

# Head-driven Phrase Structure Grammar

## Eine Einführung in HPSG

Seminararbeit im Rahmen des Seminars  
„Syntaxtheorien und computerlinguistische Praxis“  
bei Prof. Dr. M. Hess, lic.phil. G. Schneider und lic.phil. S. Clemenide  
(Erster Teil)

Juliette Huber  
Krummenlandstrasse 21  
5107 Schinznach-Dorf  
Tel.: 056/443 20 38  
durian@freesurf.ch

Davina Rodgers  
Kapfstrasse 9  
8032 Zürich  
Tel.: 01/422 49 50  
DavinaRodgers@access.unizh.ch

Mai 2000

---

## Inhaltsverzeichnis

<b>1. EINLEITUNG</b>	<b>4</b>
<b>1.1. VORLÄUFERTHEORIEN</b>	<b>4</b>
1.1.1. CATEGORIAL GRAMMAR	4
1.1.2. GENERALIZED PHRASE STRUCTURE GRAMMAR (GPSG)	5
1.1.3. ARC PAIR GRAMMAR	6
1.1.4. LEXICAL-FUNCTIONAL GRAMMAR (LFG)	7
<b>2. MERKMALSTRUKTUREN UND SYNTAKTISCHE KATEGORIEN</b>	<b>7</b>
<b>2.1. MERKMALE</b>	<b>7</b>
2.1.1. SORTIERTE UND TYPISIERTE MERKMALSTRUKTUREN	11
2.1.2. SORTENHIERARCHIE	12
<b>2.2. SYNTAKTISCHE KATEGORIEN</b>	<b>12</b>
2.2.1. DAUGHTERS, HEADS UND DAS MERKMAL SUBCAT	14
<b>3. GRAMMATIK</b>	<b>16</b>
<b>3.1. GRAMMATIK-PRINZIPIEN</b>	<b>16</b>
3.1.1. HEAD-FEATURE PRINZIP	17
3.1.2. SUBKATEGORISIERUNGSPRINZIP	18
3.1.3. WEITERE GRAMMATIK-PRINZIPIEN	19
3.1.3.1. Das ID-Prinzip	19
3.1.3.2. Das Marking-Prinzip	19
3.1.3.3. Das Spec-Prinzip	19
3.1.3.4. Das Nonlocal-Feature-Prinzip	20
3.1.3.5. Das Trace-Prinzip	20
3.1.3.6. Das Weak-Coordination-Prinzip	20
3.1.3.7. Das Semantics-Prinzip	20
3.1.3.8. Das Raising-Prinzip	20

<b>3.2. GRAMMATIK-REGELN</b>	<b>20</b>
3.2.1. REGELSCHEMA 1 (HEAD-SUBJECT SCHEMA)	21
3.2.2. REGELSCHEMA 2 (HEAD-COMPLEMENT SCHEMA)	22
3.2.3. REGELSCHEMA 3 (HEAD-SUBJECT-COMPLEMENT SCHEMA)	24
3.2.4. WEITERE REGELN	25
3.2.4.1. Regel 4 (Head-Marker Schema)	25
3.2.4.2. Regel 5 (Head-Adjunct Schema)	25
3.2.4.3. Regel 6 (Head-Filler Schema)	25
<b>3.3. LEXIKON</b>	<b>25</b>
3.3.1. VERERBUNGSPRINZIPIEN	26
3.3.2. LEXIKALISCHE REGELN	27
<b>3.4. ZUSAMMENFASSUNG</b>	<b>28</b>
<b>4. BEHANDLUNG EINIGER PHÄNOMENE IN HPSG</b>	<b>29</b>
<b>4.1. AGREEMENT</b>	<b>29</b>
<b>4.2. KOMPLEMENTE UND ADJUNKTE</b>	<b>30</b>
<b>4.3. SUBJEKTE UND PRÄDIKATE</b>	<b>31</b>
<b>4.4. GRAMMATISCHE FUNKTIONEN</b>	<b>32</b>
<b>4.5. WH-FRAGEN</b>	<b>32</b>
<b>4.6. SATZFRAGEN</b>	<b>34</b>
4.6.1 HILFSVERBEN	34
4.6.2. INVERSION	35
<b>5. SCHLUSSWORT</b>	<b>36</b>
<b>6. BIBLIOGRAFIE</b>	<b>37</b>
<b>LITERATUR</b>	<b>37</b>
<b>VORLESUNGEN / SKRIPTS</b>	<b>37</b>
<b>HOME PAGES</b>	<b>38</b>

## 1. Einleitung

Die Head-driven Phrase Structure Grammar, auf Deutsch Kopforientierte Phrasenstrukturgrammatik genannt, wurde erstmals 1985 von Carl Pollard präsentiert und 1987 im Buch *Information-Based Syntax and Semantics, Volume 1: Fundamentals* von Carl Pollard und Ivan Sag ausführlich vorgestellt.

Bei der Head-driven Phrase Structure Grammar (nachfolgend mit HPSG abgekürzt) handelt es sich um eine generative Grammatiktheorie aus der Familie der Unifikationsgrammatiken. Unifikationsbasierte Grammatiktheorien wie die Phrasenstrukturgrammatiken werden auch constraint-basiert (bedingungs-basiert) genannt. Diesen Theorien eigen ist, dass sie linguistische Information u.a. mit Hilfe von Merkmalstrukturen repräsentieren. Die linguistische Information lässt sich auf einer Darstellungsebene (auf der Oberflächenstruktur) als deklaratives Wissen darstellen. Diese Darstellung auf einer Satzebene wird auch monostratale Repräsentation genannt; die Struktur eines Satzes besteht aus einem einzigen Syntaxbaum.<sup>1</sup>

### 1.1. Vorläufertheorien

HPSG verbindet Elemente aus der Categorical Grammar (CG, Kategoriale Grammatik), der Generalized Phrase Structure Grammar (GPSG), der Arc Pair Grammar (APG) und der Lexical-Functional Grammar (LFG) miteinander. Zum besseren Verständnis fügen wir hier kurze Definitionen dieser soeben erwähnten Vorläufertheorien an. Wir stützen uns hierbei vor allem Bussmann (1990) und Borsley (1997).

