

Linguistische  
Arbeiten

394

Herausgegeben von Hans Altmann, Peter Blumenthal, Herbert E. Brekle,  
Gerhard Helbig, Hans Jürgen Heringer, Heinz Vater und Richard Wiese



*Stefan Müller*

# Deutsche Syntax deklarativ

Head-Driven Phrase Structure Grammar  
für das Deutsche

Max Niemeyer Verlag  
Tübingen 1999



Die Deutsche Bibliothek – CIP-Einheitsaufnahme

**Müller, Stefan:** Deutsche Syntax deklarativ : head driven phrase structure grammar für das Deutsche / Stefan Müller. – Tübingen : Niemeyer, 1999  
(Linguistische Arbeiten ; 394)

ISBN 3-484-30394-8      ISSN 0344-6727

D 291

© Max Niemeyer Verlag GmbH, Tübingen 1999

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

Gedruckt auf alterungsbeständigem Papier.

Druck: Weihert-Druck GmbH, Darmstadt

Buchbinder: Nägele Verlags- und Industriebuchbinderei, Nehren

# Inhaltsverzeichnis

Danksagung . . . . .	ix
Vorwort . . . . .	xi
1. Syntaktische Merkmale und Kategorien . . . . .	1
1.1. Konstituenten und Kategorien . . . . .	1
1.2. Aufbau von Merkmalstrukturen zur Beschreibung von Zeichen . . . . .	2
1.3. Kopfmerkmale . . . . .	4
1.4. Das Subkategorisierungsmerkmal (SUBCAT) . . . . .	7
1.5. Abkürzungen . . . . .	8
1.6. Das Subkategorisierungsprinzip . . . . .	10
1.7. Das SUBJ-Merkmal . . . . .	11
2. Dominanzschemata . . . . .	14
2.1. Flache Strukturen . . . . .	15
2.2. Binär verzweigende Strukturen . . . . .	20
2.3. Zusammenfassung . . . . .	22
3. Die Semantik innerhalb der HPSG . . . . .	24
3.1. Individuen, Sachverhalte und Situationen . . . . .	24
3.2. Die Semantik von Zeichen . . . . .	27
3.3. Das Semantikprinzip . . . . .	30
3.4. Rollen und fakultative Komplemente . . . . .	34
3.5. Abkürzungen . . . . .	36
4. Modifikation . . . . .	38
4.1. Komplemente und Modifikatoren . . . . .	38
4.2. Das Kopf-Adjunkt-Schema . . . . .	39
4.3. Kapselnde Modifikation . . . . .	42
4.4. Interpretation der Bedeutung von Adjektiven . . . . .	50
5. Markierer oder Komplementierer . . . . .	52
5.1. Markierer . . . . .	52
5.2. Das positionale <i>es</i> . . . . .	54
5.3. Subordinierende Konjunktionen . . . . .	55
5.4. Zusammenfassung . . . . .	57
6. Possessiva – das Spezifikatorprinzip . . . . .	58
7. Lexikalische Regeln und Morphologie . . . . .	63
7.1. Die Hierarchie lexikalischer Typen . . . . .	63
7.2. Morphologie . . . . .	68

8. Ein topologisches Modell des deutschen Satzes . . . . .	83
8.1. Verbstellungstypen . . . . .	83
8.2. Vorfeld, Mittelfeld und Nachfeld . . . . .	83
8.3. Die topologische Einteilung Bechs . . . . .	85
9. Nichtlokale Abhängigkeiten . . . . .	90
9.1. Bewegungsphänomene im Englischen und im Deutschen . . . . .	90
9.2. Das NONLOC-Prinzip . . . . .	92
9.3. Extraktionsverbot . . . . .	97
9.4. Einführung nichtlokaler Information . . . . .	104
9.5. Probleme . . . . .	113
10. Relativsätze . . . . .	117
10.1. Aufbau und Funktion von Relativsätzen . . . . .	117
10.2. Relativpronomina . . . . .	123
10.3. Relativsätze mit Bezugsnomen . . . . .	124
10.4. Freie Relativsätze . . . . .	131
10.5. Nichtlokale Information und Subkategorisierung . . . . .	150
10.6. Linearisierungsregeln für Relativsätze . . . . .	151
10.7. Rattenfänger-Konstruktionen ( <i>Pied Piping</i> ) . . . . .	152
10.8. Probleme . . . . .	157
11. Konstituentenreihenfolge . . . . .	160
11.1. LP-Regeln . . . . .	160
11.2. Wortstellungsdomänen . . . . .	161
11.3. Lexikalische Köpfe . . . . .	163
11.4. Komplemente und Modifikatoren im Mittelfeld . . . . .	166
11.5. Alternativen . . . . .	175
12. Extraktion ins Mittelfeld . . . . .	186
12.1. Die Phänomene . . . . .	186
12.2. Die Analyse . . . . .	190
12.3. Alternativen . . . . .	199
12.4. Zusammenfassung . . . . .	203
13. Extraposition . . . . .	205
13.1. Die Phänomene . . . . .	205
13.2. Extraposition über den NONLOC-Mechanismus . . . . .	224
13.3. Komplexe Wortstellungsdomänenbildung . . . . .	244
13.4. Alternativen . . . . .	248
13.5. Probleme . . . . .	249
13.6. Zusammenfassung . . . . .	251

14. Der Verbalkomplex . . . . .	253
14.1. Die Bildung des Verbalkomplexes . . . . .	253
14.2. Die Wortstellung im Verbalkomplex . . . . .	259
14.3. Unterbrechung des Verbalkomplexes . . . . .	263
14.4. Probleme . . . . .	268
15. Kasuszuweisung und Passiv . . . . .	270
15.1. Kasus im Deutschen – das Kasusprinzip . . . . .	270
15.2. Akkusativobjekte – das ERG-Merkmal . . . . .	285
15.3. Passiv . . . . .	288
15.4. Auswahl des Hilfsverbs in Perfektkonstruktionen . . . . .	307
15.5. Adjektivische Partizipien . . . . .	308
15.6. Kongruenzkasus . . . . .	312
15.7. Alternativen . . . . .	315
15.8. Probleme . . . . .	319
16. Kongruenz . . . . .	322
16.1. Determinatoren . . . . .	322
16.2. Adjektive . . . . .	323
16.3. Subjekt-Verb-Kongruenz . . . . .	325
16.4. Relativsätze . . . . .	327
17. Infinite Konstruktionen . . . . .	329
17.1. Die Phänomene . . . . .	329
17.2. Argumentanziehung und Kohärenz . . . . .	337
17.3. Die Lexikoneinträge . . . . .	338
17.4. Intraposition . . . . .	343
17.5. Die dritte Konstruktion . . . . .	344
17.6. Adjunkte im Mittelfeld . . . . .	348
17.7. Alternativen . . . . .	350
18. Voranstellung von Teilprojektionen . . . . .	353
18.1. Die Phänomene . . . . .	353
18.2. Die Analyse . . . . .	356
18.3. Alternativen . . . . .	364
18.4. Probleme . . . . .	374
18.5. Zusammenfassung . . . . .	380
19. Abtrennbare Präfixe . . . . .	381
19.1. Die Phänomene . . . . .	381
19.2. Die Analyse . . . . .	391
19.3. Probleme . . . . .	402

20. Bindung von Pronomina und Anaphern . . . . .	404
20.1. Die Bindungsprinzipien von Pollard und Sag . . . . .	406
20.2. Probleme mit dem Dominanzbegriff . . . . .	409
20.3. Eine nichtkonfigurationale Bindungstheorie . . . . .	410
20.4. Probleme . . . . .	412
20.5. Zusammenfassung . . . . .	421
21. Diskontinuierliche Konstituenten . . . . .	422
21.1. Wortstellung im Mittelfeld . . . . .	422
21.2. Die Komplexität des Parse-Problems . . . . .	432
21.3. Zusammenfassung . . . . .	437
22. Schlußbemerkungen . . . . .	439
22.1. Zusammenfassung . . . . .	440
22.2. Offene Fragen . . . . .	440
Abkürzungsverzeichnis . . . . .	443
Verwendete Schemata und Prinzipien . . . . .	445
Literaturverzeichnis . . . . .	455
Index . . . . .	469

# Danksagung

Ich bedanke mich hiermit bei Berthold Crysmann, Kordula De Kuthy, Gregor Erbach, Walter Kasper, Frank Keller, Tibor Kiss, Hans-Ulrich Krieger, Christoph Lehner, Detmar Meurers, Klaus Netter, Karel Oliva, Androulla Papakyriacou, Christian Pietsch, Noemi Preissner, Hans Uszkoreit und Wolfgang Wahlster für ihre kritischen Anmerkungen zu früheren Versionen des vorliegenden Buches bzw. für Anmerkungen zu aus der Arbeit an diesem Buch hervorgegangenen Aufsätzen.

Frank Keller verdanke ich so manchen sachkundigen Hinweis zum Thema Extraposition. Bei Andreas Kathol bedanke ich mich für anregende Diskussionen per Email in den Jahren 1994 und 1995.

Vielen Dank auch an Gosse Bouma, Dan Flickinger, Erhard Hinrichs, Tilman Höhle, Paul King, Tibor Kiss, Jonas Kuhn, Frank Richter, Ivan Sag, Manfred Sailer, Hartmut Schmidt, Renate Schmidt und Gertjan van Noord für anregende Diskussionen auf Konferenzen, Sommerschulen, Projekttreffen und bei anderen Gelegenheiten.

Detmar Meurers und Kordula De Kuthy haben mir 1998 sehr bei der Beschaffung von Literatur geholfen.

Bei Renate Schmidt möchte ich mich fürs Korrekturlesen und für inhaltliche Anmerkungen bedanken. Barbara Schmidt danke ich für ihre Akzeptabilitätsurteile.

Johannes Heinecke und den Mitgliedern der *German Language TeX Users Group Communication List* sei für die Hilfe bei der Lösung von so manchem kniffligen L<sup>A</sup>T<sub>E</sub>X-Problem gedankt.

Dieses Buch wurde auf einem Computer mit dem Betriebssystem Linux gesetzt. Dank auch an alle Linux-Entwickler.



## Vorwort

Dieses Buch enthält eine umfassende und detaillierte Beschreibung eines Grammatikfragments für das Deutsche. Als linguistische Theorie liegt die *Head-Driven Phrase Structure Grammar* (HPSG) zugrunde. Da alle für das Verständnis der Analysen notwendigen Voraussetzungen in den ersten Kapiteln zur Verfügung gestellt werden, ist das Buch auch als Einführung in die Theorie der *Head-Driven Phrase Structure Grammar* verwendbar. Kontrollfragen und Übungsaufgaben zu einzelnen Kapiteln findet man unter <http://www.dfki.de/~stefan/Pub/hpsg.html>. Die in diesem Buch entwickelte linguistische Theorie bildet die theoretische Grundlage des Babel-Systems<sup>1</sup> – einer Prolog-Implementation eines Systems zur Verarbeitung des in diesem Buch vorgestellten HPSG-Fragments (Müller, 1996). Die Funktion der dargestellten Prinzipien und der Beschreibungen linguistischer Phänomene wurde mit Hilfe des Systems verifiziert.

Die ersten sieben Kapitel beschäftigen sich mit den Grundbausteinen der HPSG-Theorie. Im ersten Kapitel wird auf syntaktische Merkmale und den Zusammenhang zwischen phrasalen Kategorien und ihren Köpfen eingegangen. Subkategorisierung und das damit verbundene Prinzip werden erklärt. Dominanzschemata werden in Kapitel 2 vorgestellt. Binäre und flache Strukturen werden gegeneinander abgewogen. Kapitel 3 beschäftigt sich mit der kompositionalen Semantik der HPSG. Die Semantik einzelner Zeichen wird erklärt und das Prinzip, das die Bedeutung phrasaler Zeichen bestimmt. In Kapitel 3 wird auch auf die Frage der Argumentbezeichnung eingegangen. Ich habe mich für eine stärkere Verwendung des Typkonzepts und für ein Grundinventar elementarer Rollen entschieden. Kapitel 4 beschäftigt sich mit der Modifikation. Eine Links-vor-Rechts-Regel wird vorgeschlagen, die unechte Mehrdeutigkeiten und falsche Analysen beseitigt. Nach der Erklärung der in der HPSG verwendeten Mittel zur Beschreibung der Semantik in Modifikationsstrukturen kann in Kapitel 5 gezeigt werden, daß es nicht sinnvoll ist, satzeinleitende Konjunktionen als Markierer zu beschreiben. In Kapitel 6 wird der Aufbau der Lexikoneinträge für Possessivpronomina erläutert. Es wird gezeigt, wie man mit Hilfe des SPEC-Merkmals die semantische Verbindung zum spezifizierten Nomen herstellen kann. Kapitel 7 widmet sich der Morphologie von Verben, Substantiven, Adjektiven und Determinatoren. Es wird der Ansatz von Reimann (1992) erklärt, bei dem bestimmte Flexionsmerkmale explizit im Lexikon kodiert werden. Die entsprechenden Merkmale werden aufgeführt und kurz erklärt. Außerdem wird erklärt, wie mittels einer Hierarchie lexikalischer Typen Redundanz im Lexikon vermieden werden kann. Im Gegensatz zu existierenden Ansätzen zur Beschreibung der Morphologie verwende ich lexikalische Regeln, die nicht lexikalische Zeichen einer Grundform (bei Verben gewöhnlich die Infinitivform) auf anders flektierte Zeichen abbilden, sondern statt dessen kompatible Stamm-Endungs-Paare auf flektierte lexikalische Zeichen.

Das zentrale Problem beim Schreiben von Grammatiken für die deutsche Sprache ist die vergleichsweise freie Konstituentenreihenfolge im Deutschen. So ist es nicht verwunderlich, daß die Wortstellung Gegenstand der Betrachtung in den meisten Kapiteln des vorliegenden Buches ist. Die Verarbeitung von Grammatiken für Sprachen mit freier Konstituentenreihenfolge ist in vielerlei Hinsicht problematisch. In dieser Arbeit zeige ich, daß es möglich ist,

---

<sup>1</sup> <http://www.dfki.de/~stefan/Babel/>

ein großes Fragment des Deutschen unter Berücksichtigung der Wortstellungsproblematik zu beschreiben und ein System zu dessen Verarbeitung effizient zu implementieren.

Kapitel 8–14, 17, 21 und 18 beschäftigen sich mit Wortstellungsphänomenen. Aufbauend auf dem im Kapitel 8 vorgestellten topologischen Modell des deutschen Satzes wird in Kapitel 9 die Voranstellung von Konstituenten in Verbzweitsätzen als nichtlokale Abhängigkeit mittels des NONLOCAL-Mechanismus erklärt. Verschiedene Möglichkeiten zur Einführung nichtlokaler Information werden diskutiert. So werden Vor- und Nachteile von phonologisch leeren Elementen, unären Dominanzschemata und lexikalischen Regeln sowohl aus linguistischer als auch aus computationeller Sicht diskutiert. Es wird begründet, warum ich in meiner Grammatik unäre Projektionen verwende. Lexikalische Regeln finden nur in den Fällen Anwendung, in denen die phonologische Form von Lexikoneinträgen geändert wird.

Die Möglichkeit der relativ freien Anordnung von zu einem Kopf gehörenden Konstituenten im Mittelfeld (*Scrambling*) wird im Kapitel 11 mit Hilfe der Wortstellungsdomänen erklärt. Fälle, in denen tiefer eingebettete Konstituenten frei im Mittelfeld auftauchen, können durch ein spezielles Dominanzschema beschrieben werden, das diese Art von *Scrambling* mit Hilfe nichtlokaler Abhängigkeiten erklärt (Kapitel 12). Dieses Vorgehen ist dadurch motiviert, daß die tiefer eingebetteten Konstituenten auch ins Vorfeld gestellt werden können. Sogenanntes *Preposition Stranding* wird mittels lexikalischer Regeln beschrieben.

