, haben wir gesehen, dass eine gebundene Lesart vorliegen muss (oder dass normale Annahmen, z. B. in einem fiktiven Kontext, nicht gelten). Eine gebundene Lesart ist auf jeden Fall auch vorhanden, wenn ein Idiom nicht dekomponierbar ist. Allerdings ist das Vorhandensein eines einzelnen Kennzeichens noch keine Garantie für eine ausschließlich gebundene Lesart und es müssen bei solchen Idiomem mehrere Kennzeichen gleichzeitig vorliegen.

3.2.6 Phraseoschablonen?

Ein weiteres mögliches Kennzeichen sind sogenannte Phraseoschablonen (siehe Fleischer, 1997, Abschnitt 3.4). Die Schablone *jmd./etw. ist/macht etw. wie etw.* ist solch ein Fall, belegt in folgenden Idiomem:

(148) *essen wie ein Scheunendrescher*

(149) *sich freuen wie ein Schneekönig*

(150) *dumm wie Bohnenstroh sein*

(151) *klar wie Kloßbrühe sein*

Diese „stehenden Vergleiche“ sind so beschaffen, dass das Komplement der PP stets bildlich gemeint ist.

Eine Unterart solcher Vergleiche ist die Schablone *ein ... von ...* bei Äußerungen wie *ein Ozean von einem Markt* oder *ein Betonklotz von Hotel*. Das letztgenannte Beispiel hat allerdings eine wörtliche Bedeutung.

Es gibt noch weitere Phraseoschablonen, die keinen Hinweis auf eine idiomatische Lesart beinhalten, z. B. *tot ist tot, der Wagen springt und springt nicht an, Welch eine Frau!* oder *Schritt für Schritt*.³

Man kann also das Vorhandensein einer Phraseoschablone im Allgemeinen nicht als Kennzeichen für Idiome mit nicht wörtlicher Bedeutung auffassen. Auch was die stehenden Vergleiche betrifft, kann man, da die Schablone produktiv ist, jede Menge Sätze konstruieren, die zwar genauso aussehen, aber stets wörtlich gemeint sind:

(152) *rumpeln wie ein Güterzug*

(153) *schnell wie ein Leopard*

Hier wird zwar auch verglichen – explizit durch das Wort *wie*, aber die Vergleiche sind wörtlich gemeint: Ein Güterzug rumpelt wirklich und ein Leopard ist in der Tat schnell.

Es besteht hier jedoch ein Unterschied zu den Beispielen (148) bis (151). Dort liegt eine Metaphorik zu Grunde: Ein *Scheunendrescher* (eine Maschine) kann nicht wirklich essen und wenn etwas *klar wie Kloßbrühe* ist, ist es verständlich, aber nicht durchsichtig wie es Brühe sein kann.⁴ *Bohnenstroh* ist nicht wirklich dumm und ein Tier wie der Zaunkönig – daher kommt das Wort *Schneekönig* – erweckt mit seinem Ruf auch nur den Eindruck, dass es sich freut. Wir haben es hier also nur deshalb mit einer gebundenen Lesart zu tun, weil bei diesen Vergleichen immer ein Bild mitschwingt. Damit wären wir beim nächsten Punkt.

3.2.7 Seltsame sprachliche Bilder?

Ein weiteres Kennzeichen für eine idiomatische Lesart könnte das Vorkommen eines unsinnigen oder ungewöhnlichen Bildes sein:

(154) *jdm. einen Bären aufbinden*

(155) *jdn. über den Tisch ziehen*

(156) *das Gesicht verlieren*

³Beispiele alle aus Fleischer (1997, S. 137f)

⁴Tatsächlich beruht das Idiom auf einem ironischen Widerspruch, da Kloßbrühe normalerweise milchig trüb ist.

Leider weiß oft nur ein menschlicher Sprecher, manchmal selbst nur der Muttersprachler, dass es sich hier um eine nicht wörtlich gemeinte Äußerung handelt. Einen wörtlichen Kontext kann es sogar für (156) geben: z. B. bei Charakteren in einer graphischen Computeranwendung, bei der ein Programmierfehler vorliegt und das Gesicht nicht dargestellt wird.⁵ Ein weiteres Beispiel ist der Film „Im Körper des Feindes“ („Face/Off“) mit N. Cage und J. Travolta, deren Gesichter chirurgisch vertauscht werden. Man kann hier also nicht von einem zuverlässigen Kennzeichen sprechen, da man diese Wendungen in Ausnahmefällen auch wörtlich benutzen könnte.

Einem System zur maschinellen Sprachverarbeitung kann die Absurdität eines aufgebundenen Bären nicht nahegelegt werden. Obwohl der Mensch keine Probleme hat, Bilder und Metaphern zu erkennen, ist es sehr schwer, diese Fähigkeit auf ein digitales Sprachverarbeitungssystem zu übertragen.

3.3 Zusammenfassung

Um den Datenbereich für diese Arbeit zu umreißen, haben wir definiert, was Idiome ohne freie Lesart sind. Dafür haben wir folgende formale Kennzeichen ausgemacht: das Vorkommen von unikalen Elementen, morphologischen und syntaktischen Anomalien, Veränderungen in der Valenzstruktur der Verben, Verletzungen der Selektionsbeschränkungen und fehlende Kompositionalität. Wir haben gesehen, dass Phraseoschablonen und ungewöhnliche semantische Bilder keine eindeutigen bzw. formalen Kennzeichen darstellen.

Man könnte noch weitere Kennzeichen anführen, denkt man beispielsweise an die Unfähigkeit von Idiomen, sich bestimmten syntaktischen Operationen zu unterziehen (siehe Abschnitt 2.3.4). Dies ist bei einer konkreten Instanz eines Idioms nicht auf den ersten Blick zu erkennen, weshalb wir hier von einem Kennzeichen zweiten Grades sprechen wollen.

In dieser Arbeit wollen wir uns auf die Kennzeichen ersten Grades (unikale Elemente, morphologische und syntaktische Anomalien, Valenzänderungen, Verletzungen der Selektionsbeschränkungen und fehlende Kompositionalität) verlassen, um nur solche IVPs zu untersuchen, die eine nicht wörtliche Lesart besitzen.

An dieser Stelle ist noch einmal zu betonen, dass die genannten Kennzeichen nicht für sich alleine gesehen hinreichend sind, um Phraseologismen ohne freie Lesart zu identifizieren. Es müssen oft mehrere Kennzeichen zusammen auftreten, um ein eindeutiges Urteil fällen zu können.

Wir wollen im Folgenden u. a. untersuchen, inwiefern sich diese Kennzeichen in einer HPSG-Grammatik modellieren lassen. Nur wenn dies gelingt, kann man in einem weiteren Schritt sehen, wie hilfreich die genannten Hinweise bei der automatischen Sprachverarbeitung sein können; genauer gesagt, ob man mit einem Sprachverarbeitungssystem, das nach diesen Kennzeichen sucht, auch tatsächlich idiomatische von nicht idiomatischen Wendungen unterscheiden kann. Solche verarbeitungsbezogenen Fragestellungen können jedoch in dieser Arbeit leider nicht verfolgt werden.

⁵Hinweis von Brigitte Krenn (persönliche Mitteilung)

4 Herangehensweisen in der HPSG

In Anlehnung an Ferdinand de Saussure steht das sprachliche Zeichen im Mittelpunkt der HPSG.

Nachdem wir die Eigenschaften von IVPs umrissen und uns im letzten Kapitel bewusst auf einen bestimmten Datenbereich beschränkt haben, den der IVPs ohne freie Lesart, wollen wir einer formalen Analyse dieser Ausdrücke im Rahmen der HPSG nachgehen.

In diesem Kapitel sollen folgende Ansätze vorgestellt und kritisch hinterfragt werden: die interne Selektion der Komplemente von Krenn/Erbach (1994), die externe Selektion, die in Soehn (2003) entwickelt wurde, eine Konstruktionsanalyse nach Riehemann (2001) und ein Kollokationsansatz von Sailer (2003).

4.1 Grundbausteine der HPSG

Vor die Diskussion der Herangehensweisen soll ein Kurzüberblick über grundlegende Konzepte der HPSG gestellt werden, um dem HPSG-unkundigen Leser ein Verständnis der vorliegenden Arbeit zu ermöglichen. Dieser Abschnitt kann daher von mit der HPSG vertrauten Lesern übersprungen werden.

Es ist jedoch nicht das Ziel, eine komplette HPSG-Einführung zu bieten und wir können hier nicht im Detail auf alle Standard-Merkmale dieser Grammatiktheorie eingehen. Dafür sei auf Pollard/Sag (1994) als Referenz verwiesen. Eine HPSG-Einführung mit Lehrbuchcharakter stellt Sag et al. (2003) dar. Einführungen für das Deutsche finden sich beispielsweise bei Kiss (1995b) oder Klenk (2003, Kap. 5). Diese Arbeiten verfolgen jedoch nicht das Ziel, die formalen Grundlagen im Einzelnen zu erläutern und überhaupt liegt kein Werk in deutscher Sprache vor, in dem die technischen Begrifflichkeiten des der HPSG zu Grunde liegenden Formalismus in exakter und konsistenter Weise dargestellt werden.¹

¹In Richter (1997) werden formale Grundlagen von RSRL (*Relational Speciate Re-entrant Language*), einer Beschreibungssprache für die HPSG, auf Deutsch erklärt, wobei es sich hier allerdings um eine Vorversion von RSRL handelt.

Das Wort „HPSG“ bezeichnet zwei verschiedene Dinge. Erstens ist es ein Formalismus, um eine Grammatik zu formulieren und zu formalisieren.² Zweitens ist die HPSG eine universelle Grammatiktheorie, um sprachliches Wissen zu beschreiben und zu erklären.

Es gibt verschiedene Grammatik-Implementierungsplattformen, welchen die HPSG zu Grunde liegt (siehe Einführung). Eine Rechner-Implementierung vereint sozusagen Formalismus und Grammatik, da eine bestimmte Grammatik im Formalismus implementiert wird.

Die HPSG hat zwar ihren Ursprung in den USA (vgl. Pollard/Sag, 1987 und 1994) und wurde zunächst für das Englische beschrieben, aber mittlerweile gibt es Veröffentlichungen für die verschiedensten Sprachen. Zum Deutschen siehe beispielsweise Müller (1999).

Die HPSG ist eine nicht derivationelle, beschränkungs-basierte und oberflächenorientierte Grammatiktheorie. Nicht derivationell bedeutet, dass es in der HPSG nicht die Sichtweise gibt, dass eine Struktur in eine andere überführt wird (so wie bei *move- α* in der Theorie des *Government & Binding*). Stattdessen hängen unterschiedliche Repräsentationen durch deklarative Beschränkungen (deswegen beschränkungs-basiert) voneinander ab. Oberflächenorientiert ist die HPSG deshalb, da Äußerungen so charakterisiert werden, wie sie tatsächlich erscheinen, und nicht zwischen Tiefen- und Oberflächenstruktur unterschieden wird.³

4.1.1 Der Formalismus

4.1.1.1 Die Objektebene: Perlen und Schnüre

Abstrakte Gegenstände, welche die Linguistik untersucht, wie Wörter oder Phrasen, werden im HPSG-Formalismus als Objekte verstanden. Diese Objekte haben bestimmte Eigenschaften, die das linguistische Wissen über die Objekte repräsentieren. Man modelliert den Gegenstand dieses Wissens, indem für Objekte bestimmte Merkmale definiert werden, welche ihrerseits als Werte andere Objekte haben können. Dies lässt sich mit einer Perle (dem Objekt) veranschaulichen, an der eine oder mehrere Schnüre (Merkmale) befestigt sind, an denen wiederum Perlen hängen. Ein Merkmal ist also die Ver-

²Als Beschreibungssprache, welche es erlaubt, die Ideen der HPSG in formal korrekter Weise zu erfassen, sei auch in dieser Arbeit RSRL herangezogen, welche in Richter/Sailer/Penn (1999) und Richter (2004a) motiviert und definiert wurde.

³Siehe die Zusammenfassung von A. Lima und dessen Bibliographie: http://www.uq.edu.au/english/linguistics/Working%20Papers/ananda_ling/HPSG_Summary.htm (21.10.04). Leider werden jedoch auch dort die formalen Grundlagen zu ungenau erläutert.

knüpfung zwischen zwei Perlen. Es ist Aufgabe der Theorie, nur die wohlgeformten, grammatikalisch korrekten Objekte oder – im Bild – alle korrekten Perlen-Schnur-Konstellationen einer gegebenen Sprache zu beschreiben.

4.1.1.2 Die Beschreibungsebene: Etiketten und Matrizen

Möchte man diese Objekte beschreiben, kann man sich vorstellen, an die Perlen und Schnüre je ein Etikett zu kleben und sie zu benennen. Ein Objekt wird so einer Sorte zugeordnet und ein Merkmal bekommt einen Attributsnamen. Parallel zu den Perlen und Schnüren hat somit eine Sorte eine bestimmte Anzahl an Attributen, deren Werte wiederum Sorten sind. Für eine Beschreibung dieser „Konstellationen“, der sogenannten **getypten Merkmalsstrukturen**, bieten sich die sogenannten Attribut-Wert-Matrizen (*attribute-value-matrices*, AVM) an, mit welchen man Sorten und deren Attribute in eckigen Klammern darstellt. Es folgt ein kurzes (skizzenhaftes) Beispiel zur Illustration.

<i>familie</i> NACHNAME <i>mustermann</i> EHEMANN □ <i>markus</i> EHEFRAU <i>maria</i> KINDER <i>junge</i> NAME <i>michael</i> VATER □
--

Die Attribute NACHNAME, EHEMANN, EHEFRAU und KINDER sind also für die Sorte *familie* definiert. Das Attribut NACHNAME hat als Wert die Sorte *mustermann*. Die Sorten *markus* und *maria* sind entsprechend die Werte der Attribute EHEMANN und EHEFRAU. Das Attribut KINDER hat als Wert eine Liste, die hier genau ein Element enthält. Listen sind auch als Konstellationen zu sehen, was aber durch die Schreibweise in der AVM abstrakter und kürzer ausgedrückt wird. Das Listenelement der Sorte *junge* hat wiederum zwei Attribute, wobei der Wert von VATER identisch mit dem Wert von EHEMANN ist. Dies wird durch *tags* (□) angezeigt. Wichtig dabei ist, dass der Wert nicht nur ‚gleich‘ ist, sondern es sich um das **identische Objekt** (*markus*) handelt. Man spricht hier auch von Strukturteilung (*structure sharing* oder auch *token-identity*).

Um linguistische Fakten redundanzfrei darzustellen, werden Sorten zusammengefasst und Obersorten eingeführt, so dass eine **Sortenhierarchie** aus Ober- und Untersorten entsteht. Diejenigen Sorten, die keine Untersorten mehr haben, sind maximal spezifisch. Innerhalb der Sortenhierarchie gilt die Vererbung aller Merkmale, das bedeutet, dass alle Untersorten sämtliche Attribute der Obersorte erben.

Ein linguistisches Objekt (also beispielsweise ein geäußelter Satz) ist von einer bestimmten Sorte, welche maximal spezifisch sein muss – das Objekt ist somit *sort resolved*. Dabei haben sämtliche Attribute des Objekts auch einen maximal spezifischen Wert, der seinerseits für dieses Attribut definiert (bzw. angemessen) sein muss. Das Objekt ist damit *totally well-typed*, d. h. alle und nur die für die Sorte definierten Attribute sind vorhanden.

Auf der Beschreibungsebene jedoch wird zumeist nicht ein einziges Objekt beschrieben, sondern eine Klasse von verschiedenartigen Konstellationen mit bestimmten Eigenschaften, etwa alle Verbalphrasen mit zwei Komplementen. Aus diesem Grund werden in AVMS für gewöhnlich nicht alle für diese Sorte definierten Attribute spezifiziert, es liegt somit eine Unterspezifikation vor. Einerseits erlaubt dies eine Beschreibung von verschieden konfigurierten Konstellationen, andererseits ist dies ein gebräuchliches Mittel, um AVMS übersichtlich zu halten und nur die für die Diskussion relevanten Attribute und Werte darzustellen.

Im Gegensatz zu beispielsweise französischen Veröffentlichungen zur HPSG (vgl. z. B. Abeillé, 1993, Kap. 3) behält man im Deutschen in der Regel die englischen Attribut- und Sortennamen bei.

Es gibt bei AVMS eine Reihe von Abkürzungskonventionen. Oft wird der Sortenname nicht explizit angegeben, da er sich implizit daraus ergibt, welche Attribute die Sorte hat bzw. welchem Attribut die Sorte als Wert zugeordnet ist. Zum Beispiel ist das Attribut CASE nur für die Sorte *noun* angemessen, wenn also in einer AVM CASE spezifiziert wurde, ist klar, dass es sich dabei um die Beschreibung eines *noun* handelt. Listen werden mit spitzen Klammern gekennzeichnet und deren Elemente durch Kommata getrennt. Statt des Sortennamens für eine leere Liste (*elist*) schreibt man oft $\langle \rangle$. Darüber hinaus werden Attribut- und Sortennamen üblicherweise abgekürzt. Ein entsprechendes Verzeichnis findet sich am Beginn dieser Arbeit.

Der Wert von PHON ist eine Kette phonologischer Werte, eigentlich also eine Kette von Lauten, welche man im phonetischen Alphabet (IPA) notieren müsste. Der Einfachheit halber erscheinen in den meisten Arbeiten jedoch die Wörter in Standard-Orthographie.⁴

4.1.2 Signatur und Theorie

Die HPSG als Grammatik besteht aus zwei Teilen, der Signatur und der Theorie. In der Signatur sind die Sorten und deren Merkmale spezifiziert. Die so-

⁴Zur Integration einer phonologischen Komponente in die HPSG siehe Bird/Klein (1994) und den Vorschlag von Höhle (1999).

nannten Angemessenheitsbedingungen (*appropriateness conditions*) beinhalten die Information, welche Merkmale für welche Sorten definiert sind und was für Werte sie annehmen können.

Die Theorie enthält allgemeine Prinzipien der Grammatik. Die Signatur und die Prinzipien müssen so formuliert werden, dass sie genau die grammatischen Äußerungen einer Sprache lizensieren und gleichzeitig als Beschränkungen fungieren, so dass ungrammatische Äußerungen ausgeschlossen werden.

Alle Prinzipien sind in der Regel als Implikationen formuliert, d. h. wenn ein Zeichen von einer bestimmten Sorte ist, dann gilt dafür eine bestimmte Beschränkung. Solche Prinzipien werden als Sortenbeschränkungen bezeichnet. Allerdings kann man Prinzipien auch so formulieren, dass sie nicht für bestimmte einfache Objekte gelten, sondern komplexere Antezedenzen (Sorten mit in bestimmter Weise spezifizierten Attributen) haben. Dies hängt z. T. auch von den Vorlieben des Grammatikschreibers ab.

Außerdem umfasst die Theorie ein Lexikon, in welchem sämtliche Lexikoneinträge mit ihren idiosynkratischen Eigenschaften stehen. Später wird die prominente Stellung des Lexikons in der HPSG noch deutlich. Eine gängige Art, das Lexikon zu definieren, ist das sog. WORD-PRINCIPLE bzw. WORT-PRINZIP. Es handelt sich dabei um ein Prinzip, das für jedes Wort (Sorte *word*) gilt und als Konsequenz eine Disjunktion von Beschreibungen, den Lexikoneinträgen, besitzt. Diese (ziemlich umfangreiche) Menge an Disjunkten wird dann als „das Lexikon“ bezeichnet. Wir werden darauf in Kapitel 6.2 zurückkommen.

4.1.3 Wichtige Attribute, Sorten und Prinzipien

In Abb. 4.1⁵ wird ein sprachliches Zeichen, eine Phrase, beschrieben. Der Beschreibung liegt eine Signatur zu Grunde, die in Pollard/Sag (1994) vorgestellt und in vielen weiteren Arbeiten zur HPSG verwendet wird.

Die Sorte *phrase* ist eine Untersorte von *sign*. Jedes Zeichen hat bestimmte phonologische Eigenschaften (PHONOLOGY) und ist Träger gewisser syntaktischer und semantischer Informationen (SYNSEM). Diese Attribute sind für *sign* definiert und werden automatisch von den Untersorten *word* und *phrase* geerbt.

Als Phrase hat das Zeichen in Abb. 4.1 zusätzlich noch Töchter (Attribut DAUGHTERS): eine Kopftochter (unter HEAD-DAUGHTER) und eine An-

⁵In dieser AVM wurden Vereinfachungen vorgenommen: Außer Sorten- und Attributbezeichnungen wurde auch der Zweck einiger Attribute vermerkt und erscheint in **gerader und fetter Schrift**. Solche Angaben sind normalerweise nicht Teil einer AVM.

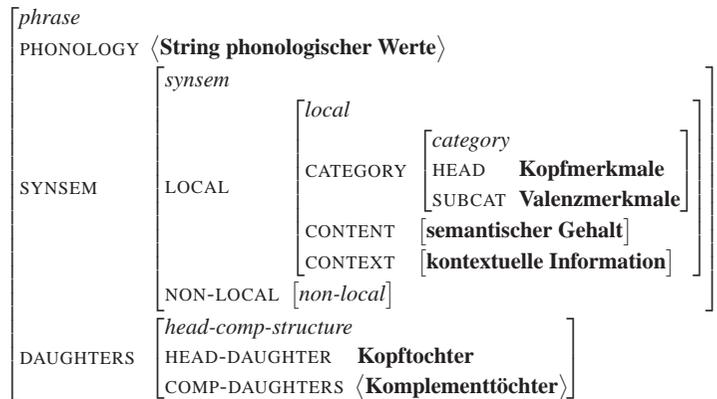


Abbildung 4.1: Eine einfache AVM

zahl von Komplementtöchtern, die auf der COMP-DAUGHTERS-Liste stehen. In Arbeiten, welche die binären syntaktischen Strukturen auch in der Merkmalsgeometrie ausdrücken, wird das Attribut COMP-DAUGHTERS durch NONHEAD-DAUGHTER ersetzt, welches als Wert nur ein Zeichen besitzen kann (vgl. z. B. Sailer, 2003). Für Zeichen der Sorte *word* ist das DAUGHTERS-Merkmal nicht definiert.

Zuletzt ist zu bemerken, dass unter SYNSEM zwischen lokalen (LOCAL) und nicht lokalen Informationen (NON-LOCAL) unterschieden wird. Lokale Informationen sind beispielsweise Kasus, Genus Verbi und die Art und Anzahl der Komplemente. Nicht lokale Information betrifft die Modellierung von nicht lokalen Abhängigkeiten und Extraktion.

Weitere Attribute, etwa für die logische Form und den idiosynkratischen Status, werden im Laufe der vorliegenden Arbeit hinzukommen. Im Appendix B findet sich eine Zusammenstellung aller Attribute und Werte, die wir gegenüber der Signatur in Pollard/Sag (1994) neu einführen. Es muss betont werden, dass es in der Fülle der Arbeiten zur HPSG unterschiedliche Annahmen bezüglich der Signatur gibt. So findet man z. B. häufig an Stelle des Merkmals SUBCAT das Attribut VALENCE mit weiteren Attributen SUBJ, COMPS und SPR. Je nach Zielsetzung der Arbeit wird der Fokus auf bestimmte Teile der Signatur gelegt und man äußert sich nie zu allen Bereichen der Merkmalsgeometrie.⁶

⁶Es gibt auch die Ansicht, dass die Signatur universale Gültigkeit haben sollte und die sprachspezifischen Besonderheiten nur in den Prinzipien und in den Lexikoneinträgen zum

Ein wichtiges Prinzip einer HPSG-Grammatik ist das KOPF-MERKMAL-PRINZIP. Es besagt, dass die Kopf-Eigenschaften entlang der Projektionslinie eines Zeichens identisch sind. Technisch ausgedrückt fordert es eine Strukturteilung zwischen den HEAD-Werten eines Zeichens und denen seiner Kopftochter.

Ein weiteres wichtiges Prinzip ist das SUBKATEGORISIERUNGS-PRINZIP, welches die Sättigung der Valenzmerkmale eines Zeichens sicherstellt. Dies geschieht dadurch, dass laut diesem Prinzip die Elemente auf der SUBCAT-Liste eines Zeichens identisch sind mit denjenigen bei der Kopftochter, jedoch abzüglich (der SYNSEM-Werte) der Komplemente dieser Kopftochter. In einer Baumstruktur wird also die SUBCAT-Liste nach oben hin kürzer, wenn sich ein Komplement mit einer Projektion des Verbs verbindet (zur Verdeutlichung siehe auch die Abb. 5.11). Das „abgesättigte Komplement“ verschwindet jeweils von der Liste.

Um Zeichen miteinander zu kombinieren, werden in der HPSG sogenannte I(MMEDIATE)-D(OMINANCE)-SCHEMAS verwendet, welche Dominanzrelationen beschreiben. Jede Phrase muss nach dem ID-PRINCIPLE genau einem ID-SCHEMA genügen. Ein Schema besteht dabei aus einer Menge von Beschränkungen, welche bestimmen, welche Zeichen zu größeren Zeichen zusammengefasst werden können. Vergleiche dazu auch die Diskussion in Abschnitt 6.2.1.

Dies soll als Kurzeinführung genügen. In dieser Arbeit werden wir überdies immer auf zu Grunde liegende Annahmen eingehen, soweit sie für die Diskussion maßgeblich sind.

4.2 Interne Selektion

Subkategorisierung bzw. Selektion der Komplemente ist einer der zentralen Mechanismen in formalen Grammatikmodellen. Er dient einerseits dem Zweck der Spezifizierung von syntaktischen und semantischen Argumenten. Zudem wird die Relation zum Selektor angegeben, was normalerweise durch die Zuweisung thematischer Rollen (Thetarollen) geschieht. Andererseits werden die Eigenschaften des selektierten Elementes eingegrenzt (Wortkategorie

Ausdruck kommen sollten. Dies wird aber nicht überall vertreten und so findet man mitunter erhebliche Unterschiede zwischen den entworfenen Signaturen. Letztendlich wird man sich erst bei der Implementierung eines Grammatikfragments auf eine bestimmte Signatur in allen ihren Einzelheiten festlegen müssen.

etc.). Bei der internen Selektion, also der Selektion durch den Kopf der Phrase, werden beide Arten von Restriktion angewandt.

In der HPSG geschieht die Selektion über Valenzlisten im CATEGORY-Wert eines Zeichens, die für Wörter lexikalisch festgelegt werden (in dieser Arbeit durch den Lexikoneintrag). Auf der SUBCAT-Liste (oder deren Entsprechungen, s. o.) sind beispielsweise Subjekte von Verben und deren Komplemente, wie direktes oder indirektes Objekt, aufgelistet.

Selektion umfasst nicht das gesamte Zeichen: Pollard und Sag argumentieren (1987, S. 143f und 1994, S. 23f), dass Selektion ein lokales Phänomen ist. Alle Abhängigkeiten, die über die Subkategorisierung behandelt werden, Komplemente und ihre grammatikalischen Relationen, Wortkategorie, Kasus, Kongruenz, semantische Rollen und Selektionsrestriktionen, haben die Eigenschaft gemeinsam, lokal zu sein. Mit ihrem Lokalitätsprinzip legen sie fest, dass ein Kopf weder die interne Struktur noch die phonologische Form seiner Argumente selektieren kann. Empirisch lässt sich das leicht begründen. So gibt es beispielsweise kein Verb, das von seinem satzwertigen Komplement fordert, ein transitives und kein intransitives Verb als Kopf seiner VP zu haben. Hier würde eine nicht lokale Subkategorisierung vorliegen. Die Merkmalsgeometrie in Pollard/Sag (1994) lässt daher nur noch die Selektion von *synsem*-Objekten zu, ohne die Eigenschaften der internen Struktur (DAUGHTERS) und der phonologischen Form (PHON).

Deshalb müssen alle für die Selektion relevanten Merkmale innerhalb des SYNSEM-Wertes repräsentiert werden.

Krenn/Erbach (1994) leisteten einen beachtlichen Beitrag zur Analyse von Idiomen, FVGs, Kollokationen und Mehrwortlexemen in der HPSG. Sie schlagen eine lexikonbasierte Repräsentation von Idiomen auf Wortebene vor. Dabei setzen sie sich jedoch bewusst über das oben motivierte Lokalitätsprinzip hinweg und behandeln Idiome mittels totaler Selektion. So definieren sie die SUBCAT-Liste des Verbs *kommen* für den Phraseologismus *in Frage kommen* wie in Abb. 4.2 dargestellt. Es muss angemerkt werden, dass hier nicht die üblichen Abkürzungskonventionen gelten und NP und PP für jeweils das ganze Zeichen stehen.

$$\left[\text{SUBCAT} \left\langle \text{NP}[\text{nom}], \text{PP} \left[\text{PHON} \langle \text{in Frage} \rangle \right] \right\rangle \right]$$

Abbildung 4.2: SUBCAT-Liste von *kommen* bei Krenn/Erbach

Hier wird ein Komplement mit einer bestimmten phonologischen Form gefordert. So etwas ist im Allgemeinen nicht ratsam, da sich die Phonologie durch Kasuszuweisung (z. B. im Passiv) ändern kann.

Dieser Tatsache tragen Krenn und Erbach mit Hilfe der Lexem-Selektion Rechnung, da Lexem-Werte unabhängig von morphologischen Eigenschaften konstant bleiben. Nehmen wir als Beispiel die Wendung *jdm. den Garaus machen*. Die Selektionsabhängigkeiten sind in Abbildung 4.3 illustriert.⁷

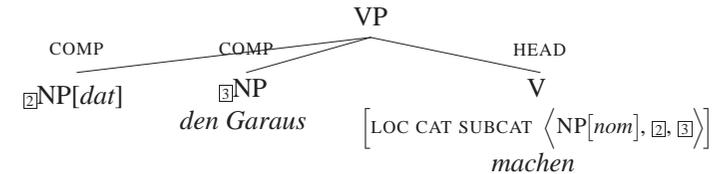


Abbildung 4.3: Interne Selektion in *jdm. den Garaus machen*

Der Lexikoneintrag von *machen* für diese Wendung wird in Abb. 4.4 skizziert, leicht optisch modifiziert nach Krenn/Erbach (1994, S. 382).

Aus dem Lexikoneintrag ist ersichtlich, dass hier auf die DTRS-Information der Akkusativ-NP zugegriffen wird. Es wird ein bestimmtes Lexem (*garaus*) selektiert, wozu die Autoren das Merkmal LEXEME im semantischen Index einführen.⁸ Sie verlangen außerdem, dass die Kopftochter ein lexikalisches Zeichen ist (in Abb. 4.4 nicht dargestellt), und schließen so eine Modifikation von *Garaus* aus.

Diese Herangehensweise scheint für *den Garaus machen* zu funktionieren. Sie ist jedoch aus mehreren Gründen problematisch.

Erstens, wie schon angesprochen, wird hier das Lokalitätsprinzip verletzt, was wir als Nachteil des Ansatzes werten.

Zweitens können unikale Elemente so nicht behandelt werden. Bei unikalenen Elementen muss man sicherstellen, dass sie nur in ihren spezifischen Kontexten auftreten. Dies kann man aber mit interner Subkategorisierung nicht erreichen. Die Subkategorisierung wird auf Grund der vielen Kombinationsmöglichkeiten der Sprache bewusst allgemein gehalten, um alle möglichen

⁷Die Abkürzungen V und NP im Baum beziehen sich auf die SYNSEM-Werte der entsprechenden Zeichen. Dies gilt für sämtliche Bäume in dieser Arbeit, außer dort, wo *sign*-Attribute explizit angegeben sind.

⁸Nebenbei bemerkt kann diese Art der Lexemselektion so nicht funktionieren. Weiter unten (Abschnitt 4.3) werden wir dieses HPSG-technische Problem lösen.

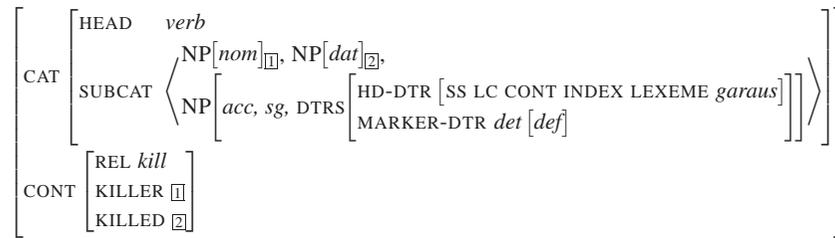


Abbildung 4.4: Skizze des SYNSEM-Wertes von *machen* in *jdm. den Garaus machen*

Kandidaten zu lizensieren. Das hat zur Folge, dass alle Informationen, die nicht unbedingt nötig sind, in den Beschreibungen unterspezifiziert bleiben. Ein unikales Element müsste als Zeichen im Lexikon vorliegen, um es subkategorisieren zu können. Man kann aber nicht ausschließen, dass ein unikales Element an anderer Stelle auftritt als dort, wo es explizit (durch Lexem-Selektion) subkategorisiert wurde.

Eine weitere Konsequenz dieser Vorgehensweise wäre, dass man für alle idiomatischen Versionen von Verben (wie *machen*) eigene Lexikoneinträge anlegen müsste. Dies ist an sich nicht weiter beunruhigend. Betrachtet man jedoch Präpositionen, die ein unikales Element subkategorisieren, werden Phrasen, in welchen eine Koordination zweier unikalener Elemente vorliegt, ungrammatisch.

(157) *aus Herzenslust* → *aus*₃ selegiert [... LEXEME *herzenslust*]

(158) *aus Leibeskräften* → *aus*₇ selegiert [... LEXEME *leibeskräfte*]

(159) *Auf dem Schulkonzert schmetterte das Mädchen ihr Lied aus?* [*Herzenslust und Leibeskräften*].

Es liegt hier eine Koordination zweier unikalener Elemente vor, die beide von *aus* selegiert werden. Welche Präposition soll man in (159) verwenden? Überdies ist es kontraintuitiv zu sagen, die intrinsische Eigenschaft (der Lexikoneintrag) der Präposition *aus* sei, dass sie *Daffke*, *Herzenslust*, *Leibeskräfte*, *Stegreif*, *freie Stücke* oder ein beliebiges nicht unikales Element subkategorisiert. Genauso kontraintuitiv ist es zu sagen, dass es für jedes der genannten Elemente ein eigenes *aus* gäbe.

Drittens kommt man mit diesem Ansatz bei Phraseologismen wie *auf dem absteigenden Ast sitzen* nicht mehr weiter, weil man zusätzlich zur Kopftochter noch deren Modifikator subkategorisieren müsste. Gerade, wenn außer diesem Adjunkt noch weitere Modifikatoren möglich sind,⁹ kommt man in Schwierigkeiten.

Ein letztes Problem stellt die Tatsache dar, dass man bei nicht unikalenen Elementen noch einen Mechanismus braucht, um die nicht idiomatische Bedeutung zu unterdrücken. Der Vorschlag von Krenn und Erbach, den Komplement-NPs leere Thetarollen zuzuweisen, funktioniert nur dann, wenn man die Semantik der NPs ignorieren kann, was sehr selten der Fall ist.

Aus den genannten Gründen genügt der Ansatz der internen Selektion von Krenn/Erbach (1994) nicht den Anforderungen einer korrekten Analyse unserer Daten.

4.3 Externe Selektion

Eine alternative Herangehensweise – Selektion in die „entgegengesetzte Richtung“ – wurde in Soehn (2003) entwickelt und in Soehn/Sailer (2003) weitergeführt.

Externe Selektion bedeutet, dass ein Komplement, also nicht der Kopf einer Phrase, den Kopf selegiert. Dies wurde für Adjunkte, Determinierer und Komplementierer bereits in der HPSG implementiert (vgl. Pollard/Sag, 1994). Die Tatsache, dass bestimmte Adjunkte nur bestimmte Köpfe modifizieren können und dass Komplementierer¹⁰ eine satzwertige Phrase selegieren, spricht dafür, dass Nichtköpfe auch auf Köpfe bestimmte Beschränkungen legen.

Dieser gut motivierte Selektionsmechanismus in der HPSG, bei dem Nichtkopftöchter die Kopftöchter auf einer Ebene selegieren, soll kurz erläutert werden. Über das Attribut SPEC[IFIED], in Pollard/Sag (1994) für *functional*-Wörter definiert, kann beispielsweise der Determinierer als Komplement in einer NP ein N' (eine Projektion des Kopfes) selegieren. In Pollard/Sag (1994) wird der SPEC-Mechanismus für Determinierer und auch für Komplementierer (*marker*) benutzt. Dabei selegiert der Komplementierer den Satz, welchen er markiert.

⁹Google, 23.10.03: [...] *die ihn des öfteren schon auf den sogenannten „absteigenden Ast“ setzten.* (http://www.retrospec.de/musik/2000_08_02_bob_dylan_tour.shtml)

¹⁰Komplementierer im Englischen werden von Pollard/Sag (1994, S. 44f) als Nichtköpfe analysiert, da die Standardanalyse aus der GB-Tradition nach Meinung der Autoren ebenso viele Fragen aufwirft wie sie beantwortet.

Außerdem gibt es für Adjunkte das MOD[IFIED]-Attribut (für *substantive*-Zeichen definiert), durch welchen der Kopf, also das modifizierte Element, selegiert wird. Zum Beispiel modifiziert *rotes* in *rotes Auto* das Nomen. Im MOD-Wert des Adjektives steht der SYNSEM-Wert des Nomens und hier hängt sich *rotes* nicht nur syntaktisch an *Auto* an, sondern liefert auch eine semantische Zusatzinformation für die ganze NP. Dabei ist zu beachten, dass ein Adjektiv wie *rot* nicht jedes beliebige Nomen modifizieren kann (**rote Zerstretheit*).

Wie ist diese Art externer Selektion beschaffen? Nichtköpfe selegieren Köpfe, wobei sie diesen verschiedene Beschränkungen auferlegen. Der Grad der Beschränkungen bewegt sich hier in einem Kontinuum. So stellen Adjunkte (MOD) an ihr modifiziertes Element einige semantische Anforderungen (s. o.), weiterhin werden Thetarollen vergeben, vgl. Rauh (1993, S. 113ff). Determinierer oder Quantoren (SPEC) vergeben keine Thetarollen und haben weniger Anforderungen an ihre Nomina, jedoch müssen bei Quantoren Nomina unterschieden werden, die eine Masse oder etwas Zählbares denotieren. Bei Markierern spielt die Semantik keine Rolle mehr.

Insgesamt ermöglicht die externe Selektion, die Kookurrenz von Konstituenten, also deren Kombinierbarkeit, zu beschränken.

Da SPEC für *functional* und MOD für *substantive* definiert ist, unterliegen sie einer komplementären Verteilung. Pollard und Sag stellen jedoch selbst fest, dass es sich bei SPEC und MOD im Grunde um denselben Mechanismus handelt:

To enable an adjunct to select its head, we introduce the head feature MODIFIED (MOD), which is analogous to the SPEC feature employed by markers and determiners for the same purpose. (Pollard/Sag, 1994, S. 55)

In Soehn (2003) wurden konsequenterweise beide Arten von externer Selektion zusammengefasst und an Stelle von SPEC und MOD ein Attribut XSEL (für *external selection*) gesetzt. Gleichzeitig wurde der Mechanismus so generalisiert, dass in jeder Phrase die Nichtkopftochter Aussagen über die Kopftochter machen kann. Diese Generalisierung geschah, ohne die interne Selektion über die Liste SUBCAT in irgendeiner Weise zu beeinträchtigen.

Die Abbildung 4.5 illustriert die vorgenommenen Änderungen gegenüber der Signatur von Pollard/Sag (1994). Der Wert des Attributs XSEL ist von der Sorte *synsem*. Da XSEL unter *head* definiert wurde, kann der Wert durch das KOPF-MERKMAL-PRINZIP perkolieren. Das Lokalisierungsprinzip wird hier also nicht verletzt, die Selektion – auch die externe – bleibt ein lokales Phänomen.

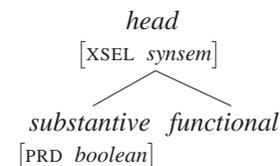


Abbildung 4.5: XSEL: Änderungen in der Signatur

Ferner wurde das nun überflüssige SPEC-PRINZIP ersetzt durch das PRINZIP DER EXTERNEN SELEKTION (PXS), das folgendermaßen lautet:

(160) PRINZIP DER EXTERNEN SELEKTION

(Version aus Soehn/Sailer, 2003):

Der SYNSEM-Wert eines Wortes ist identisch mit dem XSEL-Wert eines jeden Elementes auf seiner ARG-ST-Liste. In einer *head-adjunct-phrase* und in einer *head-marker-phrase* ist der SYNSEM-Wert der Kopftochter identisch mit dem XSEL-Wert der Nichtkopftochter.

Das Prinzip wurde bewusst nicht so formuliert, dass eine Selektionsbeziehung zwischen den Töchtern bei allen Kopfstrukturen besteht, so dass die Identität zwischen dem XSEL-Wert des Nichtkopfes und dem SYNSEM-Wert des Kopfes gefordert würde. Bei Anhebungsstrukturen kann ein *synsem*-Objekt auf mehreren SUBCAT-Listen erscheinen, während es nur auf einer ARG-ST-Liste steht.¹¹ Kopf-Füller-Strukturen sind vom PXS ausgenommen, da keine Selektionsbeziehung zwischen dem Kopf und dem Füller besteht.

Von diesen neu eingeführten Anpassungen blieb der schon bestehende Mechanismus der externen Selektion unberührt. Dies gilt vor allem für Adjunkte und Marker, die ihre Köpfe über XSEL selegieren können, was durch den zweiten Teil in (160) sichergestellt wird.

Es ist noch anzumerken, dass die Zusammenführung der beiden Selektionsmechanismen über SPEC und MOD die Merkmalsgeometrie der HPSG ökonomischer und einheitlicher macht, da die Komplementarität von SPEC und MOD aufgehoben wird. Durch die Sorten *substantive* und *functional* kann immer noch ausreichend differenziert werden.

¹¹Wir folgen Meurers (2000) und nehmen an, dass Anhebung nur über die COMPS-Liste erfolgt.

4.3.1 Selektion von Lexemen

Bevor wir uns die konkreten Daten ansehen, für die der Ansatz entwickelt wurde, soll noch erläutert werden, wie man bestimmte Lexeme selektieren kann. Wie bereits erwähnt, muss der Ansatz von Krenn/Erbach (1994) modifiziert werden, was in Soehn (2003) bereits vorgenommen wurde.

Bisher wurde i. A. davon ausgegangen, dass es eine Selektion von Lexemen gar nicht geben müsste, da alle Daten in dieser Richtung als Kollokationen behandelt werden könnten. Nun ist es aber recht plausibel zu sagen, dass beispielsweise *rümpfen* in *die Nase rümpfen* das Lexem *Nase* selektiert und kein anderes Lexem sonst. Ähnlich verhält es sich bei dem Ausdruck *die Stirn/Augenbrauen runzeln*.¹² Bei der Auswahl der Hilfsverben für die Perfektbildung im Deutschen (über AUXF) wird de facto nichts anderes als ein Lexem selektiert, siehe dazu Heinz/Matiasek (1994, S. 222). Auch die Perfektbildung im Französischen benötigt so ein Attribut (VAUX), siehe Abeillé/Godard (1996). Es gibt noch einige andere Fälle, in denen man gern auf einen Mechanismus wie Lexemselektion zurückgreifen würde, z.B. für Komplemente von Plesionymen¹³ auf syntaktischer oder morphologischer Ebene. Siehe dazu folgende Beispiele:

- (161) a) eine Entscheidung fällen/*fassen
 b) einen Beschluss fassen/*fällen

Nach Krenn/Erbach muss ein entsprechendes Merkmal in der HPSG-Signatur zwar unter SYNSEM stehen, damit es für die Subkategorisierung zugänglich ist; es kann jedoch kein HEAD-Attribut sein. Die Autoren argumentieren, dass Pronomen den gleichen LEXEME-Wert wie ihr Antezedens haben und ihrerseits keinen eigenen Wert mitbringen. Wenn ein bestimmtes Lexem an einer Stelle gefordert wird, kann an dieser Stelle ja auch ein Pronomen stehen, welches darauf referiert (*Er rümpfte die Nase. - Er rümpfte sie.*). Pronomen und ihr Antezedens teilen aber keine Kopfmerkmale miteinander, deshalb schreiben sie das LEXEME-Merkmal in den INDEX. Man muss also neben der Forderung, dass LEXEME unter SYNSEM steht, auch Fälle von Pronominalisierung adäquat behandeln können.

Eine dritte Forderung ist, dass der Wert von LEXEME perkolieren kann, da die Projektion den gleichen LEXEME-Wert behalten soll, wie das Wort selbst

¹²Fleischer (1997) nennt die beiden genannten Beispiele „singuläre Kombinationen“.

¹³Plesionyme sind zueinander partiell synonym in Bedeutung und Distribution. Wir verwenden den Ausdruck Plesionym an Stelle des kontroversen Synonym-Begriffs (vgl. Bußmann, 2002, S. 673f).

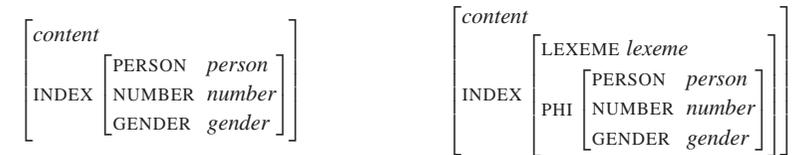


Abbildung 4.6: XSEL: Herkömmlicher und geänderter INDEX

(*Er rümpfte seine verschnupfte Nase*). Dies ist jedoch bei der Analyse von Krenn/Erbach nicht vorgesehen.

Die Vorgehensweise von Krenn/Erbach lässt noch zwei weitere Probleme erkennen. Erstens haben in der Signatur von Pollard/Sag (1994) nur *nominal-objects* einen INDEX. Ein zweites Problem ergibt sich im Zusammenhang mit den INDEX-Werten von Adjunkten. Diese übernehmen den INDEX-Wert ihres modifizierten Nomens. In Richter/Sailer (2003b) taucht das Beispiel *ein sattsam bekanntes Problem* auf. Das unikale Element *sattsam* kann nur *bekannt* modifizieren. Wenn nun *sattsam* das Lexem *bekannt*- selektieren wollte und LEXEME unter INDEX stünde, so wäre dies nicht möglich. Die INDEX-Werte von *bekanntes* und *Problem* sind ja identisch. Somit ist hier noch ein eindeutiger Hinweis darauf gegeben, dass LEXEME nicht ohne weitere Änderungen unter INDEX stehen kann.

Ein Ausweg ist, die Signatur anzupassen, um für alle Wortarten einen Index zu spezifizieren. Dabei wird das LEXEME-Attribut von den klassischen INDEX-Attributen PERSON, NUMBER und GENDER getrennt. Somit würden die genannten Probleme umgangen. Die Sorte *content* besitzt also – neben den Attributen der jeweiligen Semantiktheorie – das Attribut INDEX, das in Abb. 4.6 dargestellt wird. Für *nominal-objects* gibt es mit dieser Änderung keine weiteren Konsequenzen. Für Verben sind die PHI-Werte im Lexikoneintrag unterspezifiziert. Im LE eines jeden Adjektivs muss vermerkt werden, dass die PHI-Werte mit denen des modifizierten Nomens übereinstimmen.

Das SEMANTIK-PRINZIP wird so abgeändert, dass die INDEX-Werte immer entlang des syntaktischen Kopfes vererbt werden und der Rest (wie NUCLEUS) entlang des semantischen Kopfes. Als Werte von LEXEME seien in die Signatur neue Sorten für Lexeme eingeführt, wie z.B. *nase*, *grob*-, *komm*-, *zieh*-, etc. Für kompliziertere Fälle wie Komposita sei das Lexem des morphologischen Kopfes der LEXEME-Wert; für lexikalisierte Komposita wie *Fahrstuhl* gebe es ein eigenes Lexem. Diese Fälle spielen für die Diskussion jedoch keine Rolle. Im Lexikoneintrag kann also für jedes Wort sein LEXEME-Merkmal spezifiziert werden. Der LEXEME-Wert einer Phrase ist dann der des lexikalischen

Kopfes. Mit den beschriebenen Erweiterungen ist es nun möglich, Lexeme zu subkategorisieren. Für den Fall *rümpfen* ist also ganz einfach im LE zu vermerken, dass das nominale Komplement den LEXEME-Wert *nase* hat. Mit dem in diesem Abschnitt vorgestellten Mechanismus zur Lexemeselektion haben wir die Voraussetzungen für eine konkretere Betrachtung des Ansatzes der externen Selektion geschaffen.

4.3.2 Analyse mit XSEL

Diese Herangehensweise wurde ursprünglich für unikale Elemente entwickelt, die als Komplement einer Präposition auftreten, wie folgende Beispiele illustrieren:

(162) *Das „PC-Kummerbuch“ ist **auf** Anhieb auf grosses Interesse gestossen.*¹⁴

(163) *Die Dame sah mich an, als wäre ich nicht ganz **bei** Trost.*¹⁵

(164) *...das Mädchen saß daneben und weinte **zum** Steinerweichen.*¹⁶

Die Analyse wurde noch erweitert, um Fälle mit einzubeziehen, bei denen solche PPs mit bestimmten Verben kombiniert sind, wie bei:

(165) ***zu** Potte kommen* oder

(166) ***in** Mitleidenschaft ziehen*

Mit Hilfe des Attributs XSEL und der Lexemeselektion kann ein unikales Element die Präposition spezifizieren, mit welcher es einhergehen muss. In Abbildung 4.7 wird der LE von *Anhieb* mit dem entsprechenden XSEL-Wert skizziert. Die Identität des CONTENT-Wertes der Komplement-NP und des lexikalischen Elementes wird gefordert, um eine Modifikation von *Anhieb* zu verhindern. Abbildung 4.8 zeigt die Skizze des LE für *Brustton* in der Wendung *im Brustton der Überzeugung*, wo keine Beschränkung des CONTENT-Wertes gefordert wird. Das Attribut DEF wurde von Winhart (2002) übernommen, um die Definitheit der Präposition (*im* statt *in*) zu erzwingen. Ihr Vorschlag sei hier als eine konkrete Möglichkeit der Analyse herangezogen, um eine Unterscheidung zwischen Präpositionen mit und ohne inkorporiertem

¹⁴Züricher Tagesanzeiger, 15.11.1999, S. 67, Ressort: Computer;

¹⁵Mannheimer Morgen, 28.07.2001;

¹⁶Mannheimer Morgen, 29.06.1991, Unterhaltung;

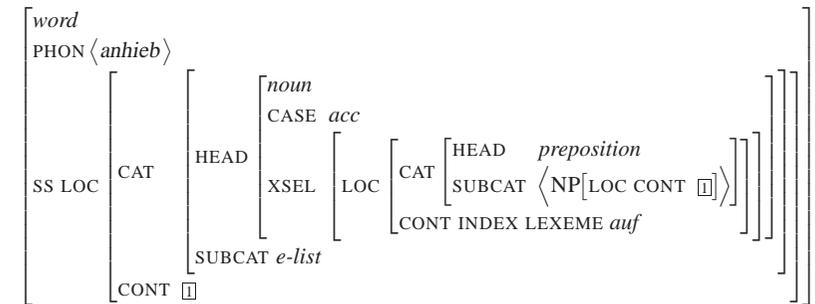


Abbildung 4.7: Skizze des LEs von *Anhieb* mit XSEL

definiten Artikel vorzunehmen. In Abbildung 4.9 wird die Interaktion des LEs mit dem PRINZIP DER EXTERNEN SELEKTION dargestellt, wo die entsprechenden Identitäten hergestellt werden.

Wie im LE von *Brustton* festgelegt, steht im XSEL-Wert die Präposition *im*. Weiterhin subkategorisiert das Nomen eine NP im Genitiv.¹⁷ Der XSEL-Wert, es handelt sich ja bei XSEL um ein Kopf-Merkmal, ist identisch mit dem XSEL-Wert der Projektion, also der NP *Brustton der Überzeugung*. Dies wird durch das KOPF-MERKMAL-PRINZIP erzwungen; die Komplement-NP der Präposition besitzt denselben XSEL-Wert wie die Kopftochter *Brustton*.

Das PRINZIP DER EXTERNEN SELEKTION erzwingt nun, dass der XSEL-Wert der Komplement-NP identisch sein muss mit dem SYNSEM-Wert der Kopftochter. Dieser SYNSEM-Wert ist der der Präposition *im*, wie im LE spezifiziert. Deshalb treten in unserem Beispiel die geforderten Identitäten auf, die Phrase ist grammatisch korrekt.

Kombinationen von PPs mit unikalen Elementen und festen Verben werden so analysiert, dass der XSEL-Wert der unikalen Komponente noch einen eingebetteten XSEL-Wert besitzt. Siehe hierzu die Abbildung 4.10, wo der Lexikoneintrag von *Potte* skizziert wird. Darin wird gefordert, dass die Präposition einen eigenen Wert des XSEL-Attributs besitzt. Da die PP (in diesem Beispiel *zu Potte*) auf der ARG-ST-Liste von *kommen* steht, wird durch das PXS erreicht, dass die PP mit dem „richtigen“ Verb zusammen auftritt. Die Distributionsidiosynkrasie von *Potte* wird also gänzlich in dessen LE spezifiziert. Die Baumstruktur in Abb. 4.11 soll dies noch einmal illustrieren.

¹⁷Die Tatsache, dass hier die eingebettete NP durch das PRINZIP DER EXTERNEN SELEKTION einen XSEL-Wert bekommt, der dem SYNSEM-Wert von *Brustton* entspricht, ist hier nicht weiter von Bedeutung.

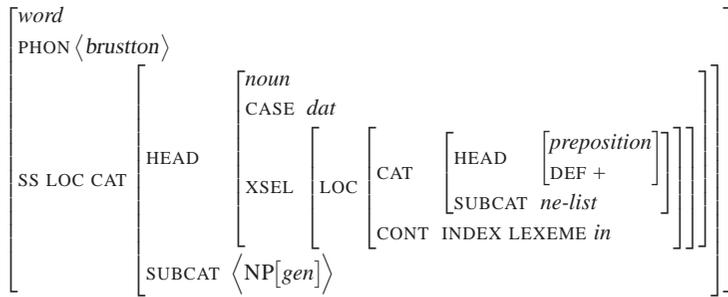


Abbildung 4.8: Skizze des LEs von *Brustton* mit XSEL

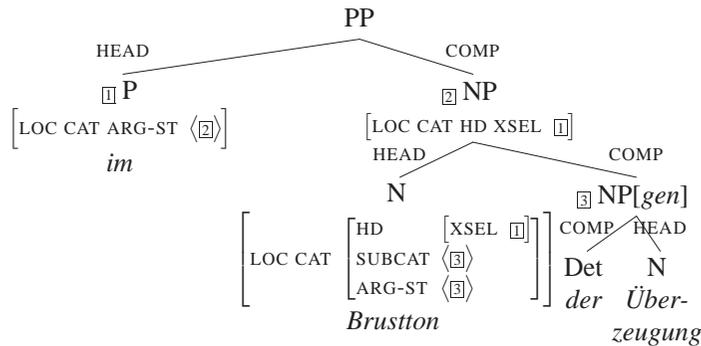


Abbildung 4.9: Beispiel für externe Selektion mit XSEL

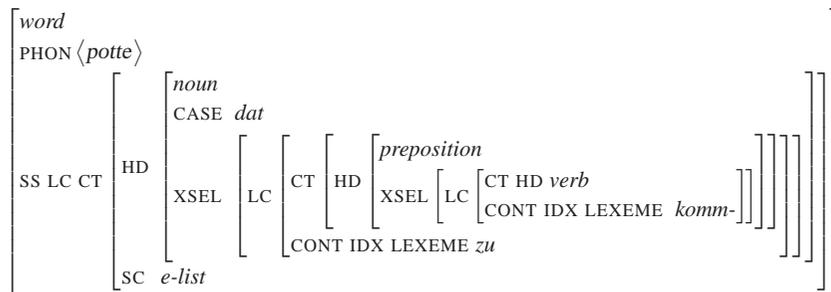


Abbildung 4.10: Skizze des LEs von *Potte* mit XSEL

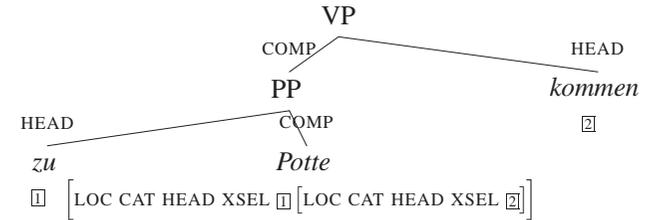


Abbildung 4.11: *zu Potte kommen* als PP mit festem Verb

Dabei macht das PXS auch korrekte Vorhersagen für Anhebungsstrukturen, so wie sie für deutsche Verbkomplexe angenommen werden (vgl. z.B. Hinrichs/Nakazawa (1989) und Abschnitt 5.3.1.2). In (167) wird die PP als Komplement des Modalverbs *sollte* realisiert.