#### 1.1.1. Categorical Grammar

Die Categorical Grammar wurde zwischen den beiden Weltkriegen von dem polnischen Logiker Ajdukiewicz entwickelt. In der Categorical Grammar hat die HPSG-Idee, dass

---

<sup>1</sup> Für die computerlinguistische Forschung sind Phrasenstrukturgrammatiken (PSG's) die wohl bedeutendsten Syntaxtheorien. Werke zur Computerlinguistik wie die von Pereira und Shieber (1987) oder Alshawi (1992) stützen sich auf PSG's.

Kategorien Informationen darüber erhalten sollen, mit welchen anderen Kategorien sie kombinierbar sind, ihren Ursprung.<sup>2</sup>

Alle Kategorien (abgesehen von einer kleinen Anzahl Basiskategorien) haben die Form  $\alpha/\beta$  oder  $\beta/\alpha$ , wobei  $\alpha$  und  $\beta$  selbst Kategorien sind. Ein Ausdruck vom Format  $\alpha/\beta$  verbindet sich mit einem ihm folgenden Ausdruck vom Format  $\beta$  zu einem Ausdruck vom Format  $\alpha$ , und ein Ausdruck vom Format  $\beta/\alpha$  verbindet sich mit einem ihm vorangehenden Ausdruck vom Format  $\beta$  zu einem Ausdruck vom Format  $\alpha$  (so kann beispielsweise VP als NP/S dargestellt werden). Diese Kategorien werden nun als mathematische Funktionen aufgefasst. Da die Symbole als arithmetische Bruchzahlen interpretierbar sind, nennt man die oben vorgestellten Verbindungsmöglichkeiten auch Kürzungsregeln ( $\alpha/\beta \times \beta = \alpha$ ). Ist das Endresultat der Kürzungsregeln S, so handelt es sich um einen wohlgeformten Satz.<sup>3</sup>

### 1.1.2. Generalized Phrase Structure Grammar (GPSG)

Wie die HPSG ist auch die GPSG eine Unifikationsgrammatik. Sie wurde 1981/1982 von Gerald Gazdar entwickelt mit dem Ziel, der Transformationsgrammatik ein Modell entgegenzustellen, das formal eingeschränkter ist und mit einer einzigen Repräsentationsebene auskommt. Der Grammatikformalismus stellt ein komplexes System von Regeln und Bedingungen dar, die die Wohlgeformtheit der lokalen Bäume und damit die Grammatikalität des ganzen Satzes bestimmen.

Die Phrasenstrukturregeln der GPSG entsprechen einer Version der X-Bar-Theorie; Metaregeln, die aus Phrasenstrukturregeln andere Phrasenstrukturregeln erzeugen, repräsentieren viele der syntaktischen Gesetzmässigkeiten, die in Transformationsgrammatiken durch Transformationen beschrieben sind (z. B. die Bildung von Passivsätzen). Jede Kategorie in der syntaktischen Struktur muss den Feature Cooccurrence Restrictions und den Feature Specification Defaults genügen. Die Merkmalsweitergabe wird gesteuert durch drei globale Bedingungen: (a) Head Feature Convention<sup>4</sup>, (b) Foot Feature Principle<sup>5</sup> und (c) Control Agreement Principle.<sup>6</sup>

---

<sup>2</sup> Dazu siehe Kap. 2.2.1.

<sup>3</sup> Borsley (1997), S. 135.

<sup>4</sup> Die Head Feature Convention entspricht dem Head Feature Principle in der HPSG; siehe Kap. 3.1.1.

### 1.1.3. Arc Pair Grammar

Die Arc Pair Grammar ist eine neuere Entwicklung der Relationalen Grammatik; letztere wurde als Modell einer Universalgrammatik von D. M. Perlmutter, P. M. Postal, D. E. Johnson u. a. entwickelt und stellt einen Gegenentwurf zur Transformationsgrammatik der sechziger Jahre dar.

Die Relationale Grammatik geht von der Grundannahme aus, dass grammatische Relationen (wie Subjekt und Objekt) eine zentrale Rolle in der Syntax natürlicher Sprachen spielen und distanziert sich somit von anderen universalgrammatischen Modellen, die von Konstituentenstrukturbegriffen ausgehen. Die Relationale Grammatik vertritt die Meinung, dass grammatische Relationen nicht weiter analysierbare, primitive Konzepte sind, und dass syntaktische Konstituentenstruktur-Repräsentationen ungeeignet sind für die Beschreibung universeller Regularitäten. Sätze werden stattdessen mittels relationaler Netzwerke analysiert; diese Netzwerke enthalten einen Satzknoten, von dem 'Bögen' für das Prädikat und dessen Argument ausgehen. Jedes abhängige Element steht auf jeder Beschreibungsebene in genau einer grammatischen Relation zu seinem regierenden Satzknoten. Die wichtigsten grammatischen Relationen sind (a) Subjekt (1-Relation), (b) direktes Objekt (2-Relation) und (c) indirektes Objekt (3-Relation). Konstituentenstruktur, lineare Abfolge und morphologische Markierung der Satzglieder werden nicht berücksichtigt.

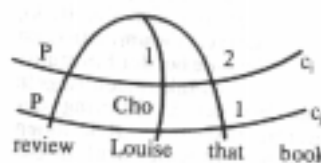


Abbildung 1

Abbildung 1 zeigt den Bogengraph für den Satz *That book was reviewed by Louise*, wobei  $c_i$  und  $c_j$  zwei syntaktische Ebenen darstellen. Die Ebene  $c_i$  stellt den 'normalen',

<sup>5</sup> Das Foot Feature Principle entspricht dem Nonlocal-Feature-Prinzip in der HPSG; siehe Kap. 3.1.3.4.

<sup>6</sup> Genauere Informationen bietet das im Rahmen dieses Seminars geschriebene Skript zur Sitzung vom 13. April 2000.

unmarkierten, entsprechenden Aussagesatz dar, während Ebene  $c_j$  die Umwandlung zum Passivsatz zeigt, wo die vorherige 1-Relation, das Subjekt, zum sogenannten Chômeur (Cho), also einem obliquen Partizipanten, demoviert wird.