Der NONLOCAL-Mechanismus wird ebenfalls zur Beschreibung von Relativsätzen (Kapitel 10) benötigt. Ein leerer Relativierer zur Beschreibung der Relativsatzsemantik und ein alternatives Relativsatzschema werden vorgestellt. Auf freie Relativsätze und Rattenfänger-Konstruktionen wird eingegangen.

Im Kapitel 13 werden verschiedene Extrapositionsphänomene und deren Beschreibung mittels NONLOC-Mechanismus diskutiert. Außerdem wird Extraposition als komplexe Bildung von Wortstellungsdomänen nach Kathol und Pollard (1995) erörtert. Die Skizze von Kathol und Pollard wird ausgearbeitet, so daß sie denselben Phänomenbereich abdeckt, der mit dem NONLOC-Mechanismus beschrieben wurde. Es wird gezeigt, daß die Analyse, die nur mit Wortstellungsdomänen arbeitet, adäquater ist, da Extraposition mit unterschiedlichen Landeplätzen einfacher zu erklären ist. Für die maschinelle Verarbeitung der Grammatik ist der Ansatz von Kathol und Pollard ebenfalls besser geeignet, da die Menge der von der Grammatik lizenzierten Zeichen nicht durch das Aufstellen von Hypothesen in bezug auf Extraposition unendlich wird.

Oberfeldumstellung und das Voranstellen von Teilprojektionen (Adjektiv- und Verbprojektionen) sind Gegenstand der Kapitel 14 bzw. 18. Im Kapitel 18 wird eine von mir entwickelte Analyse für vorangestellte Teilprojektionen vorgestellt, die wesentlich einfacher als die bekannte Analyse von Hinrichs und Nakazawa (1994b) ist. Der vorgestellte Ansatz ist auch kompatibel zu einer Passivanalyse, die keine lexikalischen Regeln zur Beschreibung des Passivs benutzen (Kapitel 15), was bei dem Ansatz von Hinrichs und Nakazawa nicht der Fall ist.

In den Kapiteln, die sich mit der Wortstellung beschäftigen, wird oft von diskontinuierlichen Konstituenten ausgegangen. Dies wird in Kapitel 21 diskutiert. Das Parsen mit diskontinuierlichen Konstituenten hat exponentiale Komplexität. In ebendiesem Kapitel wird untersucht, wie der *worst case* aussieht und ob er in der existierenden Grammatik, die durch linguistisch motivierte Bedingungen auch beim Parsen beschränkt wird, auch exponential ist. In einer ausführlichen Untersuchung werden Alternativen zu diskontinuierlichen Konstituenten beleuchtet. Eine Alternative besteht in der Aufgabe der Ordnung der Subcat-Liste.

Diese Ordnung wird aber zum Beispiel zur Zuweisung von Kasus im Deutschen gebraucht. Nominalphrasen-Argumente können in verschiedenen syntaktischen Umgebungen unter Umständen verschiedene Kasus zugewiesen bekommen. Wie man die Festlegung dieser unterschiedlichen Kasus im Lexikon und damit Redundanz vermeiden kann, wird in diesem Kapitel gezeigt. Ein Schwerpunkt des Kapitels 15 ist das deutsche Passiv. Die Analyse von Pollard (1994) (Objekt-zu-Subjekt-Anhebung) wird erklärt und geringfügig modifiziert. Zusätzlich werden lexikalische Regeln für die Erzeugung adjektivischer Partizipien angegeben. Auch Passiv-Paraphrasen und Fernpassiv werden behandelt. Im Kapitel 21 wird dann gezeigt, wie man auch ohne die Ordnung der Subcat-Liste die korrekte Kasuszuweisung sicherstellen kann. Dazu wird die Verwendung einer dynamisch erzeugten ARG-S-Liste vorgeschlagen, die auch einen Großteil der in Kapitel 9 erörterten Probleme mit Komplementextraktionsregeln löst.

Kongruenz wird in Kapitel 16 beschrieben. Es wird gezeigt, wie man Determinator-Adjektiv-Nomen- und Subjekt-Verb-Kongruenz im Lexikon beschreiben kann.

Kapitel 17 beschäftigt sich mit Kontrollstrukturen. Es wird auf Kontroll- und Anhebungsverben sowie kurz auf AcI-Konstruktionen eingegangen. Die Skopuseigenschaften von Adverbien werden erörtert. Und, wie schon erwähnt, spielt die Wortstellung auch im Zusammenhang mit Infinitivkomplementen eine große Rolle. Intraposition und Extraposition von Infinitivverbphrasen wird zusammen mit den entsprechenden Skopusphänomenen richtig beschrieben. Obwohl wesentliche Erkenntnisse aus Kiss' (1993) Arbeit in dieses Kapitel eingegangen sind, unterscheidet sich die vorgestellte Analyse von der seinigen nicht unerheblich.

Kapitel 20 schließlich ist der HPSG-Bindungstheorie gewidmet. Es werden die Probleme der Standard-HPSG-Bindungstheorie nach Pollard und Sag (1994, Kapitel 6) aufgezeigt.



# 1. Syntaktische Merkmale und Kategorien

In diesem Kapitel wird der generelle Aufbau von Merkmalstrukturen zur Beschreibung linguistischer Objekte vorgestellt. Der Kopfbegriff und das Kopfmerkmalsprinzip werden erklärt. Weiterhin wird gezeigt, wie die Valenz lexikalischer Köpfe repräsentiert wird.

## 1.1. Konstituenten und Kategorien

In natürlichen Sprachen gibt es lexikalische Einheiten, die sich bestimmten Kriterien folgend zu Phrasen kombinieren lassen. Ein Teil dieser Kriterien läßt sich mit syntaktischen Mitteln beschreiben. Lexikalische und phrasale Zeichen gehören bestimmten syntaktischen Kategorien an. Ist der Typ – also die syntaktische Kategorie einer Wortgruppe – nicht bekannt oder nicht wesentlich, spricht man von Konstituenten. Von Grewendorf (1988, Kapitel 3) werden u. a. folgende Kriterien dafür angegeben, daß eine bestimmte Wortgruppe eine Konstituente bildet:

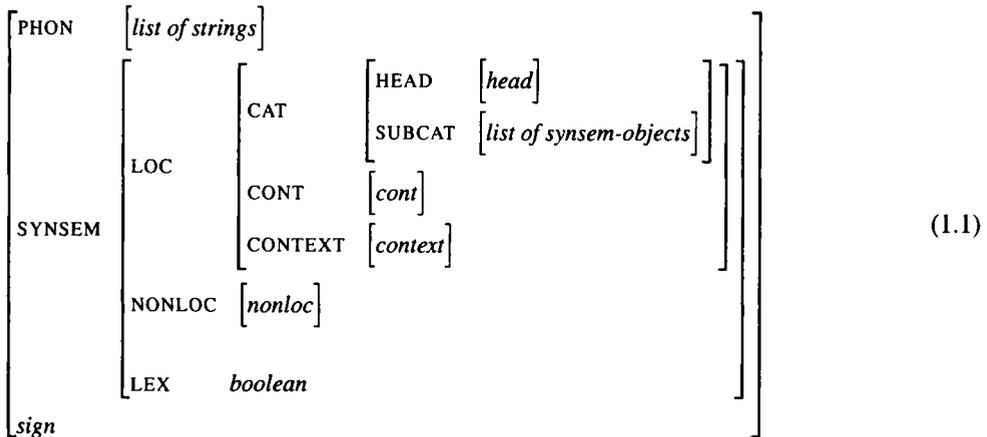
- *Substituierbarkeit* Kann man eine Wortgruppe einer bestimmten Kategorie in einem Satz gegen eine andere Wortgruppe so austauschen, daß wieder ein akzeptabler Satz entsteht, so ist das ein Indiz dafür, daß die beiden Wortgruppen Konstituenten bilden.
- *Permutierbarkeit* Wortfolgen, die man ohne Beeinträchtigung der Korrektheit des Satzes verschieben bzw. umstellen kann, bilden eine Konstituente.
- *Pronominalisierungstest* Alles, worauf man sich mit einem Pronomen beziehen kann, ist eine Konstituente.
- *Koordinationstest* Was sich koordinieren läßt, ist eine Konstituente.
- *Fragetest* Was sich erfragen läßt, ist eine Konstituente.

Die Tests liefern nicht immer konsistente Ergebnisse. Sie sind also nur als Indizien für ein Konstituentensein anzusehen.

Den Konstituenten ordnet man bestimmte Eigenschaften zu. Auf welche Weise die Eigenschaften repräsentiert werden können und wie diese Information strukturiert wird, ist Gegenstand des nächsten Abschnitts.

## 1.2. Aufbau von Merkmalstrukturen zur Beschreibung linguistischer Objekte

In der HPSG wird die Menge der Merkmale, die einem linguistischen Objekt zugeordnet werden, strukturiert. Mengen von Merkmalen, die denselben Teilaspekt eines Objekts beschreiben, werden gemeinsam repräsentiert. Die folgende Merkmalstruktur zeigt die Struktur eines Zeichens:<sup>1</sup>



Der Wert von PHONOLOGY ist eine Liste phonologischer Formen von Zeichen, also z. B. /kapita'lismʊs/ für *Kapitalismus*. Da das aber schwierig zu lesen wäre, schreibt man statt dessen *<Kapitalismus>* als PHON-Wert von *Kapitalismus*.

Unter SYNSEM steht syntaktische und semantische Information. Das Merkmal LOC hat seinen Namen, weil es sich bei Informationen unter diesem Pfad um lokale syntaktische oder semantische Informationen handelt. Im Gegensatz dazu gibt es auch nichtlokale Informationen. Diese stehen unter dem Pfad SYNSEM|NONLOC. Auf den Wert des Pfades SYNSEM|NONLOC wird in den Kapiteln 9–13 über nichtlokale Abhängigkeiten, Relativsätze und Extraposition eingegangen. Zu lokaler Information wird Information über die syntaktische Kategorie eines Zeichens (CATEGORY), Information über den semantischen Inhalt des Zeichens (CONTENT) und Information über den Kontext (CONTEXT) – also Hintergrundinformation und dergleichen – gezählt. Die Merkmale CONTENT und CONTEXT sind semantische Merkmale. Auf diese Merkmale wird in Kapitel 3 genauer eingegangen. Zu den Merkmalen, die in der Merkmalstruktur unter dem Pfad SYNSEM|LOC|CAT stehen, gehören HEAD und SUBCAT.

Das Merkmal HEAD hat als Wert eine Merkmalstruktur, die die syntaktischen Eigenschaften spezifiziert, die ein lexikalisches Zeichen mit seinen Projektionen teilt, d. h. mit den phrasalen Zeichen, deren Kopf (*head*) das entsprechende lexikalische Zeichen ist.

Das SUBCAT-Merkmal gibt Auskunft über die Valenz eines Zeichens. Sein Wert ist eine Liste, die eventuell nur teilweise spezifizierte Elemente enthält, mit denen das Zeichen kombiniert werden muß, damit eine gesättigte, d. h. in gewissem Sinne vollständige Phrase, entsteht.

Das LEX-Merkmal ist ein binäres Merkmal, das dazu dient, zwischen lexikalischen (+) Einheiten und Zeichen, die durch Kombination aus solchen entstanden sind (–), zu unter-

<sup>1</sup> Zum Begriff des sprachlichen Zeichens siehe (Saussure, 1915).

scheiden. Ich werde im Kapitel 14 im Zusammenhang mit der Analyse von Hilfs- und Modalverben noch genauer darauf eingehen.

Bei phrasalen Zeichen gibt es ein DAUGHTERS-Merkmal. Dessen Wert ist eine Merkmalstruktur, die Informationen über die Konstituenten (Töchter) enthält, aus denen sich das Zeichen zusammensetzt. Töchter der Präpositionalphrase (1.2) sind z. B. *auf* und die Nominalphrase *dem Baum*.

(1.2) auf dem Baum

Die Phrase *dem Baum* hat die Töchter *dem* und *Baum*.

Man bezeichnet *dem Baum* als Komplement der Präposition *auf*. Die Information über Komplementtöchter wird unter DTRS|COMP-DTRS repräsentiert. Der Wert von COMP-DTRS ist eine Liste von Zeichen. DTRS hat als Wert eine Merkmalstruktur vom Typ *constituent-structure*. *constituent-structure* wiederum hat die beiden Subtypen *non-headed-structure* und *headed-structure*. Ein Beispiel für Strukturen vom Typ *non-headed-structure* ist die *relativizer-structure* (siehe Kapitel 10.3.2). (1.3) ist eine partielle Beschreibung eines phrasalen Zeichens.

$$\left[ \begin{array}{l} \text{DTRS} \\ \text{COMP-DTRS} \\ \text{headed-structure} \\ \text{phrasal-sign} \end{array} \left[ \begin{array}{l} \text{HEAD-DTR} \\ \text{list-of-signs} \end{array} \right] \right] \quad (1.3)$$

Die Subtypen des Typs *headed-structure* zeigt Abbildung 1.1.

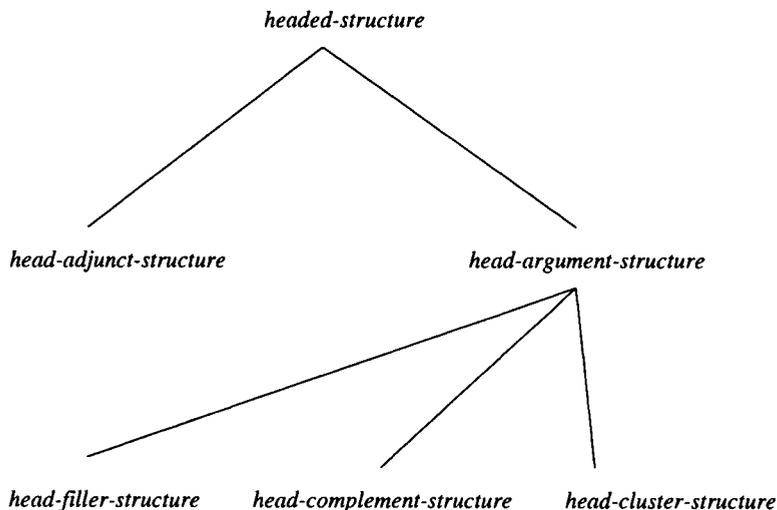


Abbildung 1.1.: Subtypen des Typs *headed-structure*

Vorerst interessiert nur der Typ *head-complement-structure*. Er besitzt genau dieselben Merkmale wie sein Supertyp *headed-structure*, besitzt jedoch im Unterschied zu den anderen Subtypen von *headed-structure* keine weiteren Merkmale.

### 1.3. Kopfmerkmale

Die Bezeichnung Kopf einer Phrase spielt in vielen syntaktischen Theorien eine Rolle (z. B. GB und GPSG). Eine Phrase enthält normalerweise ein bestimmtes Wort, das für diese Phrase von zentraler Bedeutung ist. Es bestimmt viele der syntaktischen Eigenschaften der gesamten Phrase und wird lexikalischer Kopf der Phrase (*lexical head*) genannt. Der lexikalische Kopf einer Verbphrase ist das Verb, der einer Präpositionalphrase die Präposition und der einer Nominalphrase das Nomen. Im Beispiel (1.2) ist *auf* die Kopftochter.