(167) *dass Peter zu Potte kommen sollte*

Sie taucht jedoch immer noch als Element der ARG-ST-Liste von *kommen* auf, dessen LEXEME-Wert *komm-* lautet, so wie im LE von *Potte* gefordert. Mit dieser Herangehensweise kann folgenden Möglichkeiten, in denen unikale Elemente auftreten, Rechnung getragen werden:

(168) *jdn. über den Löffel balbieren* → interne Selektion über LEXEME

(169) *jdm. einen Bären dienst erweisen* → externe Selektion über XSEL und LEXEME

(170) *futtern wie ein Scheunendrescher* → externe Selektion über Schachtelung von XSEL-Werten und LEXEME, wie bei (165)

Wie in Soehn/Sailer (2003) ausgeführt wird, ist der Ansatz sowohl lokal als auch nicht lokal. Eine Beschränkung auf Lokalität besteht darin, dass ein Nichtkopf nur auf seinen Kopf und auf ein Zeichen, das seinen Kopf selektiert (wie in *zu Potte kommen*), zugreifen kann. Dagegen kann man beispielsweise keine weiteren Komplemente eines Verbs selektieren. So ist folgende Konstellation nicht möglich, in der die NP vom unikalen Element selektiert werden würde:

(171) $[_{VP} V \mathbf{NP} [_{PP} P \underline{\text{unikales Element}}]]$

Nicht lokale Effekte der Lexemselektion ergeben sich unmittelbar daraus, dass LEXEME Teil des Indexes ist. Die Behandlung erfolgt über die Bindungstheorie der HPSG, welche auch die Auswahl des richtigen Antezedens von Pronomen und Anaphora regelt. Sailers Beispiel (persönliche Mitteilung) illustriert eine Abhängigkeit über eine Satzgrenze hinaus:

(172) *Das war ein riesiger **Schrecken**_i für mich. Und ausgerechnet du hast **ihn**_i mir ingejagt.*

Das Pronomen trägt hier den gleichen LEXEME-Wert wie sein Antezedens.

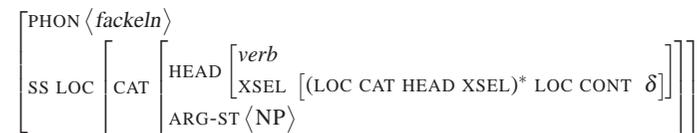
Sailer geht in Soehn/Sailer (2003) noch weiter und legt den XSEL-Wert für Zeichen fest, die entlang der Kopfprojektionslinie eines Wurzelzeichens auftreten, die also keine Nicht-Kopftöchter sind. Zum Begriff des Wurzelzeichens, der unabhängig von dieser Analyse in der Grammatik benötigt wird, siehe Richter (1997). Nach Riehemann (2001, S. 189f) wird mit *root-sign* eine Untersorte von *sign* definiert. Sailer (ebd.) führt darüber folgende Beschränkung ein:

(173) *root-sign* → [SYNSEM □ [LOC CAT HEAD [XSEL □]]]

Als Effekt von (173) ist der CONTENT-Wert einer Äußerung der XSEL-Spezifikation eines jeden Wortes innerhalb der Äußerung zugänglich, indem der XSEL-Wert des *synsem*-Objekts im XSEL-Wert rekursiv aufgesucht wird. Mit dieser erweiternden Spezifikation kann man negative Polaritätselemente wie (*nicht lange*) *fackeln* analysieren. Der LE von *fackeln* wird in Abb. 4.12 wiedergegeben.¹⁸ Allerdings wollen wir nun nicht genauer darauf eingehen und auf den Artikel von Soehn/Sailer (2003) verweisen. Es sollte nur klar werden, dass der Ansatz der externen Selektion erfolgreich zur Beschreibung unikalere Elemente herangezogen werden kann und dabei korrekte empirische Vorhersagen, wie die der Lokalität, macht. Darüber hinaus kann man damit Daten analysieren, die Richter/Sailer (2003b) bzw. Sailer/Richter (2002a) als Motivation für ihren Kollokations-Ansatz heranziehen. Zu diesem Ansatz siehe Abschnitt 4.5.

Eine Frage, die bei dieser Herangehensweise diskutiert werden muss, ist die Rolle der Adjunkte, also der modifizierenden Attribute wie Adjektive und Relativsätze. Man möchte diese ja nicht selegieren, sondern Adjunkte selegieren selbst den Modifikanten. Bei „*rotes Buch*“ z. B. selegiert nicht das Nomen sein

¹⁸Die Benutzung des Stern-Operators ist eine intuitive Darstellung und nicht Standard in der HPSG. Nach der Formalisierung in Richter (2004a) kann man die Idee dieser Notation jedoch formal ausdrücken.



* ist der Stern-Operator von Kleene: Der angegebene Pfad kann sich wiederholen, wobei δ beliebig tief eingebettet sein kann.

δ beschreibe ein *content*-Objekt gdw. es eine Negation, einen durativen Operator und den semantischen Beitrag von *fackeln* enthält, wobei die Negation und der durative Operator Skopus über diesen haben.

Abbildung 4.12: Skizze des LEs von *fackeln* mit XSEL

Adjektiv, sondern umgekehrt. In der HPSG geschah dies bisher über das MOD-Attribut, jetzt über XSEL.

Wenn bei einem Phraseologismus jedoch ein Adjunkt zwingender Bestandteil ist,¹⁹ spricht man nicht von Adjunkt, sondern von Komplement, und erweitert diesen Begriff dadurch. Praktisch kann man dies umsetzen, indem man diese Elemente über den normalen Selektionsmechanismus subkategorisiert oder indem man phrasale Lexikoneinträge vorsieht. Diese Frage wird in den nächsten Kapiteln noch einmal aufgegriffen.

Ein weiterer Kritikpunkt betrifft den Mechanismus der Lexemselektion. Lexeme sind bekanntlich Einheiten mit syntaktischen, semantischen und phonologischen Eigenschaften. Es ist daher wenig intuitiv zu sagen, dass das LEXEME-Attribut unter *content* angesiedelt sei. Wir werden uns dem Ansatz der Lexemselektion noch ausführlich in Abschnitt 5.1.2.1 annehmen.

Ein drittes Problem ist die Tatsache, dass man die XSEL-Werte schachtelt („der XSEL-Wert vom XSEL-Wert“), so wie beim LE von *Potte* oder *fackeln*, wo sogar ein Stern-Operator verwendet wird. Mit dieser Vorgehensweise wird der Ansatz zu stark, weil eine Restriktion der Einbettungstiefe nicht implementiert ist. Um die Schachtelungstiefe zu begrenzen, könnte man Beschränkungen auf Lexikoneinträge annehmen, wobei nur solche LEs zugelassen wären, bei denen maximal zwei ineinander geschachtelte XSEL-Werte spezifiziert werden. Dies wirkt jedoch sehr künstlich und wäre lediglich eine theorie-interne, aber keine empirische Art von Beschränkung. Allerdings muss man fragen, ob so etwas notwendig ist, wenn XSEL-Werte sowieso nur im Lexikon spezifiziert

¹⁹z. B. *auf die schiefe Bahn* oder *zu guter Letzt*

werden und dadurch *per se* schon beschränkt sind. Überdies liegt es nicht besonders nahe, dass beispielsweise der Präposition von *Potte* eine eigene Distributionsidiosynkrasie (als Komplement von *kommen*) „auferlegt“ wird.

Weiterhin ist problematisch, dass diese Herangehensweise bisher nur für dekomponierbare Phraseologismen herangezogen wurde. Einzelne idiomatiche Elemente mit ihrer ihnen eigenen Bedeutung wurden durch XSEL in ihrer Distribution beschränkt. Dort, wo Nichtdekomponierbarkeit vorliegt, wo also nicht klar ist, welche Bedeutung einzelne Bestandteile haben und ob man für diese überhaupt eigene Lexikoneinträge vorsehen muss, kommt man beim XSEL-Ansatz in Erklärungsnot. Dies wird vor allem daran deutlich, dass man für den XSEL-Ansatz unbedingt eigene Lexikoneinträge mit dem jeweils spezifizierten XSEL-Wert braucht. Wie wir aber weiter unten (z. B. in Abschnitt 6.2) zeigen werden, bieten sich mehrere LEs für nicht dekomponierbare Idiome (für je einen Bestandteil) kaum an.

Zuletzt ist zu bemerken, dass der ganze Mechanismus der externen Selektion (jedenfalls dort, wo die Semantik eine Rolle spielt), nur unter der Annahme funktioniert, dass sämtliche semantische Informationen lokal (also unter CONTENT) zugänglich sind. Eine Alternative zu dieser Sicht wird in Abschnitt 5.3.2.1 vorgestellt.

Soweit der Ansatz der externen Selektion und dessen wichtigste Probleme. In welcher Form man an diese Herangehensweise anknüpfen könnte, wird im entsprechenden Abschnitt 5.1.2 im folgenden Kapitel diskutiert.

4.4 Konstruktion

Bisher haben wir zwei Ansätze, die Phraseologismen auf Wortebene beschreiben, angesehen. Susanne Z. Riehemann (2001) geht in ihrem Kapitel 4 auf weitere *word-level* Ansätze ein. Außer der Analyse von Krenn und Erbach untersucht sie die Möglichkeit einer Mehrwort-Lexem-Analyse, der Subkategorisierung der phonologischen Form, einer teilweisen Selektion der Syntax und der Semantik und geht auf Herangehensweisen in der LFG und der GPSG ein. Sie führt folgende Gründe an, warum man ihrer Meinung nach eine phrasale Herangehensweise benötigt:

1. Die wörtlichen Bedeutungen der Teile eines Idioms spielen keine Rolle in der Gesamtbedeutung.

2. Idiomatiche Wörter sind nicht frei, d. h. sie kommen in ihrer idiomatiche Bedeutung nicht außerhalb des Idioms vor. Dabei führt sie auch Beispiele von unikalen Elementen an.²⁰
3. Manche Idiome sind syntaktisch nicht wohlgeformt (*close up shop, by and large*).
4. Eingeschränkte Flexibilität einiger Phraseologismen
5. Häufigeres Vorkommen und psycholinguistische Evidenz einer kanonischen Form
6. Variation von Phraseologismen und Ähnlichkeiten in den “*idiom families*”
7. Der Lokus der Abbildung von der wörtlichen auf die idiomatiche Bedeutung liegt auf phrasaler Ebene.
8. Bei nicht dekomponierbaren Phraseologismen liegt der Lokus der Semantik auf phrasaler Ebene.
9. Idiome sind mehr als bloße Kopf-Komplement-Strukturen, sie beinhalten auch Adjunkte oder ihnen fehlt ein festes bzw. internes Verb.
10. Psycholinguistische Evidenz, z. B. die schnellere Verarbeitung von Idioms gegenüber frei formulierten Phrasen²¹

Wie Riehemann (2001, S. 181) selbst einräumt, sind einige dieser Punkte nicht für jedermann einsichtig und betreffen nicht die generative Kapazität der Grammatik. Wir werden außerdem noch zeigen, dass man auch mit einem Ansatz auf Wortebene viele Punkte genauso gut berücksichtigen kann.

Sie schlägt sich aber eindeutig auf die Seite der Vertreter phrasaler Ansätze, d. h. sie überträgt Ideen aus der Konstruktionsgrammatik in die HPSG. Nachdem sie verschiedene andere Ansätze in diesem Paradigma untersucht hat (darunter die Fixierung der Phonologie, der Syntax und der Semantik), entwickelt sie ihre eigene Analyse, die im Folgenden vorgestellt werden soll.

Es sei zunächst einmal der Begriff der Konstruktion näher definiert.

²⁰Dabei wird jedoch nicht klar, wie sie diesen mit ihrem Ansatz letztendlich gerecht werden kann.

²¹Siehe jedoch Dobrovol'skij (1997, S. 14) für gegenteilige experimentelle Ergebnisse.

Der zentrale Gesichtspunkt für unsere Diskussion von Riehemanns Ansatz ist die Art, wie die idiomatischen Wörter behandelt werden. Wie bereits erläutert, gehen die Bestandteile von Redewendungen bei Riehemann auf die jeweiligen, nicht idiomatischen Gegenstücke zurück. Für die Anwendung auf unikale Elemente bedeutet das aber, dass es tatsächlich Lexikoneinträge für *Bären dienst* oder *balbieren* geben müsste. Es gibt aber von diesen Wörtern keine „nicht idiomatische“ Lesart und es kann mit unseren herkömmlichen Mitteln nichts über deren Distribution gesagt werden. Nimmt man solche Lexikoneinträge im Gegensatz dazu nicht an, scheitert Riehemanns Idee der *skeptical default unification*, da es ja keine Wörter gibt, auf die ihre *i-words* in Konstruktionen mit unikalen Elementen zurückzuführen sind.

Eine ausführliche Kritik an Riehemanns Ansatz findet sich in der Dissertation von Sailer (2003). Er geht zuerst auf den grundlegenden Unterschied zwischen zwei Strömungen innerhalb der HPSG ein. Riehemann hängt der Denkrichtung an, die Informationen über Wörter und Phrasen in der Signatur spezifiziert. Bei Riehemanns Ansatz, den er *constructional HPSG* nennt, gibt es eine große Hierarchie von *signs* mit sehr vielen Unterkategorien, die ihrerseits zum Teil mehrere Obertypen haben. Es wird eine Sorte von *sign* für jede Art von Konstruktion bzw. Satz spezifiziert. So gibt es nicht nur Sorten für normale syntaktische Kombinationsmöglichkeiten wie *head-subject-phrase*, *head-filler-structure* oder *head-specifier-structure*, sondern es wird auch für jede Redewendung eine eigene Sorte wie die *spill_beans_idiom_phrase* angenommen. Wenn nun Redewendungen in einer Kopf-Füller-Struktur oder einer Kopf-Spezifikator-Struktur auftreten können, muss man auch gemeinsame Untertypen dafür in der Signatur vorsehen. Damit wird die Signatur derart aufgebläht, dass man sich fragen muss, ob hier wirklich eine Generalisierung vorgenommen wird oder ob es nicht zu einer Duplikation von Informationen kommt, die auch schon in den Beschreibungen der Satzteile vorhanden sind. Außerdem ist es nach unserer Ansicht nicht sinnvoll, linguistische Objekte nur dadurch zu klassifizieren, ob irgendein Bestandteil eine idiomatische Wendung enthält oder nicht. Die Sortenhierarchie bildet jedoch einen wichtigen Teil der Analyse von Riehemann.

As a consequence, we can never be given the full sort hierarchy of a language, as every new word will potentially lead to the creation of new sorts. [...] nearly every adjustment of the theory will lead to a modification in the sort hierarchy as well. (Sailer, 2003, S. 307)

Das heißt, für jedes neue Wort oder jeden neuen Ausdruck in einer Sprache könnten neue Sorten eingeführt werden. Die Signatur würde sich also ändern und wir könnten sie nie vollständig angeben.

Die andere Richtung in der HPSG-Forschung, die auch Sailer einschlägt, geht dahin, dass die Signatur fest und möglichst klein bleibt, um so wenig ontologische Restriktionen wie möglich zu fordern. Man passt hier die Grammatiktheorie nur in der Art an die Sprachgegebenheiten an, dass Prinzipien der Theorie und das Lexikon verändert werden, während die Signatur intakt bleibt. Diese Richtung wird auch in dieser Arbeit verfolgt. Empirische Belege, die für eine der beiden Strömungen sprechen würden, gibt es bisher allerdings nicht.

Weiterhin kritisiert Sailer, dass die *WORDS*-Liste bei Riehemann, mit welcher man auf alle Wörter in einer Phrase direkt zugreifen kann, in keinem anderen Bereich der Grammatik benötigt wird und zeigt im Besonderen, dass das Merkmal nicht bei der Modellierung der Wortstellung Verwendung finden kann. Somit sieht er keine ausreichende Motivation für die Einführung des Merkmals (Sailer, 2003, S. 307–309).

Sailer erwähnt Unterschiede zwischen Riehemanns semantischer Repräsentation, *MRS*, und der von ihm verwendeten semantischen Darstellung und wirft das Problem nach einer exakten Formalisierung der \bar{n} -Operation auf. In seiner Fußnote 42 (S. 316) fragt Sailer nach der Motivation für Riehemanns neue Sorten, Attribute und Beschränkungen, die das Problem der Übergeneralisierung lösen sollen. Damit ist gemeint, dass in Riehemann (1997) nicht verhindert werden kann, dass *i-words* außerhalb ihrer idiomatischen Wendung auftreten könnten. Dazu erwähnt er eine Ungenauigkeit, nämlich welcher Knoten in der Baumstruktur einer idiomatischen Wendung denn nun die tatsächliche *...i_phrase* repräsentiert. Zuletzt weist er auf Riehemanns eigenes Zugeständnis hin, dass ihr Ansatz bestimmte Fälle von Pronominalisierung²³ nicht abdecken kann.

Er kommt zu dem Schluss, dass Riehemanns Ansatz keine akzeptable Idiom-Analyse innerhalb der HPSG darstellt:

This shows that the constructional approach is genuinely problematic for an account of IEs [idiomatic expressions] in HPSG. (Sailer, 2003, S. 316)

²³So etwas wie: *Eventually she spilled all the beans. But it took her a few days to spill them all.* Hier tritt im zweiten Satz das idiomatische *spill* ohne das idiomatische *beans* auf. Siehe auch Abschnitt 6.1.1.

Sailer bleibt bei dieser Kritik jedoch nicht stehen und stellt eine eigene Analyse von idiomatischen Wendungen vor, die wir uns im folgenden Abschnitt ansehen wollen.

4.5 Kollokation

Sailers Ansatz, der wieder auf der Wortebene angesiedelt ist, geht von der Vorstellung aus, dass sich bei idiomatischen Wendungen Wörter zu Kollokationen verbinden. Nach van der Wouden (1992), der in Richter/Sailer (1999a) zitiert wird, sind Kollokationen idiosynkratische Beschränkungen des kombinatorischen Potenzials lexikalischer Elemente. Das Vorkommen eines lexikalischen Elementes kann das Auftreten eines anderen verlangen oder ermöglichen. Danach besitzt also jedes Wort bestimmte Distributionseigenschaften, die es zu erfassen gilt. Sailer betrachtet das Phänomen der Kollokationen nur unter diesem Gesichtspunkt und grenzt sich damit bewusst von dem theoretisch vorbelasteten allgemeinen Kollokationsbegriff ab.

Die Analyse von Sailer findet sich unter anderem in folgenden Arbeiten: Sailer/Richter (2002a), Sailer/Richter (2002b), Richter/Sailer (2003b) und, wie schon genannt, in seiner Dissertation (2003).

Der Ansatz basiert auf der Unterscheidung zwischen intern regelmäßigen (dekomponierbaren) und unregelmäßigen (nicht dekomponierbaren) Wendungen. Für die unregelmäßigen Idiome wie *“kick the bucket”* oder *“trip the light fantastic”* nimmt Sailer phrasale Lexikoneinträge an. Was die regelmäßigen Wendungen betrifft, entwickelt er einen Mechanismus, der es erlaubt, für jedes Wort exakt zu spezifizieren, in welchen linguistischen Strukturen es auftreten kann. Aus dieser kurzen Beschreibung wird bereits klar, dass hier Elemente von beiden Ebenen, der phrasalen und der Wortebene, Verwendung finden und Sailer somit einen Mittelweg zu finden sucht.

Eine Schlüsselrolle in der Herangehensweise spielt das neu eingeführte Attribut COLL, was als Abkürzung für *Context Of Lexical Licensing* steht. Im Wert dieses Merkmals wird festgelegt, in welcher Umgebung ein Wort oder eine Phrase auftreten kann. Im Folgenden wollen wir eine Darstellung der HPSG-technischen Seite des Ansatzes liefern.

COLL wird für alle lexikalischen Zeichen eingeführt, wobei der Wert des Attributs eine Liste von Zeichen ist. Der Wert von COLL kann als Indikator dienen, mit welcher Art Zeichen man es zu tun hat: Alle dekomponierbaren Phrasen

sowie abgeleitete Wörter²⁴ haben den Wert *elist*, während alle idiomatischen Phrasen und nicht abgeleiteten Wörter den Wert *nelist* haben. Das sogenannte LEXIKON-PRINZIP garantiert, dass alle Zeichen mit [COLL *nelist*] im Lexikon aufgeführt werden.

Des Weiteren wird das COLL-PRINZIP (Abb. 4.14) formuliert, das sicherstellt, dass der Wert von COLL genau die Äußerung ist, in welcher das betreffende Wort auftritt. Formaler ausgedrückt heißt das: Für jedes Zeichen α und für jedes Zeichen β , das von α dominiert wird, dominiert jedes Element γ der COLL-Liste von β das Zeichen α .

Wenn ein Zeichen a ein Zeichen b dominiert, das einen nicht leeren COLL-Wert besitzt, dann ist das Element im COLL-Wert von b ein Zeichen c , so dass a von c dominiert wird.

$$\forall \alpha \left(\alpha[\text{sign}] \Rightarrow \left(\begin{array}{c} \forall \beta \forall \gamma \\ \left(\begin{array}{c} \text{dominate}(\alpha, \beta) \\ \wedge \beta[\text{COLL} \langle \gamma \rangle] \end{array} \right) \Rightarrow \text{dominate}(\gamma, \alpha) \end{array} \right) \right)$$

Abbildung 4.14: COLL-PRINZIP

Die formale HPSG-Beschreibungssprache RSRL lässt auch Definitionen von Relationen zu. Die rekursive Relation *dominate*, welche für das COLL-PRINZIP benötigt wird, ist folgendermaßen definiert: Zwei Zeichen x und y stehen in der Relation *dominate*(x, y), wenn x und y identisch sind (reflexiv), wenn y ein Tochterelement von x ist oder wenn ein Zeichen z ein Tochterelement von x ist und *dominate*(z, y) gilt (transitiv).

Es werden also im COLL-Wert Aussagen über den möglichen Kontext des Wortes getroffen. So kann schon im Lexikoneintrag eines Wortes die Umgebung, in der es auftritt, festgelegt werden.

²⁴Die Ableitung geschieht bei Sailer mit Hilfe von Derivationsregeln (DRs). So gibt es zum Beispiel eine Passiv-DR. Die Tatsache, dass hier nicht von lexikalischen Regeln gesprochen wird, erklärt sich dadurch, dass solche „Metaregeln“ in der HPSG schwer zu formalisieren sind, weil man sich dabei auf der Beschreibungsebene befindet. Ein Lexikoneintrag ist nämlich eine Beschränkung von Merkmalsstrukturen, beschreibt also, wie Objekte aussehen müssen. Mit DRs produziert man keine neuen Lexikoneinträge („wenn es einen LE der Form 1 gibt, dann gibt es auch einen LE der Form 2“), sondern sagt „wenn es ein Zeichen der Form 1 gibt, dann gibt es auch ein Zeichen der Form 2“ und bleibt so auf der Objektebene. Letztere Regel ist eine Derivationsregel. Für DRs zu Passiv und Komplementextraktion siehe Sailer (2003, S. 94-103), zur generellen Konzeption vgl. Meurers (1995 und 2000, Kap. 5). In dieser Arbeit werden wir ausschließlich DRs verwenden und deren Applikation in Bäumen mit dem Symbol \mapsto darstellen.

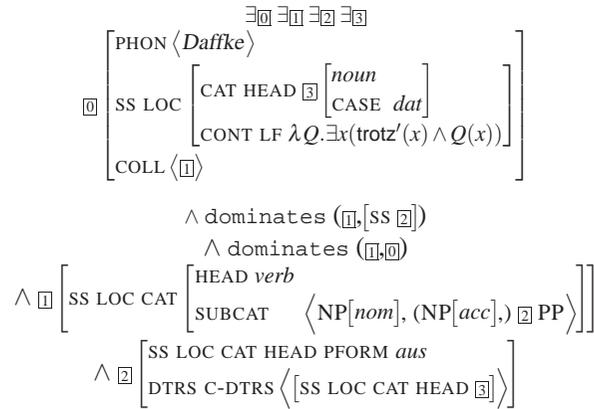


Abbildung 4.15: Skizze des LEs von *Daffke* mit COLL

Nach dieser sehr technischen Beschreibung machen wir uns die Herangehensweise am Beispiel „*etw. aus Daffke tun*“ deutlich. Im Lexikoneintrag von *Daffke* wird unter COLL ein Zeichen spezifiziert, welches das Wort selbst sowie eine PP dominiert, deren Kopftochter *aus* und deren Nichtkopftochter *Daffke* ist. Veranschaulicht wird dies in Abb. 4.15. Um das Ganze noch einmal umgangssprachlicher auszudrücken: Aus dem LE von *Daffke* geht hervor, dass dieses Wort in Sätzen auftritt, die eine Präpositionalphrase enthalten, deren Kopf *aus* ist, wobei *Daffke* das Komplement von *aus* ist. Abbildung 4.16 zeigt die PP *aus Daffke* im Satzzusammenhang. Das COLL-PRINZIP erzwingt die Identität von S' (∃) mit der im Lexikoneintrag von *Daffke* spezifizierten Phrase.

Die Semantik, in Abb. 4.16 unter dem Attribut CONT LF dargestellt, spielt hier nur insofern eine Rolle, als gezeigt wird, dass Sailer für die Bedeutungsableitung der Phrase die funktionale Applikation einsetzt.

Für nicht kompositionale Idiome führt Sailer phrasale Lexikoneinträge ein, ein Vorgehen, dem wir in Abschnitt 6.2 folgen werden.

Bei diesem Ansatz, der wie schon erwähnt auf Wortebene angesiedelt ist, werden in den Valenzlisten *synsem*-Objekte verwendet, das Lokalisierungsprinzip der Selektion bleibt somit auf den ersten Blick gewahrt. Allerdings ist innerhalb der Wörter oder der als phrasale Einheiten gespeicherten Idiome die Struktur der gesamten Äußerung verfügbar. Mit Hilfe von Relationen kann man so innerhalb dieser Strukturen an jegliche Information kommen.

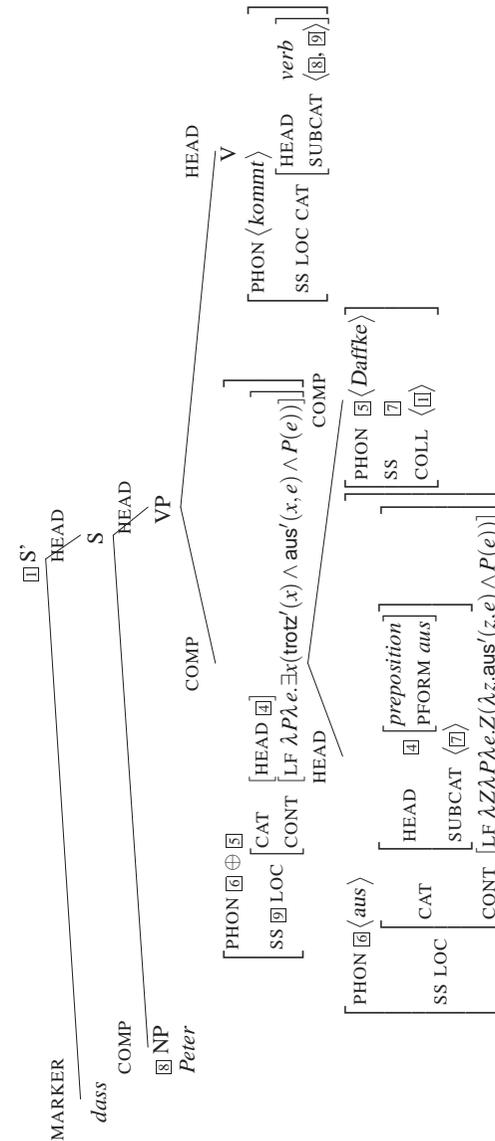


Abbildung 4.16: *dass Peter aus Daffke kommt* mit COLL

Somit ist diese Herangehensweise nicht sehr restriktiv und man kann sich klar machen, dass dieser Apparat sehr mächtig ist. Eine schwerwiegende Kritik gegen den Mechanismus liegt genau in dieser Mächtigkeit, die darin begründet ist, dass lokale Informationen an jedem Knoten in der gesamten Phrase verfügbar sind. Man könnte nicht nur Phänomene beschreiben, die mit idiosynkratischer Distribution von Wörtern zu tun haben: Wenn man alle Möglichkeiten dieses Mechanismus ausschöpft, kann man auch viele andere (vielleicht sogar alle) Bereiche der Grammatik modellieren, wie z. B. die Bindungstheorie. Man sieht auch, dass die normale Selektion über COLL machbar ist, somit wäre diese ein Spezialfall einer Kollokation.

Es wäre also wünschenswert, dass das Modul noch an geeigneter Stelle eingeschränkt wird, um es so einzusetzen zu können, dass es genau die ihm zugeordneten Phänomene behandelt.

4.6 Zusammenfassung

Wir haben in diesem Kapitel vier Herangehensweisen in der HPSG zur Analyse von Idiomen vorgestellt: die interne Selektion von Krenn/Erbach (1994), die externe Selektion von Soehn (2003), eine Konstruktionsanalyse nach Riehemann (2001) und einen Kollokationsansatz von Sailer (2003). Dabei wurde unterschieden zwischen Ansätzen, die auf Wortebene die gegebenen Anforderungen spezifizieren (die beiden Selektionsanalysen), einer Analyse im Rahmen der sogenannten *constructional* HPSG, die auf phrasaler Ebene operiert, und dem Mittelweg von Sailer, der je nach Kompositionalität der Idiome auf Wort- oder Phrasenebene arbeitet.

Es wurde deutlich, dass alle Ansätze entweder an einigen Stellen ihre Schwachpunkte haben bzw. Fragen offen lassen oder für sich genommen ungeeignet sind. Wir wollen deshalb versuchen, nach dem Motto „die Guten ins Töpfchen, die Schlechten ins Kröpfchen“, die Vorteile der einzelnen Herangehensweisen für eine Analyse von IVPs zu verwerten bzw. zu verbinden.

5 Anknüpfungs- und Ausgangspunkte

Prüft aber alles, und das Gute behaltet.

Apostel Paulus

Wir haben im vorhergehenden Kapitel vier Analysemöglichkeiten innerhalb der HPSG diskutiert. Es wurde dabei herausgestellt, dass man jede einzelne Alternative *per se* nicht so, wie sie ist, verwenden sollte. In diesem Kapitel werden deshalb Punkte herausgegriffen, an welchen man für eine akzeptable Analyse von IVPs anknüpfen kann.

Zur Erinnerung seien noch einmal die geforderten Fähigkeiten einer adäquaten Analyse aufgezählt: (i) korrekte Behandlung der Distributionsidiosynkrasien von unikalen Elementen, (ii) geeignete Darstellung von dekomponierbaren und nicht dekomponierbaren IVPs und (iii) eine linguistisch motivierte und gleichzeitig HPSG-kompatible Herangehensweise, die nicht zu mächtig ist und doch alle Fälle von IVPs abdeckt.

5.1 Anknüpfen an die Herangehensweisen

5.1.1 Konstruktion

Der Ansatz von Riehemann krankt nach unserer Sicht an einigen Punkten in einer Weise, dass nicht offensichtlich ist, wie man diese verbessern kann. Als die beiden wichtigsten sind ihre Sortenhierarchie und die „*skeptical default unification*“ zu nennen. Zum ersten Punkt: Riehemann führt neue Sorten für jeden Phraseologismus und zudem noch für jede Art von Konstruktion, in welcher er auftreten kann, ein. Es scheint kontraintuitiv, mehrere Sorten für ein einziges Idiom anzunehmen, weil es in verschiedenen Konstruktionen verwendet wird. Zum Zweiten: Abgesehen von der Tatsache, dass die Verwendung der $\hat{\pi}$ -Operation nicht kompatibel mit den hier verwendeten Mechanismen ist, kann sie für unikale Elemente nicht verwendet werden. Auch wenn man für Unikalia und Wörter mit idiomatischer Bedeutung eigene LEs vorsieht, bleibt die Frage nach der Beschränkung ihrer Distribution.

Jedoch bringt Riehemann gute Gründe für eine phrasale Analyse von nicht-dekomponierbaren Phraseologismen vor. Wir teilen ihre diesbezügliche Sicht-

weise (siehe die Punkte 3, 4, 8 und 9 in Abschnitt 4.4), da manche Idiome syntaktisch nicht wohlgeformt sind, sie eine eingeschränkte Flexibilität aufweisen, bei nicht dekomponierbaren IVPs der Lokus der Semantik auf phrasaler Ebene liegt und Idiome mehr sind als bloße Kopf-Komplement-Strukturen. Wir nehmen deshalb für bestimmte Phraseologismen phrasale Lexikoneinträge (PLEs) an. Für die konkrete Umsetzung siehe Abschnitt 6.2.

5.1.2 Externe Selektion

5.1.2.1 Lexemselektion → Listemselektion

Auch der Ansatz der externen Selektion blieb nicht unkritisiert. Wir wollen zuerst der Kritik am Mechanismus der Lexemselektion begegnen. Es kann bemängelt werden, dass das Attribut LEXEME ein Merkmal der Sorte *content* ist. Dies entspricht nicht der Intuition, da ein Lexem Informationen phonologischer, morphologischer, syntaktischer und semantischer Natur beinhaltet. Rein technisch ist dies, wie in Abschnitt 4.3 dargestellt, die beste Lösung, um den genannten drei wünschenswerten Eigenschaften – Selegierbarkeit, Perkolierbarkeit¹ und Identität des Wertes bei Pronomen und Antezedens – gerecht zu werden. Würde man das Attribut statt LEXEME nun SEMEME nennen, um den semantischen Gehalt zu betonen, wäre das nur eine kosmetische und zugleich fehlerhafte Korrektur. Oft werden ja mehrere verschiedene Sememe mit einem Formativ verbunden und es resultieren polyseme oder homonyme „Teekesselwörter“ wie *Blume*. Es wäre also nicht korrekt bzw. unmöglich, für ein Wort nur einen bestimmten Wert von SEMEME anzugeben.

Machen wir uns jedoch noch einmal klar, was die Lexemselektion in Soehn (2003) bewirkte: Es konnte damit ein ganz bestimmtes Wort selegiert werden, also ein Zeichen, das durch das WORT-PRINZIP lizenziert wird. Im nicht technischen Sinne ist ein Wort bekanntermaßen eine Einheit im Lexikon. Um solch eine Lexikoneinheit zu bezeichnen, führen Di Sciullo/Williams (1988) den Begriff **Listem** ein. Listeme können Morphe, Wörter oder auch Phrasen (siehe Abschnitt 6.2) sein. Gallmann (1990) bringt in seinem Beitrag zum Zusammenwirken von Morphologie und Syntax ebenfalls eine Definition der Begriffe „grammatisches Wort“, „Wortform“ und „Lexem“ und übernimmt selbst auch den Begriff „Listem“ von Di Sciullo/Williams für eine Einheit im Lexikon. Wir wollen daher nicht länger von Lexemselektion, sondern von Listemselektion sprechen.

¹Im Ansatz aus Soehn (2003) wurde dies durch Anpassung des SEMANTIK-PRINZIPS sichergestellt, wobei das Prinzip zum „Hybrid“ wurde, da darin Aussagen über die semantische und die syntaktische Projektionslinie gemacht wurden.

Eine andere Identifikationsmöglichkeit eines Wortes wäre sein PHONOLOGY-Wert. Jedoch kann sich dieser je nach konkreter Instanzierung in einem Satz ändern, z. B. durch das Anhängen einer Flexionsendung. Das Listem ist jedoch immer dasselbe.

Technisch lässt sich dieser neue Ansatz so realisieren, dass man unter HEAD ein Attribut LISTEME einführt. Das Attribut befindet sich damit auch unter CATEGORY, wobei der morphosyntaktische Informationsgehalt des Attributwertes betont wird. Eigentlich besteht ein Listem aus einem ganzen Zeichen, weshalb es intuitiv wäre, LISTEME als *sign*-Attribut einzuführen. Jedoch ist diese Position in der Merkmalsgeometrie für die praktische Anwendung sehr ungeschickt. Die Tatsache, dass LISTEME als Kopfmerkmal definiert wird, ermöglicht dem Wert zu perkolieren, nämlich ganz normal über das KOPF-MERKMAL-PRINZIP. Zudem ist der Wert der Selektion zugänglich.

Um auch die Möglichkeit der Pronominalisierung zu gewährleisten, müssen ein paar Änderungen vorgenommen werden: Es wird eine Beschränkung in die Theorie eingeführt, die besagt, dass alle koindizierten Pronomen auch den LISTEME-Wert ihres Antezedens übernehmen. Im Lexikoneintrag von Pronomen muss daher der Wert für dieses Attribut unterspezifiziert sein und schlicht *listeme* lauten. Wenn es zu einer Koindizierung kommt, ist der LISTEME-Wert identisch mit dem des Antezedens. Jedoch darf man nicht übersehen, dass ein Pronomen als solches auch ein individuelles Listem darstellt und daher auch einen individuellen LISTEME-Wert braucht. Es genügt also nicht, für alle nicht koindizierten Pronomen einen allgemeinen Wert (z. B. *pronominal*) anzunehmen, da dies der Idee eindeutiger LISTEME-Werte widerspräche. Aus diesem Grund ist es erforderlich, für Pronomen ein eigenes Merkmal einzuführen, dessen Wert die listemspezifische Information beinhaltet. Wir nehmen also eine eigene Subsorte von *head* für Pronomen an, die zusätzlich zu LISTEME noch das Merkmal PRO-LISTEME mit Werten wie *sie*, *ihr* etc. trägt. Eine Motivation, dieses neue Attribut einzuführen, ist sicherlich zuerst theorieinterner Natur. Empirisch lässt sich die Einführung von PRO-LISTEME und die Unterspezifikation von LISTEME teilweise rechtfertigen, da Pronomen mit ihrer syntaktischen Eigenschaft des Stellvertreters eine große „Anpassungsfähigkeit“ aufweisen. Sie richten sich in der Theorie somit nicht nur semantisch nach dem Antezedens, sondern auch in den Werten ihrer syntaktischen Merkmale.²

Die erforderliche Beschränkung ist informal in (174) dargestellt. Um ein kon-

²Koindizierung zwischen Pronomina und Antezedenzien ist im gegenwärtigen HPSG-Entwurf gleichbedeutend mit Identität von Kongruenzmerkmalen. Wie es zu einer korrekten semantischen Interpretation der Pronomen kommt, d. h. welche semantischen Bedeutungskomponenten auf das Pronomen übertragen werden und wie dies erfolgen kann, ist nicht geklärt.

ketes Beispiel vorweg zu nehmen, sei auf (242) auf Seite 170 verwiesen. Die vorliegende Beschränkung wird benötigt, um sicherzustellen, dass das Pronomen denselben LISTEME-Wert wie *Armutszeugnis* hat.³

(174) PRONOMEN-LISTEM-BESCHRÄNKUNG (informal):

Wenn ein Pronomen mit einem bestimmten Antezedens koindiziert ist, übernimmt es auch dessen LISTEME-Wert. Ist dies nicht der Fall, ist der LISTEME-Wert des Pronomens identisch mit dem PRO-LISTEME-Wert.

Formal ist die Beschränkung in (175)⁴ ausgeführt. Die erste Implikation entspricht genau der Bedeutung des ersten Satzes in der informalen Version. Die zweite Implikation ist gegenüber der Version in (174) etwas exakter. Sie bedeutet: Genau dann wenn der LISTEME-Wert des Pronomens identisch ist mit dem PRO-LISTEME-Wert, so gibt es in dem betrachteten Zeichen kein *nominal-objekt*, das mit dem Pronomen koindiziert ist.

(175) PRONOMEN-LISTEM-BESCHRÄNKUNG (formal):

$$u\text{-sign} \rightarrow \forall_{[1]} \forall_{[2]} \forall_{[3]} \left(\begin{array}{l} \exists_{[4]} \exists_{[5]} \\ \left(\left[\begin{array}{l} \text{SS LC} \left[\begin{array}{l} \text{CAT HD} \left[\begin{array}{l} \text{LISTEME} \ [3] \end{array} \right] \\ \text{CONT} \left[\begin{array}{l} \text{nom-obj} \\ \text{INDEX} \ [4] \ \text{ref} \end{array} \right] \end{array} \right] \right] \wedge \left[\begin{array}{l} \text{SS LC} \left[\begin{array}{l} \text{CAT HD} \left[\begin{array}{l} \text{LISTEME} \\ \text{PRO-LISTEME} \ [5] \end{array} \right] \\ \text{CONT} \left[\begin{array}{l} \text{ppro} \\ \text{INDEX} \ [4] \end{array} \right] \end{array} \right] \right] \end{array} \right) \end{array} \right) \rightarrow_{[2]} \left[\text{SS LC CAT HD} \left[\text{LISTEME} \ [3] \right] \right) \end{array} \right)$$

$$u\text{-sign} \rightarrow \forall_{[1]} \forall_{[2]} \forall_{[3]} \exists_{[4]} \left(\begin{array}{l} \exists_{[5]} \\ \left(\left[\begin{array}{l} \text{SS LC} \left[\begin{array}{l} \text{CAT HD} \left[\begin{array}{l} \text{LISTEME} \\ \text{PRO-LISTEME} \ [5] \end{array} \right] \\ \text{CONT} \left[\begin{array}{l} \text{ppro} \\ \text{INDEX} \ [4] \ \text{ref} \end{array} \right] \end{array} \right] \right] \end{array} \right) \leftrightarrow \neg_{[6]} \left(\left[\begin{array}{l} \text{SS LC} \left[\text{CONT} \left[\begin{array}{l} \text{nom-obj} \\ \text{INDEX} \ [4] \end{array} \right] \right] \right] \right) \end{array} \right)$$

Zusammenfassend gesagt ist bei diesem Ansatz das Attribut LISTEME mit dem Wert *listeme* unter HEAD hinzugekommen. Untersorten von *listeme* sind ato-

³Allerdings weist T. Höhle (persönliche Mitteilung) darauf hin, dass in Fällen, in denen Idiomteile allein auftreten (bzw. in unterschiedlichen Phrasen einer Äußerung), ein Nomen wie *Armutszeugnis* sehr viel semantische Information beinhaltet (*Armutszeugnis* eviziert sozusagen das *ausstellen*). Deshalb ist zu fragen, ob hier eine normale Koindizierung ausreicht und man auf die Identität der LISTEME-Werte (und damit auf die Beschränkung) verzichten kann. Eine genauere Untersuchung gestaltet sich dabei schwierig, da Daten wie in (242) selten auftreten.

⁴Zur Sorte *u-sign* siehe den betreffenden Absatz auf Seite 108.

mare Objekte wie z. B. *Nase, grob, kommen, ziehen, sie* etc., also keine Lautketten wie bei der PHON-Liste.

Der Einfachheit halber nehmen wir LISTEME-Werte auf der Wortebene an und beschreiben damit sämtliche lexikalische Wörter.⁵ Homonyme und polyseme Wörter, die zwar das gleiche Schriftbild haben, tragen dennoch einen unterschiedlichen LISTEME-Wert. Der Wert dieses Attributs gilt also als eindeutiger „Identifizierer“ eines Wortes, was auch aus verarbeitungstechnischer Sicht sinnvoll ist.⁶ Wir benötigen solch einen eindeutigen Identifizierer bei einer Selektion, die auf ein bestimmtes Wort beschränkt ist, bei Präpositionen und wenn bestimmte Wörter beispielsweise von der Passivierung (s. u.) ausgeschlossen werden sollen. Wir verwenden im Folgenden deshalb bei Sorten des Typs *listeme* Indizes in der gleichen Weise, wie in Printwörterbüchern dadurch verschiedene Lemmata unterschieden werden.

Eine ähnliche Identifikation gab es auch schon bei Pollard/Sag (1994) für Präpositionen mit dem Attribut PFORM. Es wurde dort ausschließlich für semantisch leere Argument-PPs verwendet, deren Köpfe Präpositionen wie *to* oder *of* sind, um diese zu identifizieren. Da mit Hilfe von LISTEME eine Identifizierung gewährleistet ist, kann auf PFORM verzichtet werden.

Damit haben wir eine linguistisch angemessenere Herangehensweise gegenüber der Lexemselektion entwickelt. Das Attribut LEXEME wird also nicht mehr benötigt und gilt fortan nicht mehr als Merkmal von INDEX.

An dieser Stelle sollen keine weiteren Angaben zur Listemselektion gemacht werden, da die Funktionalität für unsere Zwecke ausreichend skizziert wurde. Über mögliche Konkretisierungen und deren Implikationen muss an anderer Stelle weiter nachgedacht werden.

5.1.2.2 Weitere Kritik am XSEL-Ansatz

Einer der Kritikpunkte am XSEL-Ansatz war die Frage nach den Adjunkten. Es wurde schon im vorigen Kapitel dafür argumentiert, obligatorische „Adjunkte“ wie Komplemente zu behandeln. Falls man dies nicht möchte (auf Grund

⁵Die meisten Wörter setzen sich kompositional aus Morphemen zusammen, weshalb man sich auch LISTEME-Werte auf der Morphemebene vorstellen kann und somit der Wortbildung gerechter wird. Es gibt jedoch eine Reihe von Komposita wie beispielsweise *beiwohnen* („teilnehmen“) mit einer nicht kompositionalen Semantik. Diese bräuchten einen eigenen Lexikoneintrag mit eindeutigem LISTEME-Wert. Eine Alternative dazu wäre, ein Modul für idiosynkratische Konstruktionen auf der Morphemebene anzunehmen.

⁶Tilman Höhle (pers. Mitt.) entwirft ein ähnliches Attribut wie LISTEME und nennt es LEXICAL-IDENTITY. Er argumentiert, dass es bei Einheiten im Lexikon so viele idiosynkratische Phänomene gibt, dass man Wörter trotz sehr ähnlicher syntaktischer oder semantischer Eigenschaften dennoch gezielt auseinander halten muss.

eines engeren Komplementbegriffs), wäre ein anderer Ausweg, im Lexikon eintrag Aussagen über die Abhängigkeitsstruktur zu machen. In der Tat gibt es in der HPSG-Merkmalgeometrie das Attribut DEPS unter VALENCE, welches als Wert eine Liste von Zeichen hat, die mit dem entsprechenden Zeichen in einer Abhängigkeitsrelation stehen, vgl. hierzu u. a. die Arbeiten von De Kuthy/Meurers (2000), Bouma (2000), Bouma et al. (2000) und Avgustinova (2001).

In der DEPS-Liste kann man ein mögliches Adjunkt aufführen und dies mit Hilfe der Listenselektion genau spezifizieren. Detmar Meurers (persönliche Mitteilung) weist darauf hin, dass eine solche Spezifikation im Lexikon eintrag jedoch einer Selektion gleichkommt.

Allerdings kann man in Frage stellen, ob Adjunkte in Phraseologismen extra selektiert werden müssen. Ist es nicht vielmehr so, dass ein modifiziertes Nomen in einem Idiom erst durch das Adjektiv zu seiner idiomatischen Bedeutung kommt? So bedeutet *Ast* in „auf dem aufsteigenden *Ast*“ etwas anderes als den Auswuchs eines Baumes. Ähnlich verhält es sich mit (*schiefe*) *Bahn*, (*freien*) *Stücken* etc.

Kommen in einer Äußerung zusätzliche Adjektive vor, kann man diese mit den idiomatischen Adjektiven nicht vertauschen.

(176) *Hobbygärtner Reisinger erhielt eine Sch[r]eibe - „ohne Inventarnummer“, wie Neuwirth versicherte - und einen Zwetschenbaum: Damit er jedes Jahr ein großes blaues Wunder erleben kann.*⁷

Ungrammatisch wäre: **großes, blaues Wunder* oder **blaues großes Wunder*. Dies verdeutlicht, dass eine idiomatische Adjektiv-Nomen-Kombination eine Einheit bildet. Wir wollen deshalb solche Fälle mit Hilfe phrasaler Lexikon einträge analysieren, welche diese Einheiten als Ganzes beschreiben. Dort, wo ein Modifizierer optional ist, also ein wirkliches Adjunkt vorliegt wie bei „[*fröhliche*] *Urstände feiern*“, kann man für eine Modifikation als regulären semantischen Vorgang argumentieren.

Die weiteren Probleme, nämlich die mangelnde Analyse für nicht dekomponierbare IVPs und die Grundannahme einer ausschließlich lokalen Semantik, sind jedoch schwerwiegender. Auch wenn man den ersten Punkt durch phrasale Lexikon einträge beheben kann,⁸ bleibt das zweite Manko. Eine mögliche Erweiterung wäre, als Wert von XSEL das ganze *sign* zu deklarieren. Dies würde den Ansatz jedoch zu mächtig und unrestriktiv machen und zudem das Prinzip der lokalen Selektion verletzen. Schon die Erweiterung von Soehn/Sailer (2003) birgt eine große Mächtigkeit in sich.

⁷Oberösterreichische Nachrichten, 05.07.2000

⁸In Abschnitt 6.2 tun wir genau dies, benötigen dazu jedoch ein weiteres Attribut.

Wir müssen deshalb, im Gegensatz zur Auffassung in Soehn (2003), die Konsequenz ziehen, externe Selektion nicht als geeignetes Mittel zur IVP-Analyse anzusehen. Deshalb muss jedoch der Ansatz als solcher nicht völlig verworfen werden, da man für die im Abschnitt 4.3 definierten Fälle einen solchen Mechanismus braucht und die Theorie durch XSEL an Stelle von SPEC und MOD immer noch ökonomischer ist. Wir revidieren aber das PXS und machen die nicht strukturelle Generalisierung rückgängig.

(177) PRINZIP DER EXTERNEN SELEKTION (endgültige Version):

Hat in einer *headed-phrase* die MARKER-DTR, die ADJUNCT-DTR oder die COMP-DTR einen XSEL-Wert der Sorte *synsem*, so ist dieser identisch mit dem SYNSEM-Wert der Kopftochter.

Der Wert von XSEL sei *xsel-synsem* mit den Untersorten *synsem* oder *none*, wobei im Standardfall ein Zeichen [XSEL *none*] ist. Damit muss zwar eine nur theorieintern motivierte Sorte eingeführt werden (vgl. *mod-synsem*), es wird aber so verhindert, dass der XSEL-Wert von uneingebetteten Zeichen arbiträr ist. Das Attribut wird in Abb. 5.1 zusammen mit LISTEME dargestellt.

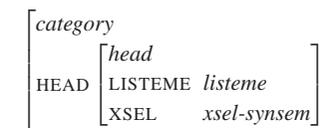


Abbildung 5.1: Neue Attribute von *head*

Mit dem XSEL-Ansatz können, so unser Vorschlag, die nicht idiomatischen Fälle externer Selektion behandelt werden. Neben der Selektion der Nomina durch ihre modifizierenden Adjektive (ehemals durch MOD) seien zwei Anwendungsfälle kurz skizziert.

Erstens kann in einer DP, in welcher der Determinierer das Nomen intern selektiert, das Nomen den Determinierer extern selektieren. Als Attribut kommt dabei weder SPEC in Frage (Nomina sind nicht von der Sorte *functional*) noch MOD, da keine semantische Modifikation stattfindet. Unser Attribut XSEL erweist sich hier als sehr adäquat. Zweitens sei die Auswahl der Hilfsverben *haben* oder *sein* bei der Perfektbildung erwähnt.⁹ Diesmal sind es keine Nomina, sondern Verben, die über das Attribut XSEL andere Verben selektieren.

⁹Dazu gibt es zwar Regeln und syntaktisch-semantische Kriterien, dennoch bleiben ein paar idiosynkratische Fälle als Ausnahmen bestehen (T. Höhle, pers. Mitt.).

Diese Anwendungsgebiete sollen genügen, um die Generalisierung über SPEC und MOD nicht nur nochmals zu motivieren, sondern für deren Beibehaltung zu sprechen. Siehe dazu auch die Zusammenfassung in Kapitel 6.4.

Nach der eben geführten Diskussion bleiben für eine Analyse idiomatischer Wendungen nur noch die Werkzeuge eines phrasalen Lexikoneintrages und der Listenselektion übrig. Wir wenden uns deshalb nun dem COLL-Mechanismus zu, wollen diesen in seiner Mächtigkeit beschränken und ihn für eine Analyse unserer Daten heranziehen.

5.1.3 Kollokation

In Abschnitt 4.5 wurde erläutert, wie der COLL-Ansatz von Sailer (2003) funktioniert. Die Kritik an diesem Kollokationsmodul bestand darin, dass es zu mächtig und unrestringiert ist. So kann man damit z.B. auch ausdrücken, dass ein bestimmtes Wort in einem Satz auftaucht, der mit einem bestimmten Phonem beginnt. Es steht also fest, dass der Ansatz einerseits sehr vielversprechend ist, aber dass andererseits seine Mächtigkeit beschränkt werden muss.

5.2 Ein neuer alter COLL-Ansatz

In Richter/Sailer (1999a) und (1999b) war der in Abschnitt 4.5 beschriebene Ansatz etwas anders aufgebaut und viel restringierter. In diesen Arbeiten war der Wert des COLL-Attributes eine Liste von Barrieren (z.B. eine sogenannte *complete clause*), an die bestimmte Lizenzierungsbedingungen geknüpft waren. Diese bestanden nur aus Restriktionen des CONTENT-Wertes. Zur Illustration siehe Abbildung 5.2.

$$\left[\text{COLL} \left\langle \left[\begin{array}{l} \text{barrier} \\ \text{LICENSER content} \end{array} \right], \dots \right\rangle \right]$$

Abbildung 5.2: COLL-Liste in Richter/Sailer (1999a)

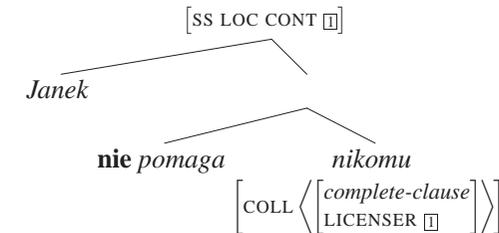
Über ein Prinzip, das COLLOCATION SCHEMA, wurde bewirkt, dass der CONTENT-Wert einer Phrase, die eine Barriere darstellt, identisch ist mit dem Wert des Attributs LICENSER, das zu einem Element auf der COLL-Liste eines Wortes gehört.

In Richter/Sailer (1999a) werden auf diese Weise die Distributionseigenschaften von Negationswörtern im Polnischen analysiert, welche eine Negations-

Kongruenz eingehen. Zum Beispiel verlangt in (178) das Wort *nikomu* einen *negative marker nie* (Beispiel 57 im betreffenden Artikel).

(178) *Janek nie pomaga nikomu.*
 Janek-NOM NM half niemandem-DAT
 ‘Janek half niemandem.’

Das Element *nikomu* kann die Präsenz des semantischen Beitrags eines *negative markers* in seinem COLL-Wert fordern:



In Richter/Sailer (1999b) werden mit denselben Mitteln die Satznegation und die Negations-Kongruenz im Französischen modelliert.

Man sieht an dieser Variante, dass sie sehr viel beschränkter ist im Vergleich zu Sailer (2003). Es gibt Barrieren, die eine Grenze für den Einfluss eines Lizenzierungsparameters darstellen. Man kann mit Hilfe der Barrierspezifikation sozusagen modular den Einflussbereich der geforderten Lizenzierung steuern. Auf Satzebene sind nun nicht mehr alle lokalen Informationen verfügbar. Zudem sprechen die beiden Autoren auch nur von den CONTENT-Werten. Diese Variante der Kollokationstheorie ist sehr vielversprechend für unsere Zwecke. Wir wollen sie jedoch nicht unverändert übernehmen, sondern in der Weise abwandeln, dass wir als Wert des Lizenzierungsmerkmals nicht *content*-Objekte, sondern *local*-Objekte annehmen, da wir ja auch etwas über Werte von LISTEME (in CATEGORY) aussagen wollen. Vor der Darstellung dieser technischen Voraussetzungen werden wir auf den Barrierenbegriff genauer eingehen. Im Kapitel 6 wird illustriert werden, wie man den neuen Ansatz konkret umsetzen kann.

5.2.1 Syntaktische Barrieren

Der Begriff der syntaktischen Barrieren kommt aus der GB-Tradition. Die Barrirentheorie vereint in sich Rektion und Bindung, da es in beiden Bereichen um die Begrenzung syntaktischer Bereiche geht. Die einflussreichste Arbeit ist ‘*Barriers*’ von Noam Chomsky (1986), wobei dieses Werk nicht mehr als

ein Denkanstoß für spätere, erfolgreichere Implementierungen dieser Ideen sein konnte. So wird der Leser „in ein Begriffs- und Definitionswirrwarr [verstrickt] und stellt schnell fest, dass es sich eben nicht um eine fertige Theorie handelt“ (Brandt et al., 1999, S. 257). Eine der später darauf aufbauenden Arbeiten stellt z. B. Haegeman (1994) dar, die in Kapitel 10 eine Diskussion der Barrieren bietet, welche sich auf die Arbeit Chomskys stützt.

Sternefeld (1991) hat sich (u. a.) der Problematiken angenommen und die Barrierentheorie weiterentwickelt.

Wie Sternefeld einleitend darstellt, gewinnt die GB umso mehr an Erklärungskraft, je enger die begriffliche Verzahnung der einzelnen Module (wie Phrasenstruktur, Rektion, Bindung, pro-Modul) ist. Wenn dieselben Konzepte in unterschiedlichen Modulen gelten, so machen sie die Gesamtheorie nicht nur ökonomischer, sondern deren „Richtigkeit“ wird auch stärker untermauert. Der Begriff der Barriere bildet einen Beitrag zu dieser Verzahnung.

Grenzen oder Barrieren sind in folgenden Modulen der GB relevant: in der Theorie der Kasuszuweisung bzw. Rektion, der Theorie der leeren Kategorien (ECP) und der Theorie der Bewegung (*bounding theory*).