#### 1.1.4. *Lexical-Functional Grammar (LFG)*

Die LFG ist eine generative Grammatiktheorie aus der Klasse der Unifikationsgrammatiken, die Ende der siebziger Jahre von J. Bresnan und R. Kaplan entwickelt wurde. Sie misst den grammatischen Relationen Subjekt, direktes/indirektes Objekt u. a. grosse Bedeutung bei. Die LFG geht davon aus, dass viele der syntaktischen Gesetzmässigkeiten, die in der Transformationsgrammatik durch Transformationen beschrieben wurden, lexikalischer Natur sind und daher auch nur im Lexikon repräsentiert werden können. Die LFG unterscheidet zwei Ebenen der syntaktischen Repräsentation; die C-Struktur<sup>7</sup> und die F-Struktur<sup>8</sup>, die parallel von den Phrasenstrukturregeln erzeugt werden.<sup>9</sup>

## 2. Merkmalstrukturen und Syntaktische Kategorien

Kapitel 2 liefert eine Einführung in die grundlegenden Beschreibungsmittel von HPSG, die Merkmalstrukturen. Zuerst soll auf die Darstellung von Merkmalstrukturen im Allgemeinen eingegangen werden (Repräsentation, Notation, hierarchische Gliederung), danach auf die für HPSG charakteristischen typisierten Merkmalstrukturen.

### 2.1. Merkmale

In der HPSG sind alle linguistischen Einheiten durch Merkmalstrukturen (feature structures) repräsentiert. Der Einsatz von Merkmalstrukturen ist für Unifikationsgrammatiken sehr typisch, und in HPSG wird die Phrasenstruktur selbst

---

<sup>7</sup> Die C-Struktur entspricht dem HPSG-Merkmal DTRS; siehe Kapitel 2.2.1.

<sup>8</sup> Die F-Struktur entspricht dem HPSG-Merkmal SUBCAT; siehe *ibid.*

<sup>9</sup> Für ausführlichere Erklärungen steht das Skript zur Sitzung vom 6. April 2000 zur Verfügung.

innerhalb einer Merkmalstruktur dargestellt. Diese Merkmalstrukturen werden nach Pollard / Sag (1988) *Zeichen* genannt und beinhalten

1. phonologische (PHON),
2. syntaktische (SYN) sowie
3. semantische (SEM)

Informationen über Konstituenten.<sup>10</sup> ‘Zeichen’ wird in diesem Kontext im Sinne Ferdinand de Saussures gebraucht und meint die Relation zwischen Form (signifiant) und Inhalt (signifié) eines linguistischen Ausdrucks. Als Zeichen betrachtet werden in HPSG jedoch nicht nur Wörter, Phrasen und Sätze, sondern auch ganze Diskurse. Die HPSG unterscheidet sich von anderen Syntaxtheorien dadurch, dass sie nur geordnete Merkmalstrukturen verwendet. Es werden keine anderen Beschreibungsmittel eingesetzt (wie beispielsweise in GPSG, deren ID-Schemata in Form von Phrasenstrukturregeln repräsentiert werden). Lexikoneinträge, Regeln und Prinzipien haben folglich gleichermassen die Form von geordneten Merkmalstrukturen.

Häufig werden für die Darstellung von Merkmalstrukturen Matrizen (attribute-value matrices oder AVMs) gewählt.

Bsp.:

```
-           -
| cat NP     |
| agr [ num singular ] |
-           -
```

Man findet jedoch unter den Repräsentationsmitteln auch gerichtete Merkmalsgraphen (directed acyclic graph, DAG). Folgender DAG erfasst die Struktur des englischen Pronomens *she*.<sup>11</sup>

<sup>10</sup> Siehe dazu Kapitel 2.2.

<sup>11</sup> Pollard / Sag (1994), S. 17.