Die folgende Merkmalstruktur ist ein Ausschnitt der Merkmalstruktur, die die Präpositionalphrase in (1.2) repräsentiert.

$$\left[ \begin{array}{l} \text{PHON } \langle \textit{auf dem Baum} \rangle \\ \\ \text{DTRS} \left[ \begin{array}{l} \text{HEAD-DTR} \left[ \begin{array}{l} \text{PHON } \langle \textit{auf} \rangle \\ \textit{lexical-sign} \end{array} \right] \\ \\ \text{COMP-DTRS} \left\langle \left[ \begin{array}{l} \text{PHON } \langle \textit{dem Baum} \rangle \\ \text{DTRS } \dots \\ \textit{phrasal-sign} \end{array} \right] \right\rangle \\ \\ \textit{head-complement-structure} \end{array} \right] \\ \textit{phrasal-sign} \end{array} \right] \quad (1.4)$$

Allgemein gilt für lexikalische Köpfe und deren Projektionen das folgende Prinzip:

#### Prinzip 1 (Kopfmerkmalprinzip (*Head Feature Principle*))

$$\left[ \text{DTRS} \left[ \textit{headed-structure} \right] \right] \Rightarrow$$

$$\left[ \begin{array}{l} \text{SYNSEM|LOC|CAT|HEAD} \quad \boxed{1} \\ \text{DTRS|HEAD-DTR|SYNSEM|LOC|CAT|HEAD} \quad \boxed{1} \end{array} \right]$$

Das bedeutet: Wenn ein Zeichen eine Kopftochter hat, teilt es die Kopfmerkmale mit ihr. Die Strukturteilung wird durch die  $\boxed{1}$  repräsentiert. Prinzipien drücken Bedingungen aus. Wenn eine Merkmalstruktur mit der linken Seite der Implikation unifiziert, so muß sie auch mit der rechten Seite der Implikation unifizieren.

Mit instantiierten Kopfmerkmalen sieht die Merkmalstruktur (1.4) wie folgt aus:

$$\left[ \begin{array}{l}
 \text{PHON} \quad \langle \text{auf dem Baum} \rangle \\
 \text{SYNSEM} \quad \left[ \text{LOC|CAT|HEAD} \quad \boxed{1} \right] \\
 \text{DTRS} \quad \left[ \begin{array}{l}
 \text{HEAD-DTR} \quad \left[ \begin{array}{l}
 \text{PHON} \quad \langle \text{auf} \rangle \\
 \text{SYNSEM} \quad \left[ \text{LOC|CAT|HEAD} \quad \boxed{1} \quad [\text{prep}] \right] \\
 \text{lexical-sign}
 \end{array} \right] \\
 \text{COMP-DTRS} \quad \left\langle \left[ \begin{array}{l}
 \text{PHON} \quad \langle \text{dem Baum} \rangle \\
 \text{DTRS} \quad \dots \\
 \text{phrasal-sign}
 \end{array} \right] \right\rangle \\
 \text{head-complement-structure}
 \end{array} \right] \\
 \text{phrasal-sign}
 \end{array} \right] \quad (1.5)$$

Das *Head Feature Principle* entspricht dem der GPSG, und auch die grammatischen Kategorien ähneln denen der GPSG. Zu jeder Kategorie gehört ein Typ: *noun*, *verb*, *adj*, *prep*, *det*, *adv*, *relativizer*. Diese Typen sind Subtypen des Typs *head*.

Eine Merkmalstruktur vom Typ *noun* hat zusätzlich zu den Merkmalen, die sie von ihrem Supertyp erbt, das Merkmal CASE für den Kasus des Nomens bzw. der Nominalphrase.

$$\left[ \begin{array}{l}
 \text{CAS} \quad \text{case} \\
 \text{noun}
 \end{array} \right] \quad (1.6)$$

Eine Merkmalstruktur vom Typ *verb* hat zusätzlich zu den Merkmalen vom Typ *head* folgende Merkmale:

$$\left[ \begin{array}{l}
 \text{VFORM} \quad \text{vform} \\
 \text{INITIAL} \quad \text{boolean} \\
 \text{AUX} \quad \text{boolean} \\
 \text{SUBJ} \quad \left[ \text{list of synsem-objects} \right] \\
 \text{verb}
 \end{array} \right] \quad (1.7)$$

*verb* hat die beiden Subtypen *fin-verb* und *non-fin-verb*.

$$\left[ \begin{array}{l}
 \text{VFORM} \quad \text{fin} \\
 \text{INITIAL} \quad \text{boolean} \\
 \text{AUX} \quad \text{boolean} \\
 \text{SUBJ} \quad \langle \rangle \\
 \text{fin-verb}
 \end{array} \right] \quad \left[ \begin{array}{l}
 \text{VFORM} \quad \text{non-fin} \\
 \text{INITIAL} \quad \text{boolean} \\
 \text{AUX} \quad \text{boolean} \\
 \text{SUBJ} \quad \left[ \text{list of synsem-objects} \right] \\
 \text{AUXF} \quad \text{auxf} \\
 \text{FLIP} \quad \text{boolean} \\
 \text{ERG} \quad \left[ \text{list of synsem-objects} \right] \\
 \text{non-fin-verb}
 \end{array} \right] \quad (1.8)$$

VFORM kann einen der folgenden Werte haben:

<i>fin</i>	gebeugte Verbformen
<i>bse</i>	die Infinitivform
<i>inf</i>	die Infinitivform mit zu
<i>ppp</i>	die Partizipform, die sowohl für Perfekt- als auch für Passivkonstruktionen benutzt wird

Die verschiedenen Werte von VFORM spielen bei der Auswahl von Komplementen eine Rolle. *Glauben* kann z. B. eine *inf*-Verbphrase verlangen (*Er glaubte zu träumen.*). *Sehen* dagegen verlangt ein verbales Komplement in der *bse*-Form (*Er sah den Verbrecher weglaufen.*). AUX ist ein binäres Merkmal, das bei Hilfsverben (*Auxiliaries*) den Wert + und bei anderen den Wert – hat. AUXF kann die Werte *haben* und *sein* haben, je nachdem welches Hilfsverb zu dem entsprechenden Verb gehört (siehe Kapitel 15.4). Nur bei Verben mit VFORM *ppp* hat dieses Merkmal einen spezifizierten Wert, was die weitere Untergliederung des Typs *non-fin-verb* rechtfertigen würde. Der Übersichtlichkeit halber verzichte ich hier darauf. AUXF hat bei Verbformen mit von *ppp* verschiedenem VFORM-Wert den Wert *none* und wird nicht aufgeführt. INITIAL ist ein Merkmal, das bei Verbinitialstellung den Wert + und sonst den Wert – hat (siehe Kapitel 11). Das Merkmal FLIP wird ebenfalls zur Beschreibung der Wortstellung benötigt. Es wird in Kapitel 14 erklärt. Das Merkmal SUBJ enthält bei infiniten Verben, also bei Verben mit einem von *fin* verschiedenen VFORM-Wert, das Subjekt des Verbs, so es eins hat. Bei finiten Verben und bei subjektlosen ist der Wert des SUBJ-Merkmals die leere Liste. Im Abschnitt 1.7 wird darauf etwas genauer eingegangen. Das ERG-Merkmal wird zur Beschreibung des Passivs gebraucht. Es wird im Kapitel 15.2 erklärt.

Präpositionen haben die Kopfmerkmale PFORM und CASE. Der Wert von PFORM entspricht im allgemeinen der orthographischen Form der Präposition.<sup>2</sup> Der Wert des Kasusmerkmals muß bei der Bildung von Präpositionalphrasen mit dem der Nominalphrase übereinstimmen.

- (1.9) a. auf dem Baum  
b. auf den Installateur

Verben spezifizieren den Kasus und den PFORM-Wert ihrer präpositionalen Komplemente:

- (1.10) a. Er wartet auf den Installateur.  
b. # Er wartet auf dem Installateur.<sup>3</sup>

<sup>2</sup> Eine Ausnahme bilden z. B. Präpositionen wie *vom*, bei denen eine Verschmelzung mit einem Artikel stattgefunden hat. Der PFORM-Wert von *vom* ist *von*.

<sup>3</sup> Das # bedeutet, daß der Satz zwar nicht ungrammatisch ist, aber nicht die diskutierte Konstruktion vorliegt bzw. daß der Satz eine andere Lesart hat.

(1.10b) hat natürlich eine Lesart, in der der Satz grammatisch ist. In dieser Lesart ist *auf dem Installateur* jedoch nicht Komplement des Verbs, sondern ein Modifikator (zur Modifikation siehe Kapitel 4).

$$\left[ \begin{array}{ll} \text{PFORM} & \text{pform} \\ \text{CAS} & \text{case} \\ \text{PRD} & \text{boolean} \\ & \text{prep} \end{array} \right] \quad (1.11)$$

Das PRD-Merkmal hat für prädikativ verwendete Präpositionalphrasen (1.12) den Wert + und sonst den Wert –.

(1.12) Die Kanne ist in der Küche.

Für Merkmalstrukturen vom Typ *adj* gibt es ebenfalls das PRD-Merkmal:

$$\left[ \begin{array}{ll} \text{PRD} & \text{boolean} \\ & \text{adj} \end{array} \right] \quad (1.13)$$

Prädikative Adjektive werden in Kopulakonstruktionen wie (1.14) benutzt.

(1.14) Maria ist schön.

Sie haben keine Endung und auch keine Kasus-, Genus- und Numerusmerkmale. Der Wert ihres PRD-Merkmals ist +. Nichtprädikative Adjektive sind Adjektive, die Nomina modifizieren können. Für nicht prädikativ gebrauchte Adjektive (*die schöne Frau*) ist PRD–.

Das Kopfmerkmal von Determinatoren hat folgende Struktur:

$$\left[ \begin{array}{ll} \text{DTYPE} & \text{dtype} \\ \text{NUM} & \text{num} \\ \text{CAS} & \text{case} \\ \text{GEN} & \text{gen} \\ & \text{det} \end{array} \right] \quad (1.15)$$

Dabei ist der Wert von DTYPE die Determinatorenklasse, zu der der entsprechende Determinator gehört.<sup>4</sup> *Das* und *ein* gehören unterschiedlichen Determinator Klassen an, was man daran erkennen kann, daß sie mit verschieden flektierten Adjektiven auftreten: *ein kluges Mädchen, das kluge Mädchen*.

Es gibt eine kleine Gruppe deutscher Nomina, die Determinatoren eines bestimmten Typs selektieren. Beispiele sind *Beamter, Bekannter, Gesandter, Verwandter* (*ein Beamter, aber der Beamte*).

Bei Nomina und Adjektiven sind Numerus, Person und Genus und DTYPE keine Kopfmerkmale. Wie die Übereinstimmung dieser Merkmale innerhalb von Phrasen trotzdem erreicht wird, wird im Kapitel 16 über Kongruenz erklärt.

#### 1.4. Das Subkategorisierungsmerkmal (SUBCAT)

Die Subkategorisierung oder Valenz eines lexikalischen oder phrasalen Zeichens wird durch eine Liste beschrieben. Diese Liste enthält Spezifikationen syntaktischer und semantischer

<sup>4</sup> Vergleiche (Pollard und Sag, 1994, S. 372)

Eigenschaften anderer Zeichen, mit denen das betreffende Zeichen kombiniert werden muß, um eine vollständige Phrase zu bilden. Zum Beispiel gibt es verschiedene Arten von Verben (intransitive, transitive und ditransitive), die ein, zwei oder drei Komplemente verlangen, um eine gesättigte Verbphrase zu bilden.

intransitiv	Karl schläft.	Schläft Karl?
transitiv	Karl liebt Anna.	Liebt Karl Anna?
ditransitiv	Karl gibt Anna ein Buch.	Gibt Karl Anna ein Buch?

Eine gesättigte Verbphrase mit VFORM *fin*, deren Bestandteile bestimmten Ordnungsprinzipien folgen, ist ein Satz.

Nomina wie *Buch* verlangen einen Determinator als Komplement. Eigennamen hingegen sind normalerweise gesättigt.

Die Liste der Komplemente eines Zeichens steht unter dem Pfad SYNSEM|LOC|CAT|SUBCAT. In der sogenannten Subcat-Liste befinden sich Merkmalstrukturen vom Typ *synsem*. Diese Merkmalstrukturen sind meist nur teilweise spezifiziert. Bei gesättigten Zeichen ist der Wert des Merkmals SUBCAT die leere Liste.

Der Kopf einer Phrase wird auch Regens, seine Komplemente Regime genannt. Der Kopf regiert seine Komplemente.

(1.16) Die Frau liebt den Mann.

In (1.16) regiert *liebt* die Nominalphrasen *die Frau* und *den Mann*.

## 1.5. Abkürzungen

Mit den jetzt zur Verfügung stehenden Mitteln kann man bereits Teilspezifikationen der traditionell benutzten Grammatiksymbole angeben:

Symbol	Teilspezifikation	Beispiel
DET	$\left[ \text{LOC CAT} \left[ \begin{array}{l} \text{HEAD} \quad [det] \\ \text{SUBCAT} \quad \langle \rangle \end{array} \right] \right]$	jeder ein der
$\bar{N}$	$\left[ \text{LOC CAT} \left[ \begin{array}{l} \text{HEAD} \quad [noun] \\ \text{SUBCAT} \quad \langle \text{DET} \rangle \end{array} \right] \right]$	Frau schöne Frau
NP	$\left[ \text{LOC CAT} \left[ \begin{array}{l} \text{HEAD} \quad [noun] \\ \text{SUBCAT} \quad \langle \rangle \end{array} \right] \right]$	Karl eine Frau eine schöne Frau

Symbol	Teilspezifikation	Beispiel
VP	$\left[ \text{LOC CAT} \left[ \text{HEAD} \left[ \text{SUBJ} \langle [] \rangle \right] \right] \right]$ $\left[ \text{SUBCAT} \langle \rangle \right]$	den Mantel umzufärben Maria zu lieben einen Hund füttern
S	$\left[ \text{LOC CAT} \left[ \text{HEAD} \left[ \text{VFORM} \textit{fin} \right] \right] \right]$ $\left[ \text{SUBCAT} \langle \rangle \right]$	Karl schläft. Schlief Karl?

Diese Symbole werden im weiteren auch als Abkürzungen innerhalb von Merkmalstrukturen benutzt. Hinzu kommen komplexe Symbole für stärker spezifizierte Phrasen:

Symbol	Teilspezifikation	Beispiel
NP[ <i>nom</i> ]	$\left[ \text{LOC CAT} \left[ \text{HEAD} \left[ \text{CAS} \textit{nom} \right] \right] \right]$ $\left[ \text{SUBCAT} \langle \rangle \right]$	der Mann du
VP[ <i>inf</i> ]	$\left[ \text{LOC CAT} \left[ \text{HEAD} \left[ \text{VFORM} \textit{inf} \right] \right] \right]$ $\left[ \text{SUBCAT} \langle \rangle \right]$	den Mantel umzufärben

Je nach Kontext werden diese Abkürzungen für Zeichen oder *synsem*-Objekte verwendet. Wenn von Zeichen die Rede ist, muß man den Pfad LOC|CAT durch SYNSEM|LOC|CAT ersetzen.

Mit diesen Abkürzungen kann man jetzt folgende Beispiele für Werte der Subcat-Liste angeben:

schläft  $\langle \text{NP}[\textit{nom}] \rangle$   
 zwingt  $\langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}], \text{VP}[\textit{inf}] \rangle$   
 gab  $\langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}], \text{NP}[\textit{dat}] \rangle$   
 auf  $\langle \text{NP}[\textit{dat}] \rangle$

Es sei darauf hingewiesen, daß die Reihenfolge der Elemente der Subcat-Liste nichts mit der Oberflächenreihenfolge der Komplemente, d. h. mit der Reihenfolge, in der die Komplemente wirklich im Satz vorkommen, zu tun hat. Die Elemente der Subcat-Liste sind nach einer Obliqueness-Hierarchie geordnet, wie sie in (Pollard und Sag, 1987, S. 120) angegeben ist

(siehe auch Kapitel 20, S. 406). Wozu man eine so geordnete Subcat-Liste braucht, wird im Kapitel 20 über die Bindungstheorie der HPSG erklärt.