- Maximale Projektionen, d. h. Phrasen wie VP oder PP, sind Grenzen für die Kasuszuweisung und blockieren die Rektionsbeziehung.
- Das *Empty Category Principle* stellt sicher, dass Spuren entweder von einer lexikalischen Kategorie regiert oder dass sie antezedensregiert werden. Auch die Antezedensrektion ist nie phrasenüberschreitend.
- Bewegungstransformationen werden durch das Subjanzprinzip beschränkt, das besagt, dass Bewegung höchstens einen Grenzknoten überschreiten darf.
Beispielsweise darf im Englischen nichts aus einer komplexen NP oder einer *wh*-Phrase extrahiert werden (d. h. über zwei Grenzknoten hinweg). Komplexe NPs und *wh*-Phrasen bilden also sogenannte „Extraktions-Inseln“.

Die Kernaussage der Barrierentheorie vereinheitlicht diese Tatsachen und besagt, dass (Antezedens-)Rektion keine und Bewegung höchstens eine Barriere überschreiten darf.

Nun drängt sich die Frage nach der konkreten Definition einer Barriere auf, die keinesfalls leicht zu beantworten ist. Da die Grenzen für Bewegung und Rektion nicht dieselben sind, geht Sternefeld den Weg, erst die Grenzknoten für Subjanz zu bestimmen und danach den (eingeschränkteren) Bereich der

Grenzknoten für Kasuszuweisung durch Hilfhypothesen zu erfassen. Kapitel 3 bis 7 diskutieren Chomskys Ansatz, definieren seine Begrifflichkeiten und nehmen empirische Daten unter die Lupe. Am Ende von Kapitel 7 werden zehn Mängel der Theorie aufgelistet, darunter wohl das Hauptargument, dass Chomskys Anspruch auf Vereinheitlichung über weite Strecken unerfüllt bleibt. Sternefeld vergleicht in den darauf folgenden beiden Kapiteln die Barrierentheorie Chomskys mit den Weiterentwicklungen von Mark Baker und Luigi Rizzi. Am Ende seines Buches entwickelt er zusammen mit Gereon Müller eine Variante, die besonders die Probleme des Deutschen berücksichtigt und die Vereinheitlichung vollständig durchführt.

Der Diskussionsverlauf von Sternefeld (1991) sei an dieser Stelle nicht wiedergegeben. Die Datenlage ist schlichtweg zu komplex, als dass man die Ergebnisse hier kurz und übersichtlich darstellen könnte. Zudem dreht sich die Argumentation um GB-interne Begriffe, welche nicht vorausgesetzt werden können und deren Erklärung uns von der in dieser Arbeit verfolgten Aufgabenstellung zu weit abbringen würde.

Festzuhalten ist, dass maximale Projektionen (XPs) generell Barrieren darstellen, es sei denn, es gibt ein Schlupfloch. Das passiert dann, wenn einer XP der Barrierenstatus unter bestimmten Umständen genommen wird (was nur durch eine YP, welche die XP minimal dominiert, geschehen kann).

Auf die HPSG lässt sich die Diskussion ohnehin schlecht übertragen, da es keine IPs gibt und man sich bei CPs auch nicht einig ist.¹⁰

Zusammenfassend spielt der Barrierenbegriff in der GB bei der Rektion, der Bindungstheorie und der Extraktion bzw. Bewegung (Stichwort Inselbeschränkungen) eine Rolle und ist deshalb innerhalb der GB gut motiviert. Wir wollen uns daher die Barrieren für unsere Analyse zu Nutze machen, ungeachtet mancher unterschiedlicher Sichtweisen in GB und HPSG. So spielen gerade bei Rektion und Bindungstheorie in der HPSG Barrieren keine Rolle. Die Weiterentwicklung der Barrieren im Minimalismus, die sogenannten Phasen, kann man noch schwieriger auf die HPSG übertragen. Nichtsdestoweniger bieten uns die Barrieren ein geeignetes Mittel, um das „Blickfeld“ eines COLL-Moduls einzuschränken.

¹⁰Eine CP ist bei Pollard/Sag (1994) ein S[*comp*], d. h. eine Phrase mit dem MARKING-Wert *that* oder *for*.

5.2.2 Technische Voraussetzungen

Für unsere Kollokationstheorie brauchen wir zuerst neue Sorten und Merkmale, wofür wir die Signatur leicht verändern müssen. Diese Änderungen werden in den Abbildungen 5.3, 5.4 und 5.5 dargestellt. Abbildung 5.3 zeigt die Attribute unter *sign*.¹¹ Wir übernehmen das Attribut COLL (*Context Of Lexical Licensing*) von Richter und Sailer. Der Wert des Attributs ist eine Liste von Barrieren. Ein Zeichen, das eine nicht leere COLL-Liste besitzt, kann nur in Phrasen mit genau solchen Barrieren auftreten. Diese Barrieren werden zusammen mit bestimmten Merkmalen spezifiziert und dienen dazu, den Einfluss des COLL-Mechanismus einzuschränken. Ein idiomatisches Wort kann also nur bis zur nächsten geeigneten Barriere „gucken“ und innerhalb dieses Bereiches Aussagen über die verlangten Eigenschaften machen. *barrier*-Objekte tragen das Attribut LOCAL-LICENSER, kurz LOC-LIC, dessen Wert wiederum die LOCAL-Merkmale der entsprechenden Barriere sind.

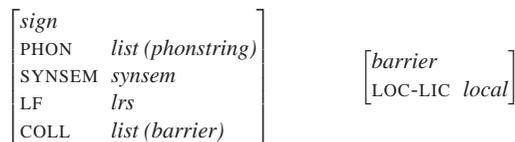


Abbildung 5.3: Neue Angemessenheitsbedingungen

Als Untersorten von *barrier* (siehe Abb. 5.4) definieren wir *complete-clause*, *utterance* und *xp* mit den maximal spezifischen Sorten *np*, *vp-ne* und *pp*.¹² Diese Untersorten beziehen sich auf bestimmte Knoten in der syntaktischen Struktur. In Abb. 5.5 definieren wir Relationen, welche die entsprechenden Knoten (NP, PP etc.) identifizieren.

Es wird intuitiv klar, dass es sich bei einer Satzgrenze (*complete-clause*) auch um eine Barriere handeln muss, beispielsweise muss das Subjekt innerhalb des Satzes realisiert werden. Diese Sorte erbt die Attribute von *barrier* und *phrase*. Diese Barriere bezieht sich auf einen vollständigen Satz, worin auch Topikalisierung stattfinden kann.

Eine Äußerung (*utterance*) bezieht sich auf eine uneingebettete Phrase. Diese hat zusätzlich zu ihrer syntaktischen Vollständigkeit (STATUS *complete*) die Eigenschaft, die kleinste Einheit mit illokutionärer Kraft (mit Merkmal

¹¹Zum Attribut LF siehe Abschnitt 5.3.2.1.

¹²Bei diesen Arten von XPs spricht sehr viel dafür, dass es sich immer um Barrieren handelt.

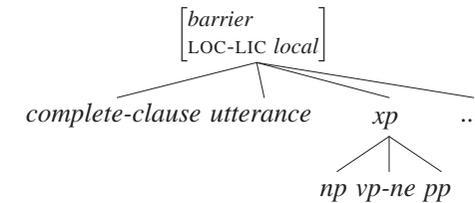


Abbildung 5.4: Sortenhierarchie für *barrier*

ILLOCUTION) zu sein. Die Attribute STATUS und ILLOCUTION werden für das Deutsche in Richter (1997, S. 68f und 136) eingeführt. Richter schafft eine Schnittstelle zu einer Illokutionstheorie: Phrasen mit illokutionärer Kraft sind stets uneingebettet, sie enthalten beispielsweise keine unabgebundenen Spuren. Um diese Unterscheidung zu ermöglichen wird die Sortenhierarchie unter *sign* erweitert und es werden *embedded-signs* von *unembedded-signs* unterschieden (siehe auch Richter, 2004b, Kap. 2.1.2). Uneingebettete Zeichen erhalten das ILLOCUTION-Attribut mit den möglichen Werten *question*, *assertion*, *exclamation* und evtl. noch anderen. Diese Sorten und Attribute sind uns für eine formale Unterscheidung der *barriers* von Nutzen. Richter (ebd.) argumentiert, dass eine feinere Unterteilung unter *sign* und die Einführung der Sorte *unembedded-sign* gerechtfertigt und notwendig sind, weil uneingebettete Zeichen als eigenständige Äußerungen ein zentrales Konzept einer Grammatiktheorie und ein empirisch beschreib- und wahrnehmbares Forschungsobjekt der Linguistik darstellen.

Die *xp*-Barrieren entsprechen allesamt eingebetteten Phrasen (*embedded-phrase*). Bei der *complete-clause* hingegen lassen wir den Zustand der Einbettung offen, da Sätze sowohl eingebettet als auch uneingebettet sein können. Alle Barrieren haben gemeinsam, dass sie Einheiten folgender Art darstellen: Entweder sind sie syntaktisch vollständig oder sie bilden semantische Einheiten. Zu den Barrieren *np* und *pp* gibt es wohl nicht sehr viel zu sagen, bei *vp-ne* jedoch ist zu bemerken, dass hier eine Barriere definiert wird, die sich auf eine ungesättigte Verbalprojektion ohne Spur bezieht („ne“ für nicht-extrahiert), wobei sich das Verb in Verbletzstellung befindet. Diese Barriere ist vor allem für solche idiomatischen Elemente relevant, die als Objekte erscheinen und nicht topikalisiert werden können. Vielleicht gibt es noch andere zu spezifizierende Barrieren, deren Bestimmung jedoch Gegenstand weiterer Forschungen ist.

Zusätzlich zu diesen Neuerungen in der Merkmalsgeometrie und im Relationsinventar benötigen wir ein geeignetes Prinzip, das die erforderlichen Identitäts-

$$\begin{aligned}
& \forall \square \left(\text{is_utterance}(\square) \leftrightarrow \left[\begin{array}{l} \square \text{ [unembedded-phrase]} \\ \text{SS} \left[\begin{array}{l} \text{STATUS complete} \\ \text{LOC CAT} \left[\begin{array}{l} \text{HEAD verb} \\ \text{SUBCAT elist} \end{array} \right] \end{array} \right] \\ \text{ILLOCUTION illocution} \end{array} \right] \right) \\
& \forall \square \left(\text{is_complete-clause}(\square) \leftrightarrow \left[\begin{array}{l} \square \text{ [phrase]} \\ \text{SS} \left[\begin{array}{l} \text{STATUS complete} \\ \text{LOC CAT} \left[\begin{array}{l} \text{HEAD verb} \\ \text{SUBCAT elist} \end{array} \right] \end{array} \right] \end{array} \right] \right) \\
& \forall \square \left(\text{is_vp-ne}(\square) \leftrightarrow \left[\begin{array}{l} \square \text{ [embedded-phrase]} \\ \text{SS} \left[\begin{array}{l} \text{STATUS incomplete} \\ \text{LOC CAT} \left[\begin{array}{l} \text{HEAD verb} \\ \text{INITIAL -} \\ \text{SUBCAT nelist} \end{array} \right] \\ \text{NLOC INH SLASH } \{ \} \end{array} \right] \end{array} \right] \right) \\
& \forall \square \left(\text{is_np}(\square) \leftrightarrow \left[\begin{array}{l} \square \text{ [embedded-phrase]} \\ \text{SS} \left[\begin{array}{l} \text{STATUS complete} \\ \text{LOC CAT} \left[\begin{array}{l} \text{HEAD noun} \\ \text{SUBCAT elist} \end{array} \right] \end{array} \right] \end{array} \right] \right) \\
& \forall \square \left(\text{is_pp}(\square) \leftrightarrow \left[\begin{array}{l} \square \text{ [embedded-phrase]} \\ \text{SS} \left[\begin{array}{l} \text{STATUS complete} \\ \text{LOC CAT} \left[\begin{array}{l} \text{HEAD prep} \\ \text{SUBCAT elist} \end{array} \right] \end{array} \right] \end{array} \right] \right)
\end{aligned}$$

Abbildung 5.5: Relationen für *barrier*-Untersorten

ten erzwingt: Das sogenannte LIZENSIERUNGSPRINZIP stellt sicher, dass eine Phrase die nächstmögliche Barriere ist und deren LOCAL-Wert dem des LOC-LIC-Wertes in der COLL-Liste eines dominierten Zeichens entspricht.

(179) LIZENSIERUNGSPRINZIP (LIP, Version für LOCAL-LICENSER):

Für jedes *barrier*-Objekt in der COLL-Liste eines Zeichens x und für jede Phrase z gilt:

der LOCAL-Wert von z ist identisch mit dem LOC-LIC-Wert genau dann, wenn z das Zeichen x dominiert, z als die spezifizierte Barriere identifiziert werden kann und z kein Zeichen y dominiert, das eine gleichwertige Barriere bildet und selber x dominiert.

Anhand des Satzes (*weil*) *er wie ein Schießhund aufpasste* soll die Wirkung des LIP verdeutlicht werden: die Identitäten in Abb. 5.6 werden durch das Prinzip lizenziert bzw. erzwungen. Das Nomen *Schießhund* fordert in COLL eine bestimmte PP sowie ein bestimmtes Verb als Auftretenskontext.

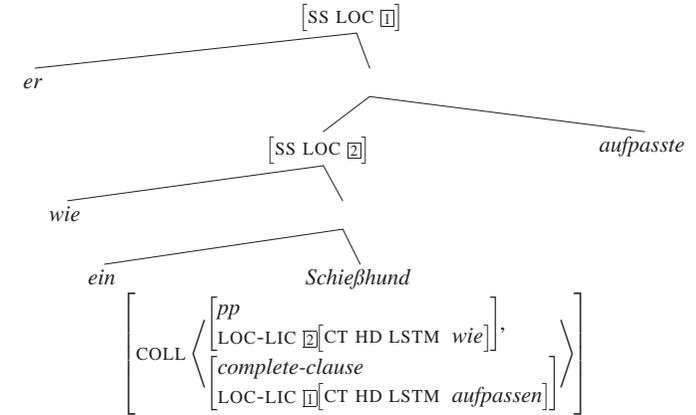


Abbildung 5.6: Beispiel für die Wirkungsweise des LIP

Formal sei das LIP beispielhaft für den Fall dargestellt, dass sich auf COLL Barrieren der Sorte *complete-clause* befinden. Für die anderen Untersorten von *barrier* muss das Prinzip entsprechend umformuliert werden (etwa mit Hilfe einer Disjunktion beim ersten Argument der Relation *member*, die auf Seite 202 definiert wird).

$$\left(\left(\begin{array}{l} \square \text{ [sign]} \\ \text{COLL } \square \text{ [nelist]} \end{array} \right) \rightarrow \right. \\
\left. \forall \square \forall \square \left(\begin{array}{l} \forall \square \forall \square \\ \left(\text{dominate} \left(\begin{array}{l} \square \text{ [phrase]} \\ \text{SS LOC } \square \end{array} \right), \square \right) \wedge \text{member} \left(\begin{array}{l} \square \text{ [complete-clause]} \\ \text{LOC-LIC } \square \end{array} \right), \square \right) \\ \wedge \text{is_complete-clause}(\square) \\ \wedge \neg \exists \square \left(\text{dominate}(\square, \square) \wedge \text{dominate}(\square, \square) \right) \\ \wedge \text{is_complete-clause}(\square) \end{array} \right) \right) \\
\rightarrow \square = \square
\end{array}$$

Eine technische Detailfrage betrifft die Zugriffsmöglichkeit der Relationen auf ihre Argumente. Wenn Relationen aufgerufen werden, können sie nur auf Objekte in der Struktur unterhalb zugreifen. Das bedeutet, wenn ein Wort eine

bestimmte Barriere verlangt, die sich ja in der Struktur über ihr befindet, kann ein Relationsaufruf wie der von `i_s_vp` auf der Ebene des Wortes nicht gelingen, da die Relation nicht auf Phrasen bzw. mögliche Barrieren zugreifen kann. Es gibt zwei Lösungsansätze. Entweder definiert man das LIP für alle *unembedded-signs*. Das hätte zur Konsequenz, dass das COLL-Modul nicht wirklich lokal beschränkt wäre, sondern, um eine Kookkurrenz sicherzustellen, immer erst ein ganzes *u-sign* „im Blick“ haben müsste. Eine andere Lösung zeigt Richter auf: Er argumentiert (bes. Richter, 2004b, Kap. 2.2.2), dass jede Entität der linguistischen Beschreibungsdomäne Teil eines uneingebetteten Zeichens sein muss, beispielweise um zu verhindern, dass eine Theorie „gestrandete Strukturen“, welche nicht Teil einer kompletten Äußerung sind, erlaubt. Durch ein Attribut `EMBEDDED`, definiert für jede Sorte, wird diese Forderung umgesetzt und somit erhält jedes eingebettete Zeichen Zugriff auf das uneingebettete Zeichen, zu dem es gehört. Damit ist auch die Voraussetzung geschaffen, dass unsere definierten Relationen von einem beliebigen eingebetteten Zeichen aus aufgerufen werden können.

Damit haben wir nun die technischen Voraussetzungen geschaffen, um eine Analyse von IVPs in Angriff zu nehmen.¹³

¹³ Mit Hilfe dieses neuen Ansatzes können auch weitere Distributionsbeschränkungen modelliert werden. Beispielsweise kann man die phonologisch begründete Verteilung von *a/an* im Englischen (siehe Asudeh/Klein, 2002) oder *le/l'* im Französischen beschreiben, indem man für *barriers* ein neues Attribut annimmt, das den PHON-Wert bestimmt. Ein weiteres mögliches Merkmal ist LF-LIC, mit dem man semantische Phänomene wie negative Polaritätselemente behandeln kann. Der Ansatz ist also „modular“ und lässt sich sowohl in vertikaler Ebene durch die Barrieren als auch in horizontaler Ebene durch verschiedene Attribute erweitern bzw. beschränken. Die *Appropriateness* von *barrier* sieht letztendlich folgende drei Attribute vor: `[PHON-LIC list (phonstring)]`, `[LOC-LIC local]` und `[LF-LIC lrs]`.

Es ist hier festzuhalten, dass kein Zugriff auf die interne Struktur, also die DTRS-Attribute, möglich ist. Eine Generalisierung des LIP gestaltet sich folgendermaßen:

(180) LIZENSIERUNGSPRINZIP (LIP, allgemeine Version):

Für jedes *barrier*-Objekt in der COLL-Liste eines Zeichens *x* und für jede Phrase *z* gilt:

- a) der PHON-Wert von *z* ist identisch mit dem PHON-LIC-Wert von *x*,
- b) der LOCAL-Wert von *z* ist identisch mit dem LOC-LIC-Wert von *x*,
- c) der LF-Wert von *z* ist identisch mit dem LF-LIC-Wert von *x*

genau dann, wenn *z* das Zeichen *x* dominiert, *z* als die spezifizierte Barriere identifiziert werden kann und *z* kein Zeichen *y* dominiert, das eine gleichwertige Barriere bildet und selber *x* dominiert.

5.3 Ausgangspunkte für die Analyse

Um die Analyse in Kapitel 6 möglichst einheitlich und umfassend anzugeben, muss vorher noch auf ein paar grundlegende Aspekte hingewiesen werden. Wir beginnen mit dem Status der Subjekte und der Frage nach der Modellierung der Verbstellung im Deutschen. Hierbei spielt der Begriff des Prädikatskomplexes eine wichtige Rolle. Weiterhin werden die Phänomene der Passivierung und Nominalisierung näher betrachtet. Anschließend wird eine Semantiktheorie diskutiert, die wir später für die Skizzierung der Lexikoneinträge und Strukturbäume heranziehen werden. Am Ende dieses Abschnitts wird erörtert, wie man Selektionsrestriktionen und nicht referenzielle Nominalphrasen in der HSPG modellieren kann.

5.3.1 Verben, Klammern und Komplexe

5.3.1.1 Das Subjekt

Es gibt für das Deutsche zahlreiche Veröffentlichungen, die das Vorhandensein einer VP in Frage stellen bzw. diskutieren, ob das Subjekt¹⁴ zur VP gehört oder nicht, siehe dazu Müller (1999), der auch zahlreiche Literaturhinweise gibt.

In der GB-Literatur gehen manche davon aus, dass Subjekte nicht subkategorisiert werden müssen, sondern am Knoten SpecIP basigeneriert werden. Dies ist allerdings für das Deutsche nicht haltbar. In der HPSG kann man dies zudem so nicht umsetzen, da es bei der Diskussion in GB darum ging, wieviel Information über die selegierten Argumente im Lexikon steht. In der HPSG ist jedoch eine Selektion nur lexikalisch möglich.

Man kann zwar tatsächlich dafür argumentieren, dass es in finiten deutschen Sätzen keine VP als Zwischenstufe gibt und dass Subjekt und Objekt(e) auf einmal gesättigt werden (vergleiche auch Oppenrieder, 1991, Kap. 1.5.3). Jedoch gibt es bei infiniten Phrasen (so wie *Maria zu kennen* in „*Hans behauptet, Maria zu kennen.*“) an der Oberfläche kein Subjekt, es liegen dort also doch strukturelle Unterschiede zwischen Subjekt und Objekt vor. Folglich muss für diese Fälle eine VP angenommen werden.

Als ein weiteres Argument für eine Sonderbehandlung von Subjekten wird angeführt, dass sie nicht interne Argumente einer IVP sind. Müller (2003a) macht einige Bemerkungen zu Subjekten als Idiombestandteilen: Einer verbreiteten Behauptung zu Folge gibt es keine Idiome, die ein Subjekt haben. Dies können wir mit folgenden Beispielen widerlegen:

¹⁴Zu Subjekten sind Nominalphrasen im Nominativ bei nicht prädikativem Gebrauch und Subjektsätze zu zählen.

(181) *Sein Thron wackelt bedenklich seit der gestrigen Wahlschlappe.*

(182) *Der Geduldsfaden reißt ihr allmählich.*

(183) *Ihm raucht förmlich der Kopf.*

(184) *Ihn sticht der Hafer.*

Müller (2003a) diskutiert Scherpenisse (1986), welcher behauptet, dass es sich bei Beispielen mit Subjekten als Idiombestandteil um sogenannte unakkusativische Verben handelt (S. 89), das heißt, dass die Subjekte keine wirklichen Subjekte, sondern zu Grunde liegende Objekte sind. Dagegen spricht aber, dass die Verben der IVPs (182) bis (184) in der wörtlichen Lesart durchaus passivierbar sind. Bei nichtakkusativischen Verben ist generell keine Passivierung möglich.¹⁵ Ein weiteres Argument von Scherpenisse (nach Müller, ebd.) ist, dass sich die Nominative in den Idiomen wie Objekte verhalten, da sie adjazent zum Verb auftreten und somit dieselben Stellungseigenschaften wie Objekte haben. Wenn man aber so argumentiert, schließt man alle Idiome aus der Diskussion aus, da Idiome im Deutschen immer bevorzugt adjazent auftreten. Damit ist dieses Argument für eine Sonderbehandlung von Subjekten hinfällig.

Um die strukturellen Unterschiede zwischen Subjekten und Objekten zu modellieren, kann man das Subjekt infiniter Verben (und nur dieses) als Kopfmerkmal in einer SUBJ-Liste repräsentieren und es nicht als Element der SUBCAT-Liste aufführen.

Bei finiten Verbformen steht das Subjekt dagegen auf der SUBCAT-Liste, da es in finiten Phrasen auch gesättigt wird. Die SUBJ-Liste ist maximal einelementig, enthält das Subjekt, und sie ist leer bei Verben, die kein Subjekt besitzen (z. B. bei *grauen* in „*Mir graut.*“ oder in „*Mit ihr ist nicht gut Kirschen essen.*“). Da es sich bei SUBJ um ein Kopfmerkmal handelt, kann die Information perkolieren und steht bei der maximalen Projektion einer infiniten Verbalphrase zur Verfügung. Dies ist für die Analyse von Kontrollkonstruktionen wichtig. Bei Meurers (1999) wird SUBJ auch für die Kasuszuweisung benötigt.

Vergleiche zu diesem Abschnitt Müller (1999, Kap. 1.7) bzw. Müller (2003b, Kap. 3.1).

¹⁵Die Tatsache, dass manche Passivierungen der Idiombeispiele fragwürdig sind, ist von anderen Idiomen bekannt. In vielen der Beispiele sind die Subjekte unbelebt, was Einfluss auf die Passivierbarkeit haben könnte.

5.3.1.2 Der Verbalkomplex

Man kann im Deutschen die Bereiche innerhalb eines Satzes in die sogenannten topologischen Felder unterteilen (siehe dazu den die Forschung zusammenfassenden Artikel von Höhle, 1986). Hierbei wird im Wesentlichen zwischen Vorfeld, linker Satzklammer, Mittelfeld, rechter Satzklammer und Nachfeld unterschieden. Ein Synonym für die rechte Satzklammer ist der Prädikats- oder Verbalkomplex.

Stefan Müller hat sich eingehend mit dem Verbalkomplex und den verschiedenen Formen der Voranstellung seiner Bestandteile beschäftigt, weshalb wir uns in diesem Abschnitt auf seine Arbeiten stützen wollen. Wichtige Vorarbeiten dazu sind: Frank (1994), Hinrichs/Nakazawa (1989; 1994), siehe dazu auch Meurers (2002), sowie Meurers (2000) und Kiss (1995a; 1995b).

In den Beispielen (185) und (186) verbinden sich die Verben zuerst zum Verbalkomplex, bevor dieser mit den Objekten kombiniert wird. Man kann zeigen, dass verbale Komplemente vor nicht verbalen gesättigt werden müssen. Damit werden bestimmte Wortstellungen wie beispielsweise in (187) ausgeschlossen.

(185) *dass Hans die Dissertation lesen wird.*

(186) *dass Hans die Dissertation lesen können wird.*

(187) **dass die Dissertation lesen Hans können wird.*

Verben bilden also einen Komplex, wobei das Matrixverb sämtliche Argumente anzieht. Dabei selegiert das Matrixverb das eingebettete Verb und dessen Argumente. Müller (erscheint) analysiert Verbalkomplexe über ein Phrasenstrukturschema, das sogenannte PRÄDIKATSKOMPLEX-SCHEMA. Er folgt damit Hinrichs/Nakazawa (1989, 1994). Allerdings ist in Müller (erscheint) das Valenzmerkmal (VCOMP in Müller, 1999 bzw. XCOMP in Müller, 2002) für die Selektion von Elementen, mit welchen ein Kopf den Verbalkomplex bildet, weggefallen.

Solche Schemata, die in Pollard/Sag (1994) *Immediate Dominance Schemas* genannt werden, dienen zur Lizenzierung von phrasalen Knoten, d. h. jeder Knoten muss einem dieser Schemata genügen. Dort wird spezifiziert, in welcher Weise sich Kopftöchter, Komplemente, Adjunkte, Marker und Füller miteinander verbinden. Das PRÄDIKATSKOMPLEX-SCHEMA beschreibt eine Phrase, deren Töchter eine *head-cluster-structure* bilden. Dort kommt der SYNSEM-Wert der ersten Nichtkopftochter (CLUSTER-DAUGHTER) zu den

Elementen der SUBCAT-Liste der Kopftochter hinzu, und die beiden Töchter verbinden sich so zu einem Prädikatskomplex. Das Schema wird in (188) dargestellt – gegenüber Müller (erscheint) leicht umformuliert und um die LISTEME-Information erweitert.

$$(188) \left[\begin{array}{l} \textit{phrase} \\ \text{SYNSEM} \left[\text{LOC CAT} \left[\text{HEAD LISTEME } \boxed{1} \right] \right. \\ \left. \text{SUBCAT } \boxed{1} \right] \\ \text{DTRS} \left[\begin{array}{l} \textit{head-cluster-struct} \\ \text{H-DTR} \left[\text{SYNSEM LOC CAT SUBCAT } \boxed{1} \oplus \langle \boxed{2} \rangle \right] \\ \text{CL-DTR} \left[\text{SYNSEM } \boxed{2} \text{ LOC CAT HEAD LISTEME } \boxed{3} \right] \end{array} \right] \end{array} \right]$$

Die Lexikoneinträge von Hilfsverben werden entsprechend spezifiziert, was am Beispiel *wird* als Futur-Hilfsverb dargestellt wird (Abb. 5.7).

$$\left[\begin{array}{l} \textit{word} \\ \text{PHON } \langle \textit{wird} \rangle \\ \text{SYNSEM LOC CAT} \left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{SUBCAT } \boxed{1} \oplus \langle \text{V[LEX+, bse, SUBCAT } \boxed{1}] \rangle \end{array} \right] \end{array} \right]$$

Abbildung 5.7: LE des Auxiliars *wird*

Die SUBCAT-Liste von *wird* besteht also aus derjenigen des Vollverbs und einem zusätzlichen verbalen Element – dem Vollverb selbst oder einem weiteren Hilfsverb im Prädikatskomplex. Das Symbol \oplus steht für die Relation *append*, mit der zwei Listen verknüpft werden.

Das Attribut LEX (unter SYNSEM) ist positiv bei lexikalischen Zeichen, d. h. bei Elementen, die noch nicht projiziert haben. Phrasen sind damit entsprechend als [LEX -] gekennzeichnet. Durch die Spezifikation des LEX-Wertes wird sichergestellt, dass in (185) *lesen* mit *wird* kombiniert wird, bevor *lesen* mit seinem Komplement verbunden wird. Da in Kopf-Komplement-Strukturen der Mutterknoten stets einen negativen LEX-Wert hat, kann die Phrase *die Dissertation lesen* nicht mit *wird* kombiniert werden. Der LEX-Wert des Mutterknotens im Prädikatskomplex-Schema bleibt unterspezifiziert (im Gegensatz zu Kopf-Komplement-Strukturen), so dass mehrere Verben wie in (186) eingebettet werden können.¹⁶

¹⁶Der Ausschluss von unechten Mehrdeutigkeiten erfolgt bei Müller (erscheint) über eine Beschränkung über die Art der Elemente, die angehoben werden können. Darauf gehen wir an dieser Stelle nicht näher ein.

$$\left[\begin{array}{l} \text{PHON } \langle \rangle \\ \text{SS} \left[\begin{array}{l} \text{LOCAL } \boxed{1} \\ \text{NONLOCAL} \left[\begin{array}{l} \text{INHERITED SLASH } \{ \boxed{1} \} \\ \text{TO-BIND SLASH } \{ \} \end{array} \right] \end{array} \right] \end{array} \right]$$

Abbildung 5.8: LE der Spur (vereinfacht)

Beim PRÄDIKATSKOMPLEX-SCHEMA ist anzumerken, dass wir darin festlegen, dass der Prädikatskomplex den LISTEME-Wert der Nichtkopftochter übernimmt. Dies hat zur Folge, dass der LISTEME-Wert eines einzelnen flektierten Verbs wie *lief* und der eines Verbalkomplexes wie *laufen können wird* identisch sind. Auch wenn *laufen* im genannten Verbalkomplex nicht der Kopf ist, ist es dennoch das wichtigste, weil sinnbestimmende Verb. Für unsere IVP-Analyse ist dies ein Vorteil, da nun eine exakte Verbselektion (von z. B. *laufen* mit Hilfe des LISTEME-Attributs) auch dann gelingen kann, wenn das Verb durch Modal- oder Hilfsverben näher bestimmt wurde.

Wir wollen im Folgenden anhand zweier Beispiele die Verbbewegung aus dem Verbalkomplex und die scheinbar mehrfache Vorfeldbesetzung illustrieren. Verbbewegung findet immer dann statt, wenn ein Verberst- oder ein Verbzweitsatz vorliegt, wenn also das Verb nicht in Verbletztposition steht.

Die Verbbewegung selbst ist ein lokales Phänomen, es handelt sich dabei nicht um eine Fernabhängigkeit. Die Konstituenten-Extraktion dagegen wird als Fernabhängigkeit analysiert, deren Behandlung in der HPSG durch drei „Bausteine“ geschieht: (i) Durch den Lexikoneintrag der Spur (Abb. 5.8) kommt der LOCAL-Wert des extrahierten Elementes in ihre INHER SLASH-Menge. (ii) Das NONLOCAL-PRINZIP bewirkt, dass der INHER SLASH-Wert nach oben perkoliert (formaler: der NONLOCAL INHER-Wert eines phrasalen Zeichens ist die Vereinigung der NONLOCAL INHER-Werte seiner Töchter abzüglich des NONLOCAL TO-BIND-Wertes der Kopftochter).

Schließlich lizenziert (iii) das KOPF-FÜLLER-SCHEMA phrasale Knoten, deren jeweilige Kopftochter den LOCAL-Wert der Füllertochter sowohl in der INHER SLASH-Menge als auch in der TO-BIND SLASH-Menge hat. Die phrasalen Knoten selbst besitzen durch das NONLOCAL-PRINZIP leere SLASH-Listen (bei einfacher Extraktion).

Bei der Voranstellung von Verben funktioniert es etwas anders. Wie schon gesagt, handelt es sich um ein lokales Phänomen, weshalb man nicht auf das NONLOCAL-PRINZIP zurückgreift. Um die Informationsweitergabe im Baum

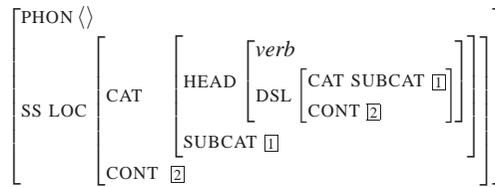


Abbildung 5.9: LE der Verbspur (vorläufige Version)

dennoch zu sichern, wurde das Merkmal DSL (für *double slash*) eingeführt, das unter HEAD steht und LOCAL-wertig ist.¹⁷ Der Lexikoneintrag der Verbspur, dargestellt in Abb. 5.9 aus Müller (erscheint, S. 21), bewirkt, dass die Valenzinformation des wegbewegten Verbs ebenfalls im DSL-Wert der Spur steht. Durch das KOPF-MERKMAL-PRINZIP ist dieser Wert an jeder Projektion der Verbspur verfügbar. Nun kommt ein weiterer Baustein dazu, nämlich eine spezielle Version des finiten Verbs, das bewegt wurde. Diese Version entsteht durch eine Derivationsregel (Abb. 5.10).¹⁸ Das so lizenzierte Verb nimmt die Projektion der Verbspur zum Argument, wobei die Kombination nur erfolgen kann, wenn die Valenzeigenschaften des Verbs mit denjenigen der Verbspur identisch sind. Die Eingabe für diese Regel sind Verben in Endstellung, was durch das *verb*-Merkmal [INITIAL -] angezeigt wird. Die Ausgabe ist ein Verb in Erststellung [INITIAL +]. Linearisierungsregeln berücksichtigen dieses Merkmal und stellen die korrekte Anordnung der Köpfe sicher. Auf Linearisierung wollen wir jedoch nicht eingehen, da es sich hierbei wieder um ein viel diskutiertes Sub-Modul in der HPSG handelt.

Die Identität der CONTENT-Werte (angezeigt durch *tag* $\boxed{4}$) bewirkt, dass bei der Kombination des Verbs in Erststellung mit der Projektion der Verbspur der semantische Beitrag der Verbspur-Projektion übernommen wird. Wegen des SEMANTIK-PRINZIPS wird dieser zum Beitrag der gesamten Konstruktion.

Wir wollen nun endlich diese abstrakt dargestellten Mechanismen anhand des Satzes „*Widmen wird er die Dissertation seinem Neffen.*“ illustrieren. Müller analysiert solche Verbzweitsätze als Fälle von Verberststellung mit zusätzli-

¹⁷*double slash* (DSL) stammt ursprünglich von Jacobson (1987), vgl. auch Borsley (1989) und Kiss (1995b, S. 223ff).

¹⁸Nach Müllers lexikalischer Regel in Müller (erscheint, S. 21). Siehe dazu auch Fußnote 24 auf Seite 93 dieser Arbeit. Für eine ausführliche Diskussion von DRs und einen formalen Expansionsmechanismus siehe Meurers (2000).

Die Derivationsregel wird in der expliziten Schreibweise dargestellt. Dafür müssten eigentlich sämtliche Identitäten zwischen IN und OUT spezifiziert werden, worauf aber zu Gunsten besserer Lesbarkeit verzichtet werden soll (vgl. dazu Meurers, 2000, S. 162f).

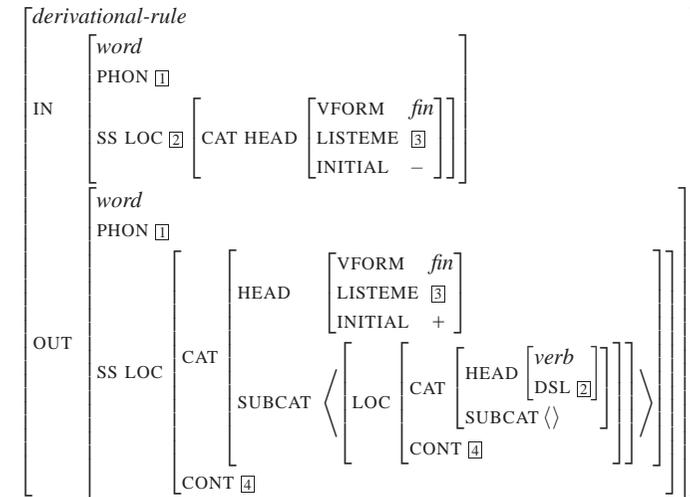


Abbildung 5.10: DR für Verberststellung (V1-DR, vorläufige Version)

cher Extraktion eines Elementes ins Vorfeld (die Füller-Tochter). Die zu Grunde liegende Struktur ist die Verbletzstellung, weshalb es von *wird* eine Spur am Satzende gibt. Die Spuren bilden am Satzende einen Prädikatskomplex. Zu finden ist der Syntaxbaum in Abbildung 5.11. Der Übersichtlichkeit halber werden folgende Abkürzungen zusätzlich zu den schon vorher verwendeten benutzt: NLOC für NONLOCAL, CT für CATEGORY, SC für SUBCAT, INH für INHERITED, TB für TO-BIND und SL für SLASH. Weiterhin wird auf das SUBJ-Merkmal bei *widmen* verzichtet – auch das Subjekt $\boxed{3}$ steht auf der SUBCAT-Liste.

In diesem Beispiel sind beide Mechanismen, Verbbewegung und Extraktion, kombiniert worden. Die Spur von *wird* ist durch den Lexikoneintrag der Verbspur lizenziert. Der DSL-Wert $\boxed{3}$ ist die SUBCAT-Liste $\boxed{2}$ des infiniten Verbs *widmen* plus der SYNSEM-Wert $\boxed{6}$ des Komplements der Verbspur. Diese Anwendung der Relation *append* wird durch das PRÄDIKATSKOMPLEX-SCHEMA (Seite 116) lizenziert. Dadurch verbinden sich die beiden Spuren zu einem Verbalkomplex, das wegbewegte Komplement ist die CLUSTER-DTR (CL). Der DSL-Wert perkoliert nach oben. Das finite Verb bekommt durch die Derivationsregel einen anderen Wert auf der SUBCAT-Liste $\boxed{7}$. Wichtig ist an dieser Stelle, dass der LOCAL-Wert $\boxed{8}$ des ursprünglichen Verbs identisch ist mit dem DSL-Wert der Verbalprojektion $\boxed{7}$. Dadurch kann die Kombinati-

Dabei müssen die Idiomteile nicht adjazent zum idiomatischen Verb stehen:

(193) [Öl][ins Feuer] dürfte auch die Ausstrahlung eines Interviews gießen, das die US-Fernsehstation ABC in der vergangenen Woche mit Elián führte.²¹

Es ist zu prüfen, ob unser COLL-Modul mit Müllers Voranstellungs-Ansatz kompatibel ist, um auch Voranstellung bei IVPs beschreiben zu können.

5.3.1.3 Komplexere Prädikate

Im vorhergehenden Abschnitt wurden Vollverben zusammen mit Hilfs- und auch Modalverben zu einem Prädikatskomplex zusammengefasst. Die entscheidende Rolle spielt dabei das Schema in (188), welches eine *head-cluster-structure* lizenziert. Zudem werden die Lexikoneinträge der Hilfs- und Modalverben so spezifiziert, dass ihre SUBCAT-Listen die Vollverben enthalten und deren Valenzmerkmale übernommen werden.

In (194) wird gezeigt, dass Verben, also Teile des Prädikatskomplexes, nicht unabhängig voneinander bewegt werden können. Die Komponenten, welche in Endstellung den Prädikatskomplex bilden, sind unterstrichen.

- (194) a) *...weil er dem Kind ein Buch geben (müssen) wird*
 b) *Ein Buch wird er dem Kind geben (müssen).*
 c) *Ein Buch geben wird er dem Kind.*
 d) *?Ein Buch geben müssen wird er dem Kind.*
 e) *Geben wird er dem Kind ein Buch.*
 f) *?Geben müssen wird er dem Kind ein Buch.*
 g) *??Geben wird er dem Kind ein Buch müssen.*
 h) **Müssen wird er dem Kind ein Buch geben.*

Man sieht in den Beispielen b) bis f), dass das finite Verb allein oder auch der gesamte Prädikatskomplex vorangestellt werden kann.²² Topikalisiert man lediglich Teile des Komplexes wie in g) und h), werden die Sätze sehr fragwürdig bzw. als ungrammatisch empfunden. Dies liegt daran, dass *müssen* das

²¹ taz, 28.03.00, S. 9;

²²Die Fragwürdigkeit von d) und f) resultiert wahrscheinlich aus der Ungebräuchlichkeit einer solchen Formulierung. Diese Ungebräuchlichkeit liegt zu einem wesentlichen Teil an der Informationsstruktur der Sätze. De Kuthy (2000) argumentiert dafür, dass man bei der Untersuchung von Wortstellungsphänomenen nicht nur syntaktische, sondern auch pragmatische Aspekte wie die Informationsstruktur berücksichtigen sollte.

Vollverb *geben* subkategorisiert und sich mit ihm verbindet. Ein weiteres empirisches Argument für den Prädikatskomplex ist die Auxiliar-Inversion (bzw. *flip*), die von Hinrichs/Nakazawa (1989) ausführlich diskutiert wird.

Der Prädikatskomplex in sich ist jedoch nicht völlig unveränderlich, da es Fälle gibt, in welchen er z.B. durch Adjunkte unterbrochen wird:

(195) *...weil man seine Illusionen hätte umso eher begraben müssen.*²³

Auch Teile des Prädikatskomplexes können bewegt werden, wie folgendes Beispiel mit einer Adjektivphrase zeigt:

(196) *Gespannt darf man darauf sein, wieweit die ‚PC-Terminals‘ Akzeptanz finden werden.*²⁴

Nicht möglich ist jedoch eine Bewegung analog zu (194 h):

(197) **Sein darf man darauf gespannt, wieweit ...*

Festzuhalten ist, dass Teile des Prädikatskomplexes eine Art Kohärenz aufweisen und man sie deshalb zu einem Komplex, der *head-cluster-structure*, zusammenfasst.

Unter Umständen können auch nominale Elemente mit in den Prädikatskomplex aufgenommen werden. Dies gilt für interne Argumente von IVPs und sicherlich auch für direkte Objekte von Funktionsverbgefügen. Auch inkorporierte Nomina kann man dazu zählen.

Stellungseigenschaften der NP bieten empirische Hinweise. In den folgenden Beispielen wurde das interne Argument des Idioms *die Leviten lesen* in den Prädikatskomplex aufgenommen. Daten dieser Art werden auch in Uszkoreit (1987, S. 106f) diskutiert. Wie in (194) sind die Komponenten, welche in Endstellung den Prädikatskomplex bilden, unterstrichen.

- (198) a) *...weil er dem Kind die Leviten lesen (müssen) wird*
 b) *Die Leviten lesen (müssen) wird er dem Kind.*
 c) *Die Leviten wird er dem Kind lesen (müssen).*
 d) *Die Leviten werden dem Kind gelesen. (Passiv)*
 e) **Lesen wird er dem Kind die Leviten.*
 f) **Lesen müssen wird er dem Kind die Leviten.*

²³(Müller, 1999, S. 264)

²⁴(Müller, 1999, S. 355)

- g) **Lesen wird er dem Kind die Leviten müssen.*
- h) **Müssen wird er dem Kind die Leviten lesen.*

Dadurch, dass die NP Teil des Prädikatskomplexes ist, erklärt sich die Ungrammatikalität der Sätze e) bis h). Zwar kann die NP bei den Beispielen b), c) und d) noch extrahiert werden (ebenso wie in 194 b), die verbalen Komponenten sind nun jedoch tiefer in den Komplex eingebettet und können nicht ohne die NP vorangestellt werden. Wir nehmen daher an, dass der neue Prädikatskomplex eine Struktur wie in Abbildung 5.14 hat und die Stellungseigenschaften daraus resultieren, dass tief eingebettete Elemente nicht alleine extrahiert werden können. Die Details zur Abbildung werden im nächsten Absatz erläutert. Wir haben damit eine „Verknüpfungshierarchie“ vorliegen. Das Vollverb, das den semantischen Kopf bildet, ist Komplement eines weiteren Vollverbs oder Hilfsverbs; diese Phrase kann dann Komplement eines weiteren Hilfsverbs oder von Modalverben sein. Anschließend kommt das interne Argument des semantischen Kopfes hinzu.

Für eine Mitaufnahme von nominalen, idiomatischen Komponenten sprechen neben den Stellungseigenschaften auch lexikalische bzw. semantische Gründe. Im Lexikoneintrag wird beispielsweise das interne Argument explizit genannt wie bei *die Leviten lesen*. Semantisch gehören die beiden Idiomteile (NP und Verb) zusammen und bilden das Idiom.

In den Prädikatskomplex werden bei idiomatischen Wendungen nur deren interne Argumente aufgenommen, nicht die wendungsexternen. Wir benötigen deshalb eine Unterscheidungsmöglichkeit zwischen beiden Arten von Argumenten. Dazu führen wir ein neues Attribut **IDIOMATIC-ARGUMENT-STATUS** ein (kurz **I-ARG-STAT**) und setzen es in der Merkmalsgeometrie unter **LOCAL**, da die Information sowohl syntaktischer als auch semantischer Natur ist. Der

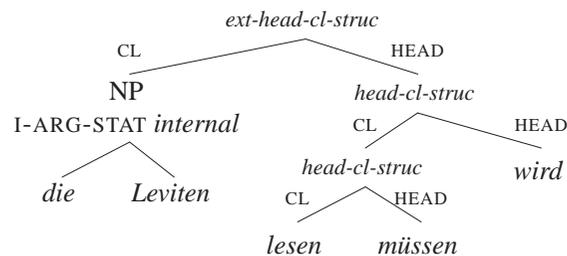


Abbildung 5.14: Strukturbaum des erweiterten Prädikatskomplexes

Wert sei *internal* bzw. *external*. Damit kann ein Verb auf seiner SUBCAT-Liste zwischen internen und externen Argumenten unterscheiden. In den nicht idiomatischen Fällen wird das Merkmal immer mit dem Wert *external* spezifiziert. Um die Verfügbarkeit der Information auf Phrasen-Ebene zu gewährleisten, definieren wir ein geeignetes Prinzip, welches analog zum **KOPF-MERKMAL-PRINZIP** funktioniert (vgl. Appendix B).

Das **PRÄDIKATSKOMPLEX-SCHEMA** (Seite 116) ist bisher nur für Komplexe mit ausschließlich verbalen Komponenten ausgelegt. Der **LISTEME**-Wert des Vollverbs, also des semantischen Kopfes, perkoliert bis zum obersten Knoten des Prädikatskomplexes. Käme ein nominaler Bestandteil hinzu, würde dessen **LISTEME**-Wert zu Gunsten des Vollverbs ebenfalls unterdrückt werden. Unser Vorschlag ist deshalb, eine Unterscheidung zwischen dem rein verbalen Komplex und seiner mit nicht verbalen Bestandteilen erweiterten Form zu treffen. Zu diesem Zweck führen wir eine neue Sorte ein, die *extended-head-cluster-structure* (kurz *ext-head-cl-struct*). Das Schema in (188) bleibt unverändert. Ein neues Schema, gezeigt in (199), lizenziert erweiterte Prädikatskomplexe. Diese beinhalten ein internes Argument.

$$(199) \left[\begin{array}{l} \text{phrase} \\ \text{SYNSEM} \left[\text{LOC CAT} \left[\text{SUBCAT } \boxed{\square} \right] \right] \\ \text{DTRS} \left[\begin{array}{l} \text{H-DTR} \left[\text{SYNSEM} \left[\text{LOC CAT SUBCAT } \boxed{\square} \oplus \langle \boxed{\square} \rangle \right] \right] \\ \text{DTRS} \quad \text{ext-head-cl-struct} \vee \text{head-cl-struct} \\ \text{CL-DTR} \left[\text{SYNSEM } \boxed{\square} \text{ LOC } \left[\text{I-ARG-STAT } \textit{internal} \right] \right] \end{array} \right] \end{array} \right]$$

Durch die beiden ID-Schemata werden Prädikatskomplexe wie in Abbildung 5.14 lizenziert.

5.3.1.4 Passivierung

In Abschnitt 2.3.4.1 wurden die Kriterien Dobrovol'skijs zur Idiom-Passivierung vorgestellt. Dort wurde u. a. stark vereinfachend gesagt, dass Idiome dann passiviert werden können, wenn das Verb in freier Verwendung passivierbar ist. Dies hängt mit der Argumentstruktur und dem Status des Verbs zusammen (transitiv, intransitiv oder ergativ), wobei der Begriff des **designierten Argumentes** von zentraler Bedeutung ist (vgl. Haider, 1986). Ein designiertes Argument liegt nur bei transitiven und intransitiven Verben vor, nämlich ihr Subjekt. Ergative Verben (auch als nichtakkusativische Verben bezeichnet) besitzen nur nicht designierte Argumente. Die Passivbildung ist nur bei den Verben möglich, welche ein designiertes Argument besitzen.

Manchmal geht bei der Passivierung jedoch wie bei (200) die idiomatische Lesart verloren – was eine idiosynkratische Eigenschaft eines jeden einzelnen IVP darstellt – oder sie erscheinen anderweitig ungrammatisch, siehe (201).

(200) *den Vogel abschießen* \Rightarrow *!Der Vogel wird (von ihr) abgeschossen.*

(201) a) *Lunte riechen* \Rightarrow *?Lunte wird (von den Dieben) gerochen.*

b) *am Hungertuch nagen* \Rightarrow *?Am Hungertuch wird genagt.*

c) *jdm. das Herz ausschütten* \Rightarrow *!/*Das Herz wird ihm ausgeschüttet.*

Das bedeutet, auch wenn eine Passivbildung des Verbs für sich grammatisch wäre, muss die Möglichkeit der Passivierung für jedes Idiom gesondert untersucht werden. Idiome mit möglicher freier Lesart können im Passiv ihre gebundene Lesart nur zum Teil behalten. Wir müssen daher durch einen geeigneten Mechanismus eine Passivbildung in unserer Herangehensweise ausschließen können.

Wir wollen uns hier nicht in einer Diskussion zum deutschen Passiv verlieren, für einen detaillierten Datenüberblick siehe Müller (2002), für eine Darstellung verschiedener HPSG-Ansätze dazu vgl. Müller (2003b). Die Passivierung richtet sich, wie gesagt, weitestgehend nach der Argumentstruktur. Festzustellen ist aber, dass trotz entsprechender Eigenschaften manche Verben „nicht gerne“ passiviert werden (z. B. *riechen*), bei *haben* in der Bedeutung „besitzen“ ist dies sogar gänzlich ungrammatisch:

(202) **Etwas wird von mir gehabt.* vs.

Etwas wird von mir besessen.

Die Passivierbarkeit scheint also idiosynkratischer zu sein als allgemein angenommen. Da es sich hier um ein Problem handelt, das die Forschung zum deutschen Passiv zu lösen hat, und es dabei nicht primär um idiomatische Wendungen geht, wollen wir an dieser Stelle nur skizzieren, wie die Passiv-Form bei nicht passivierbaren Idiomen ausgeschlossen werden kann.

Wir haben weiter oben davon gesprochen, Verben im Passiv mit Hilfe von Derivationsregeln zu lizensieren. Verfolgt man diese Sichtweise, könnte man den Input für solch eine DR so eingrenzen, dass man nicht passivierbare Wörter ausschließt. Das LISTEME-Merkmal wäre dazu ein geeigneter Kandidat. Man könnte also bei der Eingabe-Spezifikation vermerken, dass es sich beispielsweise um ein Wort handelt, dass nicht das Litem *haben* [LISTEME *-haben*] ist. Mit einer Disjunktion könnten dann weitere Wörter ausgeschlossen werden. Abbildung 5.15 zeigt die Passiv-DR aus Sailer (2003, S. 96), die uns als

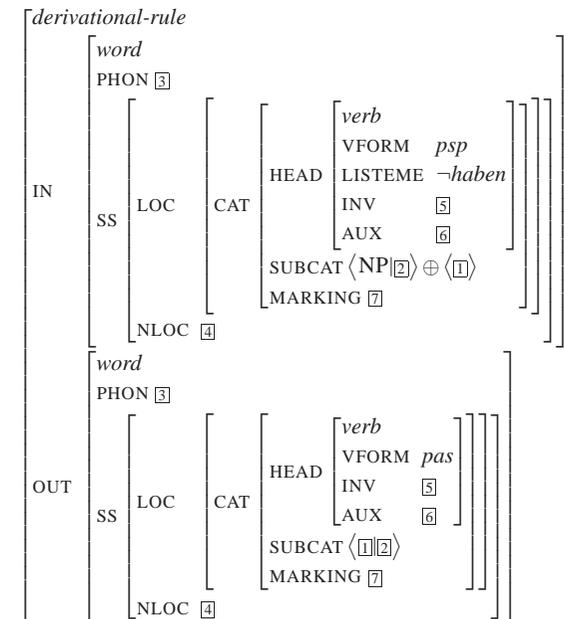


Abbildung 5.15: Passiv-DR, die *haben* ausschließt

Beispiel dienen soll, mit der entsprechenden Anpassung. Zentral bei der DR ist, dass das erste Element der SUBCAT-Liste in der Eingabe entfernt wird und das letzte Element bei der Ausgabe als erstes auf der SUBCAT-Liste steht, also zum Subjekt wird. Bei intransitiven Verben ist das zweite Element bzw. der Rest [2] einfach leer.

In neueren Arbeiten zum deutschen Passiv im Rahmen der HPSG (vgl. Müller, 2002, 2003b) nimmt man an, dass Passivierung mittels Anhebung des Objekts zum Subjekt modelliert werden kann. Der Vorteil dabei ist, dass ein einziges Partizip (der Sorte *ppp*) für Perfekt und Passiv genügt und man nicht mehr zwischen Partizip Passiv (*pas*) und Partizip Perfekt (*psp*) unterscheiden muss. Der Lexikoneintrag des Hilfsverbs (*werden*, *haben* und *sein*, evtl. noch *bekommen*) wird so gestaltet, dass je nach Konstruktion die Argumente des Partizips in geeigneter Weise vom Auxiliar angezogen werden. Die SUBCAT-Liste der Auxiliare beinhaltet also neben dem Haupt- bzw. Vollverb auch dessen subkategorisierte Elemente. Hier muss also ein Ausschluss der ungrammatischen Passivierungsfälle über die Hilfsverben erfolgen, und zwar indem man in ihren Valenzeigenschaften entsprechende Vollverben ausschließt. Dies kann entwe-

der mit Hilfe eines Attributs beim zu passivierenden Verb selbst erfolgen oder über dessen LISTEME-Wert. Geht man den letzteren Weg, werden im Lexikon-eintrag des Auxiliars alle nicht passivierbaren Listeme ausgeschlossen.²⁵

Die Listemeselektion spielt somit auch bei der Passivierung eine Rolle. Bei der Analyse dekomponierbarer IVPs werden für die idiomatischen Wörter eigene Lexikoneinträge spezifiziert. Wenn man einen solchen IVP nicht passivieren kann, können die idiomatischen Verben von der Passivierung in oben skizzierter Weise ausgeschlossen werden.

5.3.1.5 Nominalisierung

Bei IVPs treten, wie schon in Abschnitt 2.3.4.3 gesehen, Nominalisierungen auf. Dabei spielt es – wie auch bei der Passivierung – offenbar keine Rolle, ob ein IVP dekomponierbar wie bei (203) oder nicht dekomponierbar wie bei (204) ist, um sich einer Nominalisierung unterziehen zu können (Belege siehe Appendix A).

- (203) a) *etw. ausfindig machen* \Rightarrow *das Ausfindigmachen*
 b) *ins Fettnäpfchen treten* \Rightarrow *der Fettnäpfchentreter*
 c) *jdm. eine Abfuhr erteilen* \Rightarrow *die Abfuhrerteilung*
- (204) a) *Maulaffen feilhalten* \Rightarrow *das Maulaffenfeilhalten*
 b) *das Tanzbein schwingen* \Rightarrow *der Tanzbeinschwinger*
 c) *jdn. zur Strecke bringen* \Rightarrow *das Zur-Strecke-Bringen*

Wir finden mehrere Bildungsmuster vor: die Konversion eines Verbs mit Integration von Adverbien (203a), Komplementen (204a) oder anderen internen Argumenten (204c) und weiterhin die Derivation eines Verbs in ein Nomen actionis (203c) oder ein Nomen agentis (beide b), wo jeweils auch die internen Komplemente in den Nominalkopf integriert werden. Bei dieser Integration kann es dazu kommen, dass das Wort als Einheit ohne Zwischenräume geschrieben wird oder eine sogenannte Durchkupplung (204c) auftritt.²⁶

Bei der Nominalisierung fällt weiterhin auf, dass die DP-Hülle um das Komplement wegfällt (*der Tanzbeinschwinger*, nicht **der Das-Tanzbein-Schwinger*).