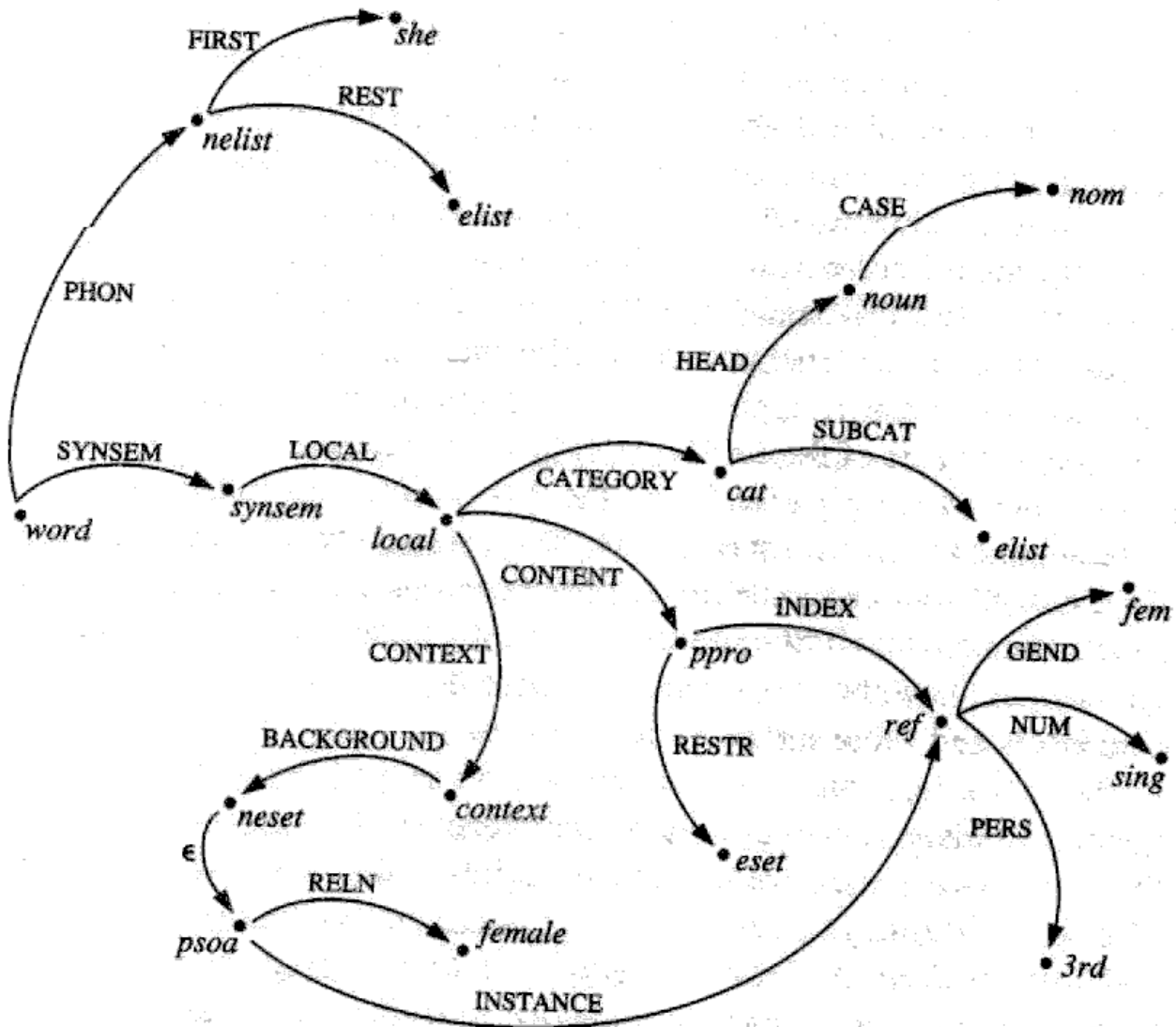


Abbildung 2

Ausserdem werden in HPSG unifikationsgrammatische Beschreibungsmittel wie Disjunktion, Negation und Implikation als Werte von Merkmalstrukturen eingesetzt. So ist es praktischer, beispielsweise bei Lexikoneinträgen für englische Verben die gleichlautenden Formen in einem Eintrag zu erfassen und mit der Information *nicht 3. pers. singular* zu versehen, anstatt für jede Person einen eigenen Eintrag zu erstellen. Die Verwendung der Disjunktion lässt sich zum Beispiel mit Angaben zum deutschen bestimmten Artikel *die* motivieren. Er kann u.a. feminin singular und feminin plural im

Nominativ und Akkusativ sein, was einem Kasusmerkmal [kasus(nom OR akk)] entsprechen würde.

Merkmale (attributes) müssen atomare Einheiten sein; Werte (values) jedoch sind atomar oder komplex, da sie wiederum aus Merkmalstrukturen bestehen können. HPSG benutzt als komplexe Erweiterungen von Merkmalstrukturen Mengen und Listen als Werte in Merkmalstrukturen. Diese unifikationsgrammatischen Beschreibungsmittel nennt man auch mengen- und listenwertige Merkmale (set values and list values); wie der Name es sagt, können diese Merkmale mehr als einen Wert besitzen. Zur Notation muss angefügt werden:

- { }: Mengenwertiges Merkmal. In einer Menge spielt die Reihenfolge der Merkmale keine Rolle.
- [ ]: Listenwertiges Merkmal. Geordnete Reihe von Merkmalen.
- <>: Notationelle Variante für listenwertige Merkmale (speziell für Valenz/Subkategorisierung).

Häufig kommt es innerhalb einer Merkmalstruktur vor, dass zwei oder mehrere Merkmale die selbe Merkmalstruktur haben wie ihr Wert. Dieses Teilen der gleichen Werte nennt man Koreferenz oder structure-sharing (auch re-entrancy). Koreferenzen werden meist durch nummerierte Kästchen dargestellt, die auch tags genannt werden.<sup>12</sup> Koreferenzen repräsentieren also Abhängigkeiten zwischen Konstituenten.

Beispiel für zwei koreferente Merkmalsmatrizen:

```
-           -
| cat S      |
| agr [1] [person 3] |
| subject [agr [1] ] |
-           -
```

ist identisch mit

```
-           -
```

---

<sup>12</sup> Wir verwenden jedoch auch die folgende Notation für Koreferenz: [1].

---

```

I cat S          |
I agr   [person 3] |
I subject [person 3] |
-                -

```

### 2.1.1. Sortierte und typisierte Merkmalstrukturen

Merkmalstrukturen können also als Menge von Merkmal-Wert Paaren gesehen werden, wobei “jedes Merkmal ein eindeutiges Symbol ist und jeder Wert entweder ein Symbol, ein Verweis oder selbst wieder eine Merkmalstruktur. Eine leere Merkmalstruktur kann als Variable interpretiert werden.”<sup>13</sup>

Eine typisierte (oder sortierte) Merkmalstruktur nun wird so definiert:

“A sorted feature structure consists of a set of features, where each feature consists of an attribute and a value. The attributes appropriate to a particular feature structure are determined by its sort.”<sup>14</sup>

Da es Abhängigkeiten zwischen gewissen Merkmalen gibt (das heisst, dass beispielsweise das Merkmal Kasus nur bei nominalen Kategorien auftritt), kann also bei bestimmten Kategorien die Einschränkung von Merkmalen gefordert werden: Man unterscheidet daher verschiedene Typen von Merkmalstrukturen.

Beispiel:

```

-                -
I cat NP          |
I agr [num singular] |
I   [person 3   ] |
-                -
nominal

```

Der Index (im Beispiel: nominal) am unteren Rand der Matrix bezeichnet den Typ (oder die Sorte) der Merkmalstruktur. Dieser Typ bestimmt, welche Merkmale innerhalb der Merkmalstruktur auftreten dürfen. Diese Typisierung bringt die Vorteile einer

---

<sup>13</sup> Volk (1998)

<sup>14</sup> Cooper (1996), S. 191.

zusätzlichen Kontrolle für die Wohlgeformtheit einer Merkmalstruktur sowie einer effizienteren Verarbeitung mit sich. In HPSG werden diese komplex erweiterten Strukturen typisierte (oder getypte) Merkmalstrukturen genannt.

Zusammenfassend lässt sich über die typisierten Merkmalstrukturen in HPSG sagen: Jeder Knoten ist mit einem Typensymbol bezeichnet, das erklärt, was für ein Objekt die Struktur modelliert. Es gibt folglich ein Typensymbol für jede Art von Konstrukt. Die Typisierung zeigt an, welche Attribute in einer Merkmalstruktur erscheinen können, da nicht alle Merkmale zwingend auch für alle Typen relevant sind.