## 1.6. Das Subkategorisierungsprinzip

Die Sättigung von Komplementen wird in der HPSG durch das Subkategorisierungsprinzip (*Subcategorization Principle* oder *Subcat Principle*) beschrieben:<sup>5,6</sup>

### Prinzip 2 (Subkategorisierungsprinzip (*Subcat Principle*))

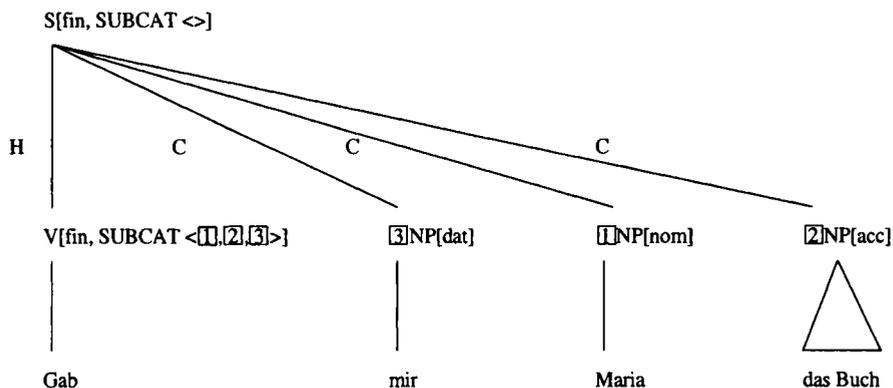
*In einem phrasalen Zeichen, dessen DTRS-Wert vom Typ headed-structure ist, ist der SUBCAT-Wert der Kopftochter die Verknüpfung der Subcat-Liste des phrasalen Zeichens mit der Liste der SYNSEM-Werte der Elemente in der COMP-DTRS-Liste.*

Abbildung 1.2 zeigt das Wirken des Subcat-Prinzips an Hand der Frage *Gab mir Maria das Buch?*. Ein H an einer Kante zeigt an, daß das Symbol am unteren Ende der Kante eine

- <sup>5</sup> Manche HPSG-Theoretiker (Vergleiche (Gunji, 1986; Hinrichs und Nakazawa, 1989b; Pollard, 1990; Engelkamp, Erbach und Uszkoreit, 1992; Nerbonne, 1994)) gehen davon aus, daß der Wert des SUBCAT-Merkmals eine Menge ist. Ein solcher Ansatz ist nur möglich, wenn gleichzeitig das Subjekt als Wert eines gesonderten SUBJ-Merkmals aufgeführt wird, da man sonst Kontrollkonstruktionen (siehe Kapitel 17) nicht beschreiben kann. Bei Kontrollkonstruktionen wird ein *structure sharing* zwischen dem Index des Subjekts bzw. Objekts des Matrixsatzes und dem nichtausgedrückten Subjekt der eingebetteten Verbphrase hergestellt. Das wäre nicht möglich, wenn die Subcat-Liste eine Menge und das Subjekt nicht explizit kenntlich gemacht wäre. Auch die Unterscheidung von strukturellem und lexikalischem Kasus (siehe Kapitel 15) benötigt eine geordnete Liste. Die Bindungstheorie (siehe Kapitel 20) macht ebenfalls von einer geordneten Liste Gebrauch. Wenn man eine Subcat-Menge annimmt, braucht man gleichzeitig eine andere Liste, die bestimmte Funktionen der Subcat-Liste erfüllt. In Kapitel 21.1.2.1 werde ich darauf noch genauer eingehen.
- <sup>6</sup> In der 87'er Version der HPSG war die Subcat-Liste eine Liste von Objekten vom Typ *sign*. Das Subcat-Prinzip konnte eleganter mit der folgenden Implikation beschrieben werden.

$$\left[ \text{DTRS} \left[ \text{headed-structure} \right] \right] \Rightarrow \left[ \begin{array}{l} \text{SYNSEM} \left[ \text{LOC|CAT|SUBCAT} \boxed{1} \right] \\ \text{DTRS} \left[ \begin{array}{l} \text{HEAD-DTR} \left[ \text{SYNSEM|LOC|CAT|SUBCAT} \boxed{1} \oplus \boxed{2} \right] \\ \text{COMP-DTRS} \boxed{2} \end{array} \right] \end{array} \right]$$

Das  $\oplus$  steht dabei für die Listenverketzung (*append*). Durch *structure sharing* wird erreicht, daß die Tochterzeichen mit den Teilspezifikationen in der Subcat-Liste unifizieren müssen. Die in der Formulierung des Subcat-Prinzips durch  $\boxed{2}$  gekennzeichnete Liste der Komplementtöchter entspricht in dem Beispiel in Abbildung 1.2 der gesamten Subcat-Liste von *gab*, d. h. alle Komplemente werden auf einmal gesättigt. Die  $\boxed{1}$  im Subcat-Prinzip entspricht dann der leeren Liste.

Abbildung 1.2.: *Gab mir Maria das Buch?*

Kopftochter ist. Ein C dagegen steht für Komplement. *Maria*, *das Buch* und *mir* sind Komplemente des Verbs *gab*. Teilspezifikationen dieser Komplemente stehen in der Subcat-Liste des Lexikoneintrags von *gab*. Alle Komplemente von *gab* werden auf einmal gesättigt. Die Subcat-Liste des entstehenden Zeichens *S[fin]* ist leer.

## 1.7. Das SUBJ-Merkmal

Zu den Subjekten zähle ich Nominalphrasen im Nominativ, wenn sie nicht prädikativ sind und auch Subjektsätze.

- (1.17) a. Die Frau kennt den Mann.  
 b. Karl ist ein Genie.  
 c. Daß Karl immer zu spät kommt, ärgert ihn.

In (1.17a) ist *die Frau* das Subjekt, in (1.17b) *Karl* und in (1.17c) der *daß*-Satz. Der Begriff „Subjekt“ ist umstritten. Reis (1982) behauptet, daß der Subjektbegriff für das Deutsche leer ist und daß man Subjekt auf Nominativ reduzieren kann. Eisenberg (1994, Kapitel 8) zeigt jedoch, daß es durchaus sinnvoll ist, auch Sätze zu Subjekten zu zählen. Genauerer zu den Subjektsätzen findet sich in Kapitel 15.3.1. Zu Subjekten und Subjektsätzen siehe auch Oppenrieder (1991).

Es gibt seit Jahren immer wieder Veröffentlichungen, in denen diskutiert wird, ob es im Deutschen eine VP gibt oder nicht (vergleiche (Haider, 1982; Webelhuth, 1985; Sternefeld, 1985b; Fanselow, 1987; Grewendorf, 1988; Webelhuth, 1990; Oppenrieder, 1991; Haider, 1993)). In Abbildung 1.2 werden alle Komplemente des Verbs auf einmal gesättigt. Es gibt also keine Zwischenstufe VP wie z. B. im Englischen (vergleiche (Pollard und Sag, 1987,

Kapitel 6.1), (Pollard und Sag, 1994, Kapitel 1.4)).<sup>7</sup> Sowohl Subjekt als auch Objekte des finiten Verbs sind Elemente der Subcat-Liste. Das heißt, es gibt keine strukturelle Unterscheidung zwischen Subjekt und Objekten. Andererseits gibt es aber infinite Verbphrasen wie z. B. *Maria zu lieben* in *Karl behauptet Maria zu lieben*. Bei infiniten Verbphrasen gibt es sehr wohl strukturelle Unterschiede zwischen Subjekt und Objekt. Genauer dazu findet man in den Kapiteln 15 und 17. Ein weiterer Vorteil der separaten Auflistung des Subjekts infinitiver Verben liegt darin, daß sowohl Verbphrasen wie die oben aufgeführten als auch finite Sätze Maximalprojektionen sind und somit ihr ähnliches syntaktisches Verhalten erklärt werden kann.

Borsley (1987) plädierte schon 1987 dafür, das Subjekt nicht in die Subcat-Liste aufzunehmen. Kiss (1993) dagegen schlug vor, nur das Subjekt infinitiver Verben in einer Liste als Kopfmerkmal zu repräsentieren. Nur in finiten Sätzen wird das Subjekt gesättigt. Deshalb ist das Subjekt auch nur bei finiten Verbformen ein Element der Subcat-Liste. Bei nichtfiniten Verbformen dagegen ist das Subjekt nicht Element der Subcat-Liste. Statt dessen gibt es ein Kopfmerkmal SUBJ, das als Wert eine maximal einelementige Liste hat. Diese Liste enthält das Subjekt bei Verben, die ein Subjekt haben, und ist leer bei Verben, die kein Subjekt haben (z. B. *grauen* in *Mir graut*).

Das Verb *lieben* hat dann in der *bse*-Form den folgenden CAT-Wert:

*lieben*:

$$\left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \text{VFORM } bse \\ \text{SUBJ } \langle \text{NP}[nom] \rangle \\ \text{verb} \end{array} \right] \\ \text{SUBCAT } \langle \text{NP}[acc] \rangle \\ \text{cat} \end{array} \right] \quad (1.18)$$

Durch das Kopfmerkmalsprinzip wird sichergestellt, daß die Information über das Subjekt in der Phrase (1.19) noch im SYNSEM-Wert – nämlich als Kopfmerkmal – enthalten ist.

(1.19) das Buch zu lesen

Das ist wichtig für Kontrollkonstruktionen (siehe Kapitel 17).

## Literatur

Auf die formalen Grundlagen von HPSG-Grammatiken kann in diesem Buch nicht eingegangen werden. Es sei darum hier auf die folgende Literatur verwiesen: Im Buch von Stuart Shieber (1986) findet man eine allgemeine Einführung in die Theorie der Unifikationsgrammatiken. Es wird ein ziemlich allgemeiner Überblick gegeben, und wichtige Grammatik-Typen (DCG, LFG, GPSG, HPSG, PATR-II) werden kurz vorgestellt. Johnson (1988) beschreibt

<sup>7</sup> Hierbei ist unwesentlich, ob man flache oder binär verzweigende Strukturen annimmt (siehe Kapitel 2). Wesentlich ist, ob das Subjekt durch ein eigenes Schema gesättigt wird.

den Formalismus ungetypter Merkmalstrukturen. Es werden Kriterien angegeben, die eine Attribut-Wert-Grammatik erfüllen muß, damit das Halteproblem für Parser entscheidbar ist. Carpenter (1992) setzt sich mit getypten Merkmalstrukturen auseinander. Die Logik, die Paul King (1994) für HPSG-Grammatiken entwickelt hat, gilt derzeit als Standard. Die Zusammensetzung der HPSG aus linguistischer Sicht ist in (Pollard und Sag, 1994, Kapitel 1.10) beschrieben.

Grewendorf (1988, Kapitel 3) motiviert, warum man Wortgruppen Kategorien zuordnet, und es werden Tests angegeben, die Indizien für solche Zuordnungen liefern.

Subkategorisierung ist sehr ausführlich in (Pollard und Sag, 1987, Kapitel 3) beschrieben. In (Pollard und Sag, 1994, Kapitel 1.4) wird nur kurz auf Subkategorisierung eingegangen. Eine gute Einführung in die Subkategorisierung findet man in (Eisenberg, 1994, Kapitel 3).

## 2. Dominanzschemata

Der Teil von Grammatiken, der unabhängig von Einzelsprachen ist, wird Universalgrammatik genannt. In der HPSG gehört auch das *Immediate Dominance Principle* (IDP) zur Universalgrammatik. Das IDP ist eine Disjunktion sogenannter Dominanzschemata. Diese Schemata sind analog zu den Grammatikregeln, *Immediate Dominance Rules* (ID-Regeln) oder  $\bar{X}$ -Schemata anderer Theorien.<sup>1</sup> Sie dienen als Beschreibung lokaler Phrasenstrukturbäume. In den letzten Jahren gab es in vielen syntaktischen Theorien einen Trend zur Lexikalisierung der Grammatik. Dabei wurde von konstruktsspezifischen Regeln zu abstrakten Dominanzschemata übergegangen. Beispiele sind die folgenden  $\bar{X}$ -Regeln:

$\bar{X}$ – Regel	mit Kategorien	
$\bar{X} \rightarrow \overline{\text{specifier } \bar{X}}$	$\bar{N} \rightarrow \overline{\text{DET } \bar{N}}$	(2.1)
$\bar{X} \rightarrow \bar{X} \overline{\text{modifier}}$	$\bar{N} \rightarrow \bar{N} \overline{\text{REL\_SATZ}}$	
$\bar{X} \rightarrow \overline{\text{modifier } \bar{X}}$	$\bar{N} \rightarrow \overline{\text{ADJ } \bar{N}}$	
$\bar{X} \rightarrow X \overline{\text{complement}}$	$\bar{N} \rightarrow N \overline{\bar{P}}$	

Eine Bedingung für die Wohlgeformtheit einer Phrase ist, daß sie einem der für eine Sprache angegebenen Dominanzschemata genügt. Da das IDP ein universelles Prinzip ist, muß es alle Dominanzschemata enthalten, die in natürlichen Sprachen vorkommen. Für Deutsch ist nur eine Teilmenge dieser Schemata zulässig. Man muß die Schemata, die nur für andere Sprachen gelten, aus der Disjunktion aller Dominanzschemata entfernen, um eine Menge der Schemata zu erhalten, die deutsche Phrasen beschreiben.<sup>2</sup>

Im folgenden werde ich zwei Möglichkeiten für die Formulierung von Dominanzschemata vorstellen. Im Abschnitt 2.1 zeige ich, wie die Dominanzschemata aussehen müßten, wenn man zulassen wollte, daß ein Kopf durch mehrere Komplemente gleichzeitig gesättigt wird. Im Kapitel 2.2 erkläre ich das Schema, das man benötigt, wenn nur jeweils ein Komplement in einem Schritt gesättigt werden soll. Probleme beider Alternativen werden diskutiert, und es wird argumentiert, daß binär verzweigende Strukturen eine stärkere Generalisierung darstellen.

In diesem Kapitel werden nur Kopf-Komplement-Strukturen diskutiert, in den folgenden Kapiteln werden jedoch noch weitere Dominanzschemata für andere Strukturen hinzukommen.

<sup>1</sup> Zur  $\bar{X}$ -Syntax siehe (Jackendoff, 1977).

<sup>2</sup> Hier stellt sich natürlich die Frage nach der Universalität dieses Prinzips. Nichts hindert einen Grammatikschreiber, für eine Sprache X ein Schema Y zu stipulieren. All diese Schemata, die eventuell nur für eine Sprache gelten, sind dann Teil des „universellen“ Prinzips. Dieses würde im Extremfall gar keine Generalisierung mehr darstellen, sondern nur eine Aufzählung verschiedener Möglichkeiten enthalten.

Zu einer diesbezüglichen Kritik an der Konzeption der Universalgrammatik innerhalb der HPSG-Theorie siehe auch (Netter, 1998, Kapitel 3.2.).