²⁵Nimmt man eine elaboriertere Hierarchie der *listeme*-Sorten an, können alle nicht passivierbaren Listeme unter einer Obersorte zusammengefasst werden.

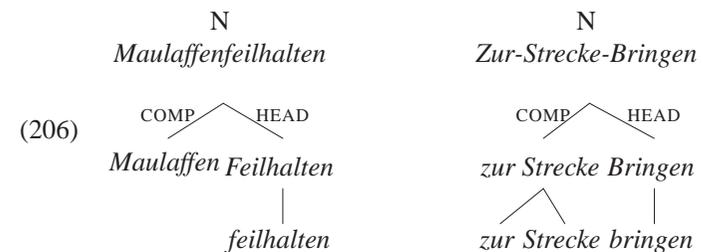
²⁶Siehe Gallmann (1985, 1990, 1997 und 1999) zu einer ausführlichen Diskussion von Nomen und Nominalisierungen, deren graphematischer Umsetzung und deren interner Strukturen.

ger) und manchmal auch die Präpositionen (*der Fettnäpfchentreter*, nicht *?der Ins-Fettnäpfchen-Treter*). In diesem Falle handelt es sich jedoch um eine Ausnahme, da normalerweise die Präpositionen der internen PPs nicht wegfällen, siehe (204c). Da wir gerade bei Ausnahmen sind: Nominalisierung ist bei IVPs ein hoch idiosynkratisches Phänomen. Beispielsweise gibt es (205a), aber nicht (205b) mit idiomatischer Lesart, obwohl es sich um eine ganz parallele Bildung handelt, die beide Male Metonymie einschließt.

- (205) a) *das Gesicht verlieren* \Rightarrow *der Gesichtsverlust*
 b) *den Kopf verlieren* \Rightarrow *! der Kopfverlust*

Ein wichtiger Punkt ist die Tatsache, dass bei der Nominalisierung von IVPs grundsätzlich nur interne Argumente integriert werden. Bei normaler Nominalisierung spielt das hingegen keine Rolle. Beispielsweise wird bei *Maulaffen feilhalten* das Akkusativobjekt integriert, bei *etw. ausfindig machen* das Adverb und nicht das Akkusativobjekt. Bei einem nicht idiomatischen Verb wie *etw. putzen* (beispielsweise *das Fensterputzen*) ist das Akkusativobjekt innerhalb der Nominalisierung, Adverbien sind es nur zum Teil (**das Gründlich-putzen*, aber: *das Synchronschwimmen*).

Was die Beschreibung der Nominalisierung betrifft, kann man feststellen, dass interne Objekte der ursprünglichen Verben zum Wortteil des nominalisierten Verbs werden. Die Objekte nehmen also keine syntaktische Position mehr ein, sondern eine morphologische. Wir folgen Reinhard (2001) und nehmen an, dass es wie auf syntaktischer Ebene auch auf morphologischer Ebene Kopf-Komplement-Strukturen sowie Kopf-Adjunkt-Strukturen gibt. So, wie z. B. *Maulaffen* ein Komplement des Verbs ist, ist es in der Nominalisierung auch ein Komplement, hier das eines Morphems. Die Analyse des syntaktischen Komplements als Wortteil, wobei dieses nicht mehr projizieren kann und keine Determinantien enthält, ist kompatibel mit den Beobachtungen von Meibauer (2003).



Reinhard gibt in ihrer Dissertation einen Überblick über die verschiedenen Ansätze zur Wortbildung. Dabei kann man zwischen drei Theorien unterscheiden: Wortbildung ausschließlich in der Syntax, Wortbildung als lexikalischem Prozess und Wortbildung als Eigenschaft des konzeptuellen Systems. Reinhard nimmt, v. a. Selkirk folgend, eine morphologische Ebene an, wobei die Wortbildung sowohl auf morphologischer als auch auf syntaktischer Ebene vor sich geht. Sie entwickelt eine morphologische Komponente für die HPSG,

die es erlaubt, Derivation, Komposition und Flexion in einem einheitlichen Rahmen zu modellieren. Insbesondere können produktive Wortbildungsprozesse und spezifische Probleme wie Fugenelemente und interne Flexion bei der Komposition, wortinitiale Kopfderivation sowie flexivische Merkmalsperkolation behandelt werden. Für die Modellierung wurde ein Konstituentenstruktur- oder wortsyntaktischer Ansatz gewählt, in dem sowohl Stämme als auch Affixe eigene Lexikoneinträge besitzen und Affixe Köpfe sind. (Reinhard, 2001, S. 202)

Sie beschreibt komplexe Morpheme bzw. Wörter mit Hilfe von *Immediate Dominance Schemas*, wobei sie diese an die Gegebenheiten der Morphologie anpasst und somit neben den normalen syntaktischen ID-Schemata nun auch morphologische ID-Schemata annimmt.

Ihre Schemata zusammen mit den von ihr eingeführten Merkmalsstrukturen für die Morphologie können die interne Struktur der Nominalisierungen beschreiben. Allerdings räumt sie ein:

Ein Standardargument gegen die Annahme einer Wortsyntax ist vor allem die Notwendigkeit der Postulierung von Null-Morphemen für die Konversion. In der hier vorgestellten HPSG-Komponente stellt dieses jedoch kein Problem dar, denn außer über die Annahme eines Null-Affixes, was in der Tat nicht erstrebenswert ist, kann die Konversion auch über eine Lexikalische Regel oder über unterspezifizierte Einträge behandelt werden. Prinzipiell scheinen beide Möglichkeiten adäquat zu sein, daher wurde hier keine Entscheidung für oder gegen eine der beiden Varianten getroffen. (Reinhard, 2001, S. 202)

Da eine Diskussion der Nominalisierung eine ganze Dissertationsschrift oder mehr füllen kann,²⁷ wollen wir uns hier auf die Konversion von IVPs beschränken. Der Standard-HPSG folgend, benutzen wir Derivationsregeln, um einen

²⁷Reinhard untersucht u. a., bei welchen semantischen Verbklassen eine Nominalisierung überhaupt möglich ist. Die Kriterien dafür sind Aktionsart, thematische Argumentstruktur,

„konvertierten“ IVP zu lizensieren. Zur Erinnerung sei erwähnt, dass Derivationsregeln folgendes bedeuten: Wenn es ein bestimmtes Objekt gibt (der IN-Wert), gibt es auch ein bestimmtes anderes Objekt (der OUT-Wert).

Es sollen an dieser Stelle jedoch noch keine konkreten Vorschläge zur Definition von DRs gemacht werden, da sonst zu viel von der IVP-Analyse vorweggenommen werden müsste bzw. es zu Redundanzen in der Darstellung käme. Zur Analyse siehe die Kapitel 6.1.2 und 6.2.2.

Gewissermaßen als Überleitung zum nächsten Themenblock wollen wir die **Semantik** von Nominalisierungen näher betrachten. Ehrlich (1991) geht in ihrem Artikel auf die Bedeutung des Wortbildungsmusters ein, wobei sie die Frage stellt, ob man Nominalisierungen als Ereignisse oder Tatsachen interpretieren soll. Wichtig ist dabei eine Unterscheidung, die bei englischen Nominalisierungen eine Rolle spielt: verbale Gerundien (207), welche Tatsachen denotieren, und nominale Gerundien (208), welche als Ereignisse interpretiert werden.

(207) *Peter's singing the Marseillaise was surprising.*

(208) *I heard Peter's singing of the Marseillaise.*

Die Frage, welcher der beiden Arten unser deutsches Nominalisierungsmuster der Konversion entspricht, ist zwar nicht leicht zu beantworten, es spricht aber einiges dafür, sich für die nominalen Gerundien bzw. die Denotation eines Ereignisses zu entscheiden. Folgende Daten, die Tests für die Unterscheidung der beiden englischen Gerundienarten beinhalten, stützen diese These. Dabei sind die Ausdrücke ... *fand zum Zeitpunkt x statt* und *dauerte x Wochen* sogenannte *narrow container*, die nur bei Ereignissen und nicht bei Tatsachen grammatisch sind, vgl. (212) und (213).

(209) **das Den-Täter-zur-Strecke-Bringen*

(210) *das Zur-Strecke-Bringen des Täters*²⁸

(211) *Das Zur-Strecke-Bringen des Täters ist unwahrscheinlich.*

(212) *Das Zur-Strecke-Bringen des Täters fand gegen Mitternacht statt.*

tur, Linkingverhalten und aspektuelle Interpretation. Daneben geht sie auf Anhebungsbeschränkungen auf lexikalischer Ebene, Realisierungsbedingungen für Argumente, Argumentvererbungsmuster etc. ein.

²⁸analog zum Unterschied *singing the Marseillaise* und *singing of the Marseillaise*

(213) *Das Ausfindigmachen des Liebhabers kostete den Detektiv zwei Wochen.*

Ehrich stellt ebenfalls drei Ansätze zur formalen Semantik von Nominalisierungen vor (Typenlogik, Ereignis-Semantik nach Cresswell und eine Ereignis-ontologie im λ -Format nach Bierwisch), kommt aber zu folgendem Schluss:

Andererseits steht eine umfassende semantische Analyse noch aus, in der die kategoriale Bedeutung der verschiedenen Nominalisierungsarten (Infinitiv- vs. Derivativenominalisierung, verbale vs. nominale Infinitive) ebenso berücksichtigt ist wie die Aspektfunktion von Nominalisierungen und der wortsemantische Beitrag, den die Verbaktionsarten zur Interpretation von Nominalisierungen leisten. [...] Eine überzeugende Analyse wird ferner die ontologische Grundfrage nach dem Status von Ereignissen (Individuen vs. Eigenschaften) eindeutig und im Hinblick auf die verschiedenen Nominalisierungsarten zu beantworten haben. (Ehrich, 1991, S. 457)

Hamm und van Lambalgen räumen der Nominalisierung eine fundamentale Rolle in ihrer Studie der Tempus- und Aspektsemantik ein, siehe Hamm/van Lambalgen (2003) bzw. noch ausführlicher ist van Lambalgen/Hamm (2004). Die Autoren entwerfen eine formale Theorie, die sich an den Ereigniskalkül (*event calculus*), wie er in der Forschung zur Künstlichen Intelligenz Anwendung findet, anlehnt. Zur Frage nach dem Status von Ereignissen schlagen sie sich auf die Seite Freges, der sie als Individuen ansieht.

Hamm/van Lambalgen bezeichnen verbale Gerundien (Tatsachen) als *fluents* bzw. *imperfect nominals* und nominale Gerundien (Ereignisse) als *event types* oder auch *perfect nominals*. Für Ereignisse führen sie ein Prädikat *happens'* ein, das als Argumente eine Nominalisierung und eine Zeitvariable enthält. So wird ausgedrückt, dass ein bestimmtes Ereignis zu einem bestimmten Zeitpunkt „passiert“. Ein *perfect nominal* ist dadurch gekennzeichnet, dass die Zeitvariable des nominalisierten Verbs gebunden wurde, weshalb man *perfect nominals* nicht intern temporal modifizieren kann.

Im Gegensatz zu den Ereignissen nimmt das „Tatsachen“-Prädikat *holdsAt'* neben der Zeitvariable ein *imperfect nominal* als Argument; eine Tatsache „gilt“ also zu einem bestimmten Zeitpunkt.

Der Umfang der in diesem Abschnitt zitierten Arbeiten und der darin gegebene Ausblick auf die Fülle der Probleme im Zusammenhang mit Nominalisierungen zwingen uns, nur die notwendigen Aspekte herauszupicken, die für

unsere Idiom-Analyse relevant sind. Dabei benötigen wir vor allem das Prädikat *happens'*, dessen Herkunft oben kurz erklärt wurde. Für eine Definition geeigneter Derivationsregeln sei auf die Kapitel 6.1.2 und 6.2.2 verwiesen.

5.3.2 Ressourcen, Restriktionen und Referenzen

5.3.2.1 Semantik: LRS

In dieser Arbeit stehen die syntaktischen Gesichtspunkte von Phraseologismen im Mittelpunkt. Bei der Analyse von IVPs und der Skizzierung von Lexikoninträgen kommen wir jedoch, wie auch bei der Nominalisierung gesehen, um einige semantische Aspekte nicht herum. Um nun Attribute und Werte, welche die Semantik eines Ausdrucks charakterisieren, schlüssig und konsistent anzugeben, müssen wir uns auf einen HPSG-Ansatz zur Semantik festlegen. In diesem Abschnitt sollen daher Grundzüge eines Semantik-Moduls dargestellt werden, soweit sie für unsere IVP-Analyse relevant sind.

Da die Semantik in Pollard/Sag (1994) nur sehr unzureichend skizziert wurde, muss man sich für ein ausgefeilteres Semantik-Modul für die HPSG-Analyse entscheiden. Wir bedienen uns in dieser Arbeit der *Lexical Resource Semantics* (LRS).

Wie in der Einführung angedeutet, geht es hier um eine formale Analyse und nicht so sehr um eine schnell umzusetzende computerlinguistische Herangehensweise. Für letzteres ist die *Minimal Recursion Semantics* (MRS, Copestake et al., 1998), mit der das System LKB (Copestake, 2002) beispielsweise arbeitet, geeigneter. In der Tat wurde MRS zuerst für die Computerlinguistik entwickelt. MRS beruht auf Techniken der unterspezifizierten Semantik und verwendet Repräsentationen, die nicht Teil einer bestimmten semantischen Theorie sind. Damit sind diese für robuste semantische Analysen geeignet, die in einem weiteren Schritt in spezifischere, modelltheoretisch interpretierbare Repräsentationen überführt werden können. Es ist einerseits der große Vorteil von LKB wie der MRS, dass dem Linguisten ein ausreichend restringiertes Grundrepertoire an Optionen geboten wird, aber andererseits werden dadurch seine analytischen Möglichkeiten in einer Weise eingeschränkt, die zwar für praktische Entwicklungen sinnvoll ist, für exemplarische Untersuchungen wie die unsrige jedoch zu restriktiv bleibt.

Mit LRS ist man dagegen freier in der Wahl der semantischen Repräsentationsprache und expliziter in der Beziehung zwischen Grammatik und semantischer Interpretation. Zudem ist bei der Entwicklung von LRS die modelltheoretische Interpretierbarkeit ein wichtiger Faktor. Es scheint deshalb so, dass LRS für unsere Absichten die bessere Wahl ist.

LRS vereint in sich Techniken der linguistischen Beschreibung im Rahmen einer beschränkungs-basierten Grammatik wie der HPSG und eine Art der semantischen Repräsentation, wie sie Semantikern aus anderen Grammatiktraditionen vertraut ist. Dadurch wird mit LRS ein Beitrag geleistet, die Semantik natürlicher Sprache in einer theoretisch adäquaten Art zu spezifizieren. Gleichzeitig ist der Ansatz auch in der Computerlinguistik brauchbar und praktikabel. Wir wollen im Folgenden die Grundlagen von LRS vorstellen und stützen uns dabei auf Richter/Sailer (2004), mit diversen Vereinfachungen. Um LRS zu benutzen, ist es unabdingbar, einige Details dieses Semantik-Moduls anzugeben, was hier nur in gebührender Kürze geschehen kann.

LRS verwendet gegenwärtig die semantische Repräsentationssprache Ty2 für natürlichsprachliche Semantik (vgl. Gallin, 1975), welche ähnliche Vorzüge wie die Intensionale Logik von Montague hat. Man bedient sich der Unterspezifikation (z.B. bei Skopus-Ambiguitäten), jedoch ist die logische Form eines Satzes ein einziger disambiguiertes Ausdruck. Richter/Sailer (2004) unterscheiden zwischen dem eigenen Semantikbeitrag eines Wortes und der Semantik einer Phrase. Die Autoren argumentieren dafür, dass die Semantik, welche die Phrase betrifft, nicht lokal verfügbar sein sollte. Gerade im Blick auf Skopus-Ambiguitäten, die ja auf Satzebene auftreten und nicht Teil der Wortsemantik sind, ist dieses Vorgehen motiviert. Dies wollen wir bei der nun folgenden Angabe der semantischen Attribute und Sorten berücksichtigen.²⁹

Die Semantik an sich wird in logischen Formen ausgedrückt, welche aus semantischen Konstanten, Variablen, Quantoren und logischen Operatoren bestehen. So ist die Bedeutung von *lesen* $\text{lesen}'(x,y)$.³⁰ Die beiden Lesarten des Satzes „Jeder Student liest ein Buch.“ finden sich in (214).

(214) Jeder Student liest ein Buch.

$$\begin{aligned} &\forall x[\text{student}'(x) \rightarrow \exists y[\text{buch}'(y) \wedge \text{lesen}'(x,y)]] \\ &\exists y[\text{buch}'(y) \wedge \forall x[\text{student}'(x) \rightarrow \text{lesen}'(x,y)]] \end{aligned}$$

Man muss nun mit Hilfe von geeigneten Attributen diese logischen Formen in die HPSG-Grammatik integrieren:

$$(215) \begin{bmatrix} \textit{lrs} \\ \text{EXTERNAL-CONTENT} & \textit{me} \\ \text{INTERNAL-CONTENT} & \textit{me} \\ \text{PARTS} & \textit{list(me)} \end{bmatrix}$$

²⁹Die Unterscheidung zwischen lokaler und nicht lokaler Semantik, die in LRS umgesetzt wird, ist auch der Grund, warum wir nicht wie beispielsweise Sailer (2003) auf LF-Ty2 zurückgreifen.

³⁰In dieser Einführung in LRS legen wir der Einfachheit halber ein extensionales Fragment zu Grunde und lassen Fragen zur Intension beiseite.

Die Sorte *Irs*, der Wert von LF unter *sign*, hat drei Attribute. Der EXTERNAL-CONTENT, kurz EXCONT oder EXC, hat den Wert *meaningful-expression*³¹, ist also ein bestimmter bedeutungstragender Ausdruck. Er beinhaltet die Bedeutung der maximalen Projektion des Kopfes. Somit ist der EXCONT-Wert einer Äußerung ihre logische Form, der EXCONT-Wert einer NP ist der Operator, der die referenzielle Variable des Kopfnomens bindet.

Im INTERNAL-CONTENT (oder INCONT bzw. INC) steht der semantische Beitrag des semantischen Kopfes mit kleinstem Skopus. Die Liste unter PARTS beinhaltet bei *utterances* genau die Elemente von EXCONT, also alle (Teil-)Ausdrücke, die zur Bedeutung beitragen. Die folgenden Spezifikationen werden für Lexikoneinträge von Wörtern angenommen:

(216) Lexikalische Spezifikation der LF-Werte:

$$\begin{array}{l} \textit{lesen}: \begin{bmatrix} \textit{lrs} \\ \text{EXC} & \textit{me} \\ \text{INC} & \boxed{1}(\text{lesen}'y)x \\ \text{PARTS} & \langle x, y, \boxed{1}, \boxed{1a}[\text{lesen}'y, \boxed{1b}[\text{lesen}'']] \rangle \end{bmatrix} \\ \\ \textit{Buch}: \begin{bmatrix} \text{EXC} & \begin{bmatrix} \textit{quantifier} \\ \text{VAR } y \end{bmatrix} \\ \text{INC} & \boxed{2}[\text{buch}'(y)] \\ \text{PARTS} & \langle y, \boxed{2}, \boxed{2a}[\text{buch}''] \rangle \end{bmatrix} \\ \\ \textit{ein}: \begin{bmatrix} \text{EXC} & \textit{me} \\ \text{INC} & \boxed{4}\exists y[\alpha \wedge \beta] \\ \text{PARTS} & \langle y, \boxed{4}, \boxed{4a}[\alpha \wedge \beta] \rangle \end{bmatrix} \end{array} \quad \begin{array}{l} \textit{Student}: \begin{bmatrix} \text{EXC} & \begin{bmatrix} \textit{quantifier} \\ \text{VAR } x \end{bmatrix} \\ \text{INC} & \boxed{3}[\text{student}'(x)] \\ \text{PARTS} & \langle x, \boxed{3}, \boxed{3a}[\text{student}''] \rangle \end{bmatrix} \\ \\ \textit{jeder}: \begin{bmatrix} \text{EXC} & \textit{me} \\ \text{INC} & \boxed{5}\forall x[\gamma \rightarrow \delta] \\ \text{PARTS} & \langle x, \boxed{5}, \boxed{5a}[\gamma \rightarrow \delta] \rangle \end{bmatrix} \end{array}$$

Im LE von *lesen* in (216) ist der EXCONT-Wert unterspezifiziert. Er muss lediglich ein Ty2-Ausdruck (*me*) sein. Der INCONT-Wert ist der semantische Beitrag des Verbs zusammen mit seinen Argumenten. Um im Term die hier angenommene Struktur explizit zu machen, wird $(\text{lesen}'y)x$ an Stelle von $\text{lesen}'(x,y)$ geschrieben. Dieser Ausdruck hat also die Teile *x*, *y*, lesen' , $\text{lesen}'y$ und $(\text{lesen}'y)x$. Die PARTS-Liste enthält genau all diese Unterausdrücke.

Die INCONT- und PARTS-Beschreibungen in den LEs der Nomen *Buch* und *Student* sind analog zu denen für *lesen*. Im EXCONT-Wert jedoch spezifizieren wir, dass es einen Quantor gibt, der die Variable im INCONT-Wert bindet.

In den LEs der Quantoren *ein* und *jeder* beschreibt der INCONT-Wert nicht eindeutig einen Ty2-Ausdruck, sondern ist eine Beschreibung, die nur Teile eines

³¹Man kann diese *meaningful-expressions* auch in AVM-Schreibweise notieren, wozu man weitere Sorten und Prinzipien benötigt. Da wir aber nur die abkürzende Schreibweise verwenden wollen, gehen wir auf diese Hintergründe nicht ein.

Ausdrucks festlegt. Im Fall *jeder* wird gesagt, dass der INCONT-Wert ein Ausdruck ist, der mit einem Allquantor über einer Variable x beginnt, die unmittelbaren Skopus über eine Implikation hat. Über die weiteren Unterausdrücke wird nichts ausgesagt. Wir benutzen griechische Kleinbuchstaben für solche un spezifizierten Teilausdrücke. Diese Zeichen sind formal nichts anderes als tags, werden aber nur für Ty2-Ausdrücke verwendet. Zur Verdeutlichung ist in (217) eine reine AVM-Notation des LF-Wertes von *jeder* dargestellt.

$$(217) \left[\begin{array}{l} lrs \\ \text{EXCONT } me \\ \text{INCONT } \boxed{S} \left[\begin{array}{l} \text{universal} \\ \text{VAR } \boxed{Sb} \text{ var} \\ \text{SCOPE } \boxed{Sa} \text{ implication} \end{array} \right] \\ \text{PARTS } (\boxed{S}, \boxed{Sa}, \boxed{Sb}) \end{array} \right]$$

Die Sorte *lrs* bekommt nun noch zwei Wohlgeformtheits-Beschränkungen: das INTERNAL-CONTENT-PRINZIP (IContP) in (218) und das EXTERNAL-CONTENT-PRINZIP (EContP) in (219).

(218) Das INTERNAL-CONTENT-PRINZIP (IContP):

In jedem Objekt der Sorte *lrs* ist der INCONT-Wert ein Element der PARTS-Liste und eine Komponente des EXCONT-Wertes.

Das IContP stellt sicher, dass der INCONT-Wert eines Wortes als Teil der PARTS-Liste erscheint. Infolgedessen ist die Semantik eines Elementes mit kleinstem Skopus auch ein semantischer Beitrag zur gesamten logischen Form. Darüber hinaus wird gefordert, dass der INCONT-Wert ein Teilausdruck des EXCONT-Wertes sein muss. Dies bindet die syntaktischen und semantischen Eigenschaften eines Zeichen näher aneinander: Der EXCONT-Wert ist in jeder Projektion des Kopfes identisch und wird als logische Form eines Zeichens interpretiert. Durch das IContP ist die Semantik des lexikalischen Kopfes auch Teil der logischen Form der Projektion.

(219) Das EXTERNAL-CONTENT-PRINZIP (EContP):

1. In jeder Phrase ist der EXCONT-Wert der Nichtkopftochter ein Element der PARTS-Liste der Nichtkopftochter.
2. In jeder Äußerung ist jeder Teilausdruck des EXCONT-Wertes der Äußerung ein Element ihrer PARTS-Liste und jedes Element dieser PARTS-Liste ist ein Teilausdruck des EXCONT-Wertes.

Das EContP legt weitere Beschränkungen auf den EXCONT-Wert eines Zeichens. In der ersten Klausel (für Phrasen) wird festgelegt, dass der EXCONT-Wert einer vollständigen Kopfprojektion (entweder ein Komplement oder eine Äußerung – Klausel 2) ein Element auf der PARTS-Liste des Zeichens ist. In Kombination mit dem IContP gibt uns Klausel 1 eine generelle Spezifikation des EXCONT-Wertes einer vollständigen Kopfprojektion: Es muss ein Ausdruck sein, der Teil des semantischen Beitrages (auf PARTS) eines dominierten Zeichens ist, und der Ausdruck muss die Semantik des lexikalischen Kopfes als Teilausdruck enthalten.

In Klausel 2 wird sichergestellt, dass kein semantischer Beitrag eines Elementes in der Äußerung verloren geht.

Das nächste wichtige Prinzip ist das LRS-PROJEKTIONSPRINZIP, dargestellt in (220). Es stellt sicher, dass die Beiträge der einzelnen Zeichen zur Gesamtbedeutung alle berücksichtigt werden.

(220) LRS-PROJEKTIONSPRINZIP:

In jeder *headed-phrase*

1. sind die EXCONT-Werte des Kopfes und des Mutterknotens identisch,
2. sind die INCONT-Werte des Kopfes und des Mutterknotens identisch³² und
3. der PARTS-Wert beinhaltet alle und nur die Elemente der PARTS-Liste der Töchter.

Nun folgt noch die Definition des SEMANTIK-PRINZIPS (SP), illustriert in (221)³³. Das Prinzip spezifiziert Beschränkungen darauf, wie die Bedeutungskomponenten von unterschiedlichen syntaktischen und semantischen Töchtern kombiniert werden können. Bei Richter/Sailer (2004) ist das LRS-PROJEKTIONSPRINZIP im SEMANTIK-PRINZIP enthalten. Die (sinnvolle) Trennung der Prinzipien nehmen Penn/Richter (2004) vor.

³²Es wird hier angenommen, dass das Nomen der Kopf einer quantifizierten NP ist. Anm.: Eine quantifizierte NP ist jede „normale“ NP. Eigennamen beispielsweise sind nicht quantifizierte NPs.

³³ $Qx[\rho \circ v]$ in der Illustration ist eine Abkürzung für

$$\left[\begin{array}{l} \text{quantifier} \\ \text{VAR } x \\ \text{SCOPE } \left[\begin{array}{l} l\text{-const} \\ \text{ARG1 } \rho \\ \text{ARG2 } v \end{array} \right] \end{array} \right]$$

(221) SEMANTIK-PRINZIP:

1. Wenn der Nichtkopf ein Quantor ist, dann ist sein EXCONT-Wert von der Form $Qx[\rho \circ v]$, der INCONT-Wert des Kopfes ist ein Teil von ρ und die EXCONT-Werte beider Töchter sind identisch.
2. Wenn der Nichtkopf eine quantifizierte NP mit EXCONT-Wert der Form $Qx[\rho \circ v]$ ist, dann ist der INCONT-Wert des Kopfes ein Teil von v .
3. In einer *head-adjunct-phrase* ist der EXCONT-Wert des Nichtkopfes ein Teil des EXCONT-Wertes des Kopfes.
 - a) Ist der Nichtkopf ein intersektiver Modifikator, dann ist sein EXCONT-Wert von der Form $\alpha \wedge \beta$ und der INCONT-Wert des Kopfes ist Teil von β .
 - b) Ist der Nichtkopf kein intersektiver Modifikator, dann ist sein EXCONT-Wert von der Form $\alpha(\beta)$ und der INCONT-Wert des Kopfes ist Teil von β .
4. ...

In Abbildung 5.16 (S. 142) ist die syntaktische und semantische Struktur des ambigen Satzes in (214) dargestellt, allerdings führen wir der Einfachheit halber das Beispiel in Verbendstellung an. An den Knoten stehen die LF-Werte in AVM-Notation und, wo nötig, werden die Beschränkungen durch das SP angegeben. Zu diesem Zweck verwenden wir das Symbol \triangleleft für die Relation „ist Teil/Unterausdruck von“. Die *tags* sind die gleichen wie in den LEs in (216). An den beiden NP-Knoten verlangt Klausel 1, dass der INCONT-Wert des Kopfes ($student'(x)$ und $book'(y)$) im Restriktor des Allquantors bzw. des Existenzquantors ($\exists \triangleleft \gamma$ und $\forall \triangleleft \alpha$) steht. Bei den Knoten VP und S ist es Klausel 2, die den INCONT-Wert des verbalen Kopfes ($lesen'(x,y)$) in den Skopus der beiden Quantoren bringt: $\forall \triangleleft \beta$ am VP-Knoten und $\exists \triangleleft \delta$ am S-Knoten.

Als Effekt von EContP können wir die EXCONT-Werte der Nichtköpfe in allen Phrasen angeben: An den NP-Knoten ist der EXCONT-Wert der Determinantien identisch mit deren INCONT-Werten, da dies die einzigen Elemente auf den PARTS-Listen sind, die den INCONT-Wert als Teilausdruck beinhalten. An den VP-Knoten (und analog auch an S) muss der EXCONT-Wert der NP *ein Buch* identisch mit dem des Determinierers sein: Es ist das einzige Element auf der PARTS-Liste der NP, das die Bedingung aus dem LE von *Buch* erfüllt, d. h. dass der EXCONT-Wert ein Quantor sein muss, der die Variable y bindet.

Am S-Knoten kommt die zweite Klausel von EContP zur Anwendung: Die Ausdrücke in der PARTS-Liste spezifizieren exakt die Knoten im Ableitungsbaum der resultierenden logischen Form. Es gibt genau zwei mögliche Ableitungen, die die Bedingungen aus den Lexikoneinträgen und den anderen Beschränkungen vom SP erfüllen: Entweder ist \exists identisch mit \forall für die $\forall\exists$ -Lesart oder \forall ist identisch mit \exists für die $\exists\forall$ -Lesart.

Wir wollen es bei diesem Grundriss von LRS belassen. Es fehlt lediglich noch die Spezifikation des CONTENT-Wertes, da bisher nur das Attribut LF besprochen wurde. In Soehn (2003) wurde das Attribut INDEX für alle Wortarten definiert, außerdem wurden dort die Attribute PERSON, NUMBER und GENDER unter PHI zusammengefasst. Wir wollen dies beibehalten in dem Sinne, dass alle Wortarten die gleichen Attribute besitzen.

Die Sorte *content* hat noch Untersorten wie beispielsweise *nominal-object* oder *personal-pronoun*. Da nun alle Wortarten die gleichen Attribute besitzen, können wir im CONTENT auf die Sorte *psoa* (*parameterized-state-of-affairs*) im Sinne von Pollard/Sag (1994) verzichten, welche dort für die Verbsemantik gebraucht wurde.³⁴ Für Verben führen wir *eventuality* als Untersorte von *content* ein. Sämtliche Untersorten sind im Appendix B dargestellt.

Sailer (2004a, S. 205f) folgend, nehmen wir für *content* weiterhin die Attribute VAR und MAIN an. Die neue Merkmalsgeometrie ist in Abb. 5.17 (S. 143) dargestellt.

Unter MAIN steht die semantische Konstante des Objekts. Der Wert von VAR ist bei Nomen das referenzielle semantische Argument des MAIN-Wertes, welches später durch Quantoren gebunden wird. Während bei Verben unter VAR die Davidson'sche Ereignisvariable e steht, werden die Attribute unter PHI nicht benötigt. Man kann entweder eine erweiterte Sortenhierarchie unter PHI annehmen, um für Verben keine Attribute PERSON, NUMBER und GENDER zu haben (vgl. Abschnitt 5.3.2.3), oder anderweitig sicherstellen, dass diese Attribute bei Verben in der Grammatik keine Rolle spielen.

Die Angaben unter INDEX und MAIN sind somit lokal zugänglich, alles andere (ECONT, ICONT und PARTS) ist unter LF auf Zeichenebene. Diese Architektur birgt deshalb eine restriktivere Selektionstheorie in sich, die jedoch mit den empirischen Befunden völlig kompatibel ist. Darüber hinaus lässt sich diese Architektur mit der HPSG-Bindungstheorie³⁵ und Ansätzen zur Argumentbindung (*linking*) nahtlos vereinen. Um die Unterscheidung zwischen lokaler

³⁴Ein Beispiel für ein *psoa* (Pollard/Sag, 1994, S. 29):

<i>psoa</i>	RELATION	<i>see</i>
SEER		<i>index</i>
SEEN		<i>index</i>

³⁵Wir beziehen uns hier auf die Bindungstheorie von Pollard/Sag (1994).

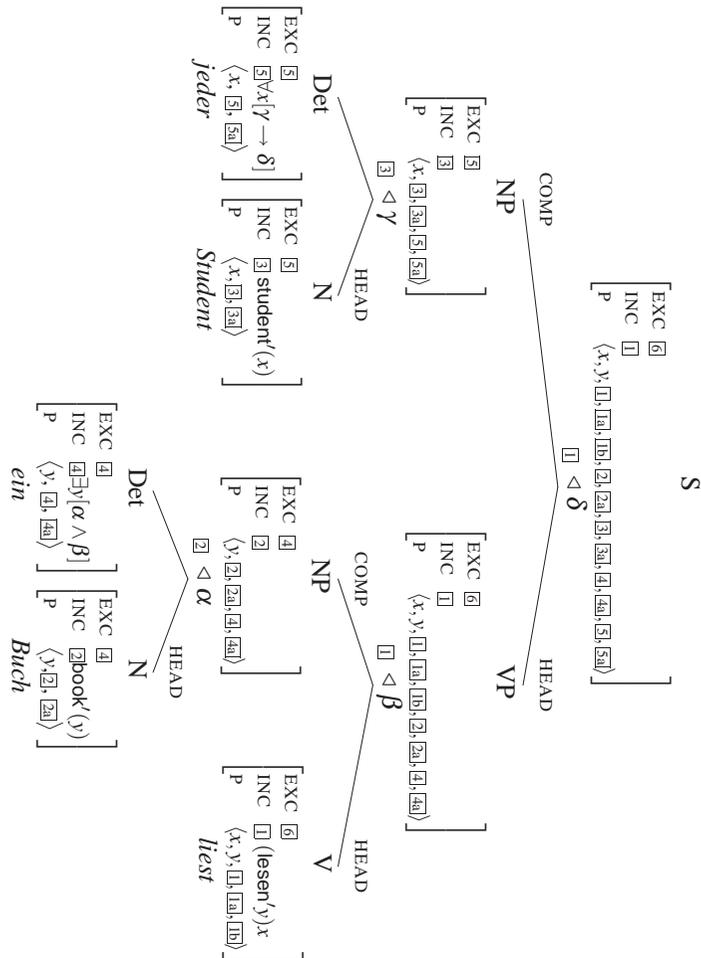


Abbildung 5.16: LRS Analyse von *jeder Student ein Buch liest*

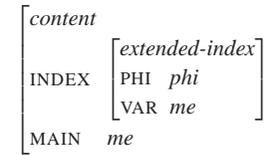


Abbildung 5.17: Merkmalsgeometrie von *content*

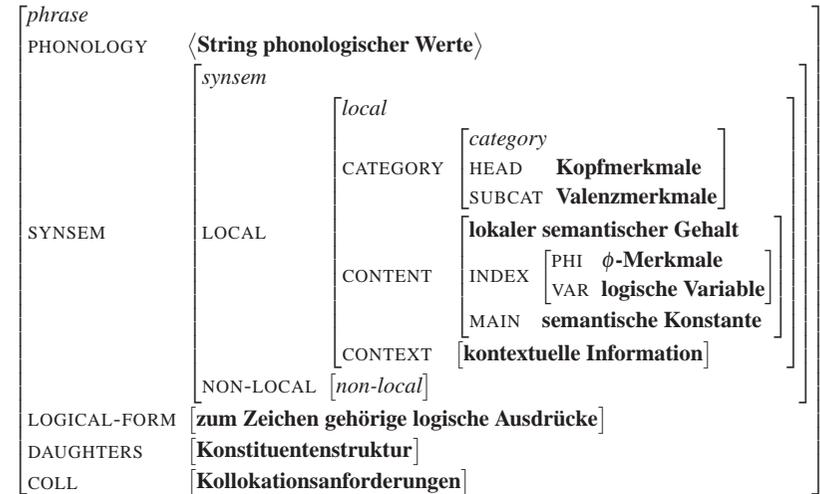
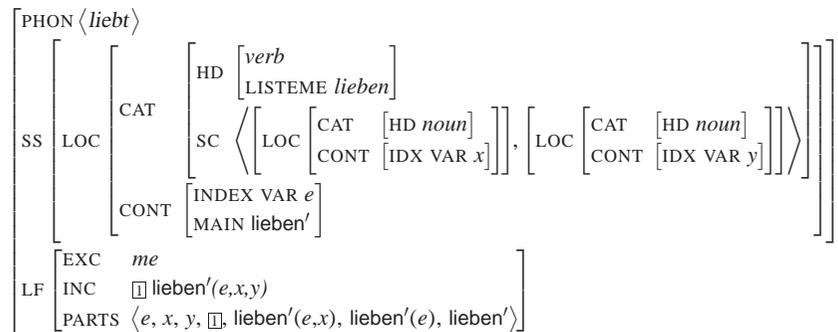


Abbildung 5.18: Lokale und nicht lokale Semantik

und nicht lokaler Semantik noch einmal zu veranschaulichen, sei die AVM in Abb. 4.1 in veränderter Form (Abb. 5.18) und zudem der gegenüber dem einführenden illustrierenden Beispiel vervollständigte LE von *liebt* (Abb. 5.19, S. 144) dargestellt.

5.3.2.2 Selektionsrestriktionen

Bisherige HPSG-Ansätze Es gibt nicht viele Veröffentlichungen über Selektionsrestriktionen in der HPSG, lediglich Nerbonne (1996) und Androutsopoulos/Dale (2000) sind uns bekannt. Androutsopoulos/Dale definieren solche Beschränkungen folgendermaßen:

Abbildung 5.19: Skizze des LEs von *liebt*

Selectional restrictions are semantic sortal constraints imposed on the participants of linguistic constructions to capture contextually-dependent constraints of interpretation. (Androutsopoulos/Dale, 2000, S. 1)

Sie heben hervor, dass Selektionsrestriktionen bei der Disambiguierung von Wortbedeutungen, syntaktischen Konstruktionen und anaphorischen Beziehungen eine große Rolle spielen. Wir haben bereits gesehen (Abschnitt 3.2.4), dass Selektionsbeschränkungen auch bei der Erkennung von Idiomen wichtig sein können. Zwar gibt es Selektionsrestriktionen auch bei Adjektiven und Nomina, wir wollen uns jedoch in der folgenden Diskussion auf Restriktionen bei Verben und ihren Argumenten beschränken.

Androutsopoulos/Dale schlagen zwei Alternativen einer HPSG-theoretischen Erfassung von Selektionsrestriktionen vor. Der erste Ansatz betrachtet die Restriktionen aus pragmatischer Sicht und bedient sich des Merkmals BACKGROUND unter CONTEXT. Dieses Merkmal reservieren Pollard/Sag für „Wohlgeformtheitsbedingungen des Äußerungskontextes“, „Präsuppositionen und konventionelle Implikaturen“ und „Angemessenheitsbedingungen“ (Pollard/Sag, 1994, S. 27 und 332; eigene Übers.), es scheinen dort also auch Selektionsbeschränkungen hinzugehören. Zumindest wird auf diese Weise eine Unterscheidung zwischen wörtlichen (im CONTENT) und nicht wörtlichen Informationen (im BACKGROUND) beibehalten. Für ihren Ansatz brauchen Androutsopoulos/Dale neben der geeigneten Information im BACKGROUND (s. u.) noch eine Komponente, die inferiert, ob eine Beschränkung verletzt wurde. Dieses „*constraint-satisfaction reasoning*“ müsste dem Parser eines Sprachverarbeitungssystems nachgeschaltet werden, da die nötigen Informa-

tionen aus einer semantischen Typenhierarchie kommen und erst mit den vorhandenen Argumenten verglichen werden müssen.

In ihrem zweiten Vorschlag, der ohne zusätzliche Inferenzkomponente auskommt, bringen Androutsopoulos/Dale die Selektionsrestriktionen im Semantik-Modul unter. Es werden Untersorten von *index* eingeführt, beispielsweise *edible*, wodurch gleich festgelegt werden kann, von welcher Art ein Komplement sein muss. Dieser Ansatz ist bei automatischer Sprachverarbeitung effizienter als der erste, hat aber den Nachteil, dass es sofort zu einem Fehlschlagen der Analyse kommt, wenn eine Restriktion nicht beachtet wurde. Da jedoch Selektionsrestriktionen keinen Effekt haben, wenn sie in bestimmte Kontexte eingebettet sind, ist dies nicht wünschenswert. Vergleiche hierzu folgende Grammatikalitätsunterschiede:

- (222) a) **Peter aß ein Motorrad.*
 b) *Peter aß kein Motorrad.*

Wenn das Komplement im Skopus einer Negation steht, ist die Äußerung wohlgeformt. Neben der Negation gibt es weitere „reparierende Kontexte“, welche die Verletzung von Selektionsrestriktionen „wieder gut machen“:³⁶

- (223) a) Modalwörter mit Negation: *Man kann/soll/darf Motorräder nicht essen*
 b) Fragen: *Isst Peter (etwa) Motorräder?*
 c) Konditionalsätze: *Wenn Peter ein Motorrad aße, würde es ihm sehr schlecht bekommen.*
 d) Nonfaktive (bzw. epistemische) Verben: *Kim versucht ein Motorrad zu essen.—Kim glaubt/träumte, dass sie ein Motorrad aß.*

Auch Nerbonne (1996) legt seiner Argumentation eine computerlinguistische Anwendung zu Grunde und bringt die Selektionsbeschränkungen ebenfalls in der Semantik unter. Auf das Problem der „reparierenden Kontexte“ geht er, wie auch Androutsopoulos/Dale im betreffenden Abschnitt, nicht ein.

Manchmal verletzt eine Äußerung sogar absichtlich die Selektionsbeschränkungen. Dies ist beispielsweise bei manchen Idiomen der Fall (siehe Abschnitt 3.2.4):

- (224) *das Gesicht verlieren* (Gesicht für Ansehen)

³⁶Chomsky (1965, p. 158) erwähnt noch meta-linguistische Relativierungen wie „*Es ist keine gute Idee, Motorräder zu essen.*“

Auch bei vielen Fällen von Metonymie (vgl. Egg, 2004) tritt eine systematische Verletzung von Selektionsrestriktionen auf. Folgende Beispiele (in Anlehnung an Egg, 2004) sollen dies illustrieren, wobei die Metonymie in Klammern beschrieben wird.

(225) *Anne stellt den Wein auf den Tisch.* (Inhalt für Behälter)

(226) *Anne stellt das Silber auf den Tisch.* (Material für Gegenstand)

(227) *Die Klarinette muss heute früher gehen.* (Instrument für Musiker)

Es ist festzuhalten, dass der semantische Ansatz nicht zufriedenstellend ist. Das Argument, dass Selektionsbeschränkungen oft missachtet bzw. systematisch übergangen werden, spricht deutlich dagegen. Die Alternative, Selektionsrestriktionen im CONTEXT, also der pragmatischen Komponente, unterzubringen, halten wir deshalb für brauchbarer, obwohl sie auf den ersten Blick nicht ohne weiteres in ein Sprachverarbeitungssystem zu implementieren sein mag. Nach der Diskussion in Kapitel 3.2.4 kann man den salomonischen Weg gehen, Selektionsrestriktionen an der Semantik-Pragmatik-Schnittstelle anzusiedeln (siehe auch Egg, 2004), und letztendlich wird genau dies in unserem Analysevorschlag (s. u.) umgesetzt.

Pragmatik als Teil der Grammatik? Wenn man über Pragmatik spricht, stellt sich zuerst einmal die Frage, ob man solche Phänomene innerhalb oder außerhalb der Grammatiktheorie behandeln möchte, also ob es gerechtfertigt ist, ein Pragmatik-Modul in einer formalen Grammatiktheorie zu haben. Die neuere Sprachwissenschaft mit ihren Unterscheidungen *langue/parole*, Sprachsystem/Sprachverwendung bzw. *competence/performance* stellt tendenziell Grammatik und Pragmatik einander gegenüber. Meibauer (2001) nennt in seiner Einführung in die Pragmatik jedoch auch Gründe, die Pragmatik als Teil der Grammatik zu sehen. So ist das Wissen über die typische Verwendung von Satzarten Teil der Grammatik. Deklarativsätze werden für Feststellungen benutzt, mit Imperativsätzen werden Aufforderungen ausgesprochen. Römer (1996) unterstreicht die Modularität des gesamten Sprachsystems, geht dabei auf das Ineinandergreifen von Syntax, Semantik und Pragmatik ein und beschreibt die wichtigsten Teilbereiche der drei Module. Außerdem legt sie dar, dass man das Lexikon als ein komplexes Gebilde betrachten kann, welches aus mehreren Sublexika und Regeln besteht. Daraus kann man den Schluss ziehen, durchaus auch pragmatische Informationen im Lexikon unterbringen.

Auch nach Harnish/Farmer (1984) ist die Pragmatik Teil des Sprachsystems, welches sie in Syntax, Semantik und Pragmatik unterteilen. In der HPSG-Merkmalgeometrie ist diese Dreiteilung durch die Attribute CATEGORY, CONTENT und CONTEXT wiederzufinden. Nach der Vorstellung von Harnish/Farmer umfasst die Pragmatik vor allem die Phänomene der Referenz und der Sprechakte. Es ist dabei nicht ersichtlich, wie z. B. konversationelle Implikaturen behandelt werden können.

Hier schließt sich die allgemeine Frage nach der inneren Modularität der Pragmatik an. Wenn im Sprachsystem nur ein Teil der Pragmatik seinen Platz findet, muss es noch andere Pragmatik-Module geben, die sich außerhalb befinden. Meibauer (2001) stellt in Kap. 5 und 9 verschiedene Ansätze vor. Er zitiert Kasher (1991), wonach es vier Bereiche zu unterscheiden gibt:³⁷

- Kernpragmatik (Wissen über Basis-Sprechakttypen wie Behauptung, Befehl und Frage)
- Konversationspragmatik (Wissen über Sprecherwechsel, Sequenzierung, Reparatur)
- Zentralpragmatik (Wissen über konversationelle Implikaturen, indirekte Sprechakte, Höflichkeit, Register und Stil)
- Schnittstellen-Pragmatik (Wissen über die Integration sprachlicher Daten mit anderen Informationen, d. h. Deixis)

Natürlich gibt es auch Gegner dieser Sichtweise, die jegliche Modularität der Pragmatik leugnen. Nach Meibauer ist in dieser Situation ein Fortschritt nur durch sorgfältige empirische Untersuchungen zu erhoffen, wobei die Interpretation der gewonnenen Ergebnisse wiederum verschiedenen theoretischen Grundannahmen unterworfen ist. Eine spekulative Schlussfolgerung bei Kasher ist, dass die ersten beiden Bereiche eng mit der Grammatik gekoppelt sind und in der linken Hirnhälfte (bei Rechtshändern) verarbeitet werden, und die letzten beiden Bereiche mehr zur allgemeinen Kognition gehören, welche in der rechten Hirnhälfte angesiedelt ist.

Pragmatik in der HPSG Wir wollen aber nicht zu weit von unserer Fragestellung abkommen und uns deshalb ansehen, wie das Pragmatik-Modul in der

³⁷Kasher nennt die vier Bereiche ‚Module‘, wobei er einen eigenen Modulbegriff definiert, auf den wir hier nicht weiter eingehen.

HPSG aussieht. Wie weiter oben schon erwähnt, werden pragmatische Merkmale unterhalb von CONTEXT platziert. Bei Pollard/Sag (1994) gibt es dort zwei Attribute, dargestellt in Abb. 5.20.³⁸

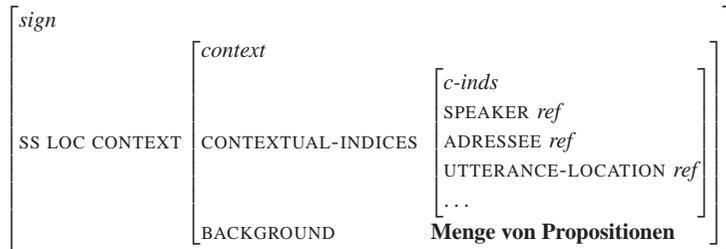


Abbildung 5.20: Merkmalsgeometrie unter CONTEXT

Die Werte von C-INDS bezeichnen die kontextuellen Anker einer Äußerung, beispielsweise Zeiger auf Sprecher, Adressat usw.

Die „Hintergrund“-Informationen in BGR umfassen Präsuppositionen und konventionelle Implikaturen.³⁹ Diese Informationen werden als Propositionen, *psaos* (*parameterized-states-of-affairs*), repräsentiert, die zahlreiche Untertypen mit jeweils unterschiedlichen Attributen besitzen. Dazu definieren Pollard/Sag das PRINZIP DER KONTEXTUELLEN KONSISTENZ, das besagt, dass der BGR-Wert einer Phrase die Vereinigung der BGR-Werte ihrer Töchter darstellt.⁴⁰ Der Lexikoneintrag eines Namens wie *Anne*, skizziert in Abb. 5.21, soll den Gebrauch des CONTEXT verdeutlichen. Dieser Lexikoneintrag beinhaltet die Beschränkung, dass man den Eigennamen *Anne* nur benutzen kann, wenn der intendierte Referent mit einem Index „3. Person Singular feminin“ *Anne* heißt (vgl. Pollard/Sag, 1994, S. 27).

Green (1996) schreibt ausführlicher über die Repräsentation pragmatischer Restriktionen. Sie deckt auf, dass eine Referenz wie bei der Namensge-

³⁸Wir verwenden fortan folgende Abkürzungen: CONTEXT (CTXT), BACKGROUND (BGR), CONTEXTUAL-INDICES (C-INDS).

³⁹Man beachte, dass solche Informationen nach Kasher nicht in die Grammatik gehören. Es ist also eine Uneinheitlichkeit in der Betrachtung der Pragmatik-Module zu beobachten.

⁴⁰Pollard/Sag räumen selbst ein (1994, S. 333), dass es sich bei diesem Prinzip um eine zu einfache Generalisierung der Vererbung von Präsuppositionen handelt. Es gibt bestimmte Ausdrücke (z.B. *wenn...*, *dann...*), welche die Vererbung systematisch blockieren. Die Autoren verweisen auf einige Theorien, welche diese Vererbungsphänomene erfassen (siehe I. Heim: *On the Projection Problem for Presuppositions*. In: Barlow et.al. 1983), unternehmen aber keinen Versuch, die Erkenntnisse in ihren HPSG-Entwurf mit aufzunehmen. Sie weisen darauf hin, dass eventuelle Änderungen sich auch auf die semantische Repräsentation auswirken müssten.

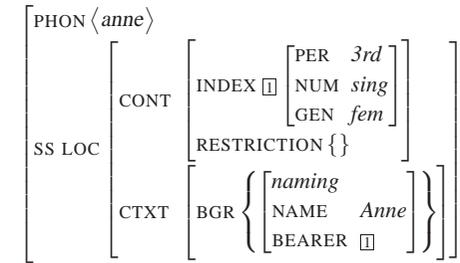


Abbildung 5.21: LE eines Namens nach Pollard/Sag (1994)

bung über den CONTEXT geschieht, aber die Instanzierung eines Objekts im CONTENT vorgenommen wird (nach Pollard/Sag, 1994, S. 26), siehe Abb. 5.22.

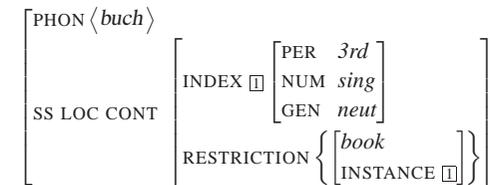


Abbildung 5.22: Instanzierung eines Objekts nach Pollard/Sag (1994)

Sie merkt in diesem Zusammenhang auch an, dass es sich bei *index* um ein abstraktes linguistisches Objekt handelt (mit den Attributen PERSON, NUMBER und GENDER) und dass bei der *book*-Relation in Wirklichkeit nicht das *index*-Objekt beschränkt werden soll, sondern die Entität, auf die es verweist. Dabei wird bei Pollard/Sag (1994) nicht klar, wie das funktionieren soll. Green entwickelt eine adäquatere Beschreibung von Referenz, indem sie diese gänzlich über CONTEXT modelliert, da ja in einem Gespräch auch über den Kontext klar wird, auf welches Objekt sich ein Sprecher bezieht. Green führt dazu *psaos* ein für das, was man normalerweise annimmt (*normally-believe*), und das, was zwei Gesprächspartner voneinander glauben (*mutually-believe*). Bei ihrem Ansatz ändert sie aber nichts an der Merkmalsgeometrie unter CONTEXT. Murphy (1995) entwirft eine Diskurs-Theorie für die HPSG, die Phänomene wie Fokus und die Thema-Rhema-Unterscheidung einschließt. Dafür benötigt er einige neue Attribute unter CONTEXT, wobei er die Zweiteilung in INDICES und BGR weitestgehend beibehält, aber einige neue Attribute für bekannte und

neue Information einführt. Sein Vorschlag stellt eine beachtenswerte Grundlage für die Weiterentwicklung der Diskurs-Theorie im Rahmen der HPSG dar. Weitere Phänomene, die im CONTEXT behandelt werden, sind der „höfliche Plural“ (3. Person Plural im Deutschen bzw. 2. Person Plural im Französischen) und die Höflichkeitsformen im Japanischen und Koreanischen (Pollard/Sag, 1994, S. 91-97). Für das Japanische bietet Siegel (2000) eine verbesserte und detailliertere Analyse. Sie führt u. a. auch neue *psaos* ein, die mit in der Menge unter BGR enthalten sind.

Ein neuer Vorschlag Nach diesem Einblick in Gesichtspunkte der Pragmatik und des CONTEXT-Attributs wenden wir uns nun der Modellierung der Selektionsrestriktionen in der HPSG zu. Androutsopoulos/Dale (2000) schlagen in ihrem ersten, pragmatikorientierten Ansatz vor, eigene *psaos* für semantische Merkmale wie beispielsweise *edible* einzuführen und diese hierarchisch zu ordnen. Wie weiter oben schon angedeutet, erfordert es eine spezielle Komponente, die überprüft, in welchem Verhältnis die *psaos* stehen, um auszumachen, ob Selektionsrestriktionen verletzt wurden.

Wir wollen hier einen neuen Weg gehen und – ähnlich wie beim zweiten Ansatz von Androutsopoulos/Dale – mit Sorten und Relationen arbeiten. So kann man eine Übereinstimmung ohne gesonderte Inferenz feststellen.

Da im letzten Abschnitt argumentiert wurde, dass man im CONTENT auf *psaos* verzichten kann, muss an dieser Stelle noch etwas zu den Werten in der BGR-Menge gesagt werden, auf welcher sich ja Propositionen befinden. Wenn man, so wie in dieser Arbeit vorgeschlagen, die Semantik mit Hilfe von logischen Formen beschreibt, ist es konsequent, solche logischen Formen auch für die Ausdrücke auf BGR zu verwenden. Da man dort aber nicht einfach einen logischen Ausdruck nach dem anderen auflisten kann, sondern etwas mehr Struktur braucht (vgl. auch die beschriebenen Ansätze weiter oben), schlagen wir vor, die Elemente auf BGR, so wie auch die *psaos*, welche in anderen Arbeiten verwendet werden, mit Hilfe eines Typs *bgr-condition* zu beschreiben. Die wohl am meisten verwendete Untersorte davon ist dann ein einfacher logischer Ausdruck mit einer *meaningful-expression*:⁴¹

⁴¹Anmerkung: Es handelt sich hier nicht einfach um eine Umbenennung von *psoa* nach *bgr-condition*. Die *psaos* in Pollard/Sag (1994) wurden für die Semantik von Verben verwendet und diese Verwendungsweise wurde bei Elementen im BGR manchmal stark erweitert bzw. zweckentfremdet. Während die Existenz des Typs *psoa* durch die hier verwendete Semantik nicht mehr gerechtfertigt ist, stellt die Einführung eines neuen Typs *bgr-condition* eine adäquatere Generalisierung der BGR-Inhalte dar.

$$\left[\begin{array}{l} \text{logical-expression} \\ \text{ME } me \end{array} \right]$$

Kommen wir zu unserem Vorschlag bzgl. der Selektionsrestriktionen zurück. Wir führen zwei neue Untersorten von *bgr-condition* ein, die wie folgt definiert sind:

$$\left[\begin{array}{ll} \text{sel-restr-imp} & \\ \text{ARG} & \text{extended-index} \\ \text{MUST-SATISFY} & \text{selection-sort} \end{array} \right] \quad \left[\begin{array}{ll} \text{sel-restr-stf} & \\ \text{ARG} & \text{extended-index} \\ \text{SATISFIES} & \text{selection-sort} \end{array} \right]$$

Die erste *bgr-condition* wird von Zeichen, die eine Selektionsrestriktion auferlegen⁴², als ein Element der Menge unter BGR eingeführt. Ein Verb kann beispielsweise ein Nomen mit einer bestimmten Beschränkung subkategorisieren, wie in Abb. 5.23 gezeigt: *essen* verlangt ein optionales Objekt, das essbar ist (zu den Werten von MUST-SATISFY und SATISFIES: siehe unten).