### 2.1.2. Sortenhierarchie

Merkmalstrukturen in HPSG werden hierarchisch geordnet und über eine Vererbungsstruktur geregelt. Diese Struktur wird Sorten- oder Typenhierarchie (sort hierarchy) genannt: Das heisst, wenn zwei typisierte Merkmale unifiziert werden, so ist der Typ der resultierenden Merkmalstruktur der nächst allgemeinere Typ in der Hierarchie. Ein 'Übertyp' gibt an, welche Merkmale für alle seine Untertypen gelten; dabei kann jeder Untertyp weitere Merkmale hinzufügen. Unterknoten erben also die Merkmale von ihnen übergeordneten Knoten. Dabei gelten die für Konstituentenstrukturen üblichen Gesetze der (unmittelbaren) Dominanz sowie der (unmittelbaren) Konstituenz.

Die hierarchische Ordnung von Typen erlaubt es, Generalisierungen über die einzelnen Typen zu formulieren. Dies ist bei einer Implementierung effizienzsteigernd, da die Generalisierungen die Grammatik um ein Vielfaches verkleinern.

## 2.2. Syntaktische Kategorien

Wie schon in Kapitel 2.1. erwähnt, sind syntaktische Kategorien in HPSG komplexe Einheiten (Merkmalstrukturen), welche wiederum aus kleineren Elementen (auch Merkmalstrukturen oder atomaren Merkmalen, d.h. Symbolen oder Attributen) bestehen.

Im Folgenden wird eine Konstituente in HPSG, eine Merkmalstruktur des Typs *Zeichen*, näher erklärt. Wie schon erläutert, werden Zeichen als strukturierte Komplexe

von phonologischer, syntaktischer, semantischer und ‘phrasenstrukturierter’ Information aufgefasst. Dabei gilt (erläutert sei eine kleine Auswahl):

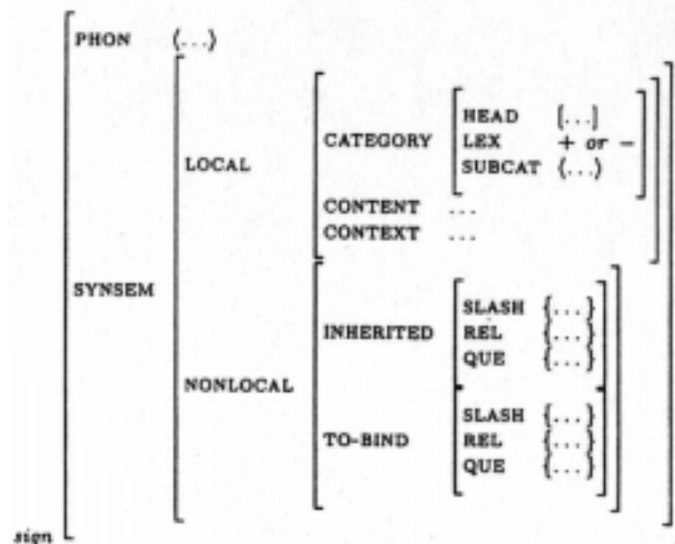


Abbildung 3

- PHON: Die phonologische Repräsentation PHON eines Zeichens besteht aus einem Wert, welcher in einer Liste die phonologische Struktur der Konstituente erfasst. Die Merkmalsrepräsentation des Lautgehalts eines Zeichens dient als Basis für phonologische und phonetische Interpretation.
- SYNSEM: Der Wert des SYNSEM Attributes kodifiziert die syntaktische und semantische Information einer Konstituente. Das Symbol SYNSEM wird weiter unterteilt in die komplexen Merkmalstrukturen LOCAL und NONLOCAL.
- LOCAL: Das Merkmal LOCAL beinhaltet CATEGORY (mit HEAD, SUBCAT und LEX), CONTENT und CONTEXT.
- CATEGORY: Zu HEAD und SUBCAT: Siehe Kapitel 2.2.1. LEX: LEX ist ein binäres Attribut, welches beschreibt, ob eine Einheit lexikalisch ist (+) oder nicht (-).
- CONTENT und CONTEXT: Beinhalten semantische Informationen.<sup>15</sup>

<sup>15</sup> Die HPSG stützt ihre Darstellung von semantischer Information auf die Situationssemantik. Die Situationssemantik basiert auf der Annahme, dass es eine einzige Welt gibt, über die gültige Aussagen über darin vorkommende Situationen gemacht werden können. Dabei gilt: „Meaning is analysed as a

- NONLOCAL: Wird bei der Behandlung ungebundener Abhängigkeitsphänomene eingesetzt. NONLOCAL enthält weiter INHERITED und TO-BIND.
- TO-BIND besteht jeweils aus SLASH (beschreibt *wh*-Abhängigkeiten), QUE (zeigt an, wo die Fragewörter in einem Satz stehen) und REL (markiert die Stellung von Relativpronomen). Die mengenwertigen Merkmale SLASH, QUE und REL sind bei Aussagesätzen mit kanonischer Wortstellung folglich leer.
- (...)

Von einer Beschreibung der weiteren Einteilungen sei hier abgesehen, da die verschiedenen Autoren (Pollard / Sag, Borsley, Cooper) in ihren Bezeichnungen und Erklärungen stark voneinander abweichen oder sich nicht auf eine Darstellung festlegen.

### 2.2.1. Daughters, Heads und das Merkmal Subcat

Da in HPSG Merkmalstrukturen sowohl als Abbildung von Satzkonstituenten als auch von lexikalischen Konstituenten eingesetzt werden können, müssen Satzkonstituenten besonders gekennzeichnet werden: Das Symbol Daughters (DTRS) erfüllt diese Funktion und benennt Zeichen, die Phrasen beschreiben. Die DTRS sind daher die einzelnen Bestandteile eines komplexen Ausdrucks (Bsp.: DTRS der NP *the house* sind ein Det *the* und ein N *house*). Meist ist das Symbol Daughters wiederum eine komplexe Merkmalstruktur und unterscheidet die HEAD-DTR von den COMP-DTRS. COMP-DTRS steht für complement daughters und ist eine Liste mit Beschreibungen von Merkmalstrukturen, in der jedes Element ein Komplement beschreibt.<sup>16</sup>

In HPSG bildet per definitionem die Verbalphrase die HEAD-DTR eines Satzes. Als Kopftochter von S wird folglich VP behandelt, nicht die Subjekts-NP. Da der Kopf als das obligatorische Element einer Phrase betrachtet wird, liefern zum Beispiel Imperativsätze, in denen kein Subjekt steht, oder Sätze mit disloziertem Subjekt (wie nachfolgendes Beispiel) Evidenz für diese Ansicht.