## 2.1. Flache Strukturen

Pollard und Sag (1994, S. 37ff) nehmen folgendes als Dominanzprinzip an:

### Prinzip 3 (Dominanzprinzip (*Immediate Dominance Principle*))

Die universell verfügbaren Optionen für eine wohlgeformte Phrase sind:

#### Schema 1 (Kopf-Komplement-Schema 1 (nach (Pollard und Sag, 1994)))

eine gesättigte Phrase, deren Kopftochter ein phrasales Zeichen ist, und die genau eine Komplementtochter hat.

$$\left[ \begin{array}{l} \text{SYNSEM} \\ \text{DTRS} \end{array} \left[ \begin{array}{l} \text{LOC|CAT|SUBCAT } \langle \rangle \\ \text{LEX } - \\ \text{HEAD-DTR } \left[ \text{SYNSEM|LEX } - \right] \\ \text{COMP-DTRS } \langle [ ] \rangle \\ \textit{head-complement-structure} \end{array} \right] \right]$$

*phrasal-sign*

oder

#### Schema 2 (Kopf-Komplement-Schema 2 (nach (Pollard und Sag, 1994)))

eine fast gesättigte Phrase, deren Subcat-Liste nur noch ein Element enthält, mit einem DTRS-Wert vom Typ *head-complement-structure*, in der die Kopftochter ein lexikalisches Zeichen ist.

$$\left[ \begin{array}{l} \text{SYNSEM} \\ \text{DTRS} \end{array} \left[ \begin{array}{l} \text{LOC|CAT|SUBCAT } \langle [ \textit{synsem} ] \rangle \\ \text{LEX } - \\ \text{HEAD-DTR } \left[ \text{SYNSEM|LEX } + \right] \\ \textit{head-complement-structure} \end{array} \right] \right]$$

*phrasal-sign*

oder

#### Schema 3 (Kopf-Komplement-Schema 3 (nach (Pollard und Sag, 1994)))

eine gesättigte Phrase mit einem DTRS-Wert vom Typ *head-complement-structure*, in der die Kopftochter ein lexikalisches Zeichen ist.

$$\left[ \begin{array}{l} \text{SYNSEM} \left[ \begin{array}{l} \text{LOC|CAT|SUBCAT } \langle \rangle \\ \text{LEX } - \end{array} \right] \\ \text{DTRS} \left[ \begin{array}{l} \text{HEAD-DTR} \left[ \begin{array}{l} \text{SYNSEM|LEX } + \end{array} \right] \\ \text{head-complement-structure} \end{array} \right] \\ \text{phrasal-sign} \end{array} \right]$$

oder

...

Die Punkte stehen für weitere Schemata, die in den folgenden Kapiteln eingeführt werden. Das Schema 1 entspricht der  $\bar{X}$ -Regel (2.2) und das Schema 2 der  $\bar{X}$ -Regel (2.3).

$$\bar{\bar{X}} \rightarrow \overline{\text{spezifizier } \bar{X}} \quad (2.2)$$

$$\bar{X} \rightarrow X \overline{\text{complement}} \quad (2.3)$$

Die linke Regelseite steht bei dieser Schreibweise für das phrasale Zeichen ohne DTRS-Wert und die rechte für eine Liste der Töchter in beliebiger Reihenfolge.

Die Abbildungen 2.1–2.3 verdeutlichen die Dominanzbeziehungen. Das Kopfmerkmalprinzip und das Subkategorisierungsprinzip sind in den Abbildungen mit enthalten.

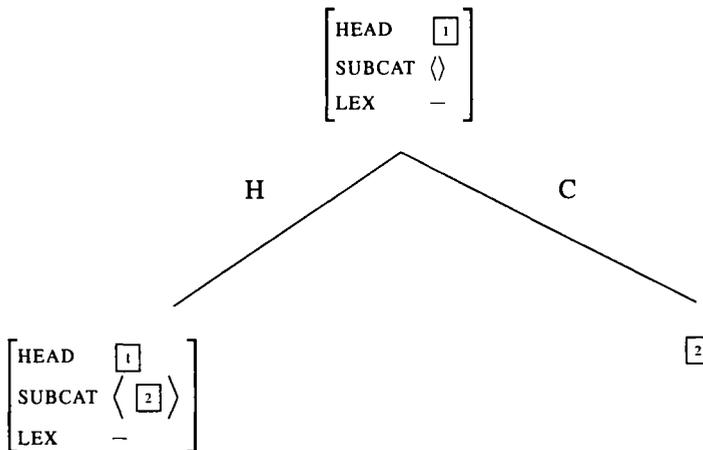


Abbildung 2.1.: Schema 1

Dabei ist zu beachten, daß Dominanzschemata nichts über die Oberflächenreihenfolge der Konstituenten aussagen. Das heißt, die Komplemente in den Abbildungen 2.1–2.3 können an der Oberfläche – in der Form, in der der Satz letztendlich gesprochen wird – durchaus vor dem Kopf und auch in anderer Reihenfolge auftauchen (in Abbildung 2.3 also z. B. [3], [2], [4]). Die Oberflächenreihenfolge wird durch Linearisierungsregeln festgelegt (siehe Kapitel 11).

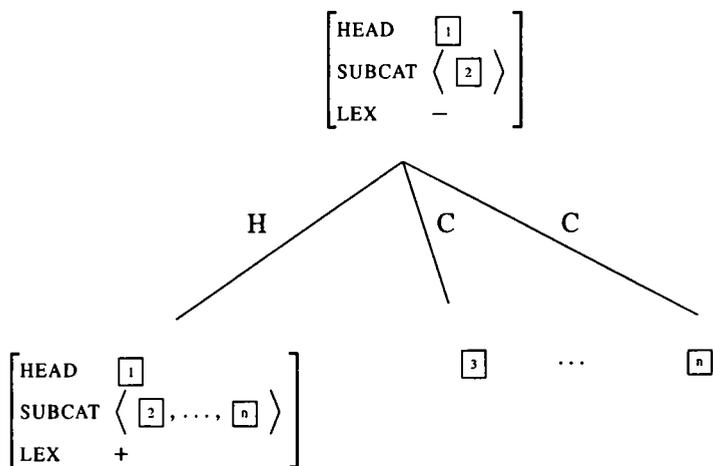


Abbildung 2.2.: Schema 2

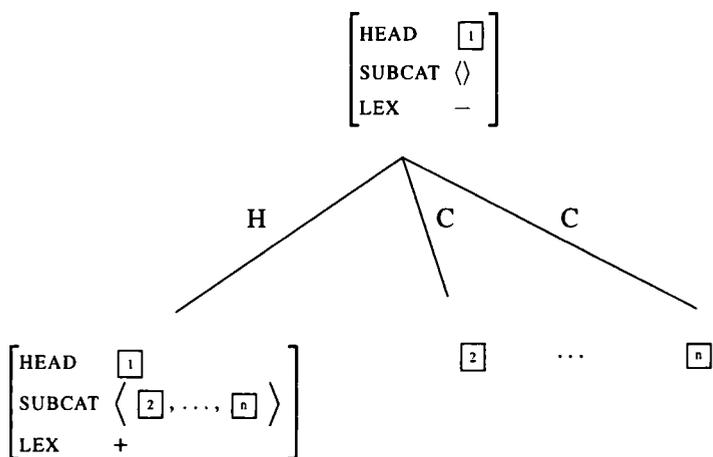


Abbildung 2.3.: Schema 3

Schema 1 und 2 lizenzieren u. a. eine Verbalprojektion, der noch das Subjekt fehlt (eine VP), und die Kombination dieser Verbalprojektion mit dem Subjekt. Zum Beispiel würde die Phrase *loves Mary* in (2.4) durch das Schema 2 lizenziert und die Kombination von *loves Mary* und *John* durch Schema 1.

(2.4) [<sub>1</sub> John [<sub>2</sub> loves Mary]].

Genauso bildet *gives Mary a book* in (2.5) eine Verbphrase, die mit dem Subjekt *John* kombiniert wird.

(2.5) [<sub>1</sub> John [<sub>2</sub> gives Mary a book]].

Fürs Englische haben Pollard und Sag (1987, S. 172) Linearisierungsregeln formuliert, die besagen, daß Komplemente, die durch das Schema 1 gesättigt wurden (also in (2.4) und (2.5) das Subjekt), links von ihrem Kopf stehen. Komplemente, die durch Schema 2 gesättigt werden, stehen rechts vom Kopf.

Diese Linearisierungsregeln gelten auch für Nominalphrasen.

(2.6) [<sub>1</sub> a [<sub>2</sub> picture of Mary]]

Hier steht der Determinator ebenfalls links vom Kopfnomen. Die Projektion *picture of Mary* wird mit Schema 2 abgeleitet. Schema 1 beschreibt wieder die Kombination mit dem letzten Element in der Subcat-Liste des Kopfes: Der Determinator *a* wird mit *picture of Mary* kombiniert.

Nach der Standardtheorie gelten die Linearisierungsregeln nur für Elemente, die vom selben Knoten dominiert werden. Das heißt, will man die Permutierbarkeit von Konstituenten unter dieser Voraussetzung erklären, so muß man annehmen, daß permutierbare Konstituenten vom selben Knoten dominiert werden.

Deshalb führen Pollard und Sag für die Analyse invertierter Sätze im Englischen und für deutsche Sätze das Schema 3 ein.<sup>3</sup>

(2.7) a. Can Kim go?

b. Liest Karl ein Buch?

<sup>3</sup> Bei Verwendung der drei Schemata bekommt man falsche Analysen, wenn man invertierte nicht von nicht invertierten Strukturen unterscheidet.

Pollard und Sag tun dies durch das Kopfmerkmal *INV*. Es wird nur in einer Fußnote (Pollard und Sag, 1994, S. 402) erwähnt. Ohne das *INV*-Merkmal wäre es möglich, mittels Schema 3 Strukturen für (i) aufzubauen.

(i) \* Loves Mary John.

Ein '\*' vor einem Satz kennzeichnet dessen Unkorrektheit. Wie Pollard und Sag (1994, S. 347) selbst feststellen, kann man *about Mary* nur mit Schema 3 ableiten, was zur Folge hätte, daß man annehmen müßte, daß es sich bei Präpositionalphrasen um invertierte Köpfe handelt.

In manchen Grammatiken für das Deutsche wird das Merkmal *INV* ebenfalls verwendet. Es hat dort aber eine andere Bedeutung. Es unterscheidet in deutschen Grammatiken für gewöhnlich Verben in Verbletzstellung von solchen in Verberststellung. Sowohl *INV-* als auch *INV+* Projektionen werden aber durch das Schema 3 lizenziert. Um Konfusion zu vermeiden, benutze ich das Merkmal *INITIAL* (siehe S. 5).

## c. Liest dieses Buch Karl?

Ein Kopf wird in einem Schritt mit all seinen Komplementen kombiniert. Würde man annehmen, daß im Deutschen genauso wie im Englischen erst eine Projektion des Verbs mit seinen Komplementen gebildet wird und daß diese dann mit dem Subjekt kombiniert wird, wäre nicht ohne weiteres zu erklären, wie das Subjekt in (2.7b) zwischen *liest* und *ein Buch* gelangen kann.

Im folgenden werde ich erklären, welche Strukturen diese Schemata in einer Grammatik des Deutschen lizenzieren würden und welche Probleme mit ihrer Verwendung einhergehen.

Schema 3 beschreibt Sätze, Adjektivphrasen und Präpositionalphrasen, entspricht also bei passender Instantiierung den Regeln in (2.8).

- (2.8)  $S \rightarrow V, \text{Comps}$     Liebt Karl Maria?  
 $\bar{A} \rightarrow \text{Comps}, A$     auf seinen Sohn stolze  
 $\bar{P} \rightarrow P, \bar{N}$             über Berlin

Schema 2 beschreibt die Bildung fast gesättigter Nominalphrasen ( $\bar{N}$ ). So wird zum Beispiel die Phrase *Bild von Maria* durch Schema 2 gerechtfertigt. Die Instantiierung des Schemas 2 ist analog zur Regel (2.9).

- (2.9)  $\bar{N} \rightarrow N, \bar{P}$     Bild von Maria

Die  $\bar{N}$ -Projektion wird benötigt, um pränominal Adjektive anschließen zu können.

- (2.10) ein [ $\bar{N}$  schönes [ $\bar{N}$  Bild von Maria]]

Genauerer dazu findet sich im Kapitel 4 über Modifikation. Schema 1 beschreibt Phrasen wie *das Bild von Maria*. *Bild von Maria* ist eine phrasale  $\bar{N}$ , die dann durch das Komplement *das* gesättigt wird. Die Instantiierung des Schemas 1 bei diesem Beispiel entspricht der Regel in (2.11).

- (2.11)  $\bar{\bar{N}} \rightarrow \bar{\text{DET}}, \bar{N}$     das Bild von Maria

Die Verwendung dieser Schemata ist nicht unproblematisch: Sie lassen für die Phrasen in (2.12) jeweils zwei Ableitungen zu.

- (2.12) a. das Bild von Maria  
 b. [<sub>1</sub> das [<sub>2</sub> Bild von Maria]]  
 c. [<sub>3</sub> das Bild von Maria]

- (2.13) a. , weil Karl der Frau das Buch gibt  
 b. [<sub>3</sub> Karl der Frau das Buch gibt]  
 c. [<sub>1</sub> Karl [<sub>2</sub> der Frau das Buch gibt]]

Für (2.12a) gibt es eine, in der alle Komplemente von *Bild* auf einmal gesättigt werden (Schema 3), und eine, in der erst die  $\bar{N}$  *Bild von Maria* gebildet wird (Schema 2) und diese dann durch *das* gesättigt wird (Schema 1).

Für (2.13a) könnte erst mit Schema 2 die VP *der Frau das Buch gibt* gebildet und diese dann mit dem Subjekt *Karl* kombiniert werden, oder es könnten wieder alle Komplemente durch Schema 3 gesättigt werden.

Reape (1994, S. 156) spezifiziert deshalb die Komplementtochter in Schema 1 als von einer Nominativ-Nominalphrase verschieden, um damit die Entstehung von Zeichen mit einer internen Struktur der Form [<sub>S</sub> NP[*nom*] VP] zu verhindern. Das löst das Problem jedoch nicht, da es im Deutschen subjektlose Verben gibt, und man für diese analoge unechte Mehrdeutigkeiten bekommt, wenn man nicht die Existenz phonologisch leerer Subjekte annehmen will.

(2.14) a. , weil ihm vor der Prüfung graut.

b. [<sub>3</sub> ihm vor der Prüfung graut]

c. [<sub>1</sub> ihm [<sub>2</sub> vor der Prüfung graut]]

Selbst wenn man annimmt, daß es phonologisch leere Subjekte gibt, verschwindet das Problem nicht. Nimmt man nämlich an, daß das Subjekt infinitiver Verbformen nicht Element der Subcat-Liste ist (vergleiche Kapitel 1.7 und Kapitel 17), dann reicht die Spezifikation der Komplementtochter als von NP[*nom*] verschieden nicht aus. Die Phrase *das Buch der Frau zu geben* in (2.15a) könnte dann mit Schema 3 ([<sub>3</sub> das Buch der Frau zu geben]) und auch mit einer kombinierten Anwendung von Schema 1 und 2 ([<sub>1</sub> das Buch [<sub>2</sub> der Frau zu geben]]) abgeleitet werden.

(2.15) a. Karl versucht, das Buch der Frau zu geben.

b. [<sub>3</sub> das Buch der Frau zu geben]

c. [<sub>1</sub> das Buch [<sub>2</sub> der Frau zu geben]]

Man könnte, statt in Schema 1 die Verschiedenheit des Komplements von einer Nominativnominalphrase zu verlangen, die Verschiedenheit von Nominalphrasen allgemein fordern. Damit wären die Fälle (2.13)–(2.15) aus der Welt geschafft. Die Fälle in (2.12) allerdings nicht. Alternativ könnte man die Köpfe in Schema 1 und 2 als nominal und den Kopf in Schema 3 als nichtnominal festlegen. Damit hätte man allerdings für Nominalstrukturen sehr spezifische Dominanzschemata, was der generellen HPSG-Philosophie zuwiderläuft: Zum einen ist die Kombination von Köpfen und Komplementen dann nicht mehr konsequent lexikongesteuert, sondern durch spezifische Regeln bestimmt, und zum anderen stellt sich die Frage nach dem universellen Charakter des Dominanzprinzips, wenn die Instantiierung für die Einzelsprachen Deutsch und Englisch so extrem unterschiedlich ausfällt.

## 2.2. Binär verzweigende Strukturen

Es ist möglich, die Schemata 1 bis 3 zu einem einzigen Schema zusammenzufassen.<sup>4</sup> Das folgende Schema stellt in bezug auf Kopf-Komplement-Strukturen eine wesentlich stärkere Generalisierung dar. Die im vorigen Abschnitt erwähnten Probleme entfallen.