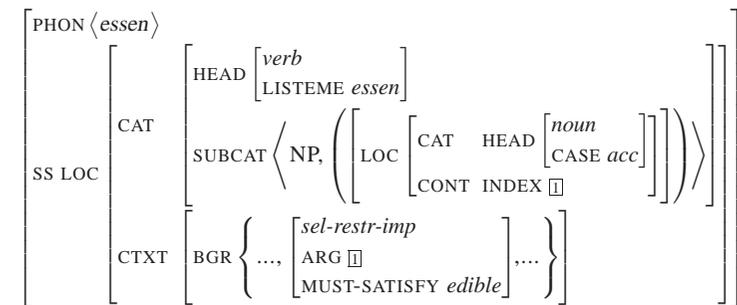


Abbildung 5.23: Skizze des LEs von *essen* mit Selektionsrestriktion

Nomen wie *Apfel* erfüllen⁴³ diese Bedingung, weil sie im LE eine entsprechende Bedingung *sel-restr-stf* in ihrer BGR-Menge definiert haben, siehe Abb. 5.24.

Eine Phrase wie „... isst *Äpfel*“ würde dann wie in Abb. 5.25 (S. 153) gezeigt aussehen. Das von unserem Vorschlag unabhängige PRINZIP DER KONTEXTUELLEN KONSISTENZ bewirkt die Vererbung der *bgr-conditions* auf den jeweiligen BGR-Mengen.

Wir benötigen zudem ein Prinzip, welches sicherstellt, dass die beiden Werte von MUST-SATISFY (M-STF) und SATISFIES (STF) in der CTXT BGR-Menge kompatibel sind. Kompatibel bedeutet in diesem Zusammenhang, dass der

⁴²*sel-restr-imp* für *imposed*

⁴³*sel-restr-stf* für *satisfied*

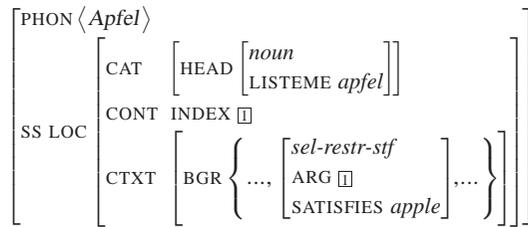


Abbildung 5.24: Skizze des LEs von *Apfel* mit SATISFIES-Wert

STF-Wert beim Argument von *essen* identisch mit dem M-STF-Wert beim Verb selbst ist oder dass der STF-Wert in der semantischen Ontologie eine Unter-einheit des M-STF-Wertes ist. Das Verb verlangt also lediglich etwas Essbares, während es sich beim Komplement um einen Schweinebraten oder eine Banane handeln kann.

Das Prinzip soll nur Phrasen lizensieren, deren M-STF- und STF-Werte kompatibel sind – allerdings nur, wenn sich das Argument bzw. die Proposition außerhalb des Skopus einer Negation, eines Konditional- oder eines Fragesatzes befindet. Wie oben schon gezeigt, „reparieren“ diese Kontexte den Effekt einer Verletzung der Restriktionen.

(228) GÜLTIGKEITSPRINZIP DER SELEKTIONSRESTRIKTIONEN (GPSR, vorläufig):

Gibt es in einer Phrase x ein Zeichen z , ein Verb v (wobei z ein Argument von v ist) und eine Proposition p , welche durch v und seine Argumente ausgedrückt wird, und

stehen weder die logische Form, die mit z assoziiert ist, noch diejenige, die mit p assoziiert ist, im Skopus eines Negationsoperators, eines Konditionaloperators oder eines Frageoperators, noch sind sie Argument eines nonfaktiven Verbs,

dann muss der STF-Wert eines Elements der Sorte *sel-rest-stf* in der CTXT BGR-Menge von x mit dem M-STF-Wert eines Elements der Sorte *sel-restr-imp* kompatibel sein, wobei die ARG-Werte von *sel-rest-stf* und *sel-restr-imp* identisch sein müssen.

Wie kann nun diese Kompatibilität formal erfasst werden? Der Wert von MUST-SATISFY bzw. SATISFIES ist von der Sorte *selection-sort*. Diese besitzt eine endliche Zahl an Untersorten wie *abstract*, *physical*, *artifact*, *animate*, *edible*, ... Diese Untersorten entsprechen Einheiten einer semantischen Onto-

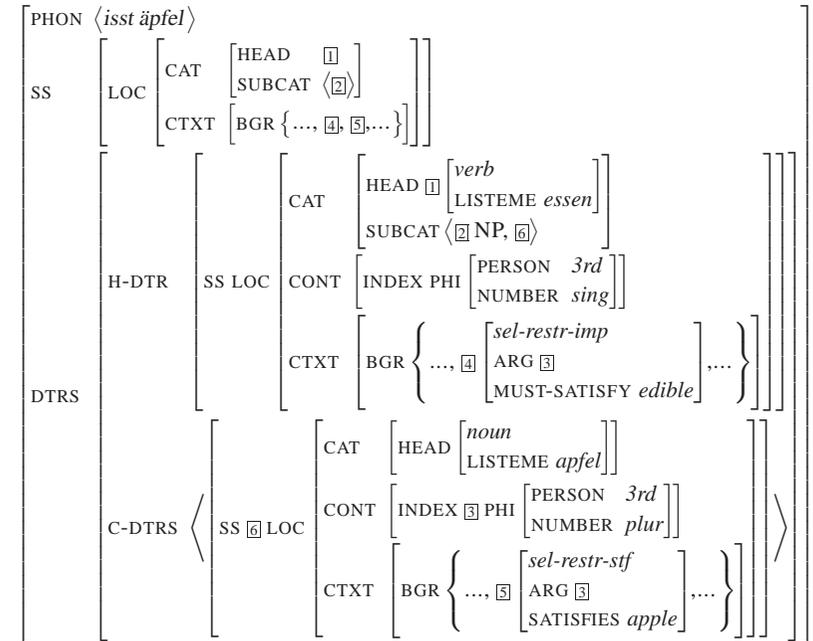


Abbildung 5.25: Skizze einer Phrase mit Selektionsrestriktionen

logie, ähnlich etwa wie in WordNet⁴⁴ oder GermaNet⁴⁵. In Appendix B wird eine semantische Ontologie grob skizziert, welche eine multiple Vererbung (Untereinheiten erben von mehreren Obereinheiten) einschließt. Die Relationen, die in dieser Ontologie bestehen und die man zwischen den Untersorten von *selection-sort* herstellen möchte, kann man in einer Liste erfassen. Die Liste ist der Wert eines neuen Attributs HIERARCHY, welches wir für die Sorte *unembedded-sign* (siehe Abschnitt 5.2.2) definieren – alle empirischen Zeichen sind ja letztlich Teil eines uneingebetteten Zeichens (s.o., Abschnitt 5.2.2). Die HIERARCHY-Liste enthält Paare von *selection-sort* Untersorten, die in der Ist-Relation (*is_a*) stehen. Formal gesehen handelt es sich dabei um eine partielle Ordnung der Elemente unter *selection-sort*. Folgendes Prinzip beschreibt die Liste und definiert sie als Wert des HIERARCHY-Attributs für jedes uneingebettete Zeichen.

⁴⁴Siehe dazu Christiane Fellbaum, Hrsg. (1998): *Wordnet: An Electronic Lexical Database*. Bradford Books, The MIT Press.

⁴⁵Siehe dazu <http://www.sfs.nphil.uni-tuebingen.de/lsd/>

(229) SELEKTIONS-HIERARCHIE-PRINZIP (skizziert):

$$u\text{-sign} \rightarrow \left[\text{HIERARCHY} \left\langle \dots, \begin{bmatrix} is_a \\ \text{ARG1 } animate \\ \text{ARG2 } animate \end{bmatrix}, \begin{bmatrix} is_a \\ \text{ARG1 } animate \\ \text{ARG2 } person \end{bmatrix}, \begin{bmatrix} is_a \\ \text{ARG1 } animate \\ \text{ARG2 } animal \end{bmatrix}, \dots \right\rangle \right]$$

Das Attribut HIERARCHY soll hier nicht in der Weise verstanden werden, dass diese zugegebenermaßen große Hierarchie ein Teil jedes uneingebetteten Zeichens wäre, sondern es soll damit ausgedrückt werden, dass der Sprecher stets Zugriff auf diese Hierarchie hat und sie verwenden kann. Mit der Aufnahme von HIERARCHY in die *Appropriateness* von *u-sign* in unserer Theorie kann die Grammatikalität jedes uneingebetteten Zeichens für sich alleine festgestellt werden.

Kompatibilität der *selection-sorts* bedeutet also, dass zwischen den Werten von MUST-SATISFY und SATISFIES solch eine Ist-Relation bestehen muss. Diese Relation kann mehrere Zwischenstufen beinhalten: Während das Verb *essen* ein Argument der Art *edible* verlangt, ist ein Apfel ein essbares Obst und damit *edible*. Ein Beispiel, bei dem es um konkrete Weinsorten wie Veltliner oder Muskateller geht, die Weine sind, wobei Weine zu den trinkbaren Stoffen gezählt werden, macht deutlich, wie komplex solch eine Ontologie ist.⁴⁶ Da die Ist-Relation transitiv ist, werden solche „Zwischenstufen“ mit erfasst. Man kann aus folgendem Grund nicht mit einer Sortenhierarchie operieren: Bei allen Objekten sind die Sorten maximal spezifisch und man kann somit keine Aussagen über Sorten machen, die sich nicht an den Blättern des Hierarchiebaumes befinden. Jedoch möchte man Generalisierungen, wie *essen* selegiert alles Essbare, ausdrücken können. Mit unserer Lösung hat man stets Zugriff auf alle Sorten (unter *selection-sort* sind ja alle maximal spezifisch) und kann gleichzeitig hierarchische Beziehungen und Generalisierungen modellieren.

Wir haben hiermit einen Vorschlag gemacht, wie Selektionsrestriktionen in die HPSG integriert werden können, wobei wir auch den Effekt der „reparierenden Kontexte“ berücksichtigt haben. Die Restriktionen werden im Lexikon eintrag der Verben festgelegt, die Komplemente müssen ihnen genügen. Wenn ein möglicherweise inkompatibles Komplement im Skopus eines „reparierenden“ Operators steht, ist der Gesamtsatz deshalb nicht ungrammatisch, was durch das GPSR erzielt wird.⁴⁷

⁴⁶An dieser Stelle muss man zugeben, dass es sehr leicht ist, solche Ontologien zu fordern und sie ansatzweise zu skizzieren. Die eigentliche Arbeit, solch eine Ontologie (nicht nur im Rahmen der HPSG) zu erstellen und sämtliche theoretischen und empirischen Probleme zu berücksichtigen, würde ein groß angelegtes Projekt erfordern.

⁴⁷Ein Argument, das man noch anführen könnte, ist, dass eine Verletzung von Selektions-

Wie sieht es nun mit den oben schon angesprochenen Metonymien aus? Die Auflösung einer Metonymie verlangt ein gewisses Maß an Weltwissen, z. B. dass ein Objekt aus bestimmten Teilen besteht, also dass ein Tisch Beine hat oder ein Computer einen Monitor besitzt. Dazu kommen noch das Wissen um das Material, aus dem ein Objekt ist, beispielsweise dass Kerzen aus Wachs sind, und eine große Menge an weiteren Beziehungen zwischen Objekten. Im oben angegebenen Beispiel (225) muss man wissen, dass trinkbare Flüssigkeiten normalerweise in Behältern aufbewahrt werden, die man auf den Tisch stellen kann.

Wir haben für die Behandlung von Selektionsrestriktionen schon eine semantische Hierarchie in Form der HIERARCHY-Liste eingeführt. Auf diese Liste wurden sogenannte Ist-Beziehungen aufgenommen: Ein Wort **ist** ein Zeichen, eine Banane **ist** essbar usw. Man kann nun weitere Beziehungen wie „Teil-von“ oder „besteht-aus“ festlegen. Solche Relationen (*has_part(x,y)* oder *consists_of(x,y)*), haben wie *is_a* zwei Argumente, welche jeweils auf eine Sorte verweisen. Für die Umsetzung bieten sich zwei Möglichkeiten an. Erstens könnte man weitere Listen definieren, welche die Sorten unter *selection-sort* enthalten. Alle derartigen Listen bestünden dann aus sämtlichen gültigen Tupeln einer bestimmten Beziehung wie *is_a* oder *has_part* in der weiter oben skizzierten Form. Zweitens könnte man über die Elemente der HIERARCHY-Liste die gewünschten Beziehungen direkt als Relationen definieren. Diese Möglichkeit hätte den Vorteil, dass man Generalisierungen über die Elemente auf HIERARCHY formulieren kann. So haben alle Personen einen Vater und eine Mutter und diese Tatsache muss man nicht für alle Unterelemente von *person* explizit ausdrücken. Diese Generalisierungen ergäben sich bei der Variante mit verschiedenen Listen nur indirekt und müssten separat aufgezählt werden. Es scheint also, dass einer direkten relationalen Definition der Vorzug gegeben werden sollte. Wir passen das GPSR an und formulieren die endgültige Version:

restriktionen durch bestimmte Kontexte außer Kraft gesetzt wird. Beispielsweise ist folgender Satz völlig grammatisch, wenn man sich verschiedene Objekte aus Schokolade vorstellt:

(230) *Peter hat einen Fußball und ein Keyboard gegessen.*

Es gibt also Kontexte, welche gegebene Fakten systematisch außer Kraft setzen und in denen beispielsweise ein essbarer Fußball oder in Science-Fiction-Filmen ein Raumschiff mit „Warpantrieb“ völlig bedenkenlos verwendet werden. Um dies zu modellieren, müsste man unter CTXT noch mehr an Struktur annehmen und etwa zwischen einem „Standard-“ und einem „aktuellen“ Kontext unterscheiden. Ferner wären Relationen nötig, welche gewisse Standardannahmen (wie Fußbälle sind nicht essbar) je nachdem im aktuellen Kontext zuließen bzw. neue Annahmen (Raumschiffe fliegen mit Warp 9) einführen könnten.

(231) GÜLTIGKEITSPRINZIP DER SELEKTIONSRESTRIKTIONEN
(GPSR, endgültig):

Gibt es in einem uneingebetteten Zeichen x ein Zeichen z , ein Verb v (wobei z ein Argument von v ist) und eine Proposition p , welche durch v und seine Argumente ausgedrückt wird, und stehen weder die logische Form, die mit z assoziiert ist, noch diejenige, die mit p assoziiert ist, im Skopus eines Negationsoperators, eines Konditionaloperators oder eines Frageoperators, noch sind sie Argument eines nonfaktiven Verbs, dann muss der STF-Wert eines Elements der Sorte *sel-rest-stf* in der CTXT BGR-Menge von x mit dem M-STF-Wert eines Elements der Sorte *sel-restr-imp* in der *is_a*-Relation auf der HIERARCHY-Liste von x stehen oder sich in einer weiteren Relation befinden, wobei die ARG-Werte von *sel-rest-stf* und *sel-restr-imp* identisch sein müssen.

Damit werden nun alle Fälle, in welchen eine Verletzung von Selektionsrestriktionen akzeptabel ist, lizenziert.⁴⁸

5.3.2.3 Referenzialität

Wie schon im Abschnitt 2.3.4.6 angesprochen wurde, sind manche NPs, die Teil eines Idioms sind, nicht referenziell.⁴⁹ Damit ist gemeint, dass sie nicht als Antezedens eines Pronomens fungieren können. Hingegen sind alle nicht idiomatischen NPs referenziell, können also zum Referieren benutzt werden. Wie u. a. auch Chur (1993) hervorhebt, ist hier die Unterscheidung zwischen referenziell und referierend zu treffen.

Nicht referierend sind NPs in beispielsweise folgenden syntaktischen Umgebungen:

- NPs im Skopus von Quantoren: *Jeder Anwalt vertritt seinen Mandanten_i. *Er_i benötigt rechtlichen Beistand.*

⁴⁸Es ist offensichtlich, dass in dieser Diskussion das weite Feld der semantischen Ontologien, der Repräsentation von Weltwissen und beispielsweise auch das der Trennung zwischen Weltwissen und linguistischem Wissen unberücksichtigt geblieben ist. Eine tiefer gehende Klärung dieser Fragen würde jedoch den Rahmen der Arbeit sprengen und auch am thematischen Ziel vorbeigehen, sie stellt aber durchaus ein Desiderat für weitere Untersuchungen dar.

⁴⁹Referenzialität betrifft ganze Klassen von Ausdrücken, z. B. Konjunktionen, Artikel, einige idiomatische NPs etc.

- NPs im Skopus von Negation: *Es ist nicht wahr, dass Karl gestern ein Auto_i gewonnen hat. *Es_i ist blau.*
- Prädikative NPs: *Maria ist ein Glückspilz. Alles scheint ihr/*ihm in den Schoß zu fallen.* Diese NPs sind sogenannte Kennzeichnungen, vgl. hierzu Wilhelm (1997, S. 13, Fußn. 1) und Neale (1990).
- NPs in einem Adjunkt: *Der Präsident von Frankreich_i war gestern in Berlin. *Es_i hat/*Dort_i leben fast 60 Mio. Einwohner.*

Man muss also unterscheiden, ob eine NP gerade nicht referierend ist oder ob sie generell nicht referenziell ist. Die Eigenschaft, referierend zu sein, hängt, wie gesehen, von der syntaktischen Einbettung ab und ist ein systematisches Phänomen.

Referenzialität ist jedoch davon unabhängig, d. h. idiomatische Wörter wie *Schur*, *Bauklötze* oder *Tapet*, sind nie referenziell, egal, in welcher syntaktischen Umgebung sie auftreten (von der Tatsache, dass manche syntaktischen Umgebungen bei IVPs ungrammatisch sind, einmal abgesehen).

Bei nicht dekomponierbaren Idiomen sind NPs nicht referenziell, da der NP als solcher keine eigene Bedeutung zugewiesen wird. Die Bedeutung wird lediglich für die Phrase als Ganze angegeben.⁵⁰ Darüber hinaus gibt es NPs, die als Kennzeichen fungieren (*klar wie Kloßbrühe*) und die schon aus diesem Grund nicht referierend sind.⁵¹ Mit der Eigenschaft, keine individuelle Semantik zu besitzen, kann die fehlende Referenzialität eines Großteils der idiomatischen NPs erklärt werden.

Jedoch bleiben idiosynkratische Fälle. So gibt es dekomponierbare Phraseologismen, die eine NP enthalten, welche nicht referenziell ist:

(232) *Über ihr Geschenk hat Susi Bauklötze gestaunt. *Sie waren hoch aufgetürmt.*

(233) *Auf diese Äußerung hin schwoll ihm die Zornader an. *Sie war gut sichtbar.*

(234) *Hans wurde von seinem Kollegen beim Chef in Misskredit_i gebracht. *Er_i war schwer wieder auszumergen.*

⁵⁰Diese NPs sind allerdings referierend, wenn sie außerhalb des Idioms auftreten, vgl. *Löffel in jdn. über den Löffel balbieren.*

⁵¹Dazu kommt, dass solche PPs wie *wie Kloßbrühe* ebenfalls nicht dekomponierbar sind.

Wir werden uns hier nicht mit der Analyse von nicht referierenden NPs auseinandersetzen, da uns dies von der eigentlichen Fragestellung zu weit weg-führen würde. Was einen HPSG-Ansatz zu nicht referenziellen NPs betrifft, führen wir zu diesem Zweck eine neue Sorte *non-referential'* ein, die folgen-dermaßen in die Sortenhierarchie eingebettet wird:

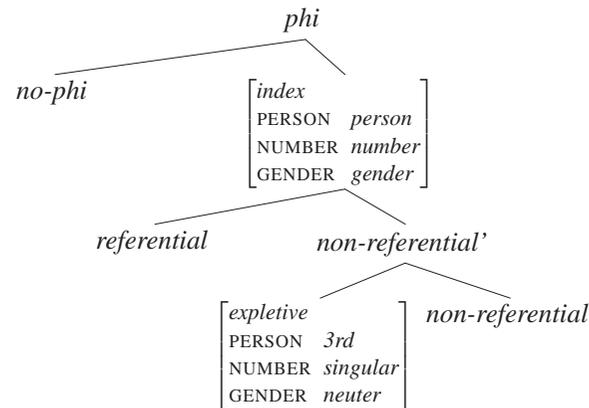


Abbildung 5.26: Sortenhierarchie unter *phi*

Der Wert von *PHI* unter *CONT INDEX* sei von der Sorte *phi*, welche die Untersorten *no-phi* (für Verben, deren ϕ -Werte keine Rolle spielen) und *index* besitzt. Objekte der Sorte *index* können vom Typ *referential* oder *non-referential'* sein. *referential* sind alle referenziellen Ausdrücke, auch nicht idiomatische NPs, wobei es keinen Unterschied macht, ob eine NP gerade referierend ist oder nicht. *non-referential'* sind alle Expletiva (z.B. das Wetter-*es*) und alle nicht referenziellen (*non-referential*) Ausdrücke, wie Determinantien, Adjektive und eben auch nicht referenzielle NPs.

Ist nun eine NP von der Sorte *non-referential*, kann sie nicht mit einem Pronomen koindiziert werden. Pronomen verlangen als Antezedens eine NP[*INDEX referential*] und diese Anforderung kann von einer nicht referenziellen NP nicht erfüllt werden. Allerdings tragen auch solche NPs die Kongruenzmerkmale *PERSON*, *NUMBER* und *GENDER*, wodurch die Kongruenz von Adjektiven und Determinantien sichergestellt ist.

5.4 Zusammenfassung

Zu Beginn dieses Kapitels wurde untersucht, welche Anknüpfungspunkte die in Kapitel 4 beschriebenen Herangehensweisen bieten.

Von der Konstruktionsanalyse haben wir die Idee übernommen, syntaktisch komplexe Lexikoneinträge einzuführen, um nicht dekomponierbare Phraseologismen zu beschreiben.

Der Ansatz der Externen Selektion lieferte uns ein wichtiges Werkzeug, das der Listenselektion. Ausgehend von der Lexemselektion (Soehn, 2003) haben wir dabei einen linguistisch angemesseneren Weg vorgeschlagen und den Begriff „Listem“ in die HPSG-Grammatik eingeführt. Weiterhin haben wir das Merkmal *XSEL* als Vereinheitlichung der Selektionsattribute *SPEC* und *MOD* übernommen.

War das Kollokationsmodul, wie es im vorigen Kapitel vorgestellt wurde, zu unrestriktiv, konnten wir es mit Berücksichtigung einer früheren Version des Ansatzes an unseren Problembereich anpassen. Das Attribut *COLL* diente dabei zur Spezifikation bestimmter Knoten in der Satzstruktur, den Barrieren, um dort zu überprüfen, ob die Distributionsbeschränkungen eines idiomatischen Wortes nicht verletzt wurden. Der Ansatz erwies sich als restriktiver und sogar modular begrenzt- und erweiterbar. Man kann, sozusagen „vertikal“, durch Spezifikation der Barrieren die Satzstruktur in unterschiedlicher Tiefe mit einbeziehen, von einer PP bis hin zur kompletten Äußerung. „Horizontal“ erweiterbar ist der Ansatz in dem Sinne, dass unterschiedliche Attribute zur Distributionsbeschränkung herangezogen werden können: die lokalen syntaktischen oder semantischen Merkmale, die phonologische Form oder die logische Form bzw. die nicht lokale Semantik. Es wurden dafür der Barrierenbegriff erläutert und die technischen Voraussetzungen für eine Analyse definiert.

Ferner wurden weitere Themen angesprochen, die mittelbar die Analyse betreffen. Deren Diskussion soll die Analyse im folgenden Kapitel übersichtlicher machen, da wir an dieser Stelle schon einige Grundlagen geschaffen haben. So wurde die Behandlung des Subjekts und des Verbalkomplexes im Deutschen zusammen mit einer Vorstellungsanalyse von Stefan Müller vorgestellt. Der Verbalkomplex wurde dann erweitert, um wendungsinterne Argumente mit aufzunehmen. Weiterhin wurden Passivbildung und Nominalisierung angesprochen und wir diskutierten, wie man IVPs, welche für diese beiden Operationen defektiv sind, von Passivierung und Konversion ausschließen kann.

Überdies wurde kurz in die semantische Repräsentation LRS eingeführt, mit deren Hilfe wir den semantischen Beitrag in unseren Lexikoneinträgen beschreiben können. Darüber hinaus erfolgte ein Vorschlag zur Integration von Selektionsbeschränkungen in die HPSG und es wurde eine Möglichkeit, nicht referenzielle Nomina zu beschreiben, dargestellt.

6 COLLage der Analysen

If natural language had been designed by a logician,
idioms would not exist. Philipp N. Johnson-Laird

Wir haben in Kapitel 5.3 verschiedene Aspekte beleuchtet: den Verbkomplex, das Passiv, eine semantische Repräsentation mit LRS, Selektionsrestriktionen und Referenzialität. Diese Punkte spielen für die nun folgende Analyse eine Rolle und wurden deshalb im Voraus diskutiert.

Wir unterscheiden bei der Analyse zwischen dekomponierbaren und nicht dekomponierbaren Idiomen und werden diese auf unterschiedliche Weise modellieren müssen. Unikale Elemente bilden dabei keinen eigenen Unterpunkt. Wenn ein unikales Element eine eigene Bedeutung hat, wird es wie jedes andere idiomatische Wort innerhalb von dekomponierbaren Phraseologismen behandelt. Solche Phraseologismen analysieren wir in Kapitel 6.1: Für idiomatische Wörter werden wir eigene Lexikoneinträge vorsehen, wobei ein besonderer Fokus auf der Spezifikation des COLL-Wertes liegt. Andererseits gibt es unikale Elemente, die Teil eines nicht dekomponierbaren Idioms sind. Solche werden dann zusammen mit diesen Idiomen beschrieben und sind Teil eines phrasalen Lexikoneintrags. Diese Analyse, wie auch die Diskussion der Bestandteile von nicht dekomponierbaren IVPs, wird in Kapitel 6.2 erläutert.

6.1 Dekomponierbare IVPs

6.1.1 Analyse

Bei der Analyse dekomponierbarer Phraseologismen nehmen wir an, dass die Teile eines dekomponierbaren Idioms aus idiomatischen Wörtern und teilweise aus nicht idiomatischen Wörtern bestehen. Erstere sind zwar gleichlautend mit ihren Gegenstücken in freier Verwendung, haben aber eine eigene, idiomatische Bedeutung und bekommen deshalb auch einen separaten Lexikoneintrag.¹ Das Entscheidende an der Analyse ist, die Kookkurrenz aller zu einem

¹Teilidiomatische Phraseologismen bestehen zwar auch aus Wörtern mit nicht idiomatischer Bedeutung, wenn sich diese aber durch ihre Valenz oder andere Merkmale von der frei verwendbaren Form unterscheiden, bekommen auch sie eigene Lexikoneinträge.

Idiom gehörenden idiomatischen Wörter zu garantieren. Somit wird andererseits ein Auftreten eines idiomatischen Wortes außerhalb des zugehörigen Idioms ausgeschlossen. Es ist dabei, wie oben schon angemerkt wurde, gleichgültig, ob es sich um ein unikales Element handelt oder um ein Wort, das mit nicht idiomatischer Bedeutung auch in freier Verwendung vorkommt.

Wir wollen einen uns schon bekannten IVP beispielhaft betrachten, nämlich *Bauklötze staunen*. Der IVP besteht aus einem Verb mit der NP *Bauklötze* als internem Argument. Dieses Idiom ist insofern dekomponierbar, als dass die NP als Intensivierer fungiert und man als ihre Bedeutung „sehr“ annehmen kann. Fehlt diese NP, liegt auch kein Phraseologismus vor – das Verb ist wie *staunen* in freier Verwendung. Hinzu kommt, dass sich die NP nicht austauschen lässt (**Bausteine staunen*). Wir drücken dies durch ein Subkategorisierungsverhältnis zwischen Verb und der NP *Bauklötze* aus. Damit besitzt *staunen* im Idiom andere Valenzeigenschaften, bekommt einen separaten LE und wird von uns als *staunen₂* bezeichnet.

Semantisch verhält sich die NP wie ein Intensivierer, weswegen man es syntaktisch als Adjunkt analysieren könnte. Allerdings ist dieses „Adjunkt“ nicht beliebig (s. Bsp. o.) und es kann auch keine anderen Verben modifizieren (**Bauklötze wundern*). Die Phraseologieforschung ist sich überdies weitestgehend einig, dass es sich bei so etwas wie *Bauklötze* um ein internes Objekt, also ein Komplement, handelt. Leider gibt es dafür außer der Akkusativ-Zuweisung keine anderen empirischen Argumente. Mit der hier vorgestellten Analyse soll kein Versuch unternommen werden, eine Objekt-NP als Adjunkt zu behandeln, und wir wollen auch nicht so weit gehen, *Bauklötze* als syntaktisches Adjunkt zu betrachten. Ein Desiderat wäre der Vergleich zu sog. inkorporierten Objekten wie in *Verdacht schöpfen* und eine einheitliche Analyse dieses zu unserem Beispiel ganz ähnlichen Phänomens in der HPSG, siehe auch Fn. 8 in diesem Kapitel. Zu inkorporierten Objekten vgl. Abschnitt 5.3.1.3.

Eine Modellierung sieht nun so aus, dass man ein idiomatisch verwendetes Verb *staunen* annimmt, das sich im Wesentlichen nur durch seine Selektionsanforderungen vom freien Verb unterscheidet. Das Listem ist *staunen₂*. In Abbildung 6.1 sind die wesentlichen Informationen skizziert.² Die SUBCAT-Liste enthält zwei NPs, das Subjekt und das Akkusativ-Objekt *Bauklötze*. Hinzu kommt eine optionale PP mit dem Kopf *über*.

Nun kommt es darauf an, das Auftreten dieser Akkusativ-NP so einzuschrän-

²Bei unserer Analyse werden wir in allen LEs als PHON-Werte die Grundform des Wortes angeben und Flexion bei Seite lassen. Die Identifikation des Wortes erfolgt ohnehin nicht über die PHONOLOGY, sondern über die LISTEME. Der Einfachheit halber geben wir bei Verben die Kasus der Komplemente mit an. Für eine Theorie der Kasuszuweisung siehe beispielsweise Meurers (1999).

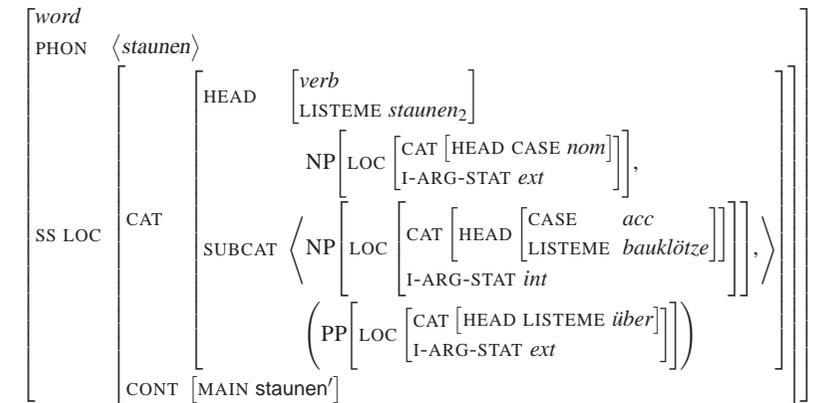


Abbildung 6.1: Skizze des LEs von *staunen* in *Bauklötze staunen*

ken, dass sie nur als Komplement des idiomatischen *staunen* auftreten kann. Dies wird durch den COLL-Wert sichergestellt.

$$\text{COLL-Wert von } \textit{Bauklötze} \text{ (vorläufig):}$$

$$\left[\text{COLL} \left\langle \left[\text{vp} \left[\text{LOC-LIC} \left[\text{CAT HEAD LISTEME } \textit{staunen}_2 \right] \right] \right] \right\rangle \right]$$

Darin wird eine VP als Barriere spezifiziert, welche als Kopfwert das idiomatische *staunen* besitzt. Durch unser LIZENSIERUNGSPRINZIP (LIP) (179, S. 110) wird nun erzwungen, dass der COLL-Wert mit der VP, in welcher *Bauklötze* auftritt, identisch ist. Genauer gesagt, der LOC-LIC-Wert des Objekts auf der COLL-Liste ist identisch mit dem LOCAL-Wert der VP. Dabei ist die VP die niedrigstmögliche in der syntaktischen Struktur. Die Wirkungsweise von COLL und dem LIP kann man mit folgendem Beispielsatz veranschaulichen.

(235) *Bei Daimler-Chrysler staunt man Bauklötze.*³

Die Struktur des Satzes zusammen mit den relevanten Informationen wird in Abb. 6.2⁴ dargestellt. Die Verbstellung wird auf die Weise lizenziert, wie sie schon im Abschnitt über den Verbalkomplex (S. 115ff) beschrieben wurde, mit Hilfe der V1-Derivationsregel (Abb. 5.10) und der Verbspur. Das interne Argument (die CLUSTER-DTR) befindet sich innerhalb des Verbalkomplexes.

³Mannheimer Morgen, 10.11.2001;

⁴Am Knoten von *Bauklötze* haben wir ausnahmsweise ein *sign*-Objekt an Stelle von *synsem* skizziert, um das COLL-Attribut ebenfalls darzustellen.

Da wir es hier mit einem dekomponierbaren IVP zu tun haben, kommt die normale semantische Bedeutungskomposition zum Tragen. Allerdings sind Adjunkte bei *Bauklötze* nicht zulässig. Aus diesem Grund spezifizieren wir die Semantik im EXCONT-Wert des Zeichens so, dass keine zusätzliche Information durch einen Modifizierer mehr zum Nomen hinzukommen kann. Eine logische Form für den Satz „*Peter staunt Bauklötze*“ könnte folgendermaßen aussehen:

$$\exists e(\exists x(\text{sehr}'(x) \wedge (\text{staunen}'(e, p) \wedge \text{intensität}'(e, x))))$$

Die Variable *e* steht für das Ereignis, dass Peter (*p*) staunt. In Worten heißt diese Repräsentation etwa: Es gibt ein *x*, *x* ist ein hohes Maß an Intensität und *e* wird mit dieser Intensität *x* ausgeführt.⁸ Die semantische LRS-Repräsentation von „(dass) *Peter Bauklötze staunt*“ ist in Abb. 6.4 angegeben. Diese Spezifikation der Semantik verhindert gleichzeitig das Auftreten eines Artikels, dessen semantischer Beitrag mit der Spezifikation inkompatibel wäre.

Zusammenfassend gesagt, kann das idiomatische Wort *Bauklötze* nur innerhalb eines Satzes, dessen Kopf das idiomatische *staunen* ist, auftreten; zudem kann beim Nomen keine Adjunktion erfolgen. Der ganze IVP hingegen kann sehr wohl mit Adjunkten auftreten, beispielsweise mit temporalen Adverbialen („*Gestern staunte man Bauklötze bei Daimler-Crysler.*“).

Wie im Beispiel (232) gezeigt, ist unser Nomen *Bauklötze* nicht referenziell. Im Lexikoneintrag wird dies durch eine entsprechende Untersorte von *index* vermerkt.

Der endgültige Lexikoneintrag von *Bauklötze* findet sich in Abbildung 6.5.

Zu diesem IVP ist zuletzt noch anzumerken, dass keine Nominalisierung und keine Passivierung stattfinden kann. Sätze wie

(237) **Bei Daimler-Chrysler gab es ein großes Bauklötze-Staunen.*

(238) **Bei Daimler-Chrysler wurden/wurde Bauklötze gestaunt.*

⁸Wie weiter oben schon erwähnt wurde, verhält sich *Bauklötze* semantisch wie ein Adjunkt. Man könnte sich eine Konstante vorstellen, welche die Semantik des Verbs *staunen'*(*x*) verändert, also etwa (*bauklötzmäßig'*(*staunen'*))(*x*). Da jedoch die Klauseln im SEMANTIK-PRINZIP (221) so formuliert wurden, dass bei Adjunkten auch syntaktische Kopf-Adjunkt-Strukturen vorliegen müssen (siehe Richter/Sailer, 2002, S. 3), können wir nicht einfach eine Objekt-NP als Adjunkt behandeln. Die Folgen einer Umformulierung des SEMANTIK-PRINZIPS in rein semantische Relationen sind schwer zu überblicken, weshalb wir davon Abstand nehmen wollen.

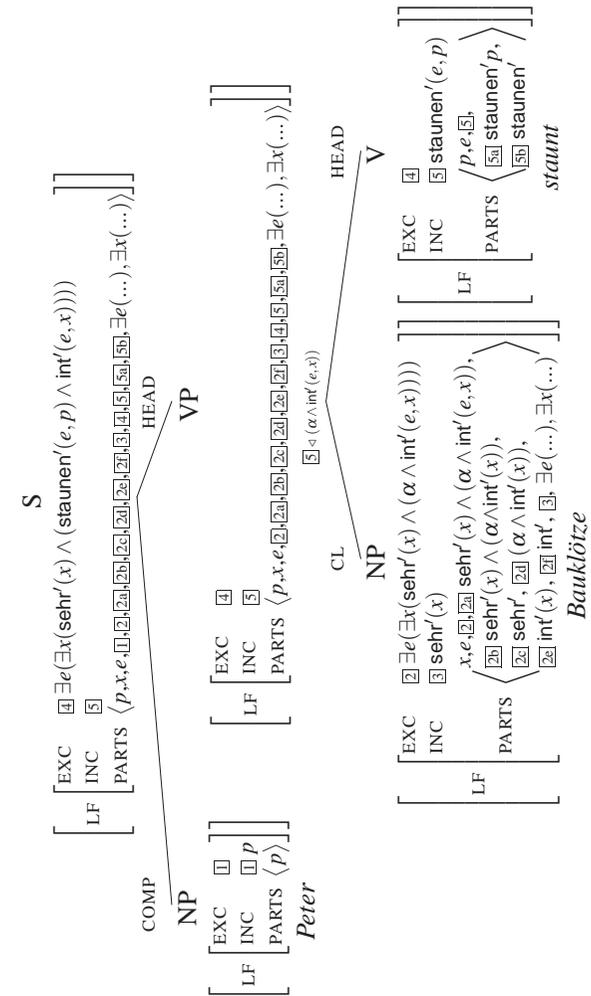
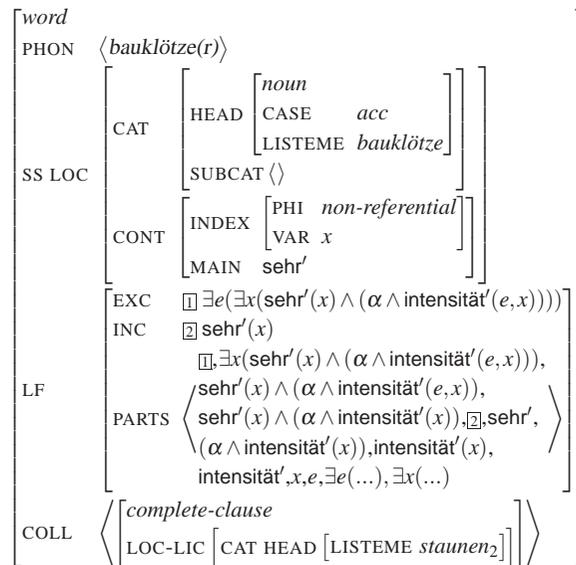


Abbildung 6.4: Semantik von (dass) *Peter Bauklötze staunt*

Abbildung 6.5: Skizze des LEs von *Bauklötze* in *Bauklötze staunen*

sind nicht belegt und erscheinen uns ungrammatisch. Deshalb müssen wir das Listem *staunen*₂ sowohl bei der Nominalisierungs-DR (s. u., Abb. 6.8) als auch bei der Passivierung ausschließen.

Wenden wir uns einem anderen Idiom zu: *jdm. einen Bären dienst erweisen*. Nach Dobrovolskij ist *Bären dienst* ein gebundenes Formativ, also ein unikales Element. Die Bedeutung des IVPs ist: „in guter Absicht etwas tun, was einem anderen, zu dessen Nutzen es gedacht war, schadet“ (Dobrovolskij, 1988, S. 88). Wir paraphrasieren das Nomen mit „gutgemeinte Tat, die schadet“ und sehen den IVP als dekomponierbar an.

Wenn man sich allerdings die Korpusdaten in COSMAS II ansieht, ist zu fragen, ob *Bären dienst* wirklich unikal ist. Nur 508 Fälle von 735 Treffern⁹ (69,1%) beinhalteten das Verb *erweisen*. Ansonsten gibt es auch Daten wie:

(239) *Ob dies nicht ein Bären dienst war?*¹⁰

Diese Datenlage lässt daher vermuten, dass *Bären dienst* ein selbstständiges Wort ist und nicht unbedingt Teil eines Idioms sein muss.¹¹ Das Verb *erweisen* kommt ebenso wie *leisten* auch mit dem morphologischen Kopf *Dienst* vor und hat in der IVP keine idiomatische Bedeutung.

Wir können daher annehmen, dass *Bären dienst* ein normales Nomen ist, welches keine zwingenden Distributionsidiosynkrasien aufweist. Diese Annahme wird durch das Verhalten des Idioms *jdm. einen Bären dienst erweisen* untermauert: Das Nomen ist referenziell¹², Adjunkte sind selten, aber möglich (siehe Appendix), eine Passivierung ist erlaubt. Eine Nominalisierung scheint ungrammatisch zu sein, was aber auch am Verb *erweisen* liegen kann. Es deutet alles darauf hin, dass sich *jdm. einen Bären dienst erweisen* syntaktisch genauso verhält wie *jdm. einen Dienst erweisen*. So konnte in COSMAS II kein

⁹Kontextanalyse mit Fenstergröße von 10 Wörtern rechts und links

¹⁰Die Presse, 09.04.1994;

¹¹*Bären dienst* ist eine selbstständige, allerdings verdunkelte bzw. undurchsichtige Metapher. Ein solches Wort wie auch andere Komposita (z.B. *Lauschangriff*) bezeichnet Duhme (1995) als „Einwortphraseologismen“, da sie sich semantisch wie Teil-Idiome verhalten. Allerdings befinden wir uns hier auf morphologischer Ebene, da Komposita den Status von Wörtern haben. Wir lassen in dieser Arbeit idiomatische Phänomene auf Wortebene unberücksichtigt und modellieren lediglich die Distributionseigenschaften.

Sailer (pers. Mitt.) weist darauf hin, dass *Bären dienst* auch eine freie Lesart hat, wenn man das Präfixoid *Bären-* als „hervorragend“ oder „bärig“ (bairisch) interpretiert. Dabei weichen die Betonungen voneinander ab (*Bären dienst* vs. *Bären dienst*, vgl. auch die beiden Lesarten von *Bären kräfte*). Nimmt man für *Bären dienst* auch eine freie Lesart an, gibt es entsprechend zwei LEs mit und ohne Distributionsbeschränkung.

¹²Ein *Bären dienst*, der der Parteibasis mächtig auf die Nerven geht. Gefunden durch Google (27.01.2004): <http://www.zdf.de/ZDFde/inhalt/31/0,1872,2085919,00.html>

analoger Eintrag für „Dienstweisung“ oder „Dienstweisen“ gefunden werden.

Von dieser Sorte gibt es noch weitere Fälle, z. B. *Armutszugnis* und *Scherflein*. Bei der IVP *sich ein Armutszugnis ausstellen* kann das Nomen auch allein vorkommen:

(240) *Wenn bis zu zehn Prozent der Jugendlichen die Schule ohne einen Abschluss verlassen, ist das ein Armutszugnis für die Gesellschaft.*¹³

Genauso verhält es sich bei *Scherflein*, welches normalerweise im IVP *sein Scherflein zu etw. beitragen* vorkommt:

(241) *Das Scherflein, das hier zu entrichten ist, bewegt sich fr 250 Seiten Manuskript zwischen 200 und 6000 Euro.*¹⁴

Die Nomina *Bärendienst* und *Scherflein* können mit verschiedenen Verben kombiniert werden. Im Gegensatz dazu hat das Verb *ausstellen* im IVP *sich ein Armutszugnis ausstellen* eine idiomatische Bedeutung. Das Idiom heißt soviel wie „einen Nachweis über die eigene Unfähigkeit erbringen“. Deshalb haben wir es bei diesem IVP nicht nur mit einem Nomen der Art *Bärendienst* zu tun; das Verb selbst ist idiomatisch und verlangt das Komplement *Armutszugnis*. Man kann nun einen neuen LE für die idiomatische Verwendung von *ausstellen* einführen, wie in Abb. 6.6 gezeigt.

Das Nomen bekommt also analog zum Fall von *Bärendienst* einen Lexikon-eintrag, der keine Distributionsidiosynkrasien enthält. Weiterhin gibt es den LE von *ausstellen*. Durch seine Subkategorisierungseigenschaften ist sichergestellt, dass überall, wo dieses Verb auftaucht, auch *Armutszugnis* als Komplement vorliegt.

Durch unser Litem-Modul lassen sich auch Pronominalisierungsfälle erfassen. Da wir durch die Beschränkung in (174) vorschreiben, dass alle Pronomen den gleichen LISTEME-Wert wie ihr Antezedens haben, sind auch die Selektionsanforderungen von *ausstellen* in folgendem Beispiel erfüllt.

(242) *Das Armutszugnis, das sich der Standort Deutschland ausgestellt hat, war schon in der Koalitionsvereinbarung auf Bundesebene mit der Finanzierung des Bundesbeitrages von 6,1 Milliarden Mark angelegt.*¹⁵

¹³Google (27.01.2004): <http://www.skh.de/pisa/index.htm?dpa/02-06-15-i.htm>

¹⁴Google (27.01.2004): <http://www.hirzel.de/universitas/archiv/eichhorn.pdf>

¹⁵Google (27.01.2004): <http://www.uvb-online.de/aufgaben/wirtschaft/positionen/2000-03-14-JahresPK.html>

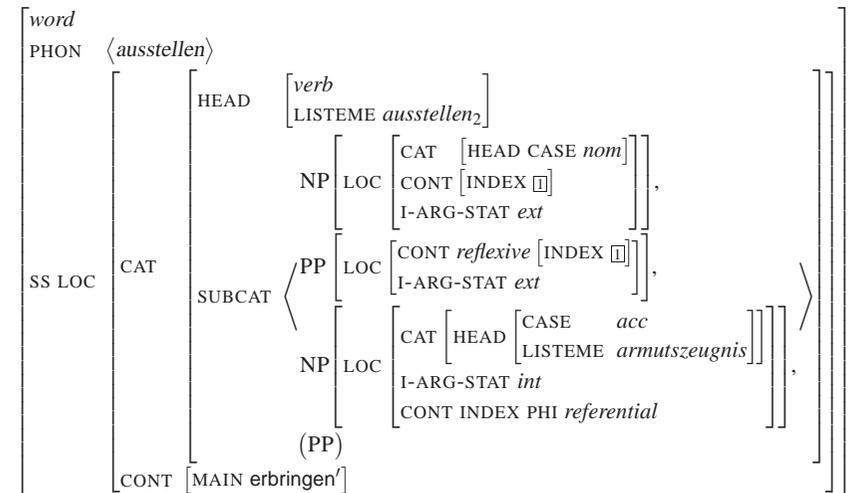


Abbildung 6.6: Skizze des LEs von *ausstellen* in *sich ein Armutszugnis ausstellen*

Nebenbei bemerkt ist unser Ansatz hier gegenüber Riehemanns Konstruktionsanalyse im Vorteil, die solche Fälle von Pronominalisierung nicht analysieren kann. Riehemann führt als Beispiel das Idiom *spill the beans* („ein Geheimnis lüften“) an. Folgender Beleg, bei dem im zweiten Satz das idiomatische *spill* ohne *beans* auftritt, bereitet Riehemanns Ansatz Probleme.

(243) *Eventually she spilled all the beans. But it took her a few days to spill them all.* (Riehemann, 2001, S. 207)

Hier referiert das Pronomen *them* auf das idiomatische *beans*. Durch die Identität des LISTEME-Wertes in unserem Ansatz sind die Anforderungen von *spill* an seine Komplemente in beiden Sätzen erfüllt. Das Antezedens von *them* seinerseits wird lizenziert durch seine COLL-Liste, in der vorgeschrieben wird, dass das idiomatische *beans* nur mit dem idiomatischen *spill* vorkommen darf.

Fassen wir noch einmal zusammen: Dekomponierbare IVPs bestehen in der Regel aus einer idiomatischen NP, die zusammen mit einem bestimmten Verb auftritt, wobei beide Teile zur idiomatischen Bedeutung des Gesamtsatzes beitragen. Wir garantieren durch unsere Analyse, dass beide Teile stets zusammen auftreten müssen: Das Verb subkategorisiert das Nomen bzw. die NP über sei-

ne SUBCAT-Liste, wobei der LISTEME-Wert des nominalen Komplements spezifiziert wird. Die NP erhält ein Objekt auf ihrer COLL-Liste, durch welches das geforderte Verb (anhand seines LISTEME-Wertes) festgelegt wird. Das LIP lizenziert alle Sätze, in denen die Anforderungen in COLL mit den tatsächlichen Eigenschaften einer Barriere übereinstimmen.

6.1.2 Syntaktische Operationen

Es steht noch die Diskussion aus, welche syntaktischen Operationen ein dekomponierbarer IVP zulässt. Die in Abschnitt 5.3.1.2 vorgestellte Verb-bewegungsanalyse verdient dabei noch einmal gesondertes Augenmerk. Bei dem Beispiel *Bauklötze staunen* haben wir in Abb. 6.2 schon eine Verb-voranstellung gehabt und sind implizit davon ausgegangen, dass der COLL-Mechanismus damit vereinbar ist. Dies soll nun explizit am Beispiel des IVPs *den Garaus machen* illustriert werden.

(244) *Den Garaus macht den Seglern die Langleinenfischerei.*¹⁶

Hier liegt eine Verbvoranstellung zusammen mit einer Extraktion der NP *den Garaus* vor. Abbildung 6.7 zeigt die Satzstruktur mit allen relevanten Informationen.¹⁷

Die Analyse der langen Abhängigkeit der Füllertochter und der Verbbewegungsmechanismus mittels Derivationsregel funktionieren so, wie weiter oben schon vorgestellt. Die Extraktion von *den Garaus* geschieht durch den Lexikoneintrag der normalen Spur und die damit erzwungene Identität des LOCAL-Wertes mit dem Element in der INHER SLASH-Menge, durch das NONLOCAL-PRINZIP und das KOPF-FÜLLER-SCHEMA.

Die Verbspur zieht die Argumente des wegbewegten Verbs an und setzt sie auf ihre SUBCAT-Liste und ihre DSL CAT SUBCAT-Liste. Das Vorkommen des DSL-Wertes entlang der Projektionslinie wird durch das KOPF-MERKMAL-PRINZIP erzwungen. Das Verb *machen* seinerseits ist Eingabe für die V1-DR

¹⁶Salzburger Nachrichten, 15.11.2000;

¹⁷Im Beispielsatz werden die Argumente des Verbs in anderer Reihenfolge realisiert als sie auf der SUBCAT-Liste erscheinen. Es gibt mehrere Möglichkeiten, dieses Problem zu lösen: Entweder man lässt (i) zu, dass ein arbiträres Element einer SUBCAT-Liste als Komplementtochter realisiert wird, (ii) man weicht die strikte Reihenfolge auf der SUBCAT-Liste auf, oder (iii) man bedient sich eines separaten Linearisierungsmechanismus. Wir wollen auf das Problem der Linearisierung bzw. des Unterschieds zwischen phänogrammatischer und tektogrammatischer Struktur nicht näher eingehen, vergleiche aber beispielsweise Kathol (2000).

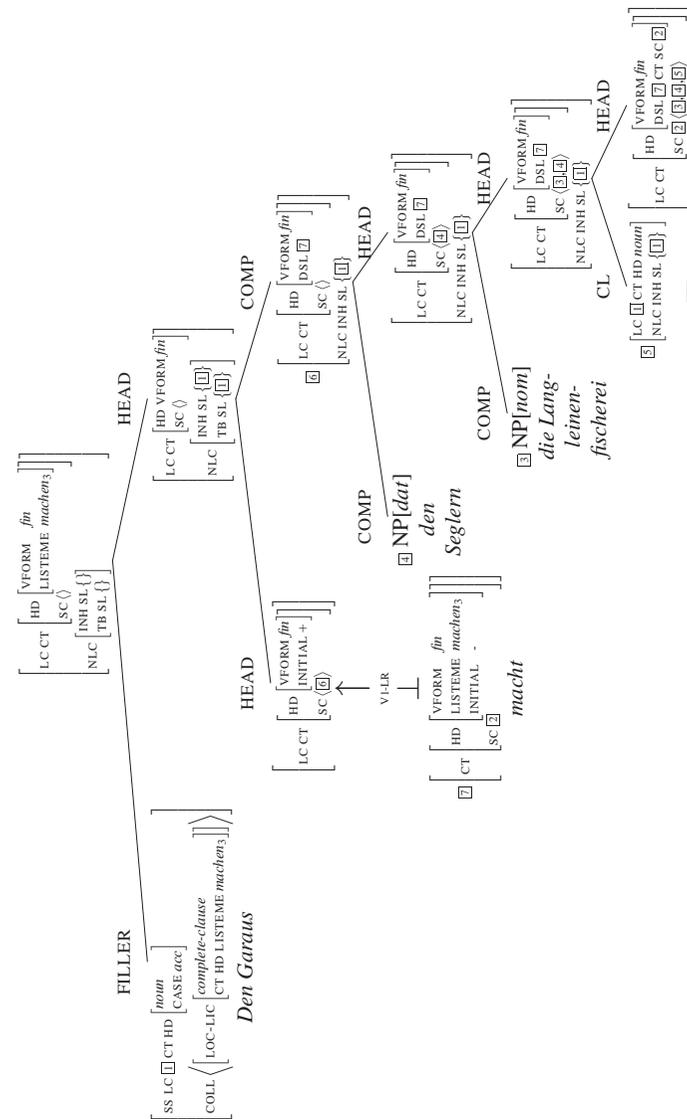


Abbildung 6.7: Analyse von *Den Garaus macht den Seglern die Langleinenfischerei*

(Abb. 5.10). Die Ausgabe subkategorisiert ein Zeichen (6), dessen DSL-Wert mit dem LOCAL-Wert der Eingabe (7) identisch ist.

Soweit die Vorstellungsmechanismen. Was unsere COLL-Analyse betrifft, sind die wichtigen Merkmale COLL bzw. LISTEME. Der Füller *den Garaus* hat einen COLL-Wert, welcher eine Barriere spezifiziert, die den LISTEME-Wert *machen₃* trägt. In der Struktur haben wir einen Knoten, welcher die minimale Barriere darstellt und dessen LISTEME-Wert genau von der geforderten Sorte ist. Also ist das LIZENSIERUNGSPRINZIP erfüllt.

Mit der Sorte *machen₃* ist ein idiomatisches Verb gemeint, das eine NP (mit SYNSEM-Wert 5) subkategorisiert, welche den LISTEME-Wert *garaus* trägt (in Abb. 6.7 nicht dargestellt). Da der LOCAL-Wert auch LISTEME mit einschließt und da die LOCAL-Werte des Füllers und der Spur identisch sind, ist die Subkategorisierungsanforderung von *machen₃* erfüllt, auch wenn *Garaus* extrahiert wurde. Weiterhin ist zu beachten, dass durch die Identität des LOCAL-Wertes 7 und des DSL-Wertes von 6 die Verbspur den korrekten LISTEME-Wert besitzt. Wir können also folgern, dass sich für dekomponierbare IVPs die (Verb)be-
wegung und der COLL-Mechanismus ohne Probleme ineinander fügen.

Was die Operation der Nominalisierung betrifft, wollen wir uns hier auf Kon-
version mit gleichzeitiger Argumentintegration beschränken, also so etwas wie:

(245) *Dampf ablassen* → *das Dampf ablassen*

Dies soll genügen, um zu zeigen, wie man einzelne IVPs von dieser Art der
Wortbildung ausschließt. Fälle wie *Abfuhrerteilung* oder *Gesichtsverlust* las-
sen wir unberücksichtigt, da eine Diskussion der morphologischen Prozesse
den Rahmen sprengen würde.

In Abb. 6.8 wird die DR für die Konversion skizziert. Da die Konver-
sion nur dann erfolgen kann, wenn es zusätzlich zum Verb noch ein in-
ternes Argument gibt, ist eine neue Art von DR erforderlich. Eine solche
dual-input-derivational-rule fordert zwei Eingaben, wobei der Eingabetyp von
IN2 nicht von der Sorte *word* sein muss. In Fällen wie *das Im-Stich-Lassen*
liegt beispielsweise eine PP als internes Argument vor. Eingabe für die DR ist
also ein Verb und ein idiomatisch internes Argument. Die Ausgabe ist einfach
die Aneinanderkettung der PHON-Werte von Argument und Verb (in dieser
Reihenfolge) mit nominalem Charakter und Genus Neutrum.