---

relation between situation types, and linguistic meaning relates utterance situation types to described situation types.“ Weitere Informationen in Borsley (1996), S. 68.

<sup>16</sup> Das Attribut DTRS entspricht ungefähr der c-Struktur in LFG.

Beispiel:

Who do you think [ \_ will win the race?]

Das Merkmal SUBCAT regelt die Subkategorisierung. Dieses Merkmal erhält nach Pollard / Sag (1994)<sup>17</sup> als Wert eine Liste mit sämtlichen Komplementen (nämlich SYNSEM-Objekten). Das Subjekt steht ebenfalls in dieser Liste. Das Merkmal SUBCAT besagt folglich, was für ein Subjekt und was für Komplemente ein Ausdruck benötigt.

Beispiel (nach Borsley<sup>18</sup>):

V[SUBCAT=< >] --> sleep

V[SUBCAT=<NP>] --> see

V[SUBCAT=<PP>] --> wait

V[SUBCAT=<NP, PP>] --> put

Man beachte, dass in dieser Repräsentation - ganz im Gegensatz zur obigen Behauptung - das Subjekt nicht in der SUBCAT-Liste erscheint. Dies liegt daran, dass in verschiedenen Varianten der HPSG für das Subjekt eine eigene SUBCAT-Liste eingeführt wird; diese Behandlung des Subjekts berücksichtigt die Tatsache, dass Subjekte gegenüber anderen Komplementen jeweils spezielle Stellungen einnehmen. Es wird in HPSG demnach manchmal mit zwei verschiedenen SUBCAT-Listen operiert, wobei in der einen (siehe oben eingeführtes Beispiel) nur mehr die Komplemente des Verbs erscheinen.

Nach Pollard / Sag (1994) wird das Subjekt in die SUBCAT-Liste aufgenommen und über das Subkategorisierungsprinzip (dazu Kapitel 3.1.2.) abgearbeitet. Im Gegensatz zu den Komplementen ist das Subjekt nicht Geschwister-Knoten zum Verb, sondern es wird in einem zweiten Schritt abgearbeitet.<sup>19</sup> Die SUBCAT-Liste ist geordnet; die Reihenfolge der darin vorkommenden Elemente wird durch den Grad

---

<sup>17</sup> Wir beziehen uns, falls nicht anders vermerkt, auf Pollard / Sag (1994), nicht auf deren erste Version einer HPSG (Pollard / Sag (1987)).

<sup>18</sup> Borsley (1997), S. 134.

<sup>19</sup> Pollard / Sag (1994), S. 33.

ihrer 'obliqueness' bestimmt, das heisst, dass das am wenigsten oblique Element zuerst kommt.<sup>20</sup>

Das Attribut SUBCAT zeigt den Grad der Sättigung eines Zeichens. Erst dann, wenn unter SUBCAT die leere Liste (< >) steht, sind alle Komplemente abgearbeitet worden; das Zeichen ist gesättigt.

### 3. Grammatik

HPSG besteht aus Grammatik-Prinzipien, Grammatik-Regeln und Lexikon-Einträgen, welche alle durch Merkmalstrukturen repräsentiert werden. Jede zulässige Konstituente wird von einer Merkmalstruktur beschrieben; diese Merkmalstruktur muss gleichzeitig von jedem Grammatik-Prinzip subsumiert werden sowie entweder von einer Grammatik-Regel (wenn die Konstituente eine Phrase ist), oder von einem Lexikon-Eintrag (wenn die Konstituente lexikalisch ist). Alle diese Merkmalstrukturen müssen notwendigerweise von der Sorte *Zeichen* sein.<sup>21</sup> HPSG besteht also aus:

1. Grammatik-Prinzipien,
2. Grammatik-Regeln (Immediate Dominance-Regeln),
3. Lexikon-Einträgen.

Da der HPSG-Formalismus stark lexikalisiert ist (das heisst, dass das Lexikon einen grossen Teil der grammatischen und syntaktischen Information enthält), sind die Grammatikregeln und -prinzipien auf ein Minimum reduziert.

#### 3.1. Grammatik-Prinzipien

HPSG besteht aus allgemeinen und - falls nicht explizit als sprachspezifisch deklariert - universell gültigen Grammatikprinzipien. Jede wohlgeformte Merkmalstruktur vom Typ *Zeichen* muss von diesen Prinzipien subsumiert werden.

---

<sup>20</sup> Dazu Cooper (1996), S. 193: „[...]Obliqueness is a function of grammatical relation: subjects are less oblique than direct objects, which are in turn less oblique than indirect objects.“

<sup>21</sup> Nach Cooper (1996), S. 192.



### 3.1.1. Head-Feature Prinzip

Das Head Feature-Prinzip stellt sicher, dass einem Kopf und seiner phrasalen Projektion die Head Features gemeinsam sind. In Borsley's Worten:

„The HEAD value of a headed phrase is identical to the HEAD value of the head daughter.“<sup>22</sup>

In Form einer Merkmalstruktur sieht das Head Feature-Prinzip folgendermassen aus:<sup>23</sup>

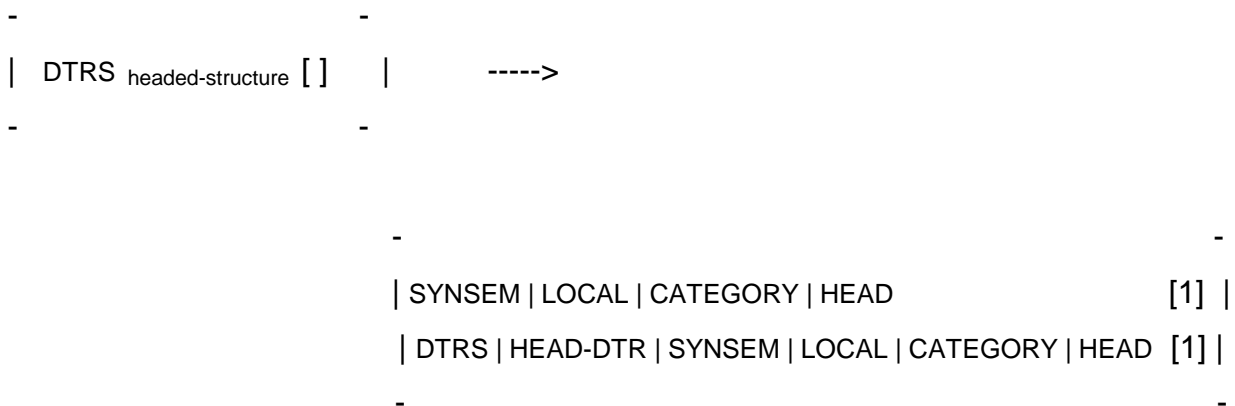


Abbildung 4

Abbildung 4 ist eine vereinfachte Darstellung: Es werden nur die Pfade zu den interessierten Werten gezeigt (vergleiche Abbildung 3). Was man also hier sieht ist der Weg über die Merkmale, die man in einer vollständigen Merkmalstruktur gehen müsste, um zu diesem bestimmten Merkmal HEAD und seinem Wert zu gelangen.