### Schema 4 (Kopf-Komplement-Schema (vorläufige Version))

$$\left[ \begin{array}{l} \text{SYNSEM|LEX -} \\ \text{DTRS } \left[ \begin{array}{l} \text{COMP-DTRS } \langle [ ] \rangle \\ \text{head-complement-structure} \end{array} \right] \\ \text{phrasal-sign} \end{array} \right]$$

Dem Schema 4 zusammen mit dem Kopfmerkmalprinzip und dem Subkategorisierungsprinzip entspricht die Abbildung 2.4. Dieses Schema besagt, daß es in Kopf-Komplement-Strukturen immer genau eine Komplementtochter gibt.

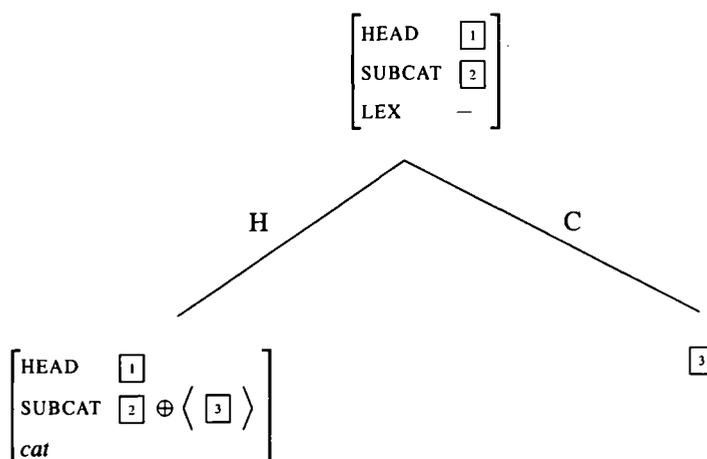


Abbildung 2.4.: Schema 4

Für den Satz (2.16) ergeben sich die Dominanzbeziehungen, die Abbildung 2.5 zeigt.

(2.16) , daß er das Buch dem Mann gab.

Gegen dieses Schema wurde lange Zeit argumentiert, da es nicht möglich zu sein scheint, Sätze wie (2.17) zu analysieren.

(2.17) , daß mir Maria das Buch gab.

Mit flachen Strukturen ist (2.17) problemlos zu analysieren, da die Dominanzschemata nichts über die Oberflächenreihenfolge von Konstituenten aussagen. Wie schon mehrfach erwähnt, wird die Reihenfolge der Konstituenten durch Linearisierungsregeln bestimmt, die von den Schemata unabhängig sind. Alle vom selben Knoten dominierten Konstituenten können beliebig permutiert werden, solange sie keine der Linearisierungsregeln verletzen.

<sup>4</sup> Dies ist nicht ganz korrekt. In Grammatiken, die lexikalische Köpfe mit spezifiziertem TO-BIND-Wert haben, muß man das Schema 4 in zwei Schemata aufspalten, wenn man annimmt, daß die linguistischen Objekte sortenaufgelöst sind.



## Literatur

In (Pollard und Sag, 1987, Kapitel 6) werden Grammatikregeln beschrieben, die den Dominanzschemata entsprechen. Etwas anders wird das Thema in (Pollard und Sag, 1994, Kapitel 1.5) behandelt.

### 3. Die Semantik innerhalb der HPSG

Ziel einer semantischen Analyse ist es, die Bedeutung bzw. Bedeutungen einer Eingabekette zu ermitteln. Es gibt viele verschiedene Ansätze in der Künstlichen Intelligenz, die alle Vor- und Nachteile haben (Aussagenlogik,  $\lambda$ -Kalkül, Situationssemantik, modal-temporale Logiken, Ereigniskalkül) und sich unterschiedlich gut für eine semantische Analyse eignen.

Pollard und Sag (1987; 1994) nutzen den HPSG-Formalismus zur Beschreibung von Sachverhalten als Teil von Situationen. Ich möchte im folgenden einige grundlegende Fakten aus (Pollard und Sag, 1987) wiederholen:

Die Situationssemantik geht davon aus, daß Dinge in einer Welt, die für eine bestimmte Gemeinschaft typisch ist, mit Hilfe eines Charakterisierungsschemas beschrieben werden. Zu solch einem Schema gehören Individuen (*Karl, Maria*), Eigenschaften (*dumm, glücklich*) und Relationen (*lieben, sterben*). Eigenschaften kann man auch als einstellige Relationen betrachten.

Innerhalb eines jeden Zeichens gibt es den Pfad SYNTAX-SEMANTICS|LOCAL|CONTENT, dessen Wert dem semantischen Hauptbeitrag des Zeichens zur beschriebenen Situation entspricht. Der semantische Beitrag eines Aussagesatzes beschreibt eine Situation, wohingegen die Beiträge von lexikalischen Zeichen Beschreibungen von Referenten, Eigenschaften oder Relationen sein können.

#### 3.1. Individuen, Sachverhalte und Situationen

Dinge von einer gewissen zeitlichen Dauer, die zur kausalen Ordnung der Welt gehören, die man wahrnehmen kann, auf die man reagieren kann, werden als Individuen bezeichnet (*Karl, die Frau*). Dazu gehören auch Konzepte (*die Angst, das Versprechen*).

Über Individuen bekannte Fakten werden mittels Relationen dargestellt. Relationen können null- (*regnen*), ein- (*schlafen, sterben*), zwei- (*lieben*), drei- (*geben*), ... stellig sein. *Kaufen* ist vierstellig: Es gibt einen Käufer, einen Verkäufer, eine Ware und einen Preis.

Pollard und Sag (1987; 1994) geben den Argumenten der Relationen Namen, die aus der Schreibweise des Verbs abgeleitet sind. Ich verwende jedoch im weiteren Rollenbezeichnungen, wie sie zum Beispiel von Kunze (1991) vorgeschlagen werden. Kunze nimmt folgendes Grundinventar an Rollen an: AGENS, PATIENS, EXPERIENCER, SOURCE, GOAL, THEMA, LOCATION, TRANS-OBJ, INSTRUMENT, MEANS und PROPOSITION.<sup>1</sup>

---

<sup>1</sup> In (Pollard und Sag, 1994, Kapitel 8.5.3) wird darauf eingegangen, daß es von Nutzen ist, Rollen wie SOURCE, GOAL und AGENS zur Verfügung zu haben. Es wird vorgeschlagen, die Rolle GEBENDER als 'Subtyp' von AGENS und die Rolle EMPFÄNGER als 'Subtyp' von GOAL zu betrachten. Mit dem innerhalb der HPSG verwendeten Formalismus ist das nicht möglich, da GEBENDER und EMPFÄNGER Bezeichnungen für Merkmale sind. Merkmale haben Werte eines bestimmten Typs, sind aber selber keine Typen. Man müßte also eine zweite Hierarchie für Merkmalsbezeichnungen einführen. Abgesehen davon, daß der Verarbeitungsaufwand steigen würde, gibt es auch Fälle, in denen man

Die Welt, über die gesprochen wird, besteht aus Situationen, in denen Individuen Eigenschaften haben und in Relationen zueinander stehen. Sachverhalte sind mögliche Zustände, in denen die Welt sich befinden kann. In der Situationssemantik wird ein solcher Zustand *state of affairs* (*soas*) genannt. Es gibt einfache und zusammengesetzte Sachverhalte. Einfache Sachverhalte bestehen aus einer Relation, einer Zuweisung der Referenten an die Rollen der Relation und einer Polarität, positiv – 1 oder negativ – 0.

$\ll \textit{schlafen, thema : karl; 1} \gg$   
 $\ll \textit{mann, instance : maria; 0} \gg$ 
(3.1)

Die aufgeführten Beispiele beschreiben Sachverhalte, sie besagen nichts über deren Gültigkeit. Sachverhalte können selbst Argumente von Relationen sein:

$\ll \textit{glauben, patiens : karl, proposition :} \ll \textit{lieben, agens : maria, patiens : karl; 1} \gg; 1 \gg$ 
(3.2)

(3.2) beschreibt den Sachverhalt, daß Karl glaubt, daß Maria ihn liebt.

$\ll \textit{scheinen, proposition :} \ll \textit{schlafen, thema : karl; 0} \gg; 1 \gg$ 
(3.3)

(3.3) beschreibt den Sachverhalt, daß Karl nicht zu schlafen scheint.

Es gibt zusammengesetzte Sachverhalte:

$\textit{and} \{ \ll \dots \gg, \ll \dots \gg \}$ 
(3.4)

$\textit{or} \{ \ll \dots \gg, \ll \dots \gg \}$ 
(3.5)

und

---

von der aus dem Verb abgeleiteten Rollenbezeichnung keine Rückschlüsse auf den 'Supertyp' (also AGENS oder GOAL oder dergleichen) ziehen kann.

Ein Beispiel dafür ist das Verb *rollen*. Die Repräsentation in Standard-HPSG ist die folgende:

$\left[ \begin{array}{l} \text{RELN} \quad \textit{rollen} \\ \text{ROLLENDER} \quad \boxed{1} \\ \text{GEROLLTES} \quad \boxed{2} \\ \textit{psoa} \end{array} \right]$	$\left[ \begin{array}{l} \text{RELN} \quad \textit{rollen} \\ \text{GEROLLTES} \quad \boxed{1} \\ \textit{psoa} \end{array} \right]$	(i)
--	--	-----

Ich verwende statt dessen die Repräsentation in (ii).

$\left[ \begin{array}{l} \text{AGENS} \quad \boxed{1} \\ \text{PATIENS} \quad \boxed{2} \\ \textit{rollen-tr} \end{array} \right]$	$\left[ \begin{array}{l} \text{THEMA} \quad \boxed{1} \\ \textit{rollen-itr} \end{array} \right]$	(ii)
--	---	------

Die jeweils vordere Struktur entspricht dem Sachverhalt von *Sisyphus rollt den Felsblock*. (den Berg hinauf), die hintere dem von *Der Fels rollt*. (den Berg hinunter). Das Merkmal GEROLLTES müßte also die 'Supertypen' THEMA und PATIENS haben. Die Information, um welchen der beiden es sich handelt, ist in den Repräsentationen in (i) nicht enthalten.

Wie dem auch sei, auch diese Repräsentation ist nicht ausreichend: Verben können ihren Argumenten mehrere Rollen zuweisen. Um das vollständig zu beschreiben, braucht man komplexe Sememrepräsentationen, wie sie in (Kunze, 1991) erklärt sind. Darauf näher einzugehen und diese Sememrepräsentationen zu verwenden, würde den Rahmen dieses Buches sprengen.

$$(\text{forall } X | \ll \dots \gg) \ll \dots \gg \quad (3.6)$$

Die Ausdrücke in dieser Logik der Sachverhalte haben eine andere Bedeutung als ihre Gegenstücke in der Prädikatenlogik der ersten Stufe (PL1). Die Bedeutung der Relationen in Sachverhalten geht über die derjenigen in der PL1 hinaus: Zwei mathematische Relationen sind identisch, falls dieselbe Menge geordneter  $n$ -Tupel Elemente dieser Relation sind. Für die Relationen in Sachverhalten gilt das nicht, da man aus dem Fakt, daß die Menge aller  $a$  für die *singen*( $a$ ) gilt, dieselben Elemente enthält, wie die Menge aller  $b$ , für die *betrunken*( $b$ ) gilt, nicht folgern kann, daß *singen* und *betrunken* dieselbe Relation ist.

Ein anderes Beispiel wird in (Pollard und Sag, 1987) gegeben: Für jedes gleichseitige Dreieck gilt, daß es gleichwinklig ist. Mathematisch gesehen ist die Relation *gleichseitiges\_dreieck* identisch mit der Relation *gleichwinkliges\_dreieck*, da die Menge der gleichseitigen Dreiecke identisch mit der Menge der gleichwinkligen Dreiecke ist.

Der Sachverhalt, daß Karl glaubt, daß das Dreieck  $t$  gleichseitig ist, wird wie folgt dargestellt:

$$\ll \text{glauben, } \textit{patients} : \textit{karl}, \quad (3.7)$$

$$\textit{proposition} : \ll \textit{gleichseitiges\_dreieck, instance} : t; 1 \gg; 1 \gg$$

(3.7) ist nicht äquivalent zu (3.8), da Karl sich der Tatsache, daß gleichseitige Dreiecke auch gleichwinklig sind, nicht bewußt sein muß.

$$\ll \text{glauben, } \textit{patients} : \textit{karl}, \quad (3.8)$$

$$\textit{proposition} : \ll \textit{gleichwinkliges\_dreieck, instance} : t; 1 \gg; 1 \gg$$

Ein weiterer Unterschied zwischen der Logik der Sachverhalte und PL1 ist, daß in der PL1 geschlossene Formeln (Formeln, die keine freien Variablen enthalten) Wahrheitswerten entsprechen.

$$a. \ll \textit{betrunken, instance} : \textit{karl}; 1 \gg \quad (3.9)$$

$$b. \textit{betrunken}(\textit{karl})$$

(3.9a) beschreibt den Sachverhalt, daß Karl betrunken ist, wohingegen (3.9b) einfach wahr oder falsch ist, je nachdem ob Karl zur Menge der Betrunkenen gehört oder nicht. Daß man mit (3.9a) mehr anfangen kann als mit (3.9b), zeigt die Tatsache, daß es in der PL1 Formeln wie *glauben*(*maria*, *betrunken*(*karl*)) nicht gibt. Erstens ist *betrunken*(*karl*) eine Formel und kein Term, also aus syntaktischen Gründen als Argument einer Relation unzulässig, und zweitens ist der obige Ausdruck auch semantischer Unsinn, da der Wert von *betrunken*(*karl*) *wahr* oder *falsch* ist, und Maria nicht *wahr* oder *falsch* glaubt, sondern den Sachverhalt, daß Karl betrunken ist.

In linguistischen Anwendungen braucht man Sachverhalte, die Variablen enthalten:

$$\ll \textit{betrunken, instance} : X; 1 \gg \quad (3.10)$$

Das PL1-Äquivalent wäre *betrunken*( $X$ ). Ein Unterschied ist jedoch, daß *betrunken*( $X$ ) erst dann einen Wert bekommt, wenn  $X$  instantiiert wird. In der Situationssemantik wird (3.10) als parametrisierter Sachverhalt verwendet und als *parametrized state of affairs* (*psoa*) bezeichnet. Ein parametrisierter Sachverhalt ist ein Sachverhalt, der Argumente besitzt, die nicht an Referenten gebunden sind. Die Wahl der Variablennamen für die Parameter ist willkürlich. Parameter können durch Angaben darüber, auf welche Menge von Referenten sie sich beziehen, eingeschränkt werden. Parametrisierte Sachverhalte mit restringierten Parametern werden wie folgt dargestellt:

$$\begin{aligned} &\ll \textit{lieben,agens} : X, \textit{patients} : Y; 1 \gg \\ X | \ll \textit{mann,instance} : X; 1 \gg, Y | \ll \textit{kind,instance} : Y; 0 \gg \end{aligned} \quad (3.11)$$

Situationen sind räumlich und zeitlich begrenzte Teile der Welt, in denen etwas vor sich geht (Referenten haben Eigenschaften oder stehen in Relationen zueinander). Man beschreibt, charakterisiert oder klassifiziert Situationen aufgrund der Sachverhalte, die in ihnen vorliegen. Man sagt, daß eine Situation einen Sachverhalt  $c$  unterstützt ( $s$  supports  $c$ ) bzw. daß  $c$  in  $s$  gilt ( $c$  holds in  $s$ ). Unterstützen ist eine Relation zwischen Situationen und Sachverhalten.

$$\ll \textit{support,situation} : s, \textit{circumstance} : c; 1 \gg \quad (3.12)$$

(3.12) bedeutet, daß  $s$   $c$  unterstützt. Man schreibt dafür auch  $s \models c$ .