In der Eingabe wird zusätzlich spezifiziert, um welches Listem es sich dabei
handeln darf. Dabei ist der Wert von LISTEME eine Disjunktion aus endlich
vielen Listemen. Fehlt ein Listem, kann dafür die Nominalisierungsregel nicht
angewendet werden.

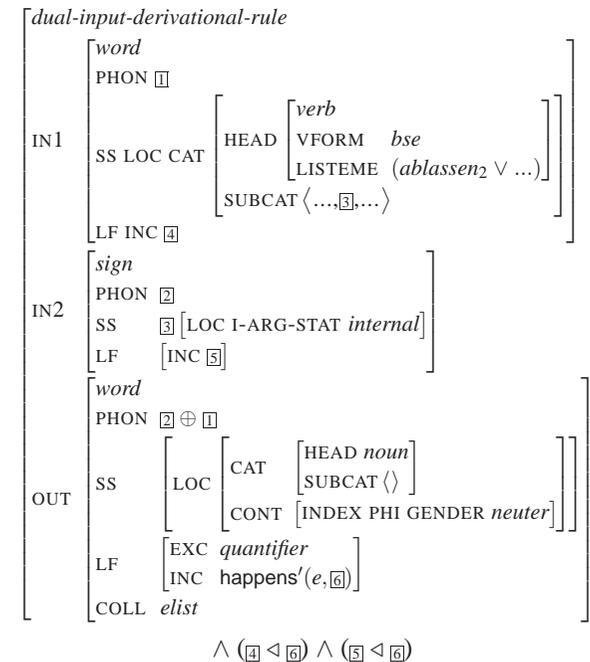
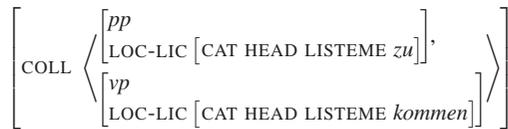


Abbildung 6.8: Nominalisierungs-DR (Konversion mit Integration des
internen Arguments)

In der DR wird, was die Semantik betrifft, festgelegt, dass sowohl die Verb-
semantik als auch die Bedeutung des Komplements in die Semantik der Nom-
inalisierung einfließen. Diese Bedeutung der Nominalisierung ist das zwei-
te Argument von *happens'*, vgl. Abschnitt 5.3.1.5. Zusätzlich wird noch ein
Quantor eingeführt, welcher die Ausgabe der DR zu einer quantifizierten NP
macht, so dass die NP vom SEMANTIK-PRINZIP erfasst wird.

6.1.3 Mehrere Elemente auf der COLL-Liste

Das LIZENSIERUNGSPRINZIP ist so formuliert, dass die LOC-LIC-Bedingun-
gen aller Elemente auf der COLL-Liste erfüllt werden müssen. Bisher hatten
wir nur jeweils ein Element auf der COLL-Liste. Es gibt jedoch auch Fälle,
in denen man zwei Elemente benötigt, um die Distribution exakt einzuschränken.
Ein Beispiel ist *Potte* im IVP *zu Potte kommen*. Das Nomen kann nur innerhalb

Abbildung 6.9: COLL-Wert von *Potte* in *zu Potte kommen*

einer PP mit der Präposition *zu* auftreten, die PP ist gleichzeitig ein Komplement von *kommen*. Die COLL-Liste von *Potte* sieht daher wie in Abb. 6.9 aus.

Ein anders geartetes Phänomen ist beim Idiom *jdm. den Garaus machen zu* beobachten. Normalerweise ist die Analyse dem vorgestellten Vorgehen analog: Man spezifiziert im LE von *Garaus* eine Barriere¹⁸ mit einem Kopf des LISTEME-Wertes *machen₃*, der seinerseits ein Objekt mit LISTEME-Wert *garaus* selegiert. Folgendes Datum kann aber damit nicht erfasst werden:

(246) *Einzig Vera Kutters im Treppenhaus gehängtes und 1999 entstandenes Negativfoto der Wiener Secession bleibt als Hinweis auf den Garaus, den die Nazis der in ihren Augen „entarteten Kunst“ machten.*¹⁹

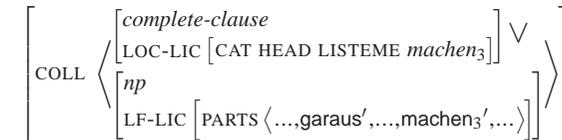
Hier modifiziert ein Relativsatz *den Garaus*, der allerdings auch das entscheidende Verb enthalten muss. Der Satz in (246) bedeutet sicherlich dasselbe wie „...bleibt als Hinweis darauf, dass die Nazis der in ihren Augen „entarteten Kunst“ den Garaus machten.“, wobei das Idiom hier in der normalen Stellung auftritt. Relativsätze wie den in (246) bezeichnet Sailer (2004c) als *propositional relative clauses*, welche eine Diskrepanz zwischen syntaktischer Struktur und Interpretation aufweisen. Sehen diese Sätze, die eine NP modifizieren, syntaktisch wie normale Relativsätze aus, so werden sie doch als Proposition interpretiert, auch wenn die NP selbst keine Interpretation als Proposition hat. Sailer (2004c) bietet eine LRS-Analyse solcher Sätze an.

Uns geht es hier um den Wert auf der COLL-Liste von *Garaus*. Wir lösen das Problem der Relativsätze mit Hilfe einer Disjunktion aus zwei möglichen Barrieren und spezifizieren folgendermaßen:

Das erste Argument der Disjunktion erfasst sämtliche gängigen Vorkommen, das zweite erlaubt Daten wie in (246). Da man an der NP-Barriere keinen Zugriff mehr auf den LISTEME-Wert von Zeichen hat, die innerhalb des Relativsatzes auftreten, muss man sich mit der semantischen Repräsentation be-

¹⁸Diese Barriere muss von der Sorte *complete-clause* sein (wie in Abb. 6.7), um Extraktion zu berücksichtigen.

¹⁹Google (04.02.04): <http://www.taz.de/pt/2001/11/30/a0123.nf/text>

Abbildung 6.10: COLL-Wert von *Garaus*

helfen. Die Semantik von *Garaus* und die der entsprechenden Version von *machen* muss Teil der Semantik einer *Garaus*-NP mit Relativsatz sein.²⁰ Wir spezifizieren dies mit Hilfe des Attributs LF-LIC und geben damit einen Ausblick auf Anwendungsmöglichkeiten, wie sie in Fußnote 13 auf Seite 112 skizziert wurden.

An dieser Stelle ist eine generelle Bemerkung zur Bewertung von Korpusdaten angebracht. Korpusbelege werden im Allgemeinen herangezogen, um die Argumentation zu untermauern oder die Bandbreite der möglichen Äußerungen aufzuzeigen. Ein Linguist verlässt sich damit nicht nur auf das eigene Sprachgefühl, sondern bezieht bei der Korpusrecherche die Verwendungweise bestimmter Ausdrücke innerhalb einer größeren Sprechergemeinschaft mit ein. Man darf selbstverständlich nicht jeden Korpusbeleg als „Beweis“ für die Akzeptabilität einer bestimmten grammatischen Erscheinung werten (so könnte man den Beleg in (246) auch als einmalig auftretenden Lapsus betrachten). Hier wird die Auffassung vertreten, dass man dem Auftreten selbst ungewöhnlicher Daten in der Diskussion schon Rechnung tragen muss. Man sollte jedoch sorgfältig den Kontext des Beleges und seine innerhalb dessen plausiblen Interpretationen berücksichtigen und auch die Tatsache im Auge behalten, dass Korpora fehlerbehaftet sein können und letztendlich nur einen Ausschnitt des tatsächlichen Sprachgebrauchs darstellen.

Damit beenden wir die Diskussion der dekomponierbaren Idiome und wenden uns den nicht dekomponierbaren IVPs zu.

²⁰Dabei muss sichergestellt werden, dass die Konstante *garaus'* auf der genannten PARTS-Liste mit der Konstante des Nomens *Garaus* identisch ist. Gleichzeitig ist entscheidend, dass dieses *garaus'* das Argument von *machen₃'* ist. Eventuell müsste man dies unter LF-LIC noch detaillieren. Genannte Charakteristika treffen allerdings auf die Semantik solcher Nomina mit Relativsatz-Adjunkt im Allgemeinen zu.

6.2 Nicht dekomponierbare IVPs

6.2.1 Analyse

Wie am Beginn dieses Kapitels schon angedeutet, definieren wir phrasale Lexikoneinträge (PLEs) für nicht dekomponierbare IVPs. Wir folgen damit Sailer (2003), der sich seinerseits u. a. auf Gazdar et al. (1985)²¹ beruft. Die Idee von syntaktisch komplexen Lexikoneinheiten findet sich z. B. auch in Di Sciullo/Williams (1988), wobei sie auch hier für idiomatische Wendungen, deren Semantik nicht kompositionell ist, postuliert werden.

Allerdings wehrte sich Weinreich (1969) innerhalb der Generativen Grammatik gegen komplexe Lexikoneinträge für Idiome wie *shoot the breeze* („plaudern“). Sein wichtigstes Argument war, dass, wenn man komplexe LEs für Idiome annimmt, die den normalen syntaktischen Prinzipien nicht gehorchen, man auch andere, vollkommen willkürliche Strukturen im Lexikon erwarten könnte. Beschreibt man idiomatische Bildungen ohne Bezug auf syntaktische Regeln, kann man nicht erklären, warum die überwältigende Zahl der Idiome syntaktisch wohlgeformt ist. Daher unterscheidet er zwischen einem Simplex-Lexikon für alle einfachen Wörter und einem Komplex-Lexikon für alle zusammengesetzten oder idiomatischen Wörter, Idiome und Sprichwörter. Er bezieht sich dabei im Wesentlichen auf Katz/Postal (1963). Die Information aus dem Komplex-Lexikon kommt erst nach der Anwendung lexikalischer Regeln und den Transformationen hinzu und kann beispielsweise regelgemäße Wortbildungen zu Gunsten von bereits vorhandenen oder „vertrauteren“ (*more familiar*) Wörtern blockieren (*inaudible* statt *unhearable*). Durch eine Idiom-Vergleichsregel, deren Anwendung obligatorisch ist, werden reguläre Ausdrücke mit gespeicherten idiomatischen Ausdrücken verglichen, woraufhin fakultativ die reguläre semantische Information mit der idiomatischen Bedeutung überschrieben wird.

Die Anwendung dieser Regel erfolgt in Weinreichs Modell bei jeder Äußerung, was die Verarbeitung im Modell sehr aufwändig und unökonomisch macht. Diesen Schwachpunkt gibt er selbst zu (Weinreich, 1969, S. 60). Darüber hinaus nimmt Weinreich trotzdem komplexe Einträge im Simplex-Lexikon an, um Einheiten wie *by and large* oder entsprechend *gang und gäbe* zu beschreiben, womit aber syntaktisch wohlgeformte und nicht wohlgeformte Idiome auf völlig verschiedene Arten im Modell behandelt werden und überdies sein Argument gegen komplexe Lexikoneinträge ausgehebelt wird.

²¹Siehe dort besonders S. 244, Anm. 33. Die Autoren klären jedoch nicht die Frage, wie konkret man PLEs im Rahmen der GPSG umsetzen kann.

Innerhalb der Psycholinguistik gibt es auch die Annahme, für Idiome ein autonomes Modul des mentalen Lexikons anzunehmen (vgl. Dobrovol'skij, 1995). Dafür sprechen zum einen komplexe Verarbeitungsmuster im Hinblick auf die bildliche Komponente von Idiomen und zum anderen das ontogenetische Argument, dass sich eine metasprachliche Kompetenz, nämlich Ausdrücke in übertragener und situativer Bedeutung zu verwenden, erst ab dem 8. Lebensjahr zeigt. Dobrovol'skij fasst zusammen,

daß nicht nur der Ausdrucksplan von Idiomen, sondern auch ihr Inhaltsplan eine Art *summerized structure* [sic!] darstellt. Memorisiert wird in der Regel nicht ein „normgerechter“ kodifizierter Wortlaut und nicht eine dazugehörige Bedeutung in der Form, wie sie in Wörterbüchern expliziert ist, sondern eine „Gestalt“, d. h. eine zusammengefaltete Struktur, die bei der Realisierung in einer Situation unterschiedlich entfaltet werden kann. (Dobrovol'skij, 1995, S. 56)

In der HPSG gibt es keine Bestrebungen, das Lexikon auf oben dargestellte Weise aufzuteilen, da in dieser Theorie der Schwerpunkt auf einer rein linguistischen Analyse und effizienter Sprachbeschreibung liegt und weniger auf einer psychologisch adäquaten Abbildung des menschlichen Sprachvermögens.²² Wir werden deshalb syntaktisch komplexe Einheiten in Form von PLEs im (allgemeinen) Lexikon speichern.

PLEs besitzen sowohl Eigenschaften von Wörtern als auch von Phrasen: Einerseits werden PLEs wie Wörter durch Lexikoneinträge lizenziert, andererseits verhalten sich PLEs wie Phrasen, da sie keine Eingabe für lexikalische Regeln sein können und syntaktische Operationen wie Topikalisierung ausgeschlossen werden, indem strukturelle Anforderungen im DTRS-Attribut festgelegt sind.

Bevor wir uns konkreten Beispielen zuwenden, soll noch darauf eingegangen werden, dass Satzteile, die durch ein PLE lizenziert werden, eine nicht kompositionelle Semantik besitzen. Dies muss durch bestimmte Prinzipien sichergestellt werden bzw. darf nicht außer Acht gelassen werden.

Den Standardannahmen in der HPSG folgend, verwenden wir *Immediate Dominance Schemas* (ID-Schemata), die normale phrasale Zeichen lizen-

²²Die Sichtweise einer „zusammengefalteten Struktur“ spricht gegen eine Analyse der Kookkurrenzen von Idiomteilen als absolute wechselseitige Determination, wie wir sie im vorigen Unterkapitel entworfen haben. Diese Sichtweise lässt sich schwerlich innerhalb einer formalen Grammatiktheorie umsetzen.

sieren. Wir benötigen ein KOPF-KOMPLEMENT-SCHEMA (KK), ein KOPF-ADJUNKT-SCHEMA (KA), ein KOPF-MARKER-SCHEMA (KM), ein KOPF-FÜLLER-SCHEMA (KF), ein PRÄDIKATSKOMPLEX-SCHEMA (PK) und ein Schema für erweiterte Prädikatskomplexe (EPK). Um nun die Anwendung eines ID-Schemas auf eine Phrase zu verhindern, welche durch einen PLE lizenziert wird, müssen wir das ID-PRINZIP folgendermaßen ändern:

(247) ID-PRINZIP (neue Version):

$$\left[\begin{array}{l} \text{phrase} \\ \text{COLL } \textit{elist} \end{array} \right] \rightarrow (\text{KK} \vee \text{KA} \vee \text{KM} \vee \text{KF} \vee \text{PK} \vee \text{EPK})$$

ID-Schemata können jetzt nur auf Phrasen angewendet werden, die eine leere COLL-Liste besitzen. Darin liegt auch schon der Schlüssel unserer Vorgehensweise: PLEs haben allesamt eine nicht leere Liste unter COLL.

Nachdem wir das ID-PRINZIP geändert haben, müssen wir mit allen Grammatikprinzipien, die bei der regulären Zusammensetzung von Zeichen eine Rolle spielen, genauso verfahren. Man muss dabei nur eine Zeile im Antezedens (diese Prinzipien sind immer als Implikation formuliert) hinzufügen, die wie folgt lautet: [COLL *elist*]. Dies betrifft das KOPF-MERKMAL-PRINZIP, das MARKING-PRINZIP, das PRINZIP DER EXTERNEN SELEKTION, das NONLOCAL-PRINZIP und das SEMANTIK-PRINZIP. Weiterhin müssten noch Prinzipien für die reguläre Verkettung der PHON-Werte und für Linearisierung entsprechend definiert werden. Beispielsweise sieht die formale Definition des KOPF-MERKMAL-PRINZIPS folgendermaßen aus:

(248) KOPF-MERKMAL-PRINZIP (neue Version):

$$\left[\begin{array}{l} \text{phrase} \\ \text{DTRS } \textit{headed-struct} \\ \text{COLL } \textit{elist} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{SS LOC CAT HEAD } \square \\ \text{DTRS H-DTR } [\text{SS LOC CAT HEAD } \square] \end{array} \right]$$

Dieses modifizierte KOPF-MERKMAL-PRINZIP sagt aus, dass Kopf-Identität nur bei regulären Phrasen gefordert ist.

In diesem Zusammenhang stellt sich die Frage, wie phrasale Lexikoneinträge und Lexikoneinträge im Allgemeinen in die Theorie eingeführt werden und wie man „falsche“ oder „unmögliche“ Einträge ausschließen kann. In Abschnitt 4.1.2 wurde schon erwähnt, dass das WORT-PRINZIP die Lizenzierung von Wörtern sicherstellt. Das Prinzip besitzt als Antezedens die Sorte *word* und als Konsequens eine große Disjunktion aller Lexikoneinträge, und zwar nur und genau der intendierten und wohlgeformten. Technisch ist also

ein Lexikoneintrag ein Disjunkt im Konsequens des WORT-PRINZIPS. Zu diesen Disjunkten kann man die PLEs jedoch nicht einfach hinzufügen, da im Antezedens nur von Objekten der Sorte *word* die Rede ist.

Sailer (2003, S. 362) nimmt eine geschickte Generalisierung vor, indem er nicht nur den PLEs den COLL-Wert *nelist* zuweist, sondern allen Lexikoneinträgen. Empirisch lässt sich dies aus dem Grund rechtfertigen, dass man damit alle die Zeichen zusammenfasst (d.h. ihnen den gleichen Wert eines Merkmals gibt), deren Eigenschaften idiosynkratischer Natur sind und sich nicht aus allgemeinen Prinzipien der Grammatik erklären lassen. Sailer (ebd.) ersetzt das WORT-PRINZIP durch das LEXIKON-PRINZIP und erfasst diese Generalisierung wie folgt: Jedes Zeichen, das einen nicht leeren COLL-Wert besitzt, muss durch einen Lexikoneintrag beschrieben werden.

(249) LEXIKON-PRINZIP:

$$\left[\begin{array}{l} \text{sign} \\ \text{COLL } \textit{nelist} \end{array} \right] \rightarrow (\text{LE}_1 \vee \dots \vee \text{LE}_m \vee \text{PLE}_1 \vee \dots \vee \text{PLE}_n)$$

Das Attribut LISTEME mit seinen individuellen Werten für jedes Element im Lexikon dient als eindeutiges Erkennungsmerkmal.²³

Durch das LEXIKON-PRINZIP ist auch die Verteilung der COLL-Werte genau festgelegt. Sobald ein Zeichen den COLL-Wert *nelist* hat, erzwingt das LEXIKON-PRINZIP, dass es durch einen LE oder PLE lizenziert wird. Alle anderen Zeichen (COLL-Wert *elist*) sind Phrasen, deren Eigenschaften aus den regulären Prinzipien folgen. Auch abgeleitete Zeichen, also solche, die durch Derivationsregeln lizenziert werden, besitzen den COLL-Wert *elist*. Deren sämtliche Eigenschaften sind ebenfalls vorhersagbar, nämlich ausgehend von der Eingabe der Regel. DRs müssen also so definiert sein, dass sie der Ausgabe den COLL-Wert *elist* zuweisen.²⁴

²³Eine alternative Vorstellung wäre, das LEXIKON-PRINZIP abzuschaffen und an dessen Stelle für jedes einzelne lexikalische Element eine Art Mini-Lexikon-Prinzip einzuführen, durch das jeder LISTEME-Wert per Implikation mit einem LE verknüpft wird. Auf die Disjunktion kann verzichtet werden, da sie schon in der Sortenhierarchie unter *listeme* vorhanden ist. Dies funktioniert jedoch nicht, da es Zeichen gibt, die den LISTEME-Wert „erben“, wie Phrasen, Pronomen oder Hilfsverben im Prädikatskomplex. Diese wären dann durch die Mini-Lexikon-Prinzipien ausgeschlossen.

²⁴Über das Attribut STORE erhält man Informationen darüber, ob ein Wort abgeleitet wurde: Ist es Ausgabe einer DR, steht auf der STORE-Liste das Eingabezeichen. Nach den eben ausgeführten Beschränkungen der COLL-Werte gilt folgende Äquivalenz, die Sailer (2003, S. 344) als STORE-COLL-PRINCIPLE bezeichnet:

$$\left[\begin{array}{l} \text{word} \\ \text{STORE } \textit{elist} \end{array} \right] \leftrightarrow \left[\begin{array}{l} \text{word} \\ \text{COLL } \textit{nelist} \end{array} \right]$$

Wir haben damit eine Unterscheidung zwischen regulären Zeichen (leere COLL-Liste) und idiosynkratischen Zeichen (nicht leere COLL-Liste) getroffen. Mit dieser Vorgehensweise haben wir gleichzeitig auch eine einheitliche Analyse für dekomponierbare und nicht dekomponierbare Idiome zur Hand. Ein nicht leerer COLL-Wert markiert eine Idiosynkrasie bzw. den idiomatischen Status eines Zeichens.

Kommen wir zu Beispielen für nicht dekomponierbare IVPs. Die Wendung *jdn. übers Ohr hauen* bedeutet „jdn. betrügen“ und ist nicht dekomponierbar. Der PLE wird in Abb. 6.11 dargestellt. Die Gesamtbedeutung *betrügen'(e,x,y)* wird für das Zeichen spezifiziert, die Teile bekommen keine eigene Bedeutung zugewiesen. Die Kopftochter ist *hauen*, die obligatorische Nichtkopftochter ist die PP *übers Ohr*.

Die Abbildung 6.11 beschreibt eine Phrase, die einige reguläre aber auch idiomatische Eigenschaften hat. Kommen wir zuerst zu den regulären Eigenschaften: Die Phrase ist eine VP mit einem verbalen Kopf. Die Kopftochter ist ein reguläres Verb, es wird jedoch ein bestimmtes *listeme* verlangt. Damit wird das dreiwertige Verb *hauen* beschrieben. Auch bei der Realisierung der Komplemente zeigt die VP reguläres Verhalten: Der SYNSEM-Wert des dritten Argumentes ist der dritte Wert auf der SUBCAT-Liste der Kopftochter (6).²⁵ Die ersten beiden SYNSEM-Werte (4 und 5) erscheinen dann qua PLE auf der SUBCAT-Liste der Phrase. Bei regulären Phrasen wäre dies eine Konsequenz aus dem SUBCAT-PRINZIP. Das PP-Komplement ist eine reguläre Phrase (der COLL-Wert ist *elist*). Es wird jedoch spezifiziert, dass die NP aus der Präposition *übers*²⁶ und dem Nomen *Ohr* bestehen muss. Ansonsten folgt die Zuweisung aller Werte innerhalb dieser Phrase aus den Lexikoneinträgen für *über* und *Ohr* und aus regulären Prinzipien.

Die idiomatische Eigenschaft dieser VP fällt sofort auf: Die Semantik der Phrase setzt sich nicht kompositionell aus den Töchtern zusammen, sondern es wird eine völlig andere Bedeutung definiert, die nichts mit den Töchtern zu tun hat. Die LRS-Prinzipien greifen hier nicht. Ebenso wenig greifen die Prinzipien der PHONOLOGY; der PHON-Wert muss explizit als Konkatenation der Töchter spezifiziert werden.

²⁵Da es sich bei der PP um ein internes Argument handelt, wird ein erweiterter Prädikatskomplex mit einer Kopf- und einer CLUSTER-Tochter beschrieben.

²⁶Das Merkmal DEF + gibt an, dass es eine definite Präposition sein muss, also *übers* und nicht *über*. Dieses Attribut geht auf Winhart (2002) zurück und wurde schon in Abschnitt 4.3.2 diskutiert.

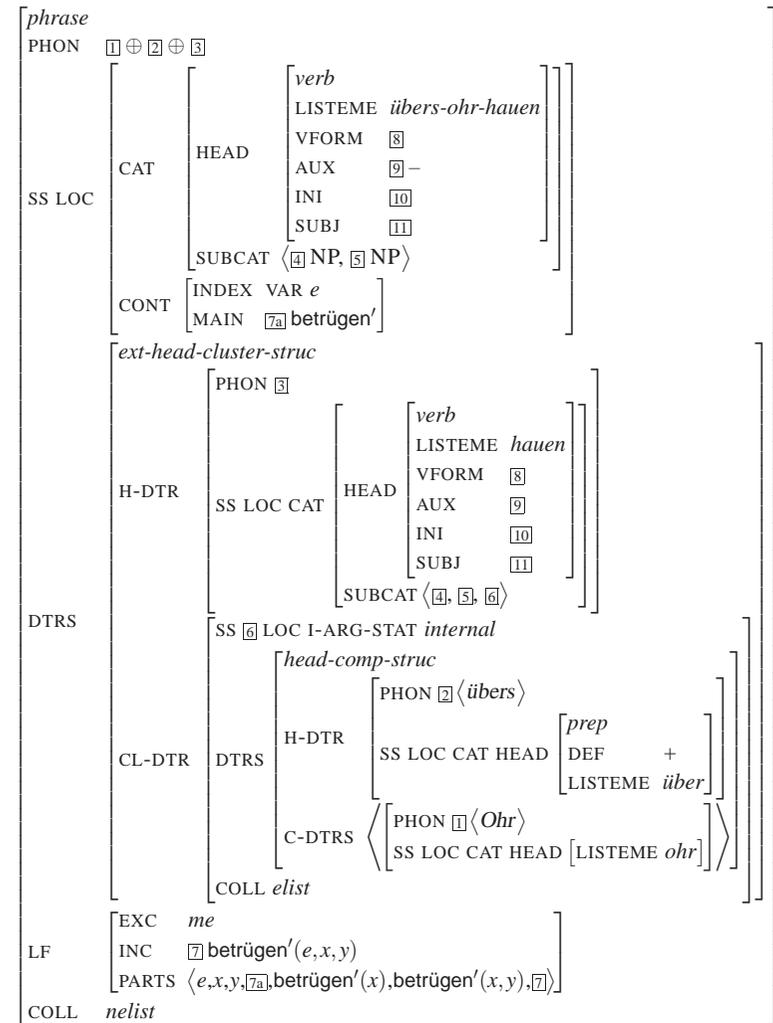


Abbildung 6.11: Skizze des PLEs von *jdn. übers Ohr hauen*

Abbildung 6.12 zeigt die Struktur von „(dass) Jens Uli übers Ohr haut“. Wie man dort sieht, kommen die einzelnen Bestandteile des Satzes erst einmal aus dem Lexikon, u. a. auch mit der freien Bedeutung. Erst auf VP-Ebene wird die kompositionelle Bedeutung durch den PLE ersetzt. Der einzige irreguläre phrasale Knoten ist somit die VP, alle anderen phrasalen Knoten folgen den normalen syntaktischen und semantischen Prinzipien. In der Abbildung wird nur der semantische Gehalt unter LF INC bzw. LF EXC angegeben und der Einfachheit halber werden die jeweils anderen LF- und CONT-Spezifikationen weggelassen. Bei der logischen Form der definiten NP wird der Maximierungsoperator Iota ι (siehe beispielsweise Krifka, 2003, bes. S. 15) benutzt; die Semantik von *übers Ohr* ist damit $\text{über}'(\iota x(\text{ohr}'(x)))$.

Die semantischen Gesichtspunkte wie Referenzialität und Modifikation wollen auch im Zusammenhang mit phrasalen Lexikoneinträgen diskutiert werden. Im Allgemeinen sind die internen Argumente in nicht dekomponierbaren IVPs **nicht referenziell**, da sie keine eigene Bedeutung besitzen. Die wenigen Fälle, die sich dennoch in Appendix A finden lassen, sind durchweg so beschaffen, dass es sich um unikale Elemente handelt. Den Fall einer referenziellen NP innerhalb eines nicht kompositionellen IVPs stellt das Beispiel *die Spendierhosen anhaben* dar.

(250) *In erstaunlichem Kontrast zu den Spendierhosen, die da die Republik Österreich angehabt hat, stand freilich eine andere Meldung über die Spesenvergütung für inländische Zeugen.*²⁷

In Abschnitt 6.2.3 werden wir sehen, dass diese unikalen Elemente einen eigenen LE bekommen und somit auch teilweise semantisch autonom sind.

Beim IVP *jdn. übers Ohr hauen* ist die NP *Ohr* nicht referenziell, d. h. es kann keine Pronominalisierung (?/!*das Ohr, über das er ihn vorhin gehauen hat*) stattfinden, ohne dass der IVP seine idiomatische Bedeutung verliert bzw. sogar dass eine ungrammatische Äußerung vorläge. Wir haben zwar im PLE dem Nomen keine Bedeutung zugewiesen, aber das Wort hat in freier Verwendung eine eigene Bedeutung, welche im konkreten Satz (vgl. Abb. 6.12) ja auch erscheint. Somit folgt die fehlende Referenzialität nicht schon automatisch nach unabhängigen semantischen Prinzipien. Da wir aber in Abb. 6.11 die NP konkret mit Hilfe des PHON-Wertes definiert haben, kann der Fall einer Pronominalisierung erst gar nicht eintreten.

Eine **interne Adjunktion** bei *jdn. übers Ohr hauen*, also eine Modifikation von *Ohr*, wäre ungrammatisch. Dies schließen wir aus, indem wir im PLE das

²⁷Neue Kronen-Zeitung, 30.06.1994;

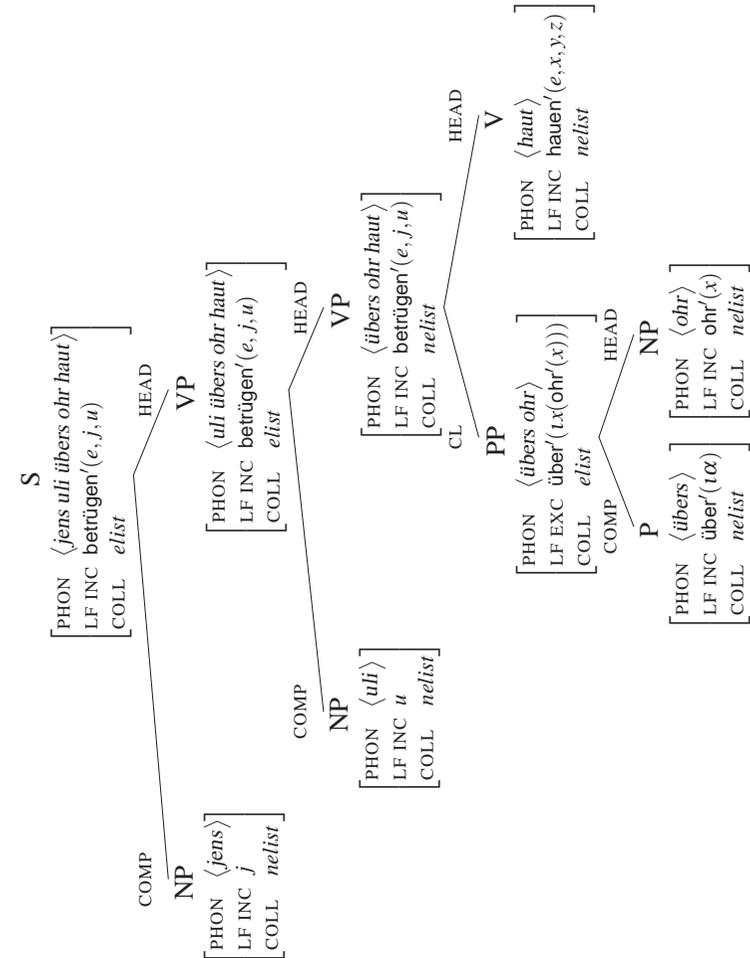


Abbildung 6.12: Struktur von (dass) Jens Uli übers Ohr haut

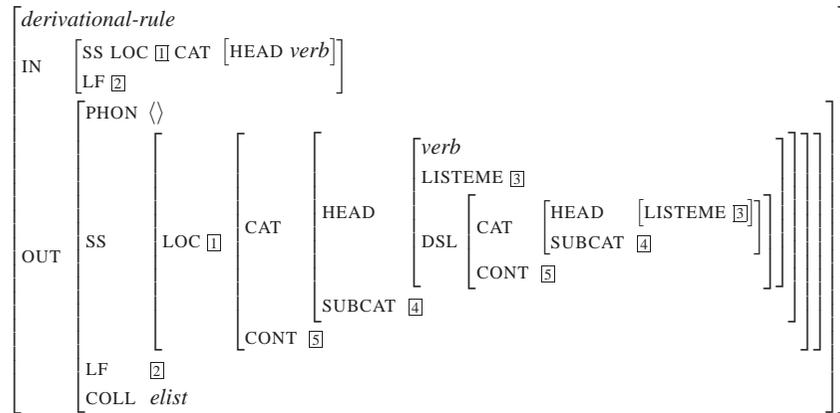


Abbildung 6.14: DR für die Verbspur

handelt, wobei der ungrammatische Fall, dass zwei unterschiedliche Verben im selben Satz Eingabe der DRs sind, ausgeschlossen wird.

Nun bleibt noch, die V1-DR (Abb. 6.15) und die MVF-DR (Abb. 6.16) entsprechend abzuändern, so dass das overte Verb keinen semantischen Beitrag mehr leistet, sondern die Semantik allein von der Projektion der Verbspur kommt. Dabei wird festgelegt, dass die LF PARTS-Liste die Werte von VAR und MAIN beinhaltet, welche wiederum von der Projektion der Verbspur kommen (Identität [4]). Das durch die DR lizenzierte Verb hat damit keinen semantischen Eigenbeitrag, weder auf lokaler noch nicht lokaler Ebene. Die Werte für EXCONT und INCONT sind unterspezifiziert. Das SEMANTIK-PRINZIP, welches die Identitäten der beiden Werte zwischen Mutterknoten und Kopftochter fordert, gibt dem vorangestellten Verb damit entsprechende eigene EXCONT- und INCONT-Werte. Diese kommen von der Komplementtochter, der Verbspur-Projektion. Auf der PARTS-Liste steht deshalb noch ein Platzhalter für den INC-Wert der Komplementtochter, um dem IContP zu genügen. Das LRS-PROJEKTIONSPRINZIP bewirkt dabei, dass der Inhalt der PARTS-Liste der Nichtkopftochter hochgereicht wird. Damit haben wir den Fall, dass die Semantik vom Nichtkopf beigestragen wird, und nicht wie sonst von der Kopftochter.

Wir müssen hierzu anmerken, dass dieses Problem, das im Zusammenhang mit PLEs aufgefallen ist, auch bei allen anderen Sätzen auftritt, sobald wir eine nicht lokale Semantik annehmen. Die gerade diskutierten Änderungen betreffen also alle vorherigen Beispiele, nicht etwa im negativen Sinn, denn

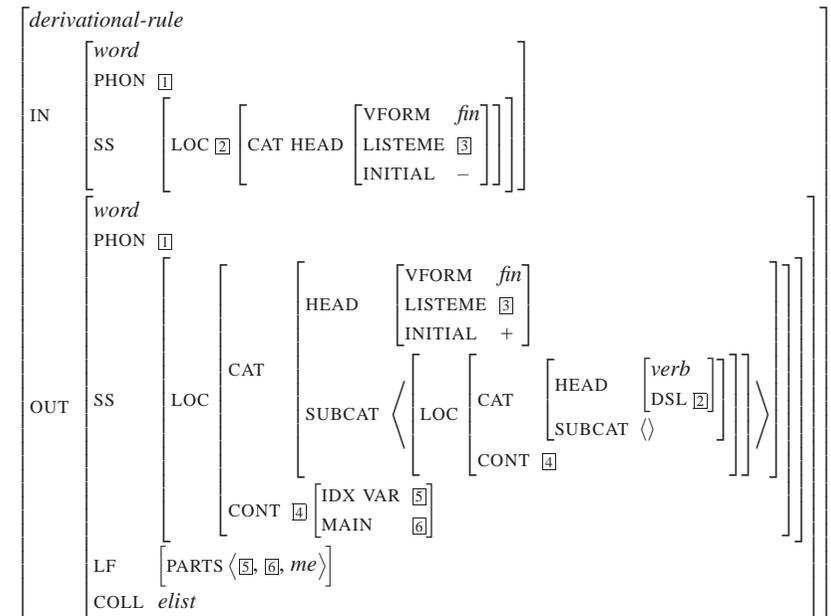


Abbildung 6.15: V1-DR (endgültige Version)

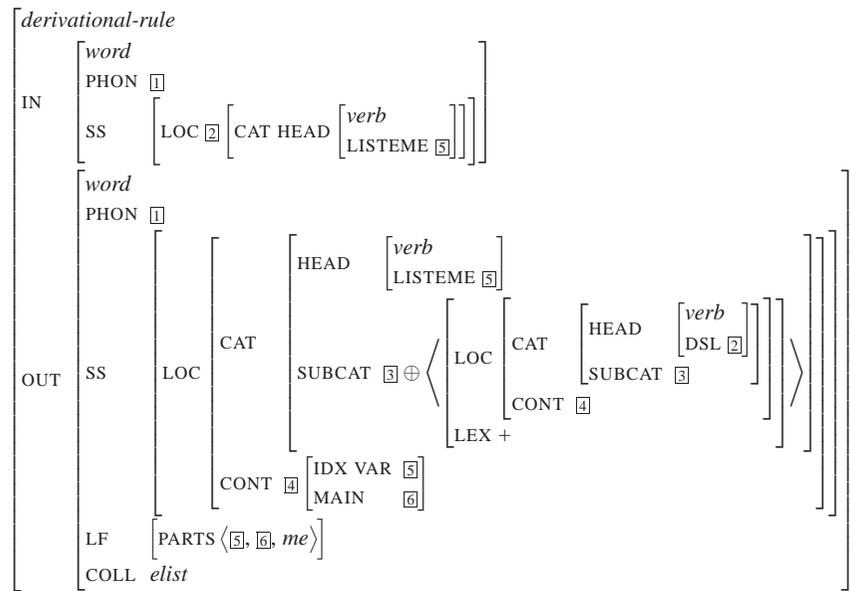


Abbildung 6.16: MVf-DR (endgültige Version)

erst mit den revidierten Fassungen der DRs kommt eine korrekte Semantik zustande, bei idiomatischen wie nicht idiomatischen Sätzen.

Einschub Für den theoretisch-technisch interessierten Leser sei an dieser Stelle noch eine Diskussion eingeschoben, die mit der „Generierung“ der Verbspur mittels einer Derivationsregel zusammenhängt. Das Problem dabei liegt in der *dominate*-Relation auf verschiedenen Ebenen der syntaktischen Struktur eines Satzes mit vorangestelltem Verb. Ein V'- oder VP-Knoten (K1), d.h. eine Verbprojektion, die noch nicht das vorangestellte Verb enthält, dominiert die Verbspur und zudem auch das Eingabe-Objekt für die DR. Ein Knoten auf noch höherer Ebene (K2) dominiert die Verbspur, das Eingabe-Objekt für die DR, das vorangestellte Verb und das Eingabe-Objekt für die V1-DR. Handelt es sich bei den Eingaben der verschiedenen DRs um unterschiedliche Objekte (allerdings mit den erforderlichen Identitäten wie dem LISTEME-Wert), ist das für die *dominate*-Verhältnisse eine unauffällige Konstellation. Sind die Eingabe-Objekte jedoch identisch, so bedeutet dies, dass der erwähnte Knoten K1 die Verbspur, das Eingabe-Objekt für die DR und damit auch das Eingabe-Objekt für die V1-DR dominiert. Damit besteht eine *dominate*-Relation zwar mit der Eingabe aber nicht mit der Ausgabe der V1-DR. Rein formal betrachtet, ist dies ein Effekt, der so nicht intendiert ist, da sich Ein- und Ausgabe von Regeln eigentlich nur in spezifischen, durch die Regel vorgegebenen Merkmalen unterscheiden sollten, nicht jedoch in ihrer Beziehung zu anderen Einheiten der Syntax. Für die Analyse bzw. die Menge der analysierbaren Sätze hat dieser Effekt u.E. keinen Einfluss. Ob man dennoch eine solche Konstellation ausschließen möchte, müssen wohl weitere Untersuchungen zeigen.

Als letzte (morpho-)syntaktische Operation sei die **Nominalisierung** erwähnt. In Abschnitt 6.1.2 haben wir eine DR für dekomponierbare Idiome skizziert. Das gleiche wollen wir nun für nicht dekomponierbare IVPs tun. Dahinter steckt folgender Gedanke: Wenn es ein Objekt gibt, welches durch einen PLE beschrieben wird, dann gibt es auch ein Objekt, das aus einer Konversion des Verbs zu einem Nomen mit integrierten internen Argumenten besteht.

Wir brauchen dafür, wie auch bei den dekomponierbaren IVPs, eine neue Art von Derivationsregel. Bei normalen DRs ist der Wert von IN und OUT stets ein Wort. Nun müssen wir aber als IN-Wert auch Phrasen zulassen, da PLEs ja Phrasen beschreiben und die Semantik bei nicht dekomponierbaren IVPs eben erst auf phrasaler Ebene hinzukommt. Genau diese Semantik wird umgeformt, um daraus eine Aktion bzw. einen Vorgang mit derselben Bedeutung wie der der Eingabephase zu machen. Als Beispiel diene uns *das Maulaffenfeilhalten*,

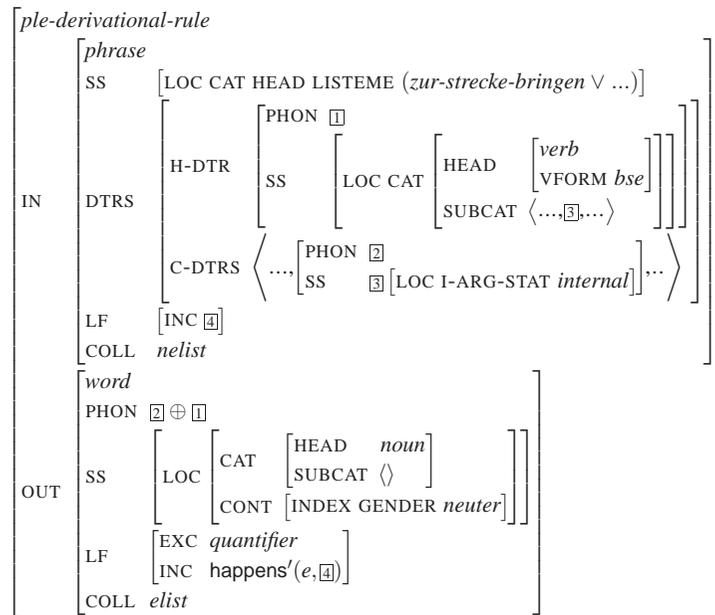


Abbildung 6.17: Nominalisierungs-DR für PLEs (Konversion mit Argumentintegration)

was analog zur entsprechenden VP *Maulaffen feilhalten* „das Gaffen“ bedeutet.

In Abb. 6.17 ist eine solche *ple-derivational-rule* skizziert. Der IN-Wert beschreibt eine idiomatische Phrase (COLL *nelist*) mit einem verbalen Kopf und einem internen Argument. Der OUT-Wert ist, parallel zur Regel in Abb. 6.8, eine Nominalisierung mit integriertem Argument. Die Semantik setzt sich hier nur aus der idiomatischen Bedeutung (4) und dem Prädikat *happens'* zusammen, wobei 4 ein Argument des Prädikats ist. Der Wert unter LISTEME ist wieder eine Disjunktion aus den Untersorten von *listeme*, welche die Nominalisierung zulassen.

Wir haben in diesem Abschnitt phrasale Lexikoneinträge eingeführt, um damit nicht dekomponierbare Idiome zu beschreiben. Die Wörter, welche die Idiombestandteile bilden, kommen aus dem Lexikon und bringen ihre eigene Semantik mit. An einem bestimmten phrasalen Knoten jedoch, der mit den Spezifikationen eines PLEs (v. a. unter DTRS) identisch ist, wird die Semantik nicht

mehr kompositional weitergereicht, sondern eine idiosynkratische Semantik wird durch den PLE lizenziert. Die Beschreibung dieses Knotens beinhaltet einen nicht leeren COLL-Wert, was verhindert, dass alle regulären Prinzipien der Grammatik greifen, beispielsweise das SEMANTIK-PRINZIP oder die lineare Verkettung der PHON-Werte.

Manche nicht dekomponierbare IVPs haben eine freie Lesart, und zwar dann, wenn die Idiomteile nur aus frei vorkommenden, an sich nicht idiomatischen Wörtern bestehen. Für solche Fälle nehmen wir zwei unterschiedliche Strukturen des jeweils gleichen Satzes an. Einerseits eine vollkommen regelgemäße Struktur mit kompositionaler Semantik, welche den normalen Grammatikprinzipien folgt, andererseits eine Struktur, die durch einen phrasalen Lexikoneintrag lizenziert wird.

Es liegt an unabhängigen semantischen Eigenschaften, wenn solche, aus an sich nicht idiomatischen Wörtern bestehende IVPs keine freie Lesart haben. Bei *einen Streit vom Zaun brechen* beispielsweise verlangen die semantischen Charakteristika einer Tätigkeit wie *etw. vom Zaun brechen* normalerweise ein konkretes Objekt (z. B. *eine Latte*) und kein Abstraktum wie *Streit*.

6.2.3 Unikale Elemente in PLEs

Im vorherigen Abschnitt haben wir gesehen, dass die Idiomteile von nicht dekomponierbaren IVPs zunächst als nicht idiomatische Wörter aus dem Lexikon kommen. Sailer (2003) betont, dass dies ein Vorteil gegenüber beispielsweise Riehemanns Ansatz ist (ebd. S. 325). Riehemann nimmt zu jedem frei verwendbaren Wort, das als Idiomteil auftauchen kann, zusätzlich ein semantisch leeres Wort an, welches dann Teil der idiomatischen Konstruktion ist. Dies wird bei PLEs nicht so gehandhabt.

Nun stellt sich die Frage, wie man unikale Elemente behandeln kann, die zwar Teil eines nicht dekomponierbaren IVPs sind, die aber sonst in keinem anderen Kontext vorkommen. Man kann für diese unikalen Elemente eigene Lexikoneinträge spezifizieren. Dies scheint auf den ersten Blick zu einer unnötigen Verdopplung von Information im Lexikon zu führen. Jedoch müssen die einzelnen Wörter irgendwie lizenziert werden, und auch empirisch lässt sich diese Vorgehensweise motivieren, da ein Sprecher diese unikalen Elemente als Wörter der Sprache identifizieren kann. Jeder kennt z. B. das Wort *Maulaffen* im IVP *Maulaffen feilhalten*. Dieses Wort hat heute keine eigene Semantik mehr. Um dem Rechnung zu tragen, ist es möglich, für unikale Elemente in nicht dekomponierbaren IVPs semantisch unterspezifizierte LEs vorzusehen. Die Semantik des IVPs wird dann vom phrasalen Lexikoneintrag an einem hö-

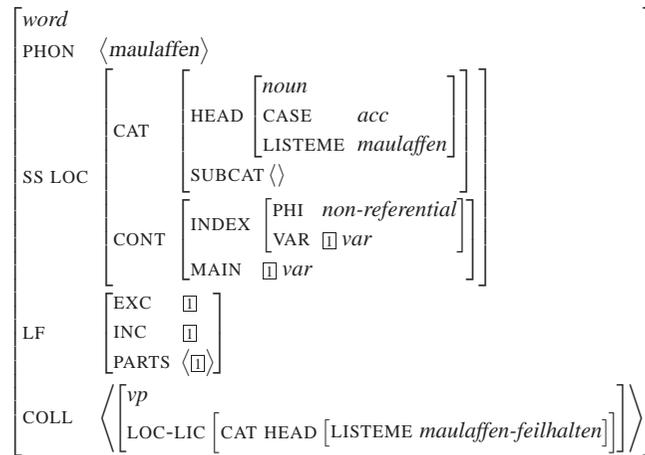


Abbildung 6.18: Skizze des LEs von *Maulaffen* in *Maulaffen feilhalten*

heren Knoten in der Struktur vergeben. Das Auftreten eines solchen unikalen Elementes wird durch den COLL-Wert beschränkt, wie bei dekomponierbaren IVPs gesehen. Beispielhaft sei der LE von *Maulaffen* in Abb. 6.18 gezeigt.³¹ Der IVP *Maulaffen feilhalten* ist deshalb ein gutes Beispiel, weil das Hauptverb *feilhalten* auch unikal ist. In dessen Lexikoneintrag schreiben wir vor, dass ein nominales Komplement mit LISTEME-Wert *maulaffen* subkategorisiert wird, um auch hier das Auftreten außerhalb des IVP auszuschließen. Die Semantik bleibt ebenso wie bei *Maulaffen* unterspezifiziert. Erst im phrasalen Lexikoneintrag für den IVP kommt die Semantik „gaffen“ hinzu.

6.2.4 Idiomatiche Adjunkte

In unserem Korpus in Appendix A findet sich eine Reihe von nicht dekomponierbaren Präpositionalphrasen, die folgende Eigenschaften haben: Erstens sehen diese Ausdrücke aus wie Vergleiche – Burger (2003) nennt sie „komparative Phraseologismen“. Syntaktisch gesehen haben sie Adjunktcharakter. Zweitens haben sie, aus semantischer Sicht, intensivierende bzw. verstärkende Bedeutung. Die unten aufgeführten Beispiele sind dabei alle teil-idiomatisch,

³¹Anmerkung zum Merkmal LF: Die Werte von VAR und MAIN sind *per definitionem* immer auf der PARTS-Liste, bleiben hier aber unterspezifiziert. Die Semantik eines unikalen Elementes ist dann genauso wie die von Pronomen, d.h. sie füllen eine Argumentstelle, leisten aber keinen semantischen Beitrag.

da die PP nicht wörtlich zu verstehen ist. Man kann diese Art der Vergleiche noch genauer spezifizieren, je nachdem wie motivierbar bzw. wie durchsichtig der Vergleich (das *tertium comparationis*) ist.³² Bei Agricola (1992, S. 29) werden diese Wendungen als **stehende Vergleiche** bezeichnet und es wird angemerkt, dass der Gesamtsinn keine Umdeutung erfährt. Daneben gibt es Idio-me, die nicht verstärkend sind, sondern das Verb in anderer Weise semantisch spezifizieren, z.B. in *dastehen wie ein Ölgötze* („unbeweglich“). Folgende Beispiele seien herausgegriffen, die Bedeutung der fett gedruckten Phrase steht jeweils dahinter.

- (256) *sich freuen wie ein Schneekönig* („sehr“)
 (257) *futtern wie ein Scheunendrescher* („sehr viel“)
 (258) *etw. für einen Appel und ein Ei kaufen* („sehr billig“)
 (259) *dumm wie Bohnenstroh sein* („sehr“)
 (260) *klar wie Kloßbrühe sein* („ganz“)

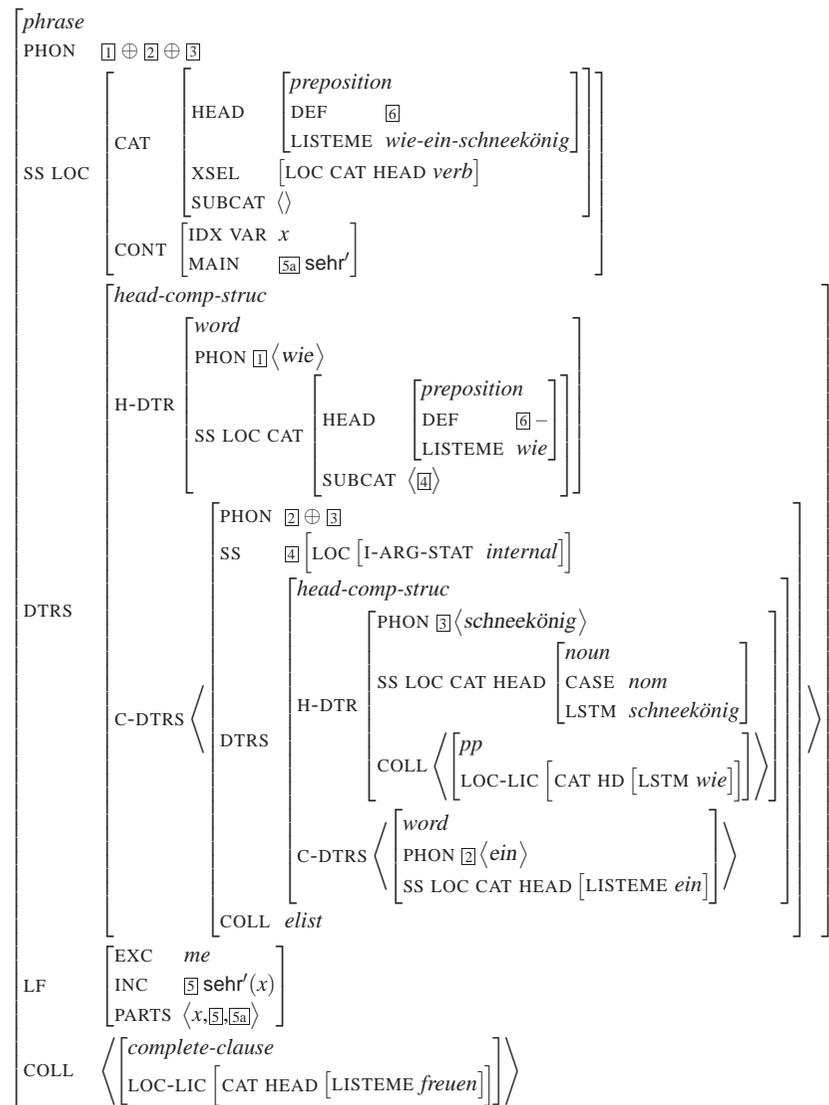
Unser Analysevorschlag behandelt diese Adjunktphrasen ähnlich wie komplette IVPs. Wir sehen einen PLE vor, in welchem die Bedeutung auf phrasaler Ebene spezifiziert wird. Wir haben es bei $[_{VP} \textit{freuen} [_{PP} \textit{wie ein Schneekönig}]]$ mit einer Kopf-Adjunkt-Struktur zu tun. Die PP besitzt deshalb einen XSEL-Wert (in anderen Arbeiten MOD), um als Adjunkttochter die Kopftochter zu selegieren. Umgekehrt selegiert das Verb nicht die PP, somit ist es kein Komplement. Um die Distribution korrekt zu erfassen, wird auf der COLL-Liste der PP der entsprechende LISTEME-Wert *freuen* spezifiziert. Dieser Vergleich (d.h. exakt dieser Wortlaut mit genau dieser Bedeutung) kann somit bei keinem anderen Verb auftreten. Der PLE für (256) ist in Abb. 6.19 dargestellt.

Da es sich bei *Schneekönig* auch um ein unikales Element handelt, sieht dessen Lexikoneintrag analog zum LE von *Maulaffen* aus. Er beinhaltet die Information, dass das Nomen nur innerhalb einer *wie*-PP vorkommen darf.

6.2.5 Syntaktisch anomale IVPs

Eine weitere Klasse von Sonderfällen, die gesonderter Aufmerksamkeit bedarf, ist die der syntaktisch anomalen IVPs, vgl. dazu Abschnitt 2.3.1, insbesondere Beispiel (33), hier wiederholt als (261).

³²Beispielsweise ist das Tier bei *flink wie ein Wiesel* wirklich flink und der Vergleich ist durchsichtig (siehe Burger, 2003, S. 45).

Abbildung 6.19: Skizze des PLEs von *wie ein Schneekönig*(261) *mit jdm. ist nicht gut Kirschen essen*

Diese anomale Konstruktion besteht zwar aus normalen Wörtern der deutschen Sprache, ist jedoch „eingefroren“. Die wörtliche Bedeutung kann man mit dieser Konstruktion nicht ausdrücken.³³ Die idiomatische Bedeutung ist undurchsichtig und eine Motivation schwer herzuleiten.³⁴

Diese „Wortkette“, bestehend aus *gut Kirschen essen*, kann nicht adjungiert werden. Ebenso wenig kann das Verb in einer anderen Form verwendet werden.

(262) **Mit ihr ist nicht gut reife Kirschen essen.*(263) **Mit ihr ist/sind nicht gute Kirschen essen.*(264) **Mit ihr ist nicht gut Kirschen zu essen.*

Auf Grund dieser Datenlage schlagen wir vor, diese „Wortkette“ als phonologische Sequenz im Lexikon zu speichern und ihr die Bedeutung „auskommen“ zuzuweisen. Die Wendung wäre dann semantisch so wie die parallele Konstruktion (265) zu analysieren.

(265) *Mit ihr ist nicht auszukommen.*

Die syntaktische Konstruktion des IVPs ist nicht nur innerhalb der eingefrorenen Sequenz anomal, sondern es fehlt auch das Subjekt. Wir wollen den Satz deshalb so analysieren, dass ein idiomatisches Verb *sein* (wir nehmen als LISTEME-Wert *sein*₅) die Sequenz *gut Kirschen essen*, eine *mit*-PP, aber kein Subjekt subkategorisiert. Der LE von *gut Kirschen essen* sieht wie in Abb. 6.20 aus.

Wir legen ferner im COLL-Wert fest, dass sich der Ausdruck im Skopus einer Negation befinden muss, da das Negationsadverb in der Wendung obligatorisch ist. Jedoch sind auch andere Negationsadverbien möglich:

(266) *Mit ihm war noch nie gut Kirschen essen.*

Wir spezifizieren dies mit Hilfe des Attributs LF-LIC, vgl. Fußnote 13 auf Seite 112.