Das Head Feature-Prinzip besagt also, dass jede Merkmalstruktur, in der das Merkmal DTRS vorkommt - und zwar mit einer Merkmalstruktur vom Typ *headed-structure* als Wert - mit der Konsequenz dieser Implikation unifiziert werden muss. Die

<sup>22</sup> Borsley (1996), S. 50.

<sup>23</sup> Anmerkung: Zur Vereinfachung der Darstellung können komplexe Merkmalstrukturen durch Pfade ersetzt werden. Bsp.: A| B| C ist gleichbedeutend mit [A[B[C]]].

Konsequenz besagt mit Hilfe von Koreferenzsymbolen, dass das HEAD-Merkmal der Mutter und das HEAD-Merkmal der Tochter (HEAD-DTR) referenz-identisch sein müssen. Eine graphische Darstellung desselben Prinzips sieht aus wie folgt (dabei steht  $\delta$  für die HEAD-DTR):<sup>24</sup>



Abbildung 5

### 3.1.2. Subkategorisierungsprinzip

Das Subkategorisierungsprinzip wird auch Valenzprinzip genannt. Es sorgt dafür, dass ein Satz (respektive ein Verb) die korrekte Anzahl und Art von Komplementen erhält. Es lässt sich wie folgt formulieren:

„Wenn ein Kopf und seine Mutter unterschiedliche Werte für [das Merkmal] SUBCAT haben, dann müssen die Kategorien, die in der SUBCAT-Liste des Kopfes, nicht jedoch die, die in der SUBCAT-Liste der Mutter stehen, mit den Schwestern des Kopfes identisch sein.“<sup>25</sup>

Cooper beschreibt das Subkategorisierungsprinzip so:

„...for any headed phrase, the subcategorization requirements of the phrase are those of the head daughter less those filled by any complement daughters.“<sup>26</sup>

Graphisch dargestellt sieht dies etwa so aus:

<sup>24</sup> Sag / Wasow (1999)

<sup>25</sup> Borsley (1997), S. 135.

<sup>26</sup> Cooper (1996), S. 193.

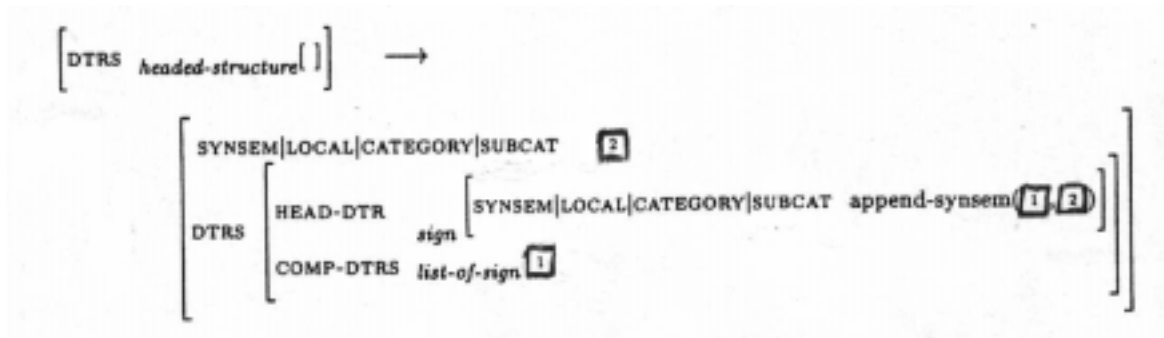


Abbildung 6

### 3.1.3. Weitere Grammatik-Prinzipien

Es existieren einige weitere universelle Grammatikprinzipien, deren einzelne Behandlung den Rahmen dieser Arbeit jedoch übersteigen würde. Es folgt daher nur eine kurze Skizze ihrer jeweiligen Verwendungszwecke.

#### 3.1.3.1. Das ID-Prinzip

Jede headed phrase muss genau einer der Immediate Dominance Regeln (ID-Regeln)<sup>27</sup> genügen.

#### 3.1.3.2. Das Marking-Prinzip

Das Marking-Prinzip kommt bei mit *that*, *for* u.ä. eingeleiteten Komplementsätzen zum Zuge.

#### 3.1.3.3. Das Spec-Prinzip

Das Spec-Prinzip wird zur Behandlung von Quantifikatoren und Possessivphrasen eingesetzt.

<sup>27</sup> Siehe dazu Kapitel 3.2.

#### 3.1.3.4. Das Nonlocal-Feature-Prinzip

Das Nonlocal-Feature-Prinzip ist die HPSG-Entsprechung zum Foot-Feature-Prinzip in GPSG und behandelt ungebundene Abhängigkeitsphänomene.

#### 3.1.3.5. Das Trace-Prinzip

Das Trace-Prinzip wird ebenfalls bei ungebundenen Abhängigkeitsphänomenen angewendet; es setzt Bedingungen für die Verwendung von Traces.

#### 3.1.3.6. Das Weak-Coordination-Prinzip

Das Weak-Coordination-Prinzip regelt das Verhältnis zwischen koordinierten Töchtern und deren Mutterknoten.

#### 3.1.3.7. Das Semantics-Prinzip

Das Semantics-Prinzip ist eine Kombination aus drei einzelnen Prinzipien: Dem Content-Prinzip, dem Quantifier-Inheritance-Prinzip und dem Scope-Prinzip. Es ist für die Semantikkomponente zuständig, die in dieser Arbeit allerdings bewusst vernachlässigt wird.