Wenn (3.12) ein Fakt ist, dann ist  $c$  ein Fakt, da es eine Situation  $s$  gibt, die  $c$  unterstützt. Die Sachverhalte, die in einer Situation  $s$  gelten, werden Fakten der Situation  $s$  genannt.

Man kann Sachverhalte nutzen, um Situationen zu beschreiben.

$$\{ \{s\} : \ll \textit{support,situation} : s, \textit{circumstance} : c; 1 \gg \} \quad (3.13)$$

(3.13) beschreibt die Menge der Situationen, in denen  $c$  gilt.

## 3.2. Die Semantik von Zeichen

Der semantische Hauptbeitrag eines Zeichens steht unter dem Pfad SYNSEM|LOC|CONT. Oft sind die Werte dieses Pfades parametrisiert. Die Parameter hängen dann vom Kontext des Zeichens in der entsprechenden Phrase ab. Sie werden z. B. durch referentielle Nominalphrasen – vereinfacht gesagt: Nominalphrasen, die sich auf bestimmte Dinge beziehen – eingeführt. Die Parameter werden mit Hilfe des Merkmals INDICES verwaltet, das unter dem Pfad SYNSEM|LOC|CONTEXT steht. Der Wert dieses Merkmals ist eine Menge und zwar die Menge aller Indizes, die im jeweiligen Zeichen enthalten sind.

### 3.2.1. Indizes

Um zu erklären, wie sich der Sachverhalt, der von einem Aussagesatz beschrieben wird, aus den semantischen Beiträgen einzelner Teile zusammensetzt, braucht man eine Merkmalstruktur, die ein Situationstheorieobjekt beschreibt, z. B. Referenten (*karl*, *anna*) und Relationen (*lieben*, *lügen*). Sachverhalte bzw. parametrisierte Sachverhalte werden durch Merkmalstrukturen vom Typ *psoa* beschrieben. *psoa* hat viele Subtypen, die den Relationen zwischen den Individuen entsprechen. Es gibt entsprechend der Stelligkeit der Relation Merkmale für die zur Relation gehörenden Rollen. Die den Sachverhalt  $\ll \textit{mann,instance} : \textit{karl}; 1 \gg$  repräsentierende Merkmalstruktur hat die Form:

$$\left[ \begin{array}{l} \textit{INST karl} \\ \textit{mann} \end{array} \right] \quad (3.14)$$

$\ll \textit{frau,instance} : \textit{Karl}; 0 \gg$  wird durch (3.15) dargestellt.

$$\left[ \begin{array}{l} \text{ARG} \left[ \begin{array}{l} \text{INST } \textit{karl} \\ \textit{frau} \end{array} \right] \\ \textit{neg} \end{array} \right] \quad (3.15)$$

Diese Notation ist natürlich keine Eins-zu-eins-Entsprechung der Darstellung in der Situationssemantik. In der HPSG findet eine Komposition<sup>2</sup> der Bedeutung von Phrasen aus der Bedeutung von Teilphrasen statt. Mit der in (3.15) angegebenen Notation ist das einfacher zu beschreiben. Insbesondere ist doppelte Negation monoton beschreibbar.

Restringierte Parameter werden durch Merkmalstrukturen vom Typ *nominal-object* (*nom-obj*) repräsentiert, der die Merkmale INDEX und RESTRICTIONS hat. Der Wert von INDEX ist eine Merkmalstruktur, die Merkmale für Person, Numerus, Genus und den Referenten enthält. RESTRICTIONS hat als Wert eine Menge von Restriktionen, die den Referenten betreffen.<sup>3</sup> Der Typ *nom-obj* besitzt die Subtypen *expletive* (*expl*) und *referential* (*ref*).

Referentielle Nominalphrasen haben einen Index vom Typ *referential*. Dieser hat die Subtypen *nonpronoun* (*npro*) und *pronoun* (*pro*). Pronomina werden in anaphorische Pronomina und Personalpronomina *ppro* unterteilt. *anaphor* (*ana*) hat die Subtypen *reflexive* (*refl*) und *reciprocal* (*recp*). Die Indizes für *Mann*, *Karl* und *Buch* sind vom Typ *npro*. *Er*, *sie*, *ihn* sind vom Typ *ppro*. *Sich* ist vom Typ *refl* und *einander* ist vom Typ *recp*. Die Subtypen des Typs *nom-obj* zeigt Abbildung 3.1.

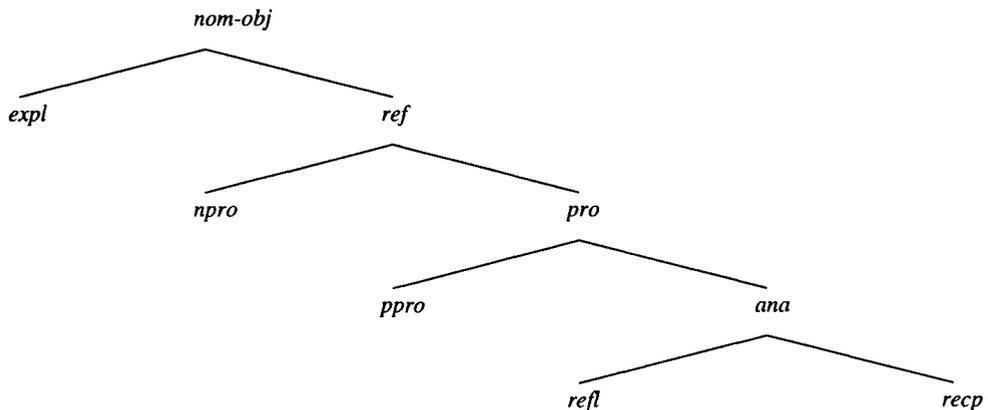


Abbildung 3.1.: Subtypen des Typs *nom-obj*

Die Struktur (3.16) beschreibt den Index für einen dummen Mann.

<sup>2</sup> Zum Begriff der Kompositionalität vergleiche auch (Cooper, 1983, Kapitel 1.3.1).

<sup>3</sup> Pollard und Sag verwenden das Merkmal REF nicht. Statt dessen ist bei ihnen die gesamte Indexstruktur Argument der Relation.

$$\left[ \begin{array}{l} \text{IND} \\ \text{RESTR} \\ \text{npro} \end{array} \left[ \begin{array}{l} \text{REF } \boxed{1} \\ \text{PER } 3 \\ \text{NUM } sg \\ \text{GEN } mas \end{array} \right] \left\{ \left[ \text{INST } \boxed{1} \right] \text{mann}, \left[ \text{INST } \boxed{1} \right] \text{dumm} \right\} \right] \quad (3.16)$$

Der Wert von IND ist für die Bindung von Pronomina und die Beschreibung von Relativsätzen und Kongruenz wichtig (siehe Kapitel 20, 10 und 16).

Expletiva haben einen Index vom Typ *expl*. Das *es* in den folgenden Sätzen ist ein Expletivpronomen, das einen Index solchen Typs besitzt.<sup>4</sup>

- (3.17) a. Es regnet.  
 b. Es gibt einen Weihnachtsmann.  
 c. Es trug ihn aus der Kurve.<sup>5</sup>  
 d. Es schüttelte mich vor Ekel.  
 e. Er bringt es bis zum Professor.  
 f. Er nimmt es mit zehn Gegnern auf.

Außer dem expletiven *es* kann auch *das* als Expletivum vorkommen, wenn Sätze entsprechend betont sind.

Merkmalstrukturen vom Typ *expl* haben die Form:

$$\left[ \begin{array}{l} \text{IND} \\ \text{RESTR } \{ \} \\ \text{expl} \end{array} \left[ \begin{array}{l} \text{REF } none \\ \text{PER } 3 \\ \text{NUM } sg \end{array} \right] \right] \quad (3.18)$$

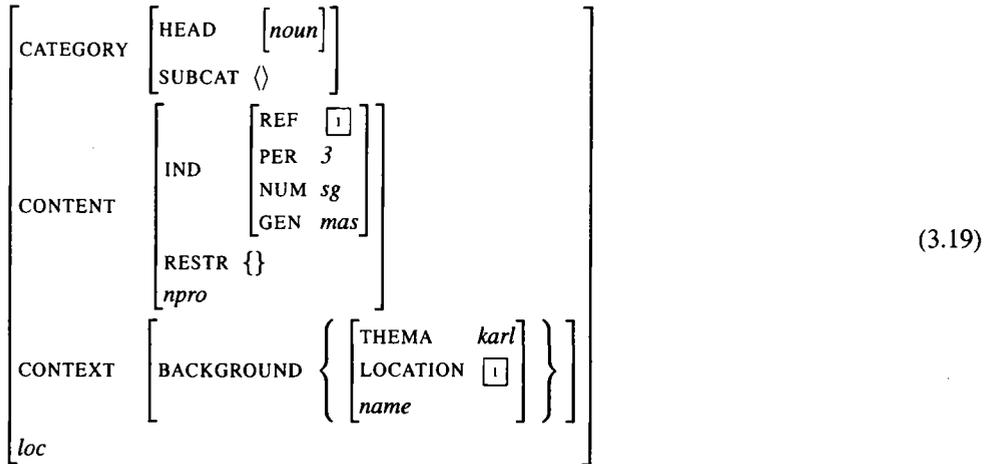
Expletivpronomen referieren nie. Der Wert von REF ist deshalb *none*. Die Menge der Restriktionen ist leer.

### 3.2.2. Hintergrundinformationen und Präsuppositionen

Mit Eigennamen kann man auf bestimmte Individuen referieren. Verwendet der Sprecher einen Eigennamen, so setzt er voraus, daß der Hörer den zum Eigennamen gehörenden Referenten ermitteln kann. Die entsprechende Information ist unter dem Pfad SYNSEM|LOC|CONTEXT|BACKGROUND in Zeichen enthalten. Das Zeichen für den Eigennamen *Karl* hat also die Form:

<sup>4</sup> Für weitere Beispiele siehe auch (Höhle, 1978, S. 24).

<sup>5</sup> (Uszkoreit, 1987, S. 147)



Der Wert des BACKGROUND-Merkmals ist eine Menge von *psaos*. Diese entsprechen Präsuppositionen und unterliegen als solche einer entsprechenden Logik. Ich nehme im folgenden vereinfachend an, daß der BACKGROUND-Wert eines phrasalen Zeichens die Vereinigung der BACKGROUND-Werte seiner Töchter ist.

In (Pollard und Sag, 1994, Kapitel 8.4) wird der Wert des Pfades SYNSEM|LOC|CONTEXT ausführlicher beschrieben, als das hier der Fall ist. Zur Merkmalstruktur vom Typ *context* gehören Merkmale wie C-INDICES und BACKGROUND. C-INDICES enthält Informationen über den Sprecher, den Adressaten und die Lokation der Äußerung. Diese Informationen sind wichtig, will man die deiktische Verwendung von Pronomina oder einen von mehreren Sprechern gesprochenen Satz korrekt erfassen.

### 3.3. Das Semantikprinzip

Wie erhält man nun die Bedeutung des Satzes *Die Frau küßt den Mann.* aus der Bedeutung seiner Wörter? Wie erhält man die Bedeutung eines phrasalen Zeichens aus der Bedeutung der lexikalischen Zeichen, aus denen sich das phrasale Zeichen zusammensetzt?

Man weiß, daß das Verb *küssen* zwei Komplemente braucht: ein Subjekt und ein Akkusativobjekt. Diese Komplemente sind Elemente der Subkategorisierungsliste, die unter dem Pfad SYNSEM|LOC|CAT|SUBCAT steht. Man weiß auch, daß die Relation *küssen* zweistellig ist, und daß die Rolle des Küssenden (AGENS) dem Subjekt und die Rolle des Geküßten (PATIENS) dem Objekt zugeordnet wird.

Zu den Nominalphrasen (*die Frau* und *den Mann*) gehören Indizes, und diese Indizes beziehen sich auf Referenten. Die Referenten (die Menge derjenigen in einer Situation, die die Eigenschaft besitzen, eine Frau bzw. ein Mann zu sein) werden durch *structure sharing* den Argumenten der Relation *küssen* zugeordnet. Das verdeutlicht die folgende Merkmalstruktur für *küßt*:

$$\left[ \begin{array}{l} \text{PHON} \\ \text{SYNSEM|LOC} \\ \textit{lexical-sign} \end{array} \left[ \begin{array}{l} \langle \textit{küßt} \rangle \\ \text{CAT} \\ \text{CONT} \end{array} \left[ \begin{array}{l} \text{HEAD} \\ \text{SUBCAT} \\ \text{AGENS} \\ \text{PATIENS} \end{array} \left[ \begin{array}{l} \left[ \begin{array}{l} \text{VFORM } \textit{fin} \\ \textit{verb} \end{array} \right] \\ \langle \text{NP[nom]} \boxed{1}, \text{NP[acc]} \boxed{2} \rangle \\ \boxed{1} \\ \boxed{2} \\ \textit{küssen} \end{array} \right] \right] \right] \right] \quad (3.20)$$

Dabei steht NP  $\boxed{1}$  für:

$$\left[ \begin{array}{l} \text{LOC} \\ \textit{local} \\ \textit{synsem} \end{array} \left[ \begin{array}{l} \text{CAT} \\ \text{CONT} \end{array} \left[ \begin{array}{l} \text{HEAD} \\ \text{SUBCAT} \\ \text{IND|REF} \end{array} \left[ \begin{array}{l} \left[ \textit{noun} \right] \\ \langle \rangle \\ \boxed{1} \\ \textit{ref} \end{array} \right] \right] \right] \right] \quad (3.21)$$

Bei NP[nom]  $\boxed{1}$  ist zusätzlich das Kopfmerkmal CAS als *nom* spezifiziert.