³³Für die wörtliche Bedeutung müsste man sagen: *mit jdm. sind Kirschen nicht gut zu essen* – was auch immer ein passender Kontext dafür wäre.

³⁴Der IVP geht auf ein mittelalterliches Sprichwort zurück: „Wer mit Herren Kirschen essen will, dem werfen sie die Stiele in die Augen“ (Duden 11, 2002).

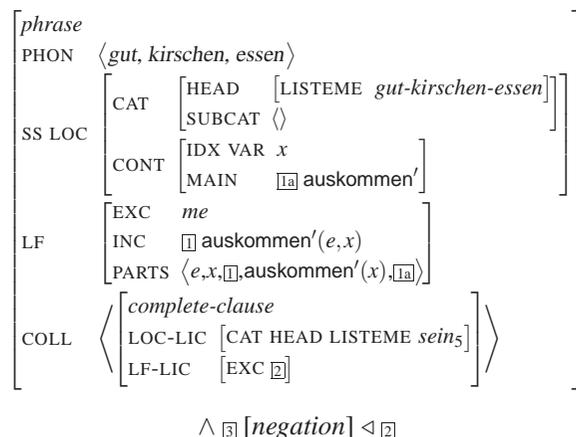


Abbildung 6.20: Skizze des PLEs von gut Kirschen essen

6.2.6 Externe Modifikation

Zum Abschluss der Diskussion phrasaler Lexikoneinträge sei noch eine andere Art Modifikation, die externe³⁵, angesprochen. Die Möglichkeit interner Modifikation, also der Adjunktion von internen Argumenten, haben wir schon beschrieben. Es kann daneben durchaus vorkommen, dass die idiomatische Bedeutung nicht dekomponierbarer IVPs als Ganze modifiziert wird. Vergleiche folgende Beispiele einer Modifikation des IVPs *das Handtuch werfen*³⁶:

(267) *Ex-Flegel John McEnroe, auch mit von der Partie, hatte das Handtuch schon früher geworfen.*³⁷

(268) *Der Rest unserer Equipe warf vorzeitig das Handtuch.*³⁸

³⁵Als externe Modifikation bezeichnet Sailer (2003) in Kap. 8.4 ein etwas anders geartetes Phänomen. Er untersucht Daten wie *kick the proverbial/social bucket*, was im Deutschen etwa *ins sprichwörtliche/gesellschaftliche Gras beißen* bedeutet. Auf diese Art von Daten sind wir schon am Ende von Kapitel 2 eingegangen.

³⁶Diese Wendung hat, im Gegensatz zur wörtlichen Lesart, bei der eine Lokativ-PP angegeben werden muss (wohin das Handtuch geworfen wurde), keine freie Lesart (Kennzeichen: Valenzänderung).

³⁷Neue Kronen-Zeitung, 22.07.1995;

³⁸Neue Kronen-Zeitung, 06.07.1996;

Hier wurde weder *das Handtuch* noch das Verb *werfen* modifiziert, sondern die Bedeutung „aufgeben“ wurde semantisch näher bestimmt bzw. der gesamte IVP modifiziert.³⁹ Man muss also die Möglichkeit der Integration weiterer semantischer Komponenten in die Gesamtbedeutung in den PLEs berücksichtigen.

Im PLE von *das Handtuch werfen* (Abb. 6.21) wurden neben der Möglichkeit der externen Modifikation auch andere, bereits besprochene Angaben gemacht. So bleibt der PHON-Wert (3) der Kopftochter (6) unterspezifiziert, da *werfen* nicht nur in flektierter Form, sondern auch mit Hilfsverben innerhalb des (erweiterten) Prädikatskomplexes auftreten kann. Weiterhin wird die geforderte Struktur so erweitert, dass zwischen dem Knoten mit der Clustertochter *das Handtuch* und dem Kopf mit LISTEME-Wert *werfen* (7) noch weitere Knoten, möglicherweise mit Adjunktöchtern, stehen können. Außerdem können noch Hilfsverben hinzukommen, welche ja auch zur Semantik des Prädikatskomplexes beitragen. Die Relation von *is-head* spielt hier die entscheidende Rolle.

Liegt eine externe Modifikation vor, behandeln wir sie folgendermaßen: Hat man beispielsweise eine VP mit dem Wortlaut *ärgerlich das Handtuch warf* (was bedeutet „ärgerlich aufgab“), muss das Adverb übernommen werden. Das Adverb hat die logische Form *ärgerlich'(A[e])*, ist also ein Operator, welcher ein Argument *A* nimmt, das mindestens die Variable *e* enthält.⁴⁰ Die logische Form der VP, unter Berücksichtigung, dass das direkte Objekt ein Definitum vom Typ *e* ist, sei damit

$$\text{ärgerlich}'(\text{werfen}'(e, x, \text{ty}(\text{handtuch}'(y))))$$

Da aber die intendierte Bedeutung die Form *ärgerlich'(aufgeben'(e,x))* hat, müssen der Term *werfen'* und dessen Subterme durch *aufgeben'* ersetzt werden. Dies kann man auf der PARTS-Liste vornehmen: Alle Elemente, welche zur wörtlichen Lesart beisteuern, werden „gelöscht“, um nur diejenigen Bestandteile zu behalten, die noch in der idiomatischen Lesart enthalten sein sollen, also die Semantik der Adjunkte und Hilfsverben. Um eine solche Listenmanipulation durchzuführen, benötigen wir die Relation *member*, die in 269 definiert ist. Ein Objekt [1] steht in der *member*-Relation mit einer Liste [2], falls [1] das erste Element auf der Liste ist oder [1] in der *member*-Relation mit dem Rest der Liste steht.

³⁹Diese Art Modifikation kann auch bei dekomponierbaren IVPs vorkommen, vgl. dazu die Adjunktion im Eintrag von *Garaus* im Appendix.

⁴⁰Alternativ könnte man das Adverb auch als intersektiv betrachten: $A[e] \wedge \text{ärgerlich}'(e)$.

(269) Die Relation *member*:

$$\forall [1] \forall [2] \left(\left(\begin{array}{l} \text{member}([1],[2]) \leftrightarrow \\ \left(\begin{array}{l} [2] \text{list} \\ \text{FIRST } [1] \end{array} \right) \vee \\ \exists [3] \left(\begin{array}{l} [2] \text{list} \\ \text{REST } [3] \end{array} \right) \wedge \text{member}([1],[3]) \end{array} \right) \right)$$

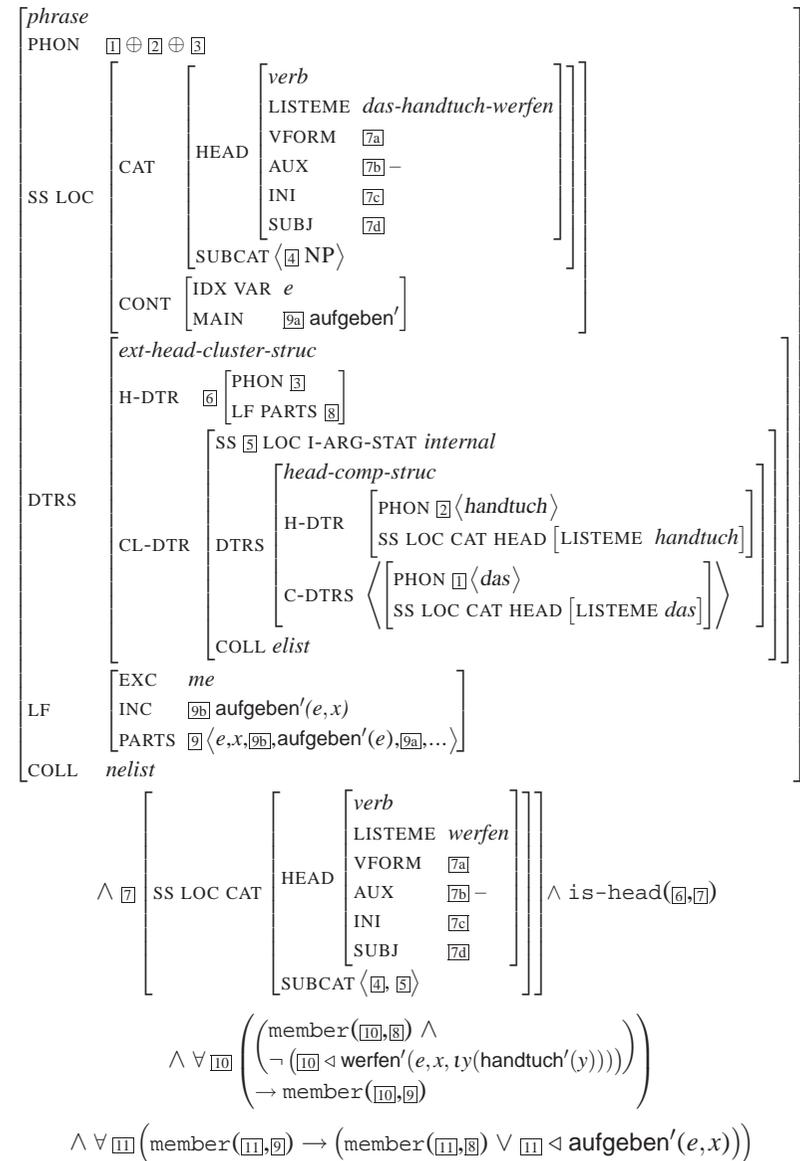
Im PLE von *das Handtuch werfen* (Abb. 6.21) wird diese Relation angewandt: Die PARTS-Liste [9] enthält alle Elemente der Liste [8] bis auf den Ausdruck *werfen'(e, x, ty(handtuch'(y)))* und sämtliche Unterausdrücke davon. Die Bedeutung des Idioms *aufgeben'(e, x)* wird vom PLE eingeführt und erscheint zusätzlich auf der PARTS-Liste [9]. Weiterhin gibt es ein Konjunkt, das die Bedingung ausdrückt, dass auf der Liste [9] nur Elemente von [8] oder *aufgeben'(e, x)* mit allen seinen Subtermen stehen darf. Zuletzt wird im PLE gefordert, dass, wenn eine Einbettungsbeziehung zwischen einem Adjunkt und der Semantik des Verbs *werfen* bestand, diese Beziehung nun auch zwischen dem Adjunkt und *aufgeben* bestehen muss. In Abb. 6.21 wird dies der Einfachheit halber in Worten ausgedrückt. Mit diesem – etwas komplizierten – Vorgehen wird erreicht, dass Adjunkte und andere semantische Komponenten zur idiomatischen Bedeutung „aufgeben“ hinzukommen können und gleichzeitig (nur und genau) die wörtliche Bedeutung des Verbs und des Komplementes unterdrückt wird.

6.3 Zusammenfassung und Diskussion

Die Analyse der IVPs im Rahmen der HPSG gliederte sich in zwei Teile. Zuerst wurden die dekomponierbaren Idiome betrachtet. Diese bestehen aus idiomatischen Wörtern, deren Distribution bzw. Kookkurrenz wir mit Hilfe des COLL-Attributes spezifizieren und so sicherstellen, dass alle Idiomteile stets gemeinsam auftreten. Ungrammatische Vorkommen von idiomatischen Wörtern werden so ausgeschlossen.

Es wurde erläutert, wie syntaktische Operationen und deren Behandlung in der HPSG mit unserem Modul interagieren und wie man einzelne Idiome korrekt von syntaktischen Transformationen, für die sie defektiv sind, ausschließen kann.

Daraufhin analysierten wir die nicht dekomponierbaren Idiome mit Hilfe von syntaktisch komplexen, phrasalen Lexikoneinträgen (PLEs), und legten fest,



∧ Wenn eine Einbettungsbeziehung zwischen einem Element auf [8] und *werfen'* bestand, so muss diese Beziehung auch zwischen demselben Element (nun auf [9]) und *aufgeben'* bestehen.

Abbildung 6.21: Skizze des PLEs von *das Handtuch werfen*

dass diese allesamt eine nicht leere Liste als Wert von COLL besitzen. Mit dieser Eigenschaft können wir PLEs von der Anwendung der Grammatik-Prinzipien zur regelgemäßen semantischen Komposition ausschließen. Darüber hinaus haben wir die Analyse der beiden Idiom-Typen vereinheitlicht, da in unserem Ansatz alle Lexikoneinträge einen nicht leeren COLL-Wert besitzen und somit von den nicht idiosynkratischen Zeichen abzugrenzen sind.

Auch bei den nicht dekomponierbaren IVPs wurde die Problematik der syntaktischen Operationen angesprochen und weiterhin sind wir auf unikale Elemente eingegangen. Idiomatiche Adjunkte bilden eine Unterklasse dieser Idiom-Gruppe. Für sie wurde ebenso eine Analyse skizziert wie für die syntaktisch anomalen IVPs. Zuletzt haben wir dargestellt, wie es zu einer externen Modifikation der Gesamtbedeutung eines nicht dekomponierbaren Idioms kommen kann.

Bei den Analysen wurden jeweils Beispiele aus unserem Korpus herangezogen, welche die zu diskutierenden Eigenschaften aufweisen. Wir haben für diese IVPs Lexikoneinträge skizziert, welche (nur) die relevanten Merkmale erwähnen. Da in der HPSG sämtliche Objekte maximal spezifische Sorten für alle ihre Attribute haben müssen, gestalten sich vollständige Beschreibungen der tatsächlich vorkommenden Objekte oftmals viel umfangreicher. Da jedoch Lexikoneinträge (nahezu) alle möglichen Vorkommen des beschriebenen Wortes (oder der Phrase) lizensieren müssen, kann und muss manches Attribut unterpezifiziert bleiben.

Die erstellten Lexikoneinträge fungieren ihrerseits auch nur als Beispiele. Es sollte deutlich geworden sein, dass man für jedes Idiom einen eigenen (P)LE definieren muss, wobei die Einträge immer etwas anders aussehen – abhängig von den idiosynkratischen Eigenschaften. Man kann also keinen der in dieser Arbeit definierten (P)LE als Muster für alle möglichen anderen Phraseologismen betrachten und unverändert damit andere Phraseologismen beschreiben. Die definierten Lexikoneinträge sollten jedoch die relevantesten Charakteristika und Schwierigkeiten beinhalten, so dass ein Grammatikschreiber (bzw. derjenige, der das Lexikon erstellt) sich bei der Definition von (P)LEs für andere Phraseologismen daran orientieren kann.

Man kann bei unserem Analysevorschlagn eine Asymmetrie beobachten. Einerseits wird ein internes Argument vom Verb direkt selegiert. Über die SUBCAT-Liste werden bis zum LISTEME-Wert alle Anforderungen spezifiziert. Andererseits kann ein Argument das idiomatiche Verb, mit dem es einhergeht, nicht selegieren, sondern legt seine Distribution über den COLL-Wert fest. Diese

Asymmetrie mag vielleicht störend wirken, ist aber nicht abwegig, wenn man die Verhältnisse von Verb und Komplement berücksichtigt. So hat das Verb eine direkte Subkategorisierungsbeziehung zu seinem Komplement. Man muss eingestehen, dass durch die Listenselektion eine semantische Beschränkung übergangen wird, die oftmals vorausgesetzt wird: Es werden nur lexikalisch-semantische Konzepte selegiert und nicht einzelne Listeme. Jedoch stellt unser Vorgehen, auch Listeme zu selegieren, eine ökonomische Erweiterung der Selektionsbeziehung dar. Bei der Wendung *jdm. einen Schrecken einjagen* scheint *einjagen* direkt das Nomen *Schrecken* zu subkategorisieren, und nicht eine NP, die das Nomen u. a. enthält.

(270) **Er hatte mir etwas eingejagt, das mir wie ein Schrecken durch Mark und Bein fuhr.*

(271) **Er hat mir etwas wie einen Schrecken eingejagt.*⁴¹

Es liegt hier also eine interne Selektionsbeziehung zwischen *einjagen* und *Schrecken* zu Grunde. Es wäre daher nicht intuitiv, mit einer Barriere zu arbeiten, etwa einer bestimmten verbalen Projektion von *einjagen*, deren Komplementtochter eine NP mit dem Kopf *Schrecken* ist.

Auf der anderen Seite ist es aus Sicht des internen Komplementes nicht klar, welcher Knoten selegiert werden soll, nachdem man die Konstituenten im Mittelfeld vertauschen bzw. es zu einer Verbalkomplex-Bildung kommen kann. Die einzige Anforderung ist, dass sich das interne Komplement irgendwo in einer VP eines bestimmten Kopfes befinden muss.

Es sieht aus diesen Gründen so aus, als ob beschriebene Asymmetrie nur ein ästhetisches Problem darstellt, sich aber aus den zu beschreibenden Daten ergibt.

In den vorigen beiden Kapiteln haben wir unsere Herangehensweise vorgestellt, welche die direkte Listenselektion von Komplementen einerseits und die Spezifikation lokaler Eigenschaften einer Barriere andererseits erlaubt. Es wurde diskutiert, welche Daten erfasst werden können, aber unser Ansatz muss sich auch der Gegenprobe unterziehen. Konfigurationen, die nicht vorkommen, sollten auch nicht beschreibbar sein. So kann ein internes Komplement bestimmte Kopf-Merkmale einer Barriere bestimmen, hat aber beispielsweise keinen Zugriff auf weitere Komplemente des Kopfes. Wir hatten in Kapitel 4.3 das Beispiel (171) eines unikalen Elementes, welches keinen Zugriff auf eine Komplement-NP des verbalen Kopfes hat. Es sei hier noch einmal wiederholt.

⁴¹Daten von Sailer (pers. Mitt.)

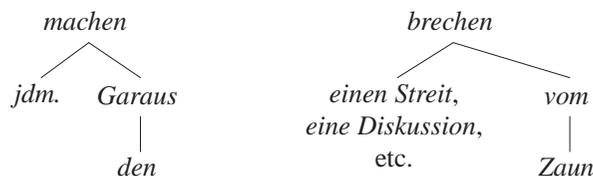
(272) [_{VP} V NP [_{PP} P [unikales Element]]]

Diese Generalisierung ist ähnlich wie schon beim Ansatz von Soehn/Sailer (2003) und ergibt sich auch in unserem Ansatz unabhängig aus der Merkmalsstruktur. Es handelt sich hier um eine Lokalitätsbeschränkung. Man hat es jedoch andererseits auch mit nicht lokalen Phänomenen wie der Pronominalisierung zu tun. Diese Fälle können wir mit Hilfe der PRONOMEN-LISTEM-BESCHRÄNKUNG (174) behandeln: Ein Relativpronomen hat denselben LISTEME-Wert wie sein Antezedenz, somit ist die Grammatikalität des folgenden Relativsatzes zu erwarten.

(273) *Ein solcher Schrecken_i, wie er ihn_i mir eingejagt hat, fährt einem wirklich durch Mark und Bein.*

Diese Beschränkungen folgen aus der Architektur unserer Herangehensweise und scheinen dabei der empirischen Datenlage Rechnung zu tragen.

O'Grady (1998) beobachtet, dass sich alle idiomatischen Teile eines Phraseologismus entlang einer Dependenzkette befinden müssen, zumindest in der kanonischen Form bzw. im Lexikoneintrag. Zur Illustration sei der Dependenzbaum von *jdm. den Garaus machen* angegeben:



Die Linie *machen–Garaus–den* ist die Dependenzkette des Idioms. Andere Zweige gehören nicht dazu. So kann man, um ein anderes Beispiel zu nehmen, *eine Diskussion* oder *einen Streit vom Zaun brechen*, das Objekt ist also nicht genau festgelegt. Dies wird im zweiten Dependenzbaum deutlich.

Jedoch kommt beispielsweise kein Idiom vor, das zwei idiomatische Nominalphrasen enthält, die beide gleichzeitig vorkommen müssen, aber Komplemente eines völlig arbiträren Verbs sind. O'Gradys *Continuity Constraint* besagt, dass alle Idiomkomponenten eine einzige Dependenzkette bilden müssen. Dies wurde in unserer Herangehensweise dadurch implementiert, dass die Selektionsbeziehung zwischen dem idiomatischen Verb und seinen Komplementen berücksichtigt wurde.

Ein Gegenbeispiel für O'Gradys Annahme scheint das Idiom „*bei jdm. ist Schmalhans Küchenmeister*“ zu sein. Diese Wendung ist in höchstem Maße

variabel, so kann *Schmalhans* mal als Subjekt, mal als (zu Grunde liegendes) Objekt auftreten und *Küchenmeister* muss nicht zwingend vorkommen. Dabei ist es aber immerhin in 109 von 193 Fällen⁴² (56,5%) vorhanden. Folgende Beispiele belegen die Variabilität:

(274) *Obwohl in den Anfängen Schmalhans Küchenmeister war, ist er auch heute noch überzeugt davon, dass er wieder die Schauspiellerei als Lebensaufgabe wählen würde.*⁴³

(275) *Lange, kalte Winter und karge Böden haben sie widerstandsfähig gemacht, ließen sie aus der Not eine Überlebenskultur entwickeln, in der Schmalhans das küchenmeisterliche Regiment führte.*⁴⁴

(276) *Ein weiterer Grund dafür, weshalb Schmalhans zum Chefkoch der Werksküche des Literaturbetriebs bestellt worden ist, hat mit der Konzentration des Buchmarkts auf einige Bestseller zu tun.*⁴⁵

(277) *Es joggten wenige Tage ins Land bis sich die Buben über Küchenmeister Schmalhans beschwerten.*⁴⁶

Diese Wendung besitzt zwar auch keine freie Lesart, ist aber sicherlich nicht zu den idiomatisierten verbalen Phraseologismen zu zählen, da das Verb nicht zwingender Bestandteil des Idioms ist. Daher fällt es nicht in unseren und wahrscheinlich auch nicht in O'Gradys Datenbereich.

Eine adäquate Beschreibung dieses Phraseologismus ist schwierig. Man könnte für *Schmalhans* eine komplexe Bedeutung annehmen („jmd., der (zuerst in der Küche) spart, wenn das Geld knapp ist“), wobei das Wort mit *Küchenmeister* in einem gewissen Kollokationsverhältnis steht.

Ein letzter Punkt, den es zu diskutieren gilt, hat mit dem Unterschied zwischen echten unikalen Elementen und anderen idiomatischen Wörtern zu tun. In der Beschreibung, die unsere Analyse nahelegt, unterscheiden sich die beiden Phänomene nicht voneinander. Beide, unikale Elemente und Wörter mit idiomatischer Bedeutung, haben jeweils eigene Lexikoneinträge, worin ihre Bedeutung und ihre Distributionsanforderungen spezifiziert sind.

⁴²COSMAS II-Recherche am 27.07.2004

⁴³St. Galler Tagblatt, 10.01.1998;

⁴⁴Mannheimer Morgen, 09.12.2002;

⁴⁵Die Presse, 24.10.1998;

⁴⁶Oberösterreichische Nachrichten, 22.04.1998;

Nun empfindet der Sprecher jedoch intuitiv einen Unterschied zwischen einem unikal Element wie *Garaus*, das in keinem anderen Kontext als im IVP *den Garaus machen* vorkommt, und einem Wort wie *Abfuhr*, welches zwar im zugehörigen IVP *jdm. eine Abfuhr erteilen* eine idiomatische Bedeutung besitzt, aber sonst auch frei (in anderer Bedeutung) vorkommen kann.

Die Diskrepanz zwischen der einheitlichen Beschreibung und der „gefühlten“ Unterschiedlichkeit ist offenkundig. Jedoch besteht diese Unterschiedlichkeit nicht so sehr im zu beschreibenden Wort selbst, sondern vielmehr in dessen Auftretenshäufigkeit einerseits und dessen eventueller Polysemie andererseits. Unikale Elemente kommen nur in bestimmten Kontexten – und damit im Durchschnitt seltener – vor, während Wörter mit idiomatischer Bedeutung schon aus anderen Kontexten bekannt sind. Darüber hinaus besteht in den meisten Fällen eine semantische Verbindung zwischen der idiomatischen und den weiteren Bedeutungen solcher Wörter, bei unikal Elementen hingegen gibt es nur eine Bedeutung. Da wir uns in dieser Arbeit auf syntaktische Merkmale konzentriert haben, mussten semantische Beziehungen zwischen Wortbedeutungen unberücksichtigt bleiben. Die Modellierung der – bisweilen sehr komplexen – semantischen Relationen zwischen Wörtern (auch im Rahmen der HPSG) muss an anderer Stelle erfolgen.

6.4 Noch einmal zum Vergleich: XSEL vs. COLL

Wir wollen den XSEL-Ansatz (Soehn, 2003) mit dem COLL-Mechanismus (aus dieser Arbeit) vergleichen. Dafür seien zunächst stichpunktartig die Vorzüge und Nachteile des Ansatzes mit XSEL zusammengefasst. Für eine genauere Diskussion, siehe die Abschnitte 4.3 und 5.1.2.

Vorteile von XSEL

- + Vereinheitlichung der HPSG-Signatur durch Zusammenführung zweier Attribute
- + Minimale Änderungen an der Merkmalsgeometrie
- + Lokalität der Selektion bleibt gewahrt.
- + Trotz lokaler externer Selektion können durch Schachtelung der XSEL-Werte größere Strukturen erfasst werden (PP+Verb).

- + Erweiterbar auf andere Bereiche, beispielsweise auf die morphologische Ebene⁴⁷

Nachteile von XSEL

- Manchmal arbiträre XSEL-Werte (z. B. bei uneingebetteten Zeichen)
- Implizite Nichtlokalität
- Uneingeschränkte Rekursion möglich durch Schachtelung der XSEL-Werte
- Ansatz nur für dekomponierbare Fälle entworfen
- Semantische Information muss lokal verfügbar sein.

Vorteile von COLL

- + Klare Trennung zwischen idiosynkratischen und nicht idiosynkratischen Zeichen
- + Erfassung dekomponierbarer und nicht dekomponierbarer Idiome, unikal Elemente und anomaler Konstruktionen
- + Änderungen an der Merkmalsgeometrie in vertretbarem Ausmaß
- + Lokalität der Selektion bleibt gewahrt.
- + Nichtlokalität wird durch Barrieren gesteuert.
- + Erweiterbar auf andere Bereiche, beispielsweise auf phonologische Ebene

Nachteile von COLL

- Teilweise mehrere Lexikoneinträge für ein einziges Idiom nötig (für jeden Bestandteil eines dekomponierbaren IVPs)
- Verwendung des komplexen Inventars an Relationen, um Flexibilität von nicht dekomponierbaren IVPs zu modellieren

⁴⁷Siehe die Diskussion von Gefügen der Form *an-PART kommen* (*angekrochen kommen*) in Soehn (2003, Kap. 6.2).

Diese Aufstellung spricht klar dafür, den COLL-Mechanismus zur Analyse von Idiomen heranzuziehen, und wir sind der Meinung, dass die Nachteile von COLL in Kauf genommen werden können. Wir wollen jedoch, wegen der positiven Eigenschaften der Generalisierung, das Attribut XSEL nicht verwerfen. Es soll dazu dienen, nicht idiomatische Fälle externer Selektion – bei Adjunkten, Determinierern und Komplementierern, welche auf lokaler Ebene angesiedelt sind – zu beschreiben. Weitere Anwendungsfälle sind beispielsweise DPs (externe Selektion des Determinierers) und die Selektion der Hilfsverben bei der Perfektbildung. In Abschnitt 5.1.2 wurde der Ansatz revidiert und den Nachteilen wurde folgendermaßen Rechnung getragen:

- Keine arbiträren XSEL-Werte durch Sorte *xsel-synsem*
- Keine Nichtlokalität durch neue Definition des PXS
- Zwar ist eine Schachtelung der XSEL-Werte grundsätzlich möglich, da sie aber im Lexikon nicht mehr vorgesehen ist, kann sie auch sonst nicht auftreten.
- Die semantische Information spielt nur eine untergeordnete Rolle. Eine Interaktion mit LRS bzw. die Integration beider Module ist möglich.

Der COLL-Mechanismus soll in Abgrenzung zu XSEL zur Beschreibung aller idiosynkratischen und idiomatischen Phänomene dienen, insbesondere zur Modellierung idiomatischer Wendungen. Aber auch andere Fälle von Idiosynkrasien können mit Hilfe des COLL-Moduls behandelt werden, beispielsweise Sandhi-Phänomene oder die Distribution von Polaritätselementen (siehe auch Fn. 13 auf Seite 112). Weitere Forschungsarbeit dazu ist in Vorbereitung, im Rahmen derer die Möglichkeiten des vorgelegten Ansatzes ermittelt werden sollen.

7 Abschließendes Fazit und ausstehende Forschung

Einige Schriftsteller haben den Bibelvers „Am Anfang war das Wort“ mit Stolz zitiert. Wichtiger scheint mir von allen Worten – das letzte Wort.

Bertolt Brecht

7.1 Zusammenfassung und Ausweitung

7.1.1 Ein lexikalistischer Ansatz

In dieser Arbeit wurde ein lexikalistischer Ansatz vorgestellt, sozusagen ein Dekompositionsansatz. Ein dekomponierbarer Phraseologismus wird hier nicht als eine Einheit bzw. Konstruktion betrachtet, sondern als eine Gesamtheit aus idiomatischen (und evtl. auch nicht idiomatischen) Teilen. Diese Sichtweise erlaubt nicht nur die Beschreibung von IVPs, sondern auch von unikalenen Elementen, deren Distribution auf einen bestimmten, aber nicht unbedingt idiomatischen Kontext beschränkt ist.

Oft werden – gerade bei dekomponierbaren Phraseologismen – gewisse Teile herausgenommen und separat von den übrigen Idiomteilen verwendet, vgl. dazu folgende Beispiele.

(278) *Ich bin bei meinem **Lebensfaden** schon ziemlich am Ende, der Junge hat noch eine weite Strecke vor sich.*¹ Eigentlich: *jdm. den Lebensfaden abschneiden*

(279) *Die Fusion der Grossbanken hat vor allem viele Sparer erschreckt, die dann ihr **Scherflein** in eine andere Bank rollen liessen.*² Eigentlich: *sein Scherflein entrichten/beitragen* (+ Konversion mit *den Rubel rollen lassen*)

¹Kleine Zeitung, 16.05.1999;

²St. Galler Tagblatt, 08.01.1998, Ressort: TB-OT;

Die Belege zeigen, dass Sprecher oft eine Intuition haben, was die Bedeutung von Idiomteilen anbelangt. Die Idiomanalyse auf Wortebene modelliert diese Intuition, da hier Idiome nicht als unteilbare Einheiten betrachtet werden. Burger (2003) erklärt dieses Phänomen mit Bezug auf die menschliche Kognition:

Ein kognitiv orientierter Ansatz ist daran interessiert, im „mental Lexikon“ eine möglichst große Dichte der Vernetzung zu erzielen. Das ist am ehesten zu erreichen, wenn die Zahl der semantisch isolierten Einheiten möglichst klein gehalten wird und wenn auch die Idiome möglichst weitgehend regulären semantischen Prozessen unterliegen. Entsprechend ist es für die Struktur des mentalen Lexikons „eleganter“ anzunehmen, daß die Komponenten des Idioms eine gewisse Autonomie besitzen, die eine Anknüpfung an ihre freie, d.h. auch „reguläre“ Verwendung ermöglicht. (Burger, 2003, S. 71)

Was die kleinere Gruppe der nicht dekomponierbaren Idiome betrifft, so tragen wir mit unserer Analyse der Tatsache Rechnung, dass diese Idiome zumeist aus regulären Wörtern der Sprache bestehen (unikale Elemente ausgenommen). In unserer Analyse kommen die Idiomteile auch aus dem Lexikon und erfahren erst auf derjenigen Ebene eine semantische Umdeutung, auf der alle notwendigen Teile, also das Verb und die internen Argumente, präsent sind. Eine Einheit bildet ein Idiom insofern, als dass es sich aus ganz spezifischen Bestandteilen zusammensetzt. Wenn nur ein einziger fehlt, kommt in der Regel keine idiomatische Lesart zu Stande. Vergleiche:

(280) *ein blaues Wunder erleben*³ – !*ein Wunder erleben*

(281) *mit allen Wassern gewaschen sein* – !*gewaschen sein*

Mit dem vorgestellten Ansatz kann man dies beschreiben, da wir im entsprechenden phrasalen Lexikoneintrag die erforderlichen Bestandteile spezifizieren. Phrasale Lexikoneinträge bilden die nicht dekomponierbaren IVPs als eine Einheit von syntaktischer Konstruktion und semantischer Idiosynkrasie ab. Damit kommen wir dem Grundgedanken der Konstruktionsgrammatik sehr nahe (vgl. Riehemann, 2001), ohne jedoch dessen nachteilige Effekte in der HPSG-Analyse in Kauf nehmen zu müssen.

³Eine Analyse für dieses Idiom wäre ein PLE für *blaues Wunder* (eine Kopf-Adjunkt-Struktur) mit der Bedeutung „große Überraschung“.

Um diese Diskussion zu vervollständigen, wollen wir noch einmal den Ansatz der Dekomposition mit dem der Konstruktion vergleichen und zu einer Synthese kommen.

Sailer (2004b) stellt den Konstruktionsansatz und die Distributionsanalyse, welche von einer Dekomposition der Wendungen ausgeht, einander gegenüber. Die Gründe, einen Konstruktionsansatz vorzuziehen, wurden schon in Kapitel 4 im Zusammenhang mit Riehemanns Analyse diskutiert. Erstens treten idiomatische Wörter nie außerhalb der Wendung auf, sie haben also keine wendungsexterne Bedeutung. Zweitens scheint ein einziger Lexikoneintrag für die ganze Wendung ökonomischer, man spart den LE für das einzelne idiomatische Wort.⁴ Drittens gibt es – außer den Analysen in Sailer (2003) und den mit diesem Ansatz verwandten Papieren von Sailer/Richter (2002a, 2002b), Soehn (2003) und der vorliegenden Arbeit – keine ausgearbeiteten Mechanismen, um Distributionsidiosynkrasien auszudrücken.

Sailer propagiert jedoch die Distributionsanalyse anhand von Wendungen mit unikalenen Elementen:

Zum einen treten alle anderen an der Wendung beteiligten Elemente in ihrer „normalen“ Bedeutung auf (bzw. in einer auch wendungsextern möglichen Bedeutung). Zum anderen ist die Verbindung des unikalenen Elements mit den übrigen Elementen der Wendung syntaktisch regulär (bzw. nicht wendungsspezifisch). Darüber hinaus ist es möglich, dem unikalenen Element eine Bedeutung zuzuweisen, so dass die Gesamtbedeutung der Wendung sich kompositionell ergibt (bzw. ohne die Annahme wendungsspezifischer Interpretationsmechanismen). (Sailer, 2004b, S. 196)

Wir teilen diese Sichtweise und wenden sie auf alle dekomponierbaren Wendungen an, auch solche ohne unikale Komponente. Des Weiteren haben wir jedoch auch Teile eines Konstruktionsansatzes übernommen, realisiert in den PLEs. Damit beschreiten wir, so wie auch schon Sailer (2003), einen Mittelweg. Sailer (2004b) betont, dass man beide Ansätze benötigt und keiner von beiden den anderen völlig ersetzen kann (S. 213f). So gibt es syntaktische Komplexe, die nicht durch eine Distributionsanalyse erfasst werden können, wie beispielsweise *mir nichts dir nichts* oder *Urlaub ist Urlaub*.

(282) *Peter weiß, wo Barthel den Most holt.*

An der Wendung in (282) zeigt Sailer (ebd.), dass der unterstrichene Teilsatz eine syntaktisch feste, nicht kompositionelle Einheit bildet und deshalb als

⁴Dennoch müssten die einzelnen Wörter irgendwie lizenziert werden.

Konstruktion zu behandeln ist (in unserem Ansatz also mit einem PLE). Andererseits ist die Verbindung mit dem Verb *wissen* syntaktisch völlig regelgemäß, wobei aber der Teilsatz ausschließlich Objekt von *wissen* sein kann, also unikal ist. Dies kann durch die Distributionsanalyse beschrieben werden, die Phrase legt also in ihrem COLL-Wert das Listem *wissen* fest.

Man sieht an diesem Beispiel, dass beide Ansätze miteinander interagieren und sich sinnvoll ergänzen.

Ein wichtiger Unterschied zu Konstruktionen und auch zum Selektionsmechanismus ist, dass (i) Distributionsbeschränkungen keinen strikten Lokalitätsbedingungen gehorchen müssen [Anm.: siehe das Konzept der Barrieren] und v. a. dass (ii) Distributionsbeschränkungen keine „generative“ Funktion haben, d. h. sie lizenzieren keine anderweitig ausgeschlossenen Verbindungen (wie Konstruktionen) oder legen die Beziehungen zwischen verschiedenen Elementen im Satz fest (wie Selektion). Ihre einzige Aufgabe ist die der Strukturbeschränkung. Wir haben es also mit einem unabhängigen Mechanismus zu tun. (Sailer, 2004b, S. 215)

7.1.2 Kennzeichen im Visier

In Kapitel 3.2 haben wir Kennzeichen für Idiomatizität vorgestellt. Es bleibt nun noch zu zeigen, wie diese Kennzeichen in unseren Ansatz einbezogen wurden.

- Die **unikalen Elemente** haben wir ausführlich diskutiert. Sie sind daran zu erkennen, dass sie Barrieren mit bestimmten Merkmalen in ihrem Lexikoneintrag spezifizieren. Diese Eigenschaft haben unikale Elemente mit allen idiomatischen Wörtern gemeinsam, die Distributionsidiosynkrasien aufweisen. Im Gegensatz zu idiomatischen Wörtern gehen unikale Elemente oft mit frei verwendbarem Kontext einher. Ein solcher Kontext ist selbst nicht idiosynkratisch in seiner Distribution (uneingeschränkte COLL-Liste im LE) und fordert das unikale Element nicht an (kein spezifischer LISTEME-Wert auf SUBCAT im LE). Idiomatische Listeme selbst hingegen werden von ihrem (idiomatischen) Kontext selektiert bzw. über COLL gefordert.
- **Morphologische Anomalien** werden im Lexikoneintrag spezifiziert, was in dieser Arbeit im PHON-Wert vorgenommen wird. Man muss außerdem entsprechende Werte in der morphologischen Komponente (un-

ter MORPH etc.) einfügen. **Syntaktisch anomale Phrasen** benötigen ihren eigenen phrasalen Lexikoneintrag (siehe Abschnitt 6.2.5).

- Änderungen in der **Valenzstruktur** gegenüber einem frei verwendbaren Verb (wie bei *staunen* in *Bauklötze staunen*) werden ebenfalls durch einen separaten Lexikoneintrag modelliert, wobei die SUBCAT-Liste von der des freien Verbs abweicht. Es konnten auf Grund unserer Korpusdaten (Appendix A) keine systematischen Beziehungen gefunden werden, welche die Etablierung von geeigneten Regeln erlauben würden.
- Informationen über **Selektionsrestriktionen** können, wie der Vorschlag in Kapitel 5.3.2.2 verdeutlicht, in die HPSG-Theorie integriert werden. Im vorliegenden Ansatz sind diese Informationen Teil der CONTEXT-Information und werden über *bgr-conditions* spezifiziert. Ein Verb kann damit bestimmen, welchen Restriktionen ein Argument genügen muss.
- Zuletzt haben wir festgestellt, dass man Nichtdekomponierbarkeit auch als Idiom-Kennzeichen zählen kann. Für diese Klasse von Idiomen wurde eine spezielle Art von Lexikoneinträgen, die PLEs, benötigt.

Zusammenfassend ist festzuhalten, dass die Idiom-Kennzeichen alle in den Lexikoneinträgen vermerkt sind. Da die HPSG eine lexikalistische Theorie ist, ist es selbstverständlich, dass auch Hinweise auf die Idiomatizität im Lexikon enthalten sind.

Wir haben in der vorliegenden Arbeit einen Ansatz vorgestellt, mit welchem Idiome ohne freie Lesart, auch solche mit unikalener Komponente, analysiert werden können. Die Schlüsselstellung bei unserem COLL-Modul nahm dabei das gleichnamige Attribut ein, durch welches man eine Distributionseigenschaft innerhalb einer bestimmten Äußerungseinheit (Barriere) spezifizieren kann. Eine solche Barriere muss je nach COLL-Wert bestimmten Anforderungen genügen, was durch das LIZENSIERUNGSPRINZIP sichergestellt wird. Zur eindeutigen Identifikation lexikalischer Elemente führten wir zudem das Attribut LISTEME ein. Hauptsächlich durch diese beiden Neuerungen gelingt es, Kookkurrenzen im Lexikon festzulegen. Weiterhin benötigt man LISTEME dafür, defektive Idiome von bestimmten Transformationen individuell auszuschließen.

7.1.3 Weitere Phraseologismen

In dieser Arbeit wurde der Datenbereich bewusst auf IVPs ohne freie Lesart eingeschränkt. Unser Ansatz ist jedoch so generell, dass man damit auch

IVPs mit freier Lesart und auch andere idiomatische Ausdrücke beschreiben kann. Als Beispiel für einen IVP mit freier Lesart sei die Wendung *Öl ins Feuer gießen* genannt. Für solche Ausdrücke nehmen wir zwei verschiedene Strukturen an, für die freie und für die gebundene Lesart. Beispiele für andere idiomatische Ausdrücke sind idiomatische nominale Phraseologismen (INPs), wie z. B. *blinder Passagier* oder *Schlag ins Gesicht*, und satzwertige Phraseologismen (Sprichwörter) wie *Morgenstund hat Gold im Mund*.

Bei all diesen Wendungen handelt es sich um Kookkurenzphänomene, wobei bestimmte Wörter in derselben Phrase vorkommen müssen, um den Phraseologismus zu bilden und die gebundene Lesart herstellen zu können.

Eine Analyse mit unserer Herangehensweise könnte für die genannten Beispiele folgendermaßen aussehen: Bei *blinder Passagier* müsste das Adjektiv in der Bedeutung „illegal“ eine NP mit dem entsprechenden Listem *passagier* über COLL fordern. Der INP *Schlag ins Gesicht* ist ein Fall für einen phrasalen Lexikoneintrag mit der Bedeutung „schwere Kränkung“. Auch für die Redewendung *Morgenstund hat Gold im Mund* gäbe es einen phrasalen Lexikoneintrag.

Es obliegt jedoch weiterführender Untersuchungen, diese anderen Unterarten idiomatischer Ausdrücke im Detail zu charakterisieren und zu beschreiben. Natürlich müsste auch eine Korpusrecherche mit einbezogen werden, um Kontexte und Modifikationsmöglichkeiten zu untersuchen. Eine der größten Schwierigkeiten bei Idiomen mit freier Lesart ist die Unterscheidung, wann eine freie und wann die gebundene Lesart auftritt. Manchmal ist die Bedeutung ambig, aber der Kontext kann in den meisten Fällen zur Disambiguierung dienen. Solche Kontextphänomene wurden in dieser Arbeit durch die Definition des Datenbereichs ausgeblendet, verdienen aber an anderer Stelle eine sorgfältige Betrachtung.

7.2 Horizonte

7.2.1 Phraseologisierung

Wie entsteht eigentlich ein Idiom? Aus einem diachronen Blickwinkel betrachtet, lässt sich ein Prozess beobachten, der Phraseologisierung genannt wird und mehrere Stadien umfasst. Zu Beginn hat man es mit einem frei formulierten Ausdruck zu tun, welcher oft das Produkt spontaner Kreativität eines Sprechers ist. Eine solche Innovation kann sich in den Köpfen (bzw. im Sprach-

gebrauch) der Sprechergemeinschaft festsetzen, es kommt zur Konventionalisierung. Dies setzt häufiges Zitieren bzw. Verwenden voraus, was u. a. dann gegeben ist, wenn derjenige, der diesen Ausdruck prägte, eine Persönlichkeit des öffentlichen Lebens ist oder wenn der Ausdruck aus einem wichtigen literarischen Werk, wie beispielsweise Goethes „Faust“ oder der Bibel, stammt. Wird diese ursprüngliche Innovation häufig benutzt, spricht man von Usualisierung, worauf eine Lexikalisierung folgen kann. Damit ist dann ein Ausdruck fester Bestandteil des Lexikons und man spricht von einem Phraseologismus. Blank (2001, S. 1596) hebt hervor, dass einer Lexikalisierung nicht unbedingt die Usualisierung vorausgehen muss. Usualisierung definiert Blank als Konvention innerhalb der Diskurs-Tradition, während Lexikalisierung eine Aufnahme ins lexikalische Repertoire einer Sprechergemeinschaft impliziert.

Hat ein Phraseologismus neben seiner wörtlichen Lesart auch noch eine übertragene oder weist er lediglich eine nicht wörtliche Lesart auf, haben wir es mit einem Idiom zu tun. Wie in Kapitel 2 angesprochen wurde, unterliegen Idiome ohne freie Lesart meist einer Demotivierung, also einem synchronen Prozess, durch welchen der Ursprung des Ausdrucks in Vergessenheit geraten ist.

Aus der synchronen Perspektive können wir in Sprachen folgendes Spektrum beobachten: frei formulierte Äußerungen, Kollokationen, Idiome mit freier Lesart (dazu zählen z. B. Sprichwörter), Idiome mit freier und gebundener Lesart und zuletzt Idiome ohne freie Lesart. Hierzu merkt O’Grady (1998) in seinem Schlusswort an, dass man Idiome auf einem Kontinuum zwischen vollständig eingefrorenen Äußerungen und normalen Konstruktionen, wie Passiv oder VPs mit zwei Objekten (*double object patterns*), ansiedeln kann. Nach dieser Sichtweise lässt sich keine scharfe Grenze zwischen Idiomen und Nicht-Idiomen ziehen. Im Allgemeinen lassen sich jedoch Äußerungen, welche nicht wörtlich zu verstehen sind oder deren Semantik nicht kompositionell ist, schon identifizieren und somit als idiomatisch klassifizieren. Dabei stellt die Metaphorik ein Problem dar, da eine Unterscheidung zwischen bildhafter und idiomatischer Lesart schwierig sein kann.

Begleitet wird der Vorgang der Phraseologisierung von einem Grammatikalisierungsprozess, bei dem die lexikalisierten Ausdrücke sozusagen einfrieren. Der Prozess zeichnet sich durch eine Annäherung an prototypische Merkmale lexikalischer Einheiten aus (vgl. Römer/Matzke, 2003, Kap. 2.2) und die Idiome verlieren die Eigenschaften einer VP in unterschiedlichem Maße. So haben wir in Kapitel 2 gezeigt, wie bestimmte grammatische Operationen nicht mehr

angewandt werden können, beispielsweise können manche Idiome nicht passiviert oder nominalisiert werden.

Was die Untersuchung von syntaktischen Restriktionen von Idiomen im Deutschen betrifft, gibt es Studien zu einzelnen Phänomenen. So untersucht Dobrovolskij (1999) die Passivierbarkeit von Idiomen und Müller (2002, Kap. 6.2.1) analysiert die Voranstellung von Idiomteilen. In den voranstehenden Kapiteln wurden in diesem Zusammenhang einige wichtige Aspekte beleuchtet. Dennoch fehlt bisher ein systematischer Überblick zu syntaktischen Restriktionen deutscher Redewendungen. Die Hierarchie der Stabilitätsgrade von Fraser (1970) im Englischen stellt ein einzelnes Phänomen in diesem Forschungszweig dar. Fraser stellt die sogenannte *frozenness hierarchy* vor, die aus sieben Stufen besteht:

Stufe 0 *completely frozen* – Bei einem solchen Idiom ist keine syntaktische Operation zulässig.

Beispiel: *trip the light fantastic*⁵

Stufe 1 *adjunction* – Eine nicht idiomatische Konstituente kann an ein internes Argument adjungiert werden.

Beispiel: *John hit the sack*.⁶ – *John's hitting of the sack*

Stufe 2 *insertion* – Ein externes Argument kann in die idiomatische VP eingefügt werden.

Beispiel: *John read the riot act to the class*.⁷ – *John read the class the riot act*.

Stufe 3 *permutation* – Die Reihenfolge zweier aufeinander folgender Konstituenten kann sich ändern, z. B. bei Fragesatzbildung.

Beispiel: *The cat has got your tongue*.⁸ – *Has the cat got your tongue?*

Stufe 4 *extraction* – Extraktion wie sie bei verschiedenen Transformationen (z. B. Passivierung) vorkommt, ist möglich.

Beispiel: *A rolling stone gathers no moss*.⁹ – *No moss is gathered by a rolling stone*.

Stufe 5 *reconstitution* – Eine Nominalisierungstransformation ist zulässig, der syntaktische Status wechselt also von VP zu NP.

⁵ „das Tanzbein schwingen“

⁶ „Hans haute sich aufs Ohr.“

⁷ „Hans las der Klasse die Leviten.“; *riot act* ist internes, *class* externes Argument.

⁸ „Du bist sprachlos.“

⁹ „Wer rastet, der rostet.“

Beispiel: *He laid down the law to his daughter*.¹⁰ – *His laying down of the law to his daughter*

Stufe 6 *unrestricted* – Es liegen keinerlei Beschränkungen vor.

Fraser kann jeder Stufe einige prototypische Idiome des Englischen zuordnen. Die Tatsache, dass diese Stufen in einer Hierarchie stehen, besagt, dass bei einem Idiom, das sich beispielsweise auf Stufe 4 befindet (also Passivierung grammatisch ist), alle Transformationen der Stufen 1 bis 3 auch zulässig sind. Aus dieser Hierarchie folgt, dass es Idiome gibt, welche Nominalisierung (Stufe 5) und damit auch Passivierung zulassen, aber keine, welche Nominalisierung, jedoch keine Passivierung zulassen. In unserem Korpus finden wir dazu einige Gegenbeispiele, gerade auch *Bauklötze staunen* und *das Handtuch werfen*.

Fraser räumt in seinem Artikel ein, dass Grammatikalitätsurteile über zulässige Transformationen sehr subjektiv sind und dass es möglich ist, dass für einige Sprecher seine Hierarchie unzutreffend ist (Fraser, 1970, S. 41).

Wollte man diese Hierarchie ins Deutsche übertragen, müsste man folgende Operationen berücksichtigen:

- Transformationen:
 - Voranstellung von internen Argumenten ins Vorfeld
 - Scrambling von internen Argumenten
- Expansionen:
 - Modifikation von internen Argumenten durch Adjektive oder Relativsätze
- Verdichtungen:
 - Nominalisierung des Verbalkomplexes
- Änderung der Satztypen
 - Deklarativ- zu Interrogativsätzen
 - Deklarativ- zu Imperativsätzen
- Negationen:
 - Negation der Konstituenten
 - Satznegation

¹⁰ „Er sprach mit seiner Tochter ein Machtwort.“

Wie schon in Kapitel 2 gezeigt wurde, ist es sehr schwierig, eine Hierarchie zwischen diesen Transformationen herzustellen, da man schnell Gegenbeispiele finden kann. Deshalb wurden in unserer Analyse die Transformationen einzeln betrachtet und Mechanismen zum Ausschluss bestimmter defektiver Idiome von diesen Transformationen bereitgestellt.¹¹

7.2.2 Generalisierung unserer Herangehensweise

Im Prozess der Phraseologisierung sind Idiome schon weitgehend fortgeschritten. So weisen sie transformationelle Defekte in unterschiedlichem Maße auf, haben eine idiomatische Lesart und haben manchmal die freie Lesart gänzlich verloren. Es gibt also bei den Wendungen, die sich in der Phraseologisierung befinden, eine gewisse Abstufung, auf der sich Idiome sehr weit oben befinden. Geht man in dieser Abstufung etwas weiter zurück, so findet man dort usuelle Wortverbindungen bzw. Kollokationen vor, die erst in jüngerer Zeit ins Blickfeld der Phraseologieforschung gerückt sind. Diese lassen sich in mehrere Subklassen unterteilen, darunter Funktionsverbgefüge (FVG), Inkorporationen und häufig gemeinsam auftretende Wörter. Auch Polaritätselemente lassen sich zu den Kollokationen zählen.

Bei **Funktionsverbgefügen** gibt es ebenfalls Beschränkungen in der Distribution der Nomina. So kommen viele davon mit ihrer FVG-internen Bedeutung nur innerhalb der FVGs vor:

(283) *Stellung in zu etw. Stellung nehmen*

(284) *Rechnung in etw. Rechnung tragen*

(285) *Betracht in in Betracht kommen*

Zumindest bei *Betracht* kann man schon von einem unikalen Element sprechen. Gegenbeispiele sind die Nomina in den auf S. 56 zitierten Wendungen *Betrieb*, *Aufführung* und *Einfluss*, da diese Nomina in derselben Bedeutung frei vorkommen.

¹¹Ob man – sicherlich nicht ohne Mühe und nicht ohne Ausnahmen zu machen – Frasers Hierarchie auf das Deutsche übertragen kann, sei dahingestellt. Falls dies gelingt und man diese in die Signatur einbauen möchte, bietet es sich an, neue Untersorten von *listeme* bereitzustellen (wie *level0*, *level1*, etc.) und die LISTEME-Werte der idiomatischen Wörter unter diesen Sorten zu subsumieren. Durch geeignete Beschränkungen ließen sich dann sämtliche Listeme einer Stufe von bestimmten Transformationen ausschließen.

Inkorporationen (v. a. inkorporierte Objekte) sind Wendungen wie *gewährleisten* oder *etw. handhaben* (siehe Gallmann, 1999). Man beachte bei diesen Beispielen die Abstufung bei der Getrennt-/Zusammenschreibung: Man sagt *ich leiste für etw. Gewähr*, aber *ich handhabe etw.* Eine Unterart der Inkorporation im weiteren Sinn ist *noun-stripping*, wie *Zeitung lesen* oder *Verdacht schöpfen*. Hier bleibt das Nomen ein eigenes syntaktisches Wort, bildet aber eine nicht projizierende Konstituente. Nach Gallmann (ebd.) handelt es sich gewöhnlich um ein Kopfadjunk.

Wörter, die „gerne“ **gemeinsam auftreten**, finden sich häufig, so beispielsweise *Zähne putzen* oder *den Rasen sprengen*. Bei diesen Beispielen fällt auf, dass man seine *Zähne* nicht *waschen* kann (im Französischen ist das genau umgekehrt) und dass man seinen *Rasen* zwar *wässern* oder *gießen*, umgekehrt aber seine Blumen nicht *sprengen* kann (jedenfalls nicht, ohne sie zu zerstören).

Fälle von **Polaritätselementen** sind *jemals* oder *(nicht lange) fackeln* (siehe Sailer/Richter, 2002a). Hier liegen nicht nur syntaktische Kookkurrenzen vor, sondern auch semantische, d. h. ein negatives Polaritätselement muss im Skopus einer Negation liegen.

Bei all diesen Daten hat man es mit Kookkurrenzphänomenen zu tun, wenn gleich ohne idiomatische Lesarten. Man könnte nun das in dieser Arbeit vorgestellte COLL-Modul auf diesen Datenbereich ausweiten, es gewissermaßen generalisieren. Es stellen sich dabei jedoch einige Fragen, beispielsweise wie oben genannte Kollokationstypen genau charakterisiert werden können. So stellt Winhart (2002) fest, dass eine exakte Definition von FVGs sehr schwierig ist. Eine andere Frage ist, ob man tatsächlich alle Kollokationen mit dem COLL-Mechanismus modellieren kann.

Oft bestehen Kollokationen jedoch aus Kookkurrenzen, die nicht zwingend sind, sondern ein Sprecher formuliert Ausdrücke lieber mit bestimmten Wörtern als mit anderen. Es zeigt sich in dieser Domäne auch deutlich, wie gut ein Nichtmuttersprachler eine Sprache beherrscht. Es geht dabei also eher um statistische Häufigkeiten als um grammatikalisch vorgeschriebene Kookkurrenzen. Es schließt sich deshalb die Frage an, wie sich dies in einer formalen Grammatiktheorie umsetzen lässt bzw. inwieweit es eine Interaktion von Grammatikregeln und statistischen Häufigkeiten („Vorlieben“) gibt und wie man diese modellieren kann. Fouvry (2003) führt in seiner Dissertation über

robuste Verarbeitung für beschränkungsbasierte Grammatikformalismen Gewichtungen für den Informationsgehalt von sprachlichen Merkmalen ein (Fouvry, ebd., Kap. 7). Er weist im Weiteren darauf hin (Kap. 9.4), dass man diese Gewichtungen durch statistische Werte ersetzen kann, um die Grammatikalitätsurteile von Sprechern in die Grammatik mit einzubeziehen. Wie man zu diesen Werten kommt, überlässt Fouvry allerdings der weiteren Forschung. Riezler (1999) beschäftigt sich in seiner Dissertation ebenfalls mit Gewichtungen, außerdem beschreibt er ein probabilistisches Modell, das in der Lage ist, die Wahrscheinlichkeiten von Kontextabhängigkeiten automatisch zu berechnen. Dieser Arbeit liegt zwar nicht explizit die HPSG zu Grunde, jedoch beschränkungsbasierte Grammatiken mit getypten Merkmalsstrukturen im Allgemeinen, welche Riezler als konkrete Grundlage für die beschriebene Art von *constraint-based logic programming* betrachtet. Brew (1995) macht einen Vorschlag für eine „stochastische HPSG“, wobei die Anwendung verschiedener Grammatikregeln durch eine zuvor erlernte statistische Häufigkeit gesteuert wird. Dieser Vorschlag müsste aber noch weiter konkretisiert werden (nach Brew v. a. in Bezug auf *re-entrancies*). Es lässt sich zusammenfassend sagen, dass noch weiterer Forschungsbedarf darin besteht, statistische Module in HPSG-Implementierungen mit einzubeziehen.