Der semantische Gehalt eines phrasalen Zeichens ist identisch mit dem semantischen Gehalt der Kopftochter des Zeichens. Diese Tatsache spiegelt das Semantikprinzip wider:

#### Prinzip 4 (Semantikprinzip – vorläufige Version)

$$\left[ \text{DTRS} \left[ \textit{headed-structure} \right] \right] \Rightarrow \left[ \begin{array}{l} \text{SYNSEM|LOC|CONT} \\ \text{DTRS|HEAD-DTR|SYNSEM|LOC|CONT} \end{array} \left[ \begin{array}{l} \boxed{1} \\ \boxed{1} \end{array} \right] \right]$$

Unter Anwendung dieses und des Subcat-Prinzips erhält man für den Satz *daß die Frau den Mann küßt*. den Baum, den Abbildung 3.2 zeigt. Der semantische Beitrag des entstehenden Zeichens für den gesamten Satz ist:

$$\left[ \begin{array}{l} \text{AGENS} \\ \text{PATIENS} \\ \textit{küssen} \end{array} \left[ \begin{array}{l} \boxed{3} \\ \boxed{5} \end{array} \right] \right] \quad (3.22)$$

Dabei ist  $\boxed{3}$  die Referenzvariable für *die Frau* und  $\boxed{5}$  die für *den Mann*. In dem Zeichen, das durch die Anwendung des Semantikprinzips in seiner vorläufigen Version entstanden ist, fehlen aber wichtige Informationen. Man hat zwar Zugriff auf die Referenzvariablen ( $\boxed{3}$  und  $\boxed{5}$ ), die Information über die Beschaffenheit der Indizes ist jedoch im entstandenen Zeichen nicht direkt verfügbar. Im Zeichen für den Satz *Die Frau küßt den Mann*. sind die Restriktionen für

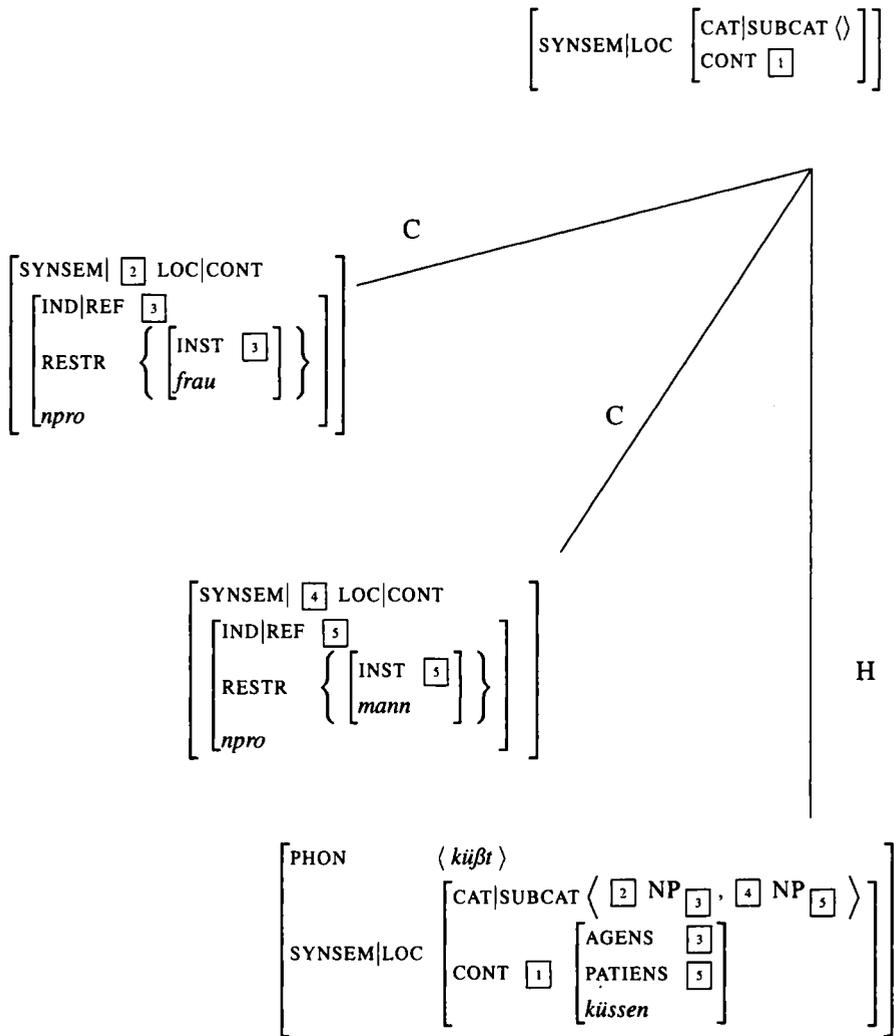


Abbildung 3.2.: die Frau den Mann küßt.

die Indizes nicht mehr auf der obersten Stufe, d. h. unter dem Pfad SYNSEM|LOC enthalten. Man könnte sie nur tief verschachtelt in den Töchterlisten wiederfinden. Die Restriktionen und auch die Information über Genus, Person und Numerus, darüber ob der Index zu einer definiten oder indefiniten Phrase gehört, ob es sich um ein Pronomen handelt und wenn ja, um was für eines, werden aber auch in phrasalen Zeichen gebraucht (siehe Kapitel 20 über Bindungstheorie). Wie erreicht wird, daß diese Information beim Zusammensetzen von komplexen Zeichen auf der obersten Ebene repräsentiert ist, wird im folgenden erklärt.

Unter dem Pfad SYNSEM|LOC|CONTEXT|INDICES steht eine Menge, die alle Indizes und die zugehörigen Restriktionen enthält, die im betreffenden Zeichen eingeführt wurden.<sup>6</sup> Diese Menge kann auch leer sein (z. B. führt eine Präposition keine Indizes ein). Bei Eigennamen enthält die INDICES-Menge gerade den CONT-Wert des Indexes, der zum Eigennamen gehört:

$$\left[ \begin{array}{l} \text{CONT} \\ \text{CONTEXT} \\ \text{loc} \end{array} \right] \left[ \begin{array}{l} \boxed{1} \\ \left[ \begin{array}{l} \text{IND} \\ \text{RESTR } \{ \} \\ \text{npro} \end{array} \right] \\ \left[ \begin{array}{l} \text{INDS } \{ \boxed{1} \} \\ \text{BACKGROUND } \left\{ \left[ \begin{array}{l} \text{THEMA } \textit{karl} \\ \text{LOCATION } \boxed{2} \end{array} \right] \right\} \\ \text{name} \end{array} \right] \end{array} \right] \left[ \begin{array}{l} \left[ \begin{array}{l} \text{REF } \boxed{2} \\ \text{PER } 3 \\ \text{NUM } \textit{sg} \\ \text{GEN } \textit{mas} \end{array} \right] \end{array} \right] \quad (3.23)$$

Die Nominalphrase *seine Frau* enthält zwei Indizes, einen für *seine* und einen für *Frau* (siehe Kapitel 6 über Possessivpronomina).

Der Wert des Merkmals INDICES eines Zeichens ist die Vereinigung der Werte des Merkmals INDICES der Tochterzeichen. Eine neue Version des Semantikprinzips folgt.<sup>7</sup>

### Prinzip 5 (Semantikprinzip – vorläufige Version)

$$\left[ \text{DTRS } \left[ \textit{headed-structure} \right] \right] \Rightarrow \left[ \begin{array}{l} \text{SYNSEM} \\ \text{DTRS} \end{array} \right] \left[ \begin{array}{l} \left[ \text{LOC} \left[ \begin{array}{l} \text{CONT } \boxed{1} \\ \text{CONTEXT|INDS } \textit{collect-inds}(\boxed{2}) \end{array} \right] \right] \\ \left[ \text{HEAD-DTR|SYNSEM|LOC|CONT } \boxed{1} \right] \end{array} \right]$$

<sup>6</sup> Quantifizierende Ausdrücke werden im vorliegenden Fragment nicht erfaßt. Wie Indizes quantifizierender Ausdrücke zu behandeln sind, muß also vorerst offen bleiben. Die Standardbehandlung von Quantoren in der HPSG halte ich für ungeeignet. Ansätze wie den von Frank und Reyle (1992) und Copestake, Flickinger und Sag (1997) sind meiner Meinung nach angemessener. Auf diese Ansätze einzugehen, würde den Rahmen dieses Buches jedoch sprengen.

<sup>7</sup> Diese Version ist nicht die endgültige Version, da die Semantik von Modifikatoren (siehe Kapitel 4) damit noch nicht beschrieben wird.

Hierbei ist *collect-inds(X)* eine Funktion, die auf die Struktur angewendet wird, die die Töchter enthält, und die Vereinigung aller Mengen der Tochterindizes liefert.

Der Wert des Pfades *SYNSEM|LOC* für das Zeichen des Satzes *Die Frau küßt den Mann.* ist also:

$$\left[ \begin{array}{l} \text{CONT} \\ \text{CONTEXT|INDS} \\ \text{loc} \end{array} \right. \left. \begin{array}{l} \left[ \begin{array}{l} \text{AGENS} \quad [3] \\ \text{PATIENS} \quad [5] \\ \text{küssen} \end{array} \right] \\ \left[ \begin{array}{l} \text{IND} \quad \left[ \begin{array}{l} \text{REF} \quad [3] \\ \text{PER} \quad 3 \\ \text{NUM} \quad \text{sg} \\ \text{GEN} \quad \text{fem} \end{array} \right] \\ \text{RESTR} \quad \left\{ \left[ \begin{array}{l} \text{INST} \quad [3] \\ \text{frau} \end{array} \right] \right\} \\ \text{npro} \end{array} \right] \\ \left[ \begin{array}{l} \text{IND} \quad \left[ \begin{array}{l} \text{REF} \quad [5] \\ \text{PER} \quad 3 \\ \text{NUM} \quad \text{sg} \\ \text{GEN} \quad \text{mas} \end{array} \right] \\ \text{RESTR} \quad \left\{ \left[ \begin{array}{l} \text{INST} \quad [5] \\ \text{mann} \end{array} \right] \right\} \\ \text{npro} \end{array} \right] \end{array} \right] \quad (3.24)$$

### 3.4. Rollen und fakultative Komplemente

Will man eine Situation beschreiben, in der gegessen wird, kann man Sätze wie *Karl ißt.* oder *Karl ißt Kuchen.* dazu benutzen. Der zweite Satz beschreibt eine Teilmenge der Situationen, die vom ersten Satz beschrieben werden, nämlich gerade die Situationen, in denen das, was Karl ißt, Kuchen ist. In jeder Essen-Situation ist jedoch etwas da, das gegessen wird. Es kann aber sein, daß es in einem Satz, der eine Essen-Situation beschreibt, nicht spezifiziert wird. Man sagt, das Objekt ist fakultativ. Fakultative Komplemente werden in der Subcat-Liste in Klammern geschrieben. Die Subcat-Liste von *essen* hat also die Form  $\langle \text{NP}[\text{nom}], (\text{NP}[\text{acc}]) \rangle$ .

$$\left[ \begin{array}{l} \text{PHON} \\ \text{SYNSEM|LOC} \\ \text{lexical-sign} \end{array} \right. \left. \begin{array}{l} \langle \text{ißt} \rangle \\ \left[ \begin{array}{l} \text{CAT} \quad \left[ \begin{array}{l} \text{HEAD} \quad \left[ \begin{array}{l} \text{VFORM} \quad \text{fin} \\ \text{verb} \end{array} \right] \\ \text{SUBCAT} \quad \langle \text{NP}[\text{nom}]_{[1]}, (\text{NP}[\text{acc}]_{[2]}) \rangle \end{array} \right] \\ \text{CONT} \quad \left[ \begin{array}{l} \text{AGENS} \quad [1] \\ \text{PATIENS} \quad [2] \\ \text{essen} \end{array} \right] \end{array} \right] \end{array} \right] \quad (3.25)$$

Dabei ist  $\langle \text{NP}[\text{nom}], (\text{NP}[\text{acc}]) \rangle$  eine Abkürzung für  $\langle \text{NP}[\text{nom}] \rangle \vee \langle \text{NP}[\text{nom}], \text{NP}[\text{acc}] \rangle$ , und  $\vee$  steht für Disjunktion.

Im Gegensatz zu *essen* steht *treten*. Der Satz *Karl tritt* beschreibt Situationen, in denen Karl einfach nur in die Luft tritt (z. B. beim Karate-Training) und Situationen, in denen Karl jemanden oder etwas tritt. Die Rolle des Getretenen ist nicht notwendigerweise vorhanden. Das kann man nur durch zwei verschiedene Lexikoneinträge beschreiben.<sup>8</sup> Deutlicher wird das noch bei einem Verb wie *entbinden* (Eisenberg, 1994, S. 81).

(3.26) Julia entbindet (Marco Polo von seinem Versprechen).

Das Objekt und die Präpositionalphrase kann zwar weggelassen werden, dann hat das Verb aber eine völlig andere Bedeutung. Für die beiden Formen von *entbinden* sind also zwei Lexikoneinträge anzunehmen.

<sup>8</sup> In (Pollard und Sag, 1987) werden Relationen, das heißt der Wert von CONT, wie folgt beschrieben:

$$\left[ \begin{array}{ll} \text{RELN} & \textit{treten} \\ \text{TRETENDER} & \boxed{1} \\ \text{GETRETENES} & \boxed{2} \\ \textit{cont} & \end{array} \right] \quad (\text{i})$$

Mit dieser Schreibweise kann *treten* mit nur einem, wenn auch sehr komplexen Lexikoneintrag, beschrieben werden:

$$\left[ \begin{array}{ll} \text{PHON} & \langle \textit{tritt} \rangle \\ \text{SYNSEM|LOC} & \left[ \begin{array}{l} \text{CAT} \left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \text{VFORM } \textit{fin} \\ \textit{verb} \end{array} \right] \end{array} \right] \\ \text{CONT} \left[ \begin{array}{ll} \text{RELN} & \textit{treten} \\ \text{TRETENDER} & \boxed{1} \end{array} \right] \end{array} \right] \end{array} \right] \wedge \quad (\text{ii}) \\ \left( \left[ \begin{array}{l} \text{SYNSEM|LOC|CAT|SUBCAT} \langle \text{NP}[\textit{nom}] \boxed{1} \rangle \end{array} \right] \vee \right. \\ \left. \left[ \begin{array}{l} \text{SYNSEM|LOC} \left[ \begin{array}{l} \text{CAT|SUBCAT} \langle \text{NP}[\textit{nom}] \boxed{1}, \text{NP}[\textit{acc}] \boxed{2} \rangle \\ \text{CONT|GETRETENES} \boxed{2} \end{array} \right] \right] \right] \end{array} \right)$$

Wie in (Pollard und Sag, 1994, Kapitel 8.5) festgestellt wurde, ist durch den Typ *cont* bzw. *circumstance* oder *psoa* die Art der möglichen Rollen einer Relation nicht eingeschränkt. Deshalb ist es besser, Relationen als Subtypen des Typs *psoa* aufzufassen. Vergleiche hierzu auch (Devlin, 1992, S. 115).

## 3.5. Abkürzungen

Symbol	Teilspezifikation
$NP[nom]_{[3,sg,fem]}$	$\left[ \begin{array}{l} \text{LOC} \\ \left[ \begin{array}{l} \text{CAT} \\ \left[ \begin{array}{l} \text{HEAD} \\ \left[ \begin{array}{l} \text{CAS } \textit{nom} \\ \textit{noun} \end{array} \right] \\ \text{SUBCAT } \langle \rangle \end{array} \right] \end{array} \right] \\ \text{CONT} \\ \left[ \begin{array}{l} \text{IND} \\ \left[ \begin{array}{l} \text{PER } 3 \\ \text{NUM } \textit{sg} \\ \text{GEN } \textit{fem} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$
$NP_{expl}$	$\left[ \begin{array}{l} \text{LOC} \\ \left[ \begin{array}{l} \text{CAT} \\ \left[ \begin{array}{l} \text{HEAD} \\ \left[ \textit{noun} \right] \end{array} \right] \\ \text{SUBCAT } \langle \rangle \end{array} \right] \\ \text{CONT} \\ \left[ \textit{expl} \right] \end{array} \right] \end{array} \right]$
$\bar{N}: \boxed{1}$	$\left[ \begin{array}{l} \text{LOC} \\ \left[ \begin{array}{l} \text{CAT} \\ \left[ \begin{array}{l} \text{HEAD} \\ \left[ \textit{noun} \right] \end{array} \right] \\ \text{SUBCAT } \langle \textit{DET} \rangle \end{array} \right] \\ \text{CONT} \\ \boxed{1} \end{array} \right] \end{array} \right]$
$NP \boxed{1}$	$\left[ \begin{array}{l} \text{LOC} \\ \left[ \begin{array}{l} \text{CAT} \\ \left[ \begin{array}{l} \text{HEAD} \\ \left[ \textit{noun} \right] \end{array} \right] \\ \text{SUBCAT } \langle \rangle \end{array} \right] \\ \text{CONT} \\ \left[ \begin{array}{l} \text{IND REF} \\ \left[ \boxed{1} \right] \\ \textit{ref} \end{array} \right] \end{array} \right] \end{array} \right]$

Je nach Kontext werden diese Abkürzungen für Zeichen oder *synsem*-Objekte verwendet. Wenn von Zeichen die Rede ist, muß man den Pfad LOC durch SYNSEM|LOC ersetzen.

## Literatur

Die Frage nach der Realität wird sehr anschaulich in Lems *Sterntagebüchern* (Lem, 1976) und im *Futurologischen Kongreß* (Lem, 1979) beantwortet. Daß sich die Existenz von Dingen, ja selbst unsere Existenz nicht beweisen läßt, kann man auch (Galouye, 1965) entnehmen. Will man Semantik betreiben, so muß man die Realität, und sei es auch nur als Arbeits-