Es wird deutlich, dass eine Ausweitung auf einen anderen Datenbereich nicht ohne weiteres vollzogen werden kann. Es stellen sich Fragen, die nicht nur den in dieser Arbeit vorgestellten Ansatz betreffen, seine Interaktion mit anderen Mechanismen (z. B. der externen Selektion) oder den neuen Datenbereich als solchen, sondern es ergeben sich auch Problemstellungen, welche die gesamte Architektur der HPSG-Grammatik tangieren. Da eine deutliche Tendenz zu beobachten ist, Kollokationen nicht unberücksichtigt zu lassen, sondern sie innerhalb der Grammatik zu beschreiben, kann man diesen Datenbereich nicht einfach ausblenden und wird sich mit den genannten Fragen auseinander setzen müssen.

Auch wenn sich eine Generalisierung nicht von selbst ergibt, so stellt unserer Meinung nach der vorgelegte Ansatz einerseits eine Verbesserung gegenüber früheren Herangehensweisen dar und ist andererseits eine Alternative zu einer konstruktionsgrammatischen Analyse. Die HPSG erlaubt dabei eine präzise Formulierung, so dass eine konstruktive Diskussion folgen kann.

Literaturverzeichnis

- Abeillé, Anne (1993). *Les nouvelles syntaxes - Grammaire d'unification et analyse du français*. Arman Colin, Paris.
- Abeillé, Anne und Godard, Danièle (1996). La Complémentation des auxiliaires français. *Langages* 122, S. 32–61.
- Agricola, Erhard (Hrsg.) (1992). *Wörter und Wendungen. Wörterbuch zum deutschen Sprachgebrauch* (Überarbeitete Neufassung der 14. Aufl.). Dudenverlag.
- Androutsopoulos, Ion und Dale, Robert (2000). Selectional Restrictions in HPSG. In *Proceedings of COLING 2000*, S. 15–20, Saarbrücken.
- Aronoff, Mark (1976). *Word Formation in Generative Grammar* (Dritter Druck 1985). MIT Press, Cambridge MA. Linguistic Inquiry Monographs.
- Asudeh, Ash und Klein, Ewan (2002). Shape Conditions and Phonological Context. In F. van Eynde, L. Hellan, und D. Beermann (Hrsg.), *Proceedings of the 8th International HPSG Conference*, S. 20–30, CSLI Publications, Stanford, CA. <http://csli-publications.stanford.edu/HPSG/2/> (24.01.05).
- Avgustinova, Tania (2001). Arguments, Grammatical Relations, and Diathetic Paradigm. In D. Flickinger und A. Kathol (Hrsg.), *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, University of California, Berkeley, S. 23–42, CSLI Publications, Stanford, CA.
- Baldauf, Christa (1997). *Metapher und Kognition: Grundlagen einer neuen Theorie der Alltagsmetapher*. Peter Lang, Frankfurt am Main.
- Bierwisch, Manfred und Schreuder, Robert (1992). From concepts to lexical items. *Cognition* 42, S. 23–60.
- Bird, Steven und Klein, Ewan (1994). Phonological analysis in typed feature systems. *Computational Linguistics* 20, S. 455–491.

- Blank, Andreas (2001). Pathways of lexicalization. In M. Haspelmath et al. (Hrsg.), *Sprachtypologie und sprachliche Universalien / Language typology and language universals: an international handbook*, Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of linguistics and communication science, S. 1596–1608, de Gruyter, Berlin.
- Borsley, Robert D. (1989). Phrase-structure grammar and the *Barriers* conception of clause structure. *Linguistics* 27(5), S. 843–864.
- Bouma, Gosse (2000). Argument Realization and Dutch R-Pronouns: Solving Bech's Problem without Movement or Deletion. In R. Cann, C. Grover und P. Miller (Hrsg.), *Grammatical Interfaces in HPSG*, S. 51–76, CSLI Publications, Stanford, CA.
- Bouma, Gosse, Malouf, Rob und Sag, Ivan (2000). Satisfying Constraints on Adjunction and Extraction. *Natural Language and Linguistic Theory*, Band 19, 2001, S. 1–65.
- Brandt, Patrick, Dettmer, Daniel, Dietrich, Rolf-Albert und Schön, Georg (1999). *Sprachwissenschaft. Ein roter Faden für das Studium*. Böhlau Studienbücher. Böhlau Verlag, Köln.
- Brew, Chris (1995). Stochastic HPSG. In *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics*, Dublin, Ireland, S. 83–88, University College.
- Burger, Harald (2003). *Phraseologie. Eine Einführung am Beispiel des Deutschen* (2., überarbeitete Aufl.). Grundlagen der Germanistik. Erich Schmidt Verlag, Berlin.
- Burger, Harald, Buhofer, Annelies, und Sialm, Ambros (1982). *Handbuch der Phraseologie*. de Gruyter, Berlin.
- Bußmann, Hadumod (Hrsg.) (2002). *Lexikon der Sprachwissenschaft* (Dritte, aktualisierte und erweiterte Aufl.). Alfred Kröner Verlag, Stuttgart.
- Cacciari, Cristina und Glucksberg, Sam (1991). Understanding idiomatic expressions: the contribution of word meanings. In Gregory B. Simpson (Hrsg.), *Understanding word and sentence*, S. 217–240, North-Holland, Amsterdam, New York.
- Cacciari, Cristina und Tabossi, Patrizia (1988). The Comprehension of Idioms. *Journal of Memory and Language* 27, S. 668–683.

- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.
- Chur, Jeannette (1993). *Generische Nominalphrasen im Deutschen. Eine Untersuchung zu Referenz und Semantik*. Max Niemeyer Verlag, Tübingen.
- Copestake, Ann (2002). *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford, CA.
- Copestake, Ann, Flickinger, Daniel, Pollard, Carl und Sag, Ivan A. (1998). *Minimal Recursion Semantics: An Introduction*. Stanford. <ftp://ftp-csli.stanford.edu/linguistics/sag/mrs.ps.gz> (24.01.05).
- De Kuthy, Kordula (2000). *Discontinuous NPs in German – A Case Study of the Interaction of Syntax, Semantics and Pragmatics*. Phil. Dissertation, Universität des Saarlandes, Saarbrücken, Germany. <http://www.sfs.uni-tuebingen.de/hpsg/archive/bibliography/papers/dekuthy-thesis.ps> (24.01.05).
- De Kuthy, Kordula und Meurers, Walt Detmar (2000). From Argument Raising to Dependent Raising. In R. Cann, C. Grover und P. Miller (Hrsg.), *Grammatical Interfaces in HPSG*, S. 97–112, CSLI Publications, Stanford, CA.
- Di Sciullo, Anna-Maria und Williams, Edwin (1988). *On the Definition of Word* (Zweiter Druck). Linguistic Inquiry Monographs. MIT Press, Cambridge, MA.
- Dobrovol'skij, Dmitrij (1988). *Phraseologie als Objekt der Universalienlinguistik*. VEB Verlag Enzyklopädie Leipzig. Linguistische Studien.
- Dobrovol'skij, Dmitrij (1995). *Kognitive Aspekte der Idiom-Semantik. Studien zum Thesaurus deutscher Idiome*, Band 8 von *Eurogermanistik*. Gunter Narr Verlag, Tübingen.
- Dobrovol'skij, Dmitrij (1997). *Idiome im mentalen Lexikon. Ziele und Methoden der kognitivbasierten Phraseologieforschung*, Band 18 von *FOKUS. Linguistisch-Philologische Studien*. Wissenschaftlicher Verlag Trier.
- Dobrovol'skij, Dmitrij (1999). Gibt es Regeln für die Passivierung deutscher Idiome? In I. Bäcker (Hrsg.), *Das Wort. Germanistisches Jahrbuch*. DAAD, Bonn.

- Dobrovolskij, Dmitrij und Piirainen, Elisabeth (1994). Sprachliche Unikalia im Deutschen: Zum Phänomen phraseologisch gebundener Formative. In *Folia Linguistica XXVIII/3-4*, S. 449–473, Mouton de Gruyter, Berlin.
- Dölling, Johannes (1992). Flexible Interpretation durch Sortenverschiebung. In I. Zimmermann und A. Stringin (Hrsg.), *Fügungspotenzen*, S. 23–62, Akademie Verlag, Berlin.
- Dölling, Johannes (1994). Sortale Selektionsbeschränkungen und systematische Bedeutungsvariation. In M. Schwarz (Hrsg.), *Kognitive Semantik/Cognitive Semantics. Ergebnisse, Probleme, Perspektiven*, S. 41–59, Gunter Narr Verlag, Tübingen.
- Duden 11 (2002). *Duden. Redewendungen. Wörterbuch der deutschen Idiomatik*, Band 11. Dudenverlag, Mannheim.
- Duhme, Michael (1995). *Lauschangriff und Rollkommando* - „Einwortphraseologismen“ in der Pressesprache am Beispiel des Nachrichtenmagazins FOCUS. In R. S. Baur und C. Chlosta (Hrsg.), *Studien zur Phraseologie und Parömiologie*, Band 6, S. 83–93, Brockmeyer, Bochum.
- Egg, Markus (2004). Metonymie als Phänomen der Semantik-Pragmatik-Schnittstelle. *metaphorik.de (ISSN 1618-2006)* 6, S. 36–53.
- Ehrich, Veronika (1991). Nominalisierungen. In A. von Stechow und D. Wunderlich (Hrsg.), *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung*, Handbücher zur Sprach- und Kommunikationswissenschaft, S. 441–458, Walter de Gruyter, Berlin.
- Feyaerts, Kurt (1994). Zur lexikalisch-semantischen Komplexität der Phraseologismen mit phraseologisch gebundenen Formativen. In C. Chlosta, P. Grzybek und E. Piirainen (Hrsg.), *Sprachbilder zwischen Theorie und Praxis*, S. 133–162, Bochum.
- Fillmore, Charles J. (1988). The Mechanisms of „Construction Grammar“. In H. S. Shelley Axmaker, Annie Jaissner (Hrsg.), *Berkeley Linguistics Society*, Band 14, S. 35–55. University of California, Berkeley.
- Fleischer, Wolfgang (1982). *Phraseologie der deutschen Gegenwartssprache*. Niemeyer, Tübingen.
- Fleischer, Wolfgang (1989). Deutsche Phraseologismen mit unikalener Komponente - Struktur und Funktion. In G. Gréciano (Hrsg.), *Europhras* 88,

- Phraséologie Contrastive, Actes du Colloque Internationale, Klingenthal-Strasbourg*, S. 117–126.
- Fleischer, Wolfgang (1997). *Phraseologie der deutschen Gegenwartssprache* (Zweite, revidierte Aufl.). Niemeyer, Tübingen.
- Fleischer, Wolfgang, Helbig, Gerhard, und Lerchner, Gotthard (Hrsg.) (2001). *Kleine Enzyklopädie – Deutsche Sprache*. Peter Lang, Frankfurt am Main.
- Frank, Anette (1994). *Verb Second by Lexical Rule or by Underspecification*. Arbeitsberichte des Sonderforschungsbereichs 340, Nr. 43, IMS Stuttgart.
- Fraser, Bruce (1970). Idioms within a Transformational Grammar. *Foundations of Language* 6, S. 22–42.
- Fouvry, Frederik (2003). *Robust processing for constraint-based grammar formalisms*. Ph. D. thesis, Graduate School, University of Essex (Colchester, UK). British Library shelf mark: DXN058463.
- Gallin, Daniel (1975). *Intensional and Higher-Order Modal Logic*. North-Holland, Amsterdam.
- Gallmann, Peter (1985). *Graphische Elemente der geschriebenen Sprache*, Band 60 von *Germanistische Linguistik*. Max Niemeyer Verlag, Tübingen.
- Gallmann, Peter (1990). *Kategoriell komplexe Wortformen. Das Zusammenwirken von Morphologie und Syntax bei der Flexion von Nomen und Adjektiv*. Germanistische Linguistik. Max Niemeyer Verlag, Tübingen.
- Gallmann, Peter (1997). Konzepte der Nominalität. In G. August, K. Blüml, D. Nerius und H. Sitta (Hrsg.), *Die Neuregelung der deutschen Rechtschreibung. Begründung und Kritik*, Germanistische Linguistik, S. 209–242, Max Niemeyer Verlag, Tübingen.
- Gallmann, Peter (1999). Wortbegriff und Nomen-Verb-Verbindungen. *Zeitschrift für Sprachwissenschaft* 18.2, S. 269–304.
- Gazdar, Gerald, Klein, Ewan, Pullum, Geoffrey und Sag, Ivan (1985). *Generalized Phrase Structure Grammar*. Harvard University Press, Cambridge, MA.
- Geeraerts, Dirk (2002). The interaction of metaphor and metonymy in composite expressions. In R. Dirven und R. Pörings (Hrsg.), *Metaphor and metonymy in comparison and contrast*, S. 435–465, Mouton de Gruyter, Berlin.

- Geeraerts, Dirk und Bakema, Peter (1993). De prismatische semantiek van idiomen en composita. In *Leuvense Bijdragen*, Band 82, S. 185–226.
- Geißler, Stefan (1994). *Lexikalische Regeln in der IBM-Basisgrammatik*. IBM Deutschland Informationssysteme GmbH. Verbmobil-Report 20.
- Gibbs, Raymond W., JR., Nayak, Nandini P., Bolton, John L. und Keppel, Melissa E. (1989). Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition* 17(1), S. 58–68.
- Götz, Thilo, Meurers, Detmar, und Gerdemann, Dale (1997). *The ConTroll Manual*. (ConTroll v.1.0 beta, XTroll v.5.0 beta). Benutzerhandbuch, Seminar für Sprachwissenschaft, Universität Tübingen. <http://www.sfs.uni-tuebingen.de/controll/code.html> (24.01.05).
- Green, Georgia M. (1996). The structure of CONTEXT: The representation of pragmatic restrictions in HPSG. In J. Yoon (Hrsg.), *Proceedings of the 5th annual meeting of the Formal Linguistics Society of the Midwest*, Studies in the Linguistic Sciences 24, S. 215–232, Urbana, IL.
- Grewendorf, Günther, Hamm, Fritz und Sternefeld, Wolfgang (1994). *Sprachliches Wissen: eine Einführung in moderne Theorien der grammatischen Beschreibung* (7. Aufl.). Suhrkamp, Frankfurt/M.
- Haegeman, Liliane (1994). *Introduction to Government and Binding Theory* (Zweite Aufl.). Blackwell Textbooks in Linguistics. Blackwell, Oxford UK & Cambridge USA.
- Haider, Hubert (1986). Fehlende Argumente: vom Passiv zu kohärenten Infinitiven. *Linguistische Berichte* 101, S. 3–33.
- Hall, Christopher (1980). *Zur transformationellen Behandlung der Nominalisierung in einer generativen Grammatik des Deutschen*. Ph. D. thesis, Institut für Philologie I der Universität Tampere.
- Hamm, Fritz und van Lambalgen, Michiel (2003). Event Calculus, Nominalisation, and the Progressive. *Linguistics and Philosophy* 26, S. 381–458.
- Harnish, Robert M. und Farmer, Ann K. (1984). Pragmatics and the modularity of the linguistic system. *Lingua* 63, S. 255–277.
- Heinz, W. und Matiasek, J. (1994). Argument Structure and Case Assignment in German. In J. Nerbonne, K. Netter, and C. Pollard (Hrsg.), *German*

- in *Head-Driven Phrase Structure Grammar*, Nummer 46 in CSLI Lecture Notes, S. 199–236, CSLI Publications, Stanford, CA.
- Hessky, Regina und Ettinger, Stefan (1997). *Deutsche Redewendungen: ein Wörter- und Übungsbuch für Fortgeschrittene*. Narr-Studienbuch, Tübingen.
- Hinrichs, Erhard W. und Nakazawa, Tsuneko (1989). Flipped Out: Aux in German. In *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, Chicago, IL, S. 193–2002.
- Hinrichs, Erhard W. und Nakazawa, Tsuneko (1994). Linearizing AUXs in German Verbal Complexes. In J. Nerbonne, K. Netter, and C. Pollard (Hrsg.), *German in Head-Driven Phrase Structure Grammar*, Nummer 46 in CSLI Lecture Notes, S. 11–37, CSLI Publications, Stanford, CA.
- Höhle, Tilman (1986). Der Begriff Mittelfeld, Anmerkungen über die Theorie der topologischen Felder. In A. Schöne (Hrsg.), *Akten des Siebten Internationalen Germanistenkongresses 1985*, Göttingen, Band 3, S. 329–340. Niemeyer Verlag, Tübingen.
- Höhle, Tilman N. (1999). An Architecture for Phonology. In R. D. Borsley und A. Przepiórkowski (Hrsg.), *Slavic in Head-Driven Phrase Structure Grammar*, S. 61–90, CSLI Publications, Stanford, CA.
- Jackendoff, Ray (1975). Morphological and Semantic Regularities in the Lexikon. *Language* 51(3), S. 639–671.
- Jackendoff, Ray (1993). *Semantic Structures* (3. Druck). Nummer 18 in Current Studies in Linguistics. MIT Press, Cambridge, MA.
- Jacobson, Pauline (1987). Phrase Structure, Grammatical Relations, and Discontinuous Constituents. In G. J. Huck und A. E. Ojeda (Hrsg.), *Discontinuous Constituency*, Nummer 20 in Syntax and Semantics, S. 27–69, Academic Press Inc., San Diego, CA.
- Janssen, Theodoor Maria Victor (1983). *Foundations and Applications of Montague Grammar*. Doktorarbeit, Mathematisch Centrum, Universität Amsterdam.
- Kasher, Asa (1991). On the pragmatic modules: A lecture. *Journal of Pragmatics* 16, S. 381–397.

- Kathol, Andreas (2000). *Linear Syntax*. Oxford University Press.
- Katz, Jerrold J. und Postal, Paul M. (1963). Semantic Interpretation of Idioms and Sentences Containing Them. In *Quarterly Progress Report*, Nummer 70, S. 275–282, Massachusetts Institute of Technology, Research Laboratory of Electronics.
- Keil, Martina (1997). *Wort für Wort. Repräsentation und Verarbeitung verbaler Phraseologismen*. Sprache und Information. Max Niemeyer Verlag, Tübingen.
- Kiss, Tibor (1995a). *Infinitive Komplementation. Neue Studien zum Verbum Infinitum*. Nummer 333 in Linguistische Arbeiten. Max Niemeyer Verlag, Tübingen.
- Kiss, Tibor (1995b). *Merkmale und Repräsentationen. Eine Einführung in die deklarative Grammatikanalyse*. Westdeutscher Verlag, Opladen.
- Klenk, Ursula (2003). *Generative Syntax*. Narr Studienbücher. Narr, Tübingen.
- Krenn, Brigitte und Erbach, Gregor (1994). Idioms and Support Verb Constructions. In J. Nerbonne, K. Netter, and C. Pollard (Hrsg.), *German in Head-Driven Phrase Structure Grammar*, Nummer 46 in CSLI Lecture Notes, S. 365–396, CSLI Publications, Stanford, CA.
- Krifka, Manfred (2003). Bare NPs: Kind-referring, Indefinites, Both, or Neither? In *Proceedings of Semantics and Linguistic Theory (SALT) 13*, erscheint.
- Lakoff, George und Johnson, Mark (1980). *Metaphors we live by*. University of Chicago Press, Chicago.
- Lang, Ewald (1994). Semantische vs. Konzeptuelle Struktur: Unterscheidung und Überschneidung. In M. Schwarz (Hrsg.), *Kognitive Semantik / Cognitive Semantics. Ergebnisse, Probleme, Perspektiven*, S. 25–40, Gunter Narr Verlag, Tübingen.
- Lascarides, Alex und Copestake, Ann (1995). Order Independent Typed Default Unification. ACQUILEX-II working paper 60.
- Levelt, Willem J. M. (1989). *Speaking: From Intention to Articulation*. MIT Press, Cambridge, MA.

- Liebert, Wolf-Andreas (2002). Bildfelder in synchroner Perspektive. In D. A. Cruse, F. Hundsnurscher, M. Job und P. R. Lutzeyer (Hrsg.), *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*, Band 1, S. 771–783, Mouton de Gruyter, Berlin, New York.
- Löbner, Sebastian (2003). *Semantik. Eine Einführung*. Walter de Gruyter, Berlin, New York.
- Markert, Katja (1999). *Metonymien – eine computerlinguistische Analyse*. Dissertation, Universität Freiburg i. Br. Veröffentlicht als: DISKI 2000, Dissertations in Artificial Intelligence.
- Meibauer, Jörg (2001). *Pragmatik. Eine Einführung* (Zweite, verbesserte Aufl.). Stauffenburg Verlag, Tübingen.
- Meibauer, Jörg (2003). Phrasenkomposita zwischen Wortsyntax und Lexikon. *Zeitschrift für Sprachwissenschaft* 22.2, S. 153–188.
- Meurers, Walt Detmar (1995). Towards a Semantics for Lexical Rules as used in HPSG. In *Proceedings of the ACQUILEX II Workshop on the Formalisation and Use of Lexical Rules*, Cambridge, UK, S. 1–20. Auch präsentiert auf der First Conference on Formal Grammar, Barcelona, 1995.
- Meurers, Walt Detmar (1999). Raising Spirits (and assigning them case). Band 43 von *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)*, S. 173–226, Rijksuniversiteit Groningen, German Department.
- Meurers, Walt Detmar (2000). *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. Phil. Dissertation, Eberhard-Karls-Universität Tübingen. Veröffentlicht als: Arbeitspapiere des SFB 340, Nr. 145.
- Meurers, Walt Detmar (2002). To flip or not to flip: On the nature of irregularities in the German verbal complex. In F. van Eynde, L. Hellan, und D. Beermann (Hrsg.), *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, S. 235–246, CSLI Publications, Stanford, CA.
- Meurers, W. Detmar, Penn, Gerald, und Richter, Frank (2002). A Web-based Instructional Platform for Constraint-Based Grammar Formalisms and Parsing. In D. Radev and C. Brew (Hrsg.), *Effective Tools and Methodologies for Teaching NLP and CL*, New Brunswick, NJ, S. 18–25. The Association

- for Computational Linguistics. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. 7.–12. Juli 2002. Philadelphia, PA.
- Müller, Gereon (1998). Beschränkungen zur Binomialbildung im Deutschen. Ein Beitrag zur Interaktion von Phraseologie und Grammatik. In *Zeitschrift für Sprachwissenschaft*, Band 16.1/2, S. 5–51, Vandenhoeck & Ruprecht.
- Müller, Gereon (2000). Idioms and Transformations. (Teil eines Vortrags bei der GGS-Konferenz in Potsdam, Juni 2000). Stand: Juli 2000. <http://www.ids-mannheim.de/gra/texte/mu51.pdf> (24.01.05).
- Müller, Stefan (1996). The Babel-System—An HPSG Prolog Implementation. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, London, S. 263–277. <http://www.cl.uni-bremen.de/~stefan/Pub/babel.html> (24.01.05).
- Müller, Stefan (1999). *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Nummer 394 in Linguistische Arbeiten. Tübingen: Max Niemeyer Verlag.
- Müller, Stefan (2000). The Passive as a Lexical Rule. In D. Flickinger und A. Kathol (Hrsg.), *Proceedings of the HPSG-2000 Conference, University of California, Berkeley*. CSLI Publications, Stanford, CA. <http://csli-publications.stanford.edu/> (24.01.05).
- Müller, Stefan (2002). *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German*. Nummer 13 in Studies in Constraint-Based Lexicalism. Center for the Study of Language and Information, Stanford, CA.
- Müller, Stefan (2003a). Deutsche Syntax deklarativ. Unveröffentlichte Erweiterungen, Entwurf vom 19.11.2003.
- Müller, Stefan (2003b). Object-To-Subject-Raising and Lexical Rule. An Analysis of the German Passive. In S. Müller (Hrsg.), *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, S. 278–297, CSLI Publications, Stanford, CA. <http://csli-publications.stanford.edu/HPSG/4/> (24.01.05).
- Müller, Stefan (erscheint). Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte*. <http://www.cl.uni-bremen.de/~stefan/Pub/mehr-vf-lb.html> (24.01.05).

- Müller, Stefan und Kasper, Walter (2000). HPSG Analysis of German. In W. Wahlster (Hrsg.), *Verbmobil: Foundations of Speech-to-Speech Translation*, Artificial Intelligence, S. 238–253. Berlin Heidelberg New York: Springer-Verlag.
- Murphy, Patrick M. (1995). *Discourse, Themes and Word Order in Head-Driven Phrase Structure Grammar*. Präsentation auf dem Spring Linguistics Colloquium 1995, University of North Carolina at Chapel Hill. <http://www.unc.edu/courses/pre2000fall/ling030/murphy/papers/slc95/theme.pdf> (24.01.05).
- Neale, Stephen (1990). *Descriptions*. A Bradford Book. MIT Press, Cambridge, MA.
- Nerbonne, John (1996). Computational Semantics – Linguistics and Processing. In S. Lappin (Hrsg.), *Handbook of Contemporary Semantic Theory*, S. 459–482. Blackwell Publishers, London.
- Netter, Klaus (1996). *Functional Categories in an HPSG for German*. Phil. Dissertation, Universität des Saarlandes. Saarbrücken Dissertations in Computational Linguistics and Language Technology.
- Newmeyer, Frederick J. (1974). The Regularity of Idiom Behaviour. *Lingua* 34, S. 327–342, North Holland Publishing Company.
- Nunberg, Geoffrey (1979). The non-uniqueness of semantic solutions: polysemy. *Linguistics and Philosophy* 3, S. 143–184.
- Nunberg, Geoffrey, Sag, Ivan A. und Wasow, Thomas (1994). Idioms. *Language* 70, S. 491–538.
- O’Grady, William (1998). The Syntax of Idioms. *Natural Language and Linguistic Theory* 16, S. 279–312.
- Oppenrieder, Wilhelm (1991). *Von Subjekten, Sätzen und Subjektsätzen*. Nummer 241 in Linguistische Arbeiten. Max Niemeyer Verlag, Tübingen.
- Palm, Christine (1997). *Phraseologie. Eine Einführung* (Zweite Aufl.). Narr Studienbücher. Gunter Narr Verlag, Tübingen.
- Pedrazzini, Sandro (1994). *Phrase Manager. A System for Phrasal and Idiomatic Dictionaries*, Band 3 von *Informatik und Sprache*. Olms Verlag, Hildesheim.

- Penn, Gerald und Richter, Frank (2004). Lexical Resource Semantics: From Theory to Implementation. In S. Müller (Hrsg.), *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, S. 423–443. CSLI Publications, Stanford, CA. <http://csli-publications.stanford.edu/HPSG/5/> (24.01.05).
- Pollard, Carl und Sag, Ivan A. (1987). *Information-Based Syntax and Semantics*, CSLI Lecture Notes, Band 1 (13). Stanford University: CSLI Publications/The University of Chicago Press.
- Pollard, Carl und Sag, Ivan A. (1994). *Head-Driven Phrase Structure Grammar*. Stanford University: CSLI Publications/The University of Chicago Press.
- Rajchštejn, Aleksandr D. (1980). *Sopostavitel'nyj analiz nemeckoj i ruskoj frazeologii*. Moskau: Vysšaja škola.
- Rauh, Gisa (1991). Prepositional forms in the lexicon: problems and suggestions. In G. Rauh (Hrsg.), *Approaches to Prepositions*, S. 169–223, Narr, Tübingen.
- Rauh, Gisa (1993). On the grammar of lexical and non-lexical prepositions in English. In C. Zelinsky-Wibbelt (Hrsg.), *The Semantics of Prepositions - From Mental Processing to Natural Language Processing*, S. 99–150, Mouton de Gruyter, Berlin, New York.
- Reinhard, Sabine (2001). *Deverbale Komposita an der Morphologie-Syntax-Semantik- Schnittstelle: ein HPSG-Ansatz*. Phil. Dissertation, Eberhard-Karls-Universität Tübingen, Seminar für Sprachwissenschaft. <http://w210.ub.uni-tuebingen.de/dbt/volltexte/2002/466/> (24.01.05).
- Reis, Marga (1995). Wer glaubst du hat recht? On So-called Extractions from Verb-Second Clauses and Verb-First Parenthetical Constructions in German. In: Rosengren, I. und Önnarfors, O. (Hrsg.), *Sprache & Pragmatik. Arbeitsberichte des Forschungsprogramms „Sprache und Pragmatik“*, Band 36, S. 27–83, Lund, Schweden.
- Richter, Frank (1997). Die Satzstruktur des Deutschen und die Behandlung langer Abhängigkeiten in einer Linearisierungsgrammatik. Formale Grundlagen und Implementierung in einem HPSG-Fragment. In E. Hinrichs, D. Meurers, F. Richter, M. Sailer und H. Winhart (Hrsg.), *Ein HPSG-Fragment des Deutschen, Teil 1: Theorie*, Arbeitspapiere des SFB 340, Nummer 95, S. 13–187, Universität Tübingen.

- Richter, Frank (1999). RSRL for HPSG. In V. Kordoni (Hrsg.), *Tübingen Studies in Head-Driven Phrase Structure Grammar*, Arbeitspapiere des SFB 340, Nummer 132, S. 74–115.
- Richter, Frank (2004a). *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar*. Phil. Dissertation (2000), Eberhard-Karls-Universität Tübingen. <http://w210.ub.uni-tuebingen.de/dbt/volltexte/2004/1203/> (24.01.05).
- Richter, Frank (2004b). *Foundations of Lexical Resource Semantics*. Habilitationsschrift, Eberhard-Karls-Universität Tübingen. Fassung vom 24.09.2004
- Richter, Frank und Sailer, Manfred (1999a). LF Conditions on Expressions of Ty2: An HPSG Analysis of Negative Concord in Polish. In R. D. Borsley und A. Przepiórkowski (Hrsg.), *Slavic in Head-Driven Phrase Structure Grammar*, S. 247–282, CSLI Publications, Stanford, CA.
- Richter, Frank und Sailer, Manfred (1999b). A Lexicalist Collocation Analysis of Sentential Negation and Negative Concord in French. In V. Kordoni (Hrsg.), *Tübingen Studies in Head-Driven Phrase Structure Grammar*, Arbeitspapiere des SFB 340, Nummer 132, S. 231–300.
- Richter, Frank und Sailer, Manfred (2002). *LRS and Welsh Negation*. Unveröffentlichtes Manuskript einer Präsentation auf dem 6th CLAIRE Workshop (26.-27.04.2002) an der Universität von Essex, Colchester, England. <http://www.sfs.uni-tuebingen.de/~fr/cards/claire02.html> (24.01.05).
- Richter, Frank und Sailer, Manfred (2003). Cranberry Words in Formal Grammar. In C. Beyssade, O. Bonami, P. Cabredo Hofherr und F. Corblin (Hrsg.), *Empirical Issues in Formal Syntax and Semantics 4*, Band 4, S. 155–171, Presses de l'Université de Paris-Sorbonne.
- Richter, Frank und Sailer, Manfred (2004). Basic Concepts of Lexical Resource Semantics. In *ESSLLI 2003 – Course Material I*, Band 5, S. 87–143, *Collegium Logicum*. Kurt Gödel Society, Wien.
- Richter, Frank, Sailer, Manfred und Penn, Gerald (1999). A Formal Interpretation of Relations and Quantification in HPSG. In G. Bouma, E. Hinrichs, G.-J. M. Kruijff und R. Oehrle (Hrsg.), *Constraints and Resources in Natural Language Syntax and Semantics*, S. 281–298, CSLI Publications, Stanford, CA.

- Riehemann, Susanne (1997). Idiomatic Constructions in HPSG. Paper presented at the *Fourth International Conference on Head-Driven Phrase Structure Grammar*, Ithaca. <http://doors.stanford.edu/~sr/idioms.ps> (24.01.05).
- Riehemann, Susanne Z. (2001). *A Constructional Approach to Idioms and Word Formation*. Ph. D. thesis, Stanford University, Stanford, CA.
- Riehemann, Susanne Z. und Bender, Emily (2000). Absolute Constructions: On the Distribution of Predicative Idioms. In S. Bird, A. Carnie, J. Haugen und P. Norquest (Hrsg.), *Proceedings of the 18th WCCFL*, S. 476–489, Cascadilla Press, Somerville.
- Riezler, Stefan (1999). *Probabilistic Constraint Logic Programming. Formal Foundations of Quantitative and Statistical Inference in Constraint-Based Natural Language Processing*. Phil. Dissertation (1998), Eberhard-Karls-Universität Tübingen. Erschienen als AIMS Report 5(1), IMS, Universität Stuttgart
- Röhrich, Lutz und Mieder, Wolfgang (1977). *Sprichwort*. Metzler, Stuttgart.
- Römer, Christine (1996). Semantik – Pragmatik – Syntax: Zum Verhältnis von semantischem, pragmatischem und syntaktischem Wissen in der Kognitiven Semantik – dargestellt am Beispiel der Modalwörter. In I. Pohl (Hrsg.), *Methodologische Aspekte der Semantikforschung* (Sonderdruck 1997), S. 125–139, Peter Lang, Frankfurt am Main.
- Römer, Christine und Matzke, Brigitte (2003). *Lexikologie des Deutschen. Eine Einführung*. Narr Studienbücher. Gunter Narr Verlag, Tübingen.
- Römer, Christine und Müller, Stefan (2003, Februar). *Syntax und Semantik idiomatisierter verbaler Phraseologismen in einem abgestuft kompositionellen Grammatikmodell*. Erstantrag auf Gewährung einer Sachbeihilfe für das gleichnamige DFG-Projekt an der Universität Jena.
- Rothweiler, Monika (2002). Spracherwerb. In J. Meibauer et al., *Einführung in die germanistische Linguistik*, Kapitel 7, S. 251–293. J. B. Metzler, Stuttgart/Weimar.
- Sabban, Annette (1998). *Okkasionelle Variationen sprachlicher Schematismen. Eine Analyse französischer und deutscher Presse- und Werbetexte*. Nummer 53 in *Romanica Monacensia*. Gunter Narr Verlag, Tübingen.

- Sag, Ivan A., Wasow, Thomas und Bender, Emily M. (2003). *Syntactic Theory. A Formal Introduction* (zweite Aufl.). Nummer 152 in *CSLI Lecture Notes*. CSLI Publications, Stanford, CA.
- Sailer, Manfred (2003). *Combinatorial Semantics and Idiomatic Expressions in Head-Driven Phrase Structure Grammar*. Phil. Dissertation (2000). Arbeitspapiere des SFB 340, Nummer 161, Eberhard-Karls-Universität Tübingen.
- Sailer, Manfred (2004a). Local Semantics in Head-Driven Phrase Structure Grammar. In O. Bonami und P. Cabredo Hofherr (Hrsg.), *Empirical Issues in Syntax and Semantics 5*, S. 197–214. Actes du Colloque de Syntaxe et Sémantique à Paris 2003, <http://www.cssp.cnrs.fr/eiss5> (24.01.05).
- Sailer, Manfred (2004b). Distributionsidiosynkrasien. Korpuslinguistische Erfassung und grammatiktheoretische Deutung. In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest*, Institut für Deutsche Sprache, Jahrbuch 2003, S. 194–221, de Gruyter, Berlin, New York.
- Sailer, Manfred (2004c). Propositional Relative Clauses in German. In S. Müller (Hrsg.), *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, S. 223–243, CSLI Publications, Stanford, CA. <http://csli-publications.stanford.edu/HPSG/5/> (24.01.05).
- Sailer, Manfred und Richter, Frank (2002a). Collocations and the Representation of Polarity. In G. Alberti, K. Balogh und P. Dekker (Hrsg.), *Proceedings of the Seventh Symposium on Logic and Language*, Pécs, S. 129–138.
- Sailer, Manfred und Richter, Frank (2002b). Not for Love or Money: Collocations! In G. Jäger, P. Monachesi, G. Penn und S. Wintner (Hrsg.), *Proceedings of Formal Grammar 2002*, S. 149–160.
- Scherpenisse, Wim (1986). The Connection Between Base Structure and Linearization Restrictions in German and Dutch. In *Europäische Hochschulschriften*, Band 47 von *Reihe XXI*. Peter Lang, Frankfurt/M.
- Siegel, Melanie (2000). Japanese Honorification in an HPSG Framework. In A. Ikeya und M. Kawamori (Hrsg.), *Proceedings of the 14th Pacific Asia Conference on Language, Information and Computation*, Waseda University International Conference Center, Tokyo, S. 289–300. Logico-Linguistic Society of Japan.

- Soehn, Jan-Philipp (2003). Von Geisterhand zu Potte gekommen. Eine HPSG-Analyse von PPs mit unikalener Komponente. Magisterarbeit, Eberhard-Karls-Universität Tübingen, Seminar für Sprachwissenschaft. <http://www.sfs.uni-tuebingen.de/hpsg/archive/bibliography/papers/majp.ps> (24.01.05).
- Soehn, Jan-Philipp und Sailer, Manfred (2003). At First Blush on Tenterhooks. About Selectional Restrictions Imposed by Nonheads. In G. Jäger, P. Monachesi, G. Penn und S. Wintner (Hrsg.), *Proceedings of Formal Grammar 2003*, S. 149–161.
- Sternfeld, Wolfgang (1991). *Syntaktische Grenzen. Chomskys Barrierentheorie und ihre Weiterentwicklungen*. Westdeutscher Verlag, Opladen.
- Steyer, Kathrin (Hrsg.) (2004). *Wortverbindungen – mehr oder weniger fest*. Institut für Deutsche Sprache, Jahrbuch 2003. Walter de Gruyter, Berlin, New York.
- Torzova, M. V. (1983). Zur Valenz der Phraseologismen. *Deutsch als Fremdsprache* 5, S. 283–287.
- Trawiński, Beata (2003). Combinatorial Aspects of PPs Headed by Raising Prepositions. In *Proceedings of the ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, Toulouse, France, S. 157–168.
- Tseng, Jesse L. (2000). *The Representation and Selection of Prepositions*. Ph. D. thesis, University of Edinburgh.
- Uszkoreit, Hans (1987). *Word Order and Constituent Structure in German*. Nummer 8 in CSLI Lecture Notes. CSLI Publications, Stanford, CA.
- van der Wouden, Ton (1992). Beprekingen op het optreden van lexicale elementen. In *De Nieuwe Taalgids*, Band 85 (6), S. 513–538.
- van Lambalgen, Michiel und Hamm, Fritz (2004). *The Proper Treatment of Events*. Blackwell Publishing.
- Veith, Werner H. (2002). *Soziolinguistik. Ein Arbeitsbuch*. Tübingen: Gunther Narr Verlag.
- Vijayashanker, K. (1993). A study of tree adjoining grammars. Dissertation, UMI Dissertation Services, Ann Arbor, MI.

- Volk, Martin (1998). The Automatic Translation of Idioms. Machine Translation vs. Translation Memory Systems. In N. Weber (Hrsg.), *Machine Translation: Theory, Applications, and Evaluation. An assessment of the state of the art*. 1. Aufl., gardez-Verlag, St. Augustin. Zitiert nach <http://www.ifi.unizh.ch/CL/volk/publications.html> (24.01.05).
- Weinreich, Uriel (1969). Problems in the Analysis of Idioms. In J. Puhvel (Hrsg.), *Substance and Structure of Language*, S. 23–81, Berkeley und Los Angeles.
- Wilhelm, Thomas (1997). Name und Gegenstand: Deutungen der paradigmatischen Beziehung zwischen Sprache und Welt. In *Europäische Hochschulschriften, Reihe XX*, Band 531, Peter Lang, Frankfurt am Main.
- Winhart, Heike (1997). Die Nominalphrase in einem HPSG-Fragment des Deutschen. In Hinrichs, E., Meurers, D., Richter, F., Sailer, M. und Winhart, H. (Hrsg.) (1997). *Ein HPSG-Fragment des Deutschen, Teil 1: Theorie*. Arbeitspapiere des SFB 340, Nummer 95, S. 319–384, Universität Tübingen.
- Winhart, Heike (2002). Funktionsverbgefüge im Deutschen, Zur Verbindung von Verben und Nominalisierungen. Unveröffentlichte Dissertation, Eberhard-Karls-Universität Tübingen.
- Wotjak, Barbara (1992). *Verbale Phraseolexeme in System und Text*. Nummer 125 in Germanistische Linguistik. Max Niemeyer Verlag, Tübingen.

A IVP-Korpus

Das Korpus, welches die empirische Grundlage für unsere Arbeit darstellt, steht als XML-Datei auf beiliegender CD-ROM zur Verfügung. Die Datei lässt sich mit den neueren Versionen der gängigen Webbrowser (z. B. Firefox, Mozilla, Internet-Explorer) öffnen.¹ Die Publikation des Korpus in digitaler Form hat den Vorteil, schneller und gezielter auf spezielle Einträge zugreifen zu können (Experten können auch Suchkriterien im Stylesheet unter `<xml:if>` definieren) und die Daten für weitere Untersuchungen zur Verfügung zu haben. Darüber hinaus befindet sich eine Online-Version im Internet unter www.soehn.net/ivp/korpus.htm, die aktuelle Korrekturen und Erweiterungen enthält.

Die Korpuseinträge sind folgendermaßen aufgebaut:

1. Es wird zunächst die Grundform (bzw. die kanonische Form) der Wendung angegeben, wobei eine alphabetische Sortierung nach dem Verb vorgenommen wurde. Bei gleichlautenden Verben wurde alphabetisch nach dem internen Argument sortiert.

Die Grundform ist diejenige, welche alle (und nur die) wendungsinternen Bestandteile enthält. Sie kann gewonnen werden, indem man aus der Summe der Korpusbelege nur die stets vorhandenen Elemente berücksichtigt. Bei Wendungen aus dem Duden haben wir den Wortlaut der dortigen Eintragung übernommen.

Die Grundform dient der eindeutigen Identifizierung eines Korpuseintrags und bildet die Ausgangsbasis für die Kennzeichnung von Varianten oder Modifikationen.

2. Die Quellenangabe: Die Wendungen stammen aus Burger (2003), dem Duden 11 (2002), Keil (1997) oder CoDII-BW.DE², einem Online-

¹Beim Kopieren der Dateien auf die Festplatte ist zu beachten, dass sich – um eine korrekte Darstellung zu gewährleisten – die beiden Dateien `ivp-liste.xml` und `ivp-style.xml` im selben Verzeichnis befinden müssen.

²CoDII-BW wurde im Teilprojekt A5 des SFB 441 an der Universität Tübingen entwickelt (<http://www.sfb441.uni-tuebingen.de/a5/codii/index.xhtm>). Es wurden v. a. Fleischer (1997), Dobrovolskij (1988) und Dobrovolskij/Piirainen (1994) als Quellen herangezogen.

Verzeichnis für gebundene Wörter. Einige Wendungen stammen auch aus Römer/Müller (2003).

3. Ein konkreter Gebrauch und dessen Kontext: Als Quelle der Belegstellen haben wir die frei verfügbaren Online-Korpora des Instituts für Deutsche Sprache in Mannheim (über COSMAS II) und die Suchmaschine GOOGLE herangezogen.
4. Die untersuchten Eigenschaften:
 - Angabe einer Paraphrase
 - Angabe, ob die Wendung einen *Marker* bzw. ein Kennzeichen aufweist (vgl. Kap. 3).
 - Angabe, ob es sich um eine metaphorische Wendung handelt.
 - Angabe, ob es sich um eine metonymische Wendung handelt.
 - Angabe, ob es sich um eine für die heutige Sprechergemeinschaft noch motivierte Wendung handelt. Motivierung und Durchsichtigkeit sind hier schwer voneinander abzugrenzen.
 - Angabe, ob die Wendung dekomponierbar ist. Wenn ja, wird die idiomatische Bedeutung der Idiomteile in Versalien spezifiziert.
 - Angabe, ob die Wendung ein unikales Element beinhaltet und welcher Kategorie es ggf. angehört. Die Bemerkung „in dieser Bedeutung“ heißt, dass es zu dem unikalen Element auch frei vorkommende Homonyme gibt, welche aber eine andere Bedeutung haben. In diesem Fall kann das Vorhandensein eines unikalen Elementes nicht als eindeutiges Idiom-Kennzeichen gewertet werden. Der Hinweis „nur Kombination“ bedeutet, dass die Teile einer nominalen Konjunktion einzeln nicht unikal sind, die Konjunktion als Ganze jedoch schon.
Manchmal stammt eine Wendung aus der CoDII-BW.DE Sammlung, wird jedoch im Folgenden ohne unikales Element eingestuft. Dies geschieht dann, wenn eine beträchtliche Zahl an Korpusbelegen ohne den „kanonischen“ Kontext existiert.
 - Eventuelle Varianten der Grundform, die häufig gebraucht werden (zur Variabilität siehe auch Dobrovol'skij, 1988, Kap. 2.3.2.2).

- Der Valenzrahmen des Verbs, untergliedert in interne und externe Argumente. Zusätzliche Abkürzungen:

subjNP	Subjekt
objNP	direktes Objekt
indObjNP	indirektes Objekt
kontr.	kontrahiert
expl.	expletiv
- 5. Syntaktische und semantische Restriktionen: Referenzialität, Passivierbarkeit, Nominalisierbarkeit, die Möglichkeit der Expansion durch interne Adjunkte und die Negierbarkeit mit „nicht“ oder „kein“ (vgl. Kap. 2). Ist eine der Möglichkeiten gegeben, wird jeweils ein Korpusbeleg angeführt. Bei der Passivierbarkeit wurde unterschieden, ob das Idiom an sich nicht passiviert werden kann oder ob dies an unabhängigen Eigenschaften des Verbs liegt. Letzterer Fall wird durch den Hinweis „Verb!“ gekennzeichnet.

B Zur HPSG-Signatur

Es folgt eine zusammenfassende Darstellung aller Sorten und Merkmale, die in dieser Arbeit neu vorgestellt bzw. vorwiegend benutzt wurden. Alle Merkmale, die es unverändert schon in Pollard/Sag (1994) gab, werden in grau dargestellt.

Merkmals-Deklaration

$$\left[\begin{array}{l} \textit{sign} \\ \text{PHON} \quad \textit{list (phonstring)} \\ \text{SYNSEM} \quad \textit{synsem} \\ \text{LF} \quad \textit{lrs} \\ \text{COLL} \quad \textit{list (barrier)} \end{array} \right] \quad \left[\begin{array}{l} \textit{phrase} \\ \text{DTRS} \quad \textit{con-struct} \end{array} \right] \quad \left[\begin{array}{l} \textit{word} \\ \text{STORE} \quad \textit{list (der-rule)} \end{array} \right]$$

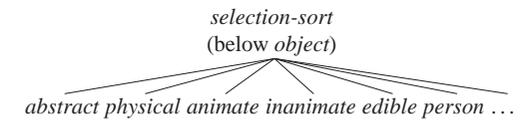
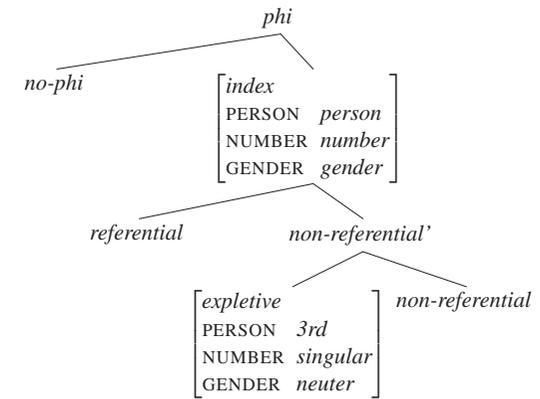
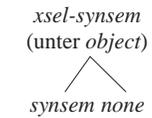
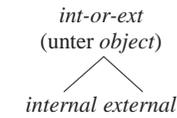
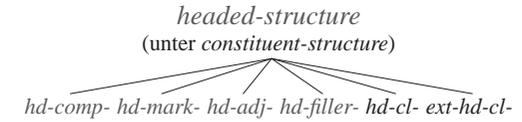
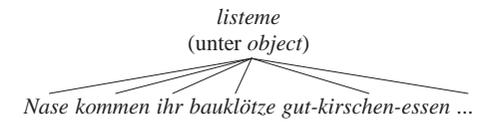
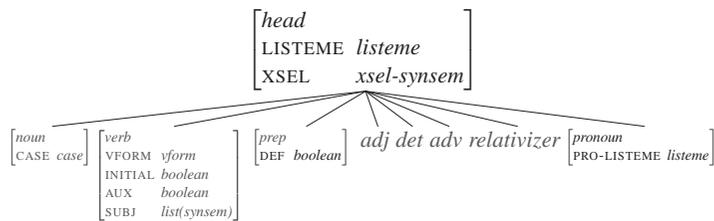
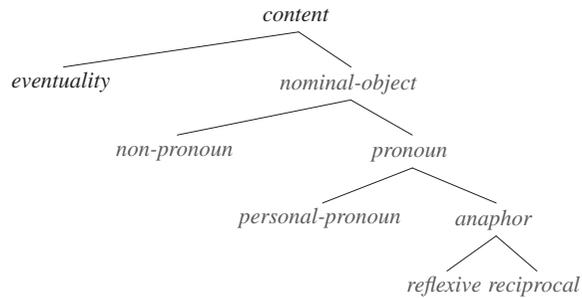
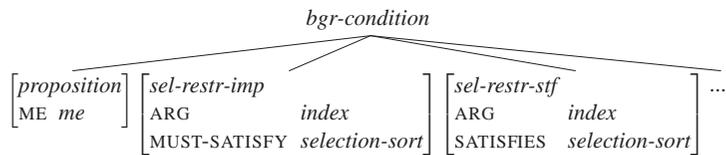
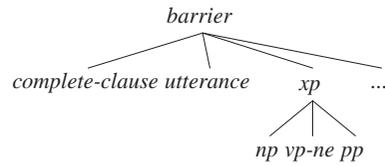
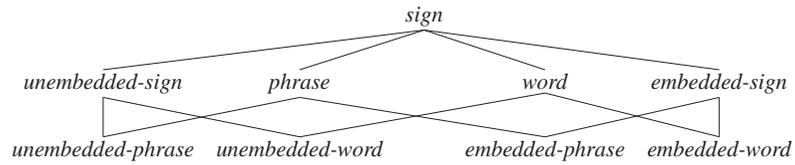
$$\left[\begin{array}{l} \textit{barrier} \\ \text{PHON-LIC} \quad \textit{list (phonstring)} \\ \text{LOC-LIC} \quad \textit{local} \\ \text{LF-LIC} \quad \textit{lrs} \end{array} \right] \quad \left[\begin{array}{l} \textit{synsem} \\ \text{LOCAL} \quad \textit{local} \\ \text{NONLOCAL} \quad \textit{nonlocal} \\ \text{LEX} \quad \textit{boolean} \end{array} \right] \quad \left[\begin{array}{l} \textit{local} \\ \text{CATEGORY} \quad \textit{category} \\ \text{CONTENT} \quad \textit{content} \\ \text{CONTEXT} \quad \textit{context} \\ \text{I-ARG-STAT} \quad \textit{int-or-ext} \end{array} \right]$$

$$\left[\begin{array}{l} \textit{content} \\ \text{INDEX} \quad \left[\begin{array}{l} \textit{extended-index} \\ \text{PHI} \quad \textit{phi} \\ \text{VAR} \quad \textit{me} \end{array} \right] \\ \text{MAIN} \quad \textit{me} \end{array} \right] \quad \left[\begin{array}{l} \textit{lrs} \\ \text{EXTERNAL-CONTENT} \quad \textit{me} \\ \text{INTERNAL-CONTENT} \quad \textit{me} \\ \text{PARTS} \quad \textit{list(me)} \end{array} \right]$$

$$\left[\begin{array}{l} \textit{derivational-rule} \\ \text{IN} \quad \textit{word} \\ \text{OUT} \quad \textit{word} \end{array} \right] \quad \left[\begin{array}{l} \textit{dual-input-derivational-rule} \\ \text{IN1} \quad \textit{word} \\ \text{IN2} \quad \textit{sign} \\ \text{OUT} \quad \textit{word} \end{array} \right] \quad \left[\begin{array}{l} \textit{ple-derivational-rule} \\ \text{IN} \quad \textit{phrase} \\ \text{OUT} \quad \textit{word} \end{array} \right]$$

Außerdem nehmen wir mit Richter (2004b) an, dass für jede Sorte das Attribut EMBEDDED mit dem Wert *unembedded-sign* definiert ist.

Sortenhierarchie



Für die Analyse relevante Prinzipien

GÜLTIGKEITSPRINZIP DER SELEKTIONSRESTRIKTIONEN (GPSR):

Gibt es in einem uneingebetteten Zeichen x ein Zeichen z , ein Verb v (wobei z ein Argument von v ist) und eine Proposition p , welche durch v und seine Argumente ausgedrückt wird, und

stehen weder die logische Form, die mit z assoziiert ist, noch diejenige, die mit p assoziiert ist, im Skopus eines Negationsoperators, eines Konditionaloperators oder eines Frageoperators, noch sind sie Argument eines nonfaktiven Verbs,

dann muss der STF-Wert eines Elements der Sorte *sel-rest-stf* auf der CTXT BGR-Menge von x mit dem M-STF-Wert eines Elements der Sorte *sel-restr-imp* in der *is_a*-Relation auf der HIERARCHY-Liste von x stehen oder sich in einer weiteren Relation befinden, wobei die ARG-Werte von *sel-rest-stf* und *sel-restr-imp* identisch sein müssen.

Das EXTERNAL-CONTENT-PRINZIP (EContP):

1. In jeder Phrase ist der EXCONT-Wert der Nichtkopftochter ein Element der PARTS-Liste der Nichtkopftochter.
2. In jeder Äußerung ist jeder Teilausdruck des EXCONT-Wertes der Äußerung ein Element ihrer PARTS-Liste und jedes Element dieser PARTS-Liste ist ein Teilausdruck des EXCONT-Wertes.

ID-PRINZIP:

$$\left[\begin{array}{l} \textit{phrase} \\ \textit{coll elist} \end{array} \right] \rightarrow \left(\begin{array}{l} \text{KOPF-KOMPLEMENT-SCHEMA} \vee \text{KOPF-ADJUNKT-SCHEMA} \vee \\ \text{KOPF-MARKER-SCHEMA} \vee \text{KOPF-FÜLLER-SCHEMA} \vee \\ \text{PRÄDIKATSKOMPLEX-SCHEMA} \vee \text{ERWEITERTER-PRÄDIKATS-} \\ \text{KOMPLEX-SCHEMA} \end{array} \right)$$

Die ersten vier Schemata sind in Pollard/Sag (1994) definiert (zu beachten ist, dass die Bedingung [COLL *elist*] im Antezedens stehen muss), die Definition des Prädikatskomplex-Schemas folgt weiter unten.

Das INTERNAL-ARGUMENT-PRINZIP (IArgP):

$$\left[\begin{array}{l} \textit{phrase} \\ \textit{DTRS headed-struct} \\ \textit{coll elist} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{SS LOC I-ARG-STAT } \square \\ \text{DTRS H-DTR } [\text{SS LOC I-ARG-STAT } \square] \end{array} \right]$$

Das INTERNAL-CONTENT-PRINZIP (IContP):

In jedem Objekt der Sorte *lrs* ist der INCONT-Wert ein Element der PARTS-Liste und eine Komponente des EXCONT-Wertes.

LEXIKON-PRINZIP:

$$\left[\begin{array}{l} \textit{sign} \\ \textit{coll nelist} \end{array} \right] \rightarrow (\text{LE}_1 \vee \dots \vee \text{LE}_m \vee \text{PLE}_1 \vee \dots \vee \text{PLE}_n)$$

LIZENSIERUNGSPRINZIP (LIP, allgemeine Version):

Für jedes *barrier*-Objekt in der COLL-Liste eines Zeichens x und für jede Phrase z gilt:

- a) der PHON-Wert von z ist identisch mit dem PHON-LIC-Wert von x ,
- b) der LOCAL-Wert von z ist identisch mit dem LOC-LIC-Wert von x ,
- c) der LF-Wert von z ist identisch mit dem LF-LIC-Wert von x

genau dann, wenn z das Zeichen x dominiert, z als die spezifizierte Barriere identifiziert werden kann und z kein Zeichen y dominiert, das eine gleichwertige Barriere bildet und selber x dominiert.

LRS-PROJEKTIONSPRINZIP:

In jeder *headed-phrase*

1. sind die EXCONT-Werte des Kopfes und des Mutterknotens identisch,
2. sind die INCONT-Werte des Kopfes und des Mutterknotens identisch und
3. der PARTS-Wert beinhaltet alle und nur die Elemente der PARTS-Listen der Töchter.

PRÄDIKATSKOMPLEX-SCHEMA:

$$\left[\begin{array}{l} \textit{phrase} \\ \textit{SYNSEM} \left[\begin{array}{l} \text{LOC CAT} \left[\begin{array}{l} \text{HEAD LISTEME } \square \\ \text{SUBCAT } \square \end{array} \right] \end{array} \right] \\ \textit{DTRS} \left[\begin{array}{l} \textit{head-cluster-struct} \\ \text{H-DTR} \left[\text{SYNSEM LOC CAT SUBCAT } \square \oplus \langle \square \rangle \right] \\ \text{CL-DTR} \left[\text{SYNSEM } \square \text{ LOC CAT HEAD LISTEME } \square \right] \end{array} \right] \\ \textit{coll elist} \end{array} \right]$$