#### 3.1.3.8. Das Raising-Prinzip

Anders als die vorhergehenden Prinzipien, die Bedingungen an Merkmalstrukturen stellen, ist das Raising-Prinzip eine Einschränkung der Form von Lexikoneinträgen. Dieses Prinzip behandelt Raising-Konstruktionen.

### **3.2. Grammatik-Regeln**

Obwohl HPSG als Phrasenstrukturgrammatik gilt, werden Regeln nicht mit formalen Rewrite Rules dargestellt wie in anderen Grammatiken. Da Phrasenstrukturbäume und phrasale Konstituenten in Merkmalstrukturen kodiert sind, werden Grammatikregeln als

Bedingungen für deren Wohlgeformtheit formuliert: Jede Grammatikregel hat die Form einer Merkmalstruktur und jede Struktur, die eine Phrase beschreibt, muss von einer Grammatikregel subsumiert werden. Die Grammatikregeln in HPSG werden Immediate Dominance-Regeln (ID-Schemata oder Phrase Structure-Regeln) genannt und dienen eigentlich als Templates<sup>28</sup> für zulässige lokale Phrasenstrukturbäume; sie lehnen sich also an die Regeln der X-bar Theorie an. Der stark lexikalische Charakter der HPSG hat zur Folge, dass nur sehr wenige, schematische Immediate Dominance Templates (Regelschemata) übrig bleiben. Tatsächlich wird die HPSG-Grammatik grösstenteils von nur zwei Regeln abgedeckt (Regeln 1 und 2, die im Folgenden dargestellt werden).

### 3.2.1. Regelschema 1 (Head-Subject Schema)

Das Regelschema 1, das der Aufgabe der Specifier-Regel aus der X-bar Theorie nachkommt, besagt, dass eine saturierte Phrase aus einem nichtlexikalischen Kopf und einem einzigen Komplement bestehen kann.

In Form einer Merkmalstruktur sieht dies wie folgt aus:<sup>29</sup>

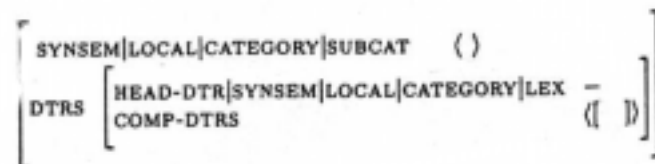


Abbildung 7

Dem Subkategorisierungsprinzip entsprechend muss der Kopf einer solchen Phrase genau ein Komplement erfordern; dieses Komplement muss mit dem Element in der Komplement-Tochter-Liste (COMP-DTRS) identisch sein.

Pollard / Sag definieren das Regelschema 1 als „a saturated ([SUBCAT < >]) phrase with DTRS value of sort *head-comp-struct* in which the HEAD-DTR value is a phrasal sign and the COMP-DTRS value is a list of length one.“<sup>30</sup> In ihrer Darstellung:<sup>31</sup>

<sup>28</sup> Templates (englisch ‘Muster’ oder ‘Schablone’) sind deklarierte Merkmalstrukturen, die über ihren Namen innerhalb anderer Merkmalstrukturen aufgerufen werden können.

<sup>29</sup> Vergleiche dazu Pollard / Sag (1994), S. 402.

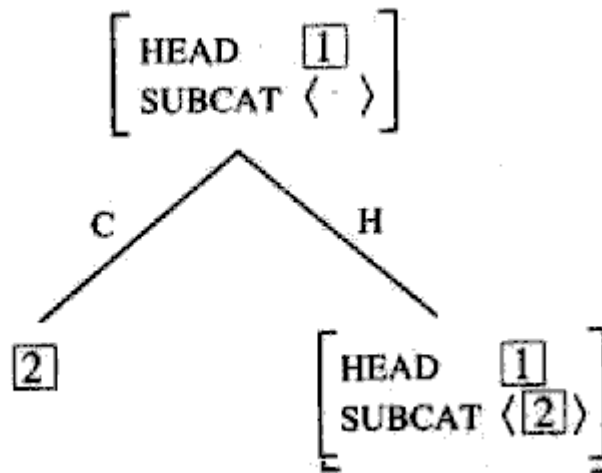


Abbildung 8

Das Regelschema 1 lizenziert folglich gesättigte Phrasen mit einer phrasalen Kopftochter und einer weiteren Tochter, die Komplement ist. Als direkte Konsequenz des Subkategorisierungsprinzips muss der Wert des SYNSEM-Merkmals der Komplement-Tochter identisch mit dem einzigen Element in der SUBCAT-Liste der Kopftochter sein; ausserdem muss dieses Element das Subjekt (das am wenigsten oblique Element der SUBCAT-Liste) des lexikalischen Kopfes sein.

Dieses Regelschema deckt die gängigen Phrasenstrukturregeln ‚S --> NP VP‘ (wobei VP der Kopf von S ist)<sup>32</sup>, ‚NP --> Det Nom‘ oder ‚NP --> NP[Gen] Nom‘ ab. In der X-Bar-Theorie wird dasselbe so ausgedrückt: ‚X‘ --> Spec X‘.

### 3.2.2. Regelschema 2 (Head-Complement Schema)

Das Regelschema 2, das der Komplement-Regel in der X-bar Theorie entspricht, besagt, dass eine Phrase aus einem (nichtinvertierten) lexikalischen Kopf zusammen mit allen seinen Komplementen ausser dem am wenigsten obliquen bestehen kann. In Form einer Merkmalstruktur wiederum sieht dieses Regelschema so aus:<sup>33</sup>

<sup>30</sup> *ibid.*, S. 38

<sup>31</sup> *ibid.*, S. 39.

<sup>32</sup> Man beachte: Das Subjekt wird hier als Komplement und somit Element der SUBCAT-Liste betrachtet.

<sup>33</sup> Cooper (1996), S. 195.

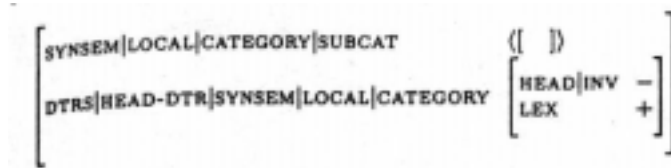


Abbildung 9

Das Regelschema 2 beschreibt eine wohlgeformte Phrase also auch so:

„An almost-saturated (SUBCAT list of length one) phrase with DTRS value of sort *head-comp-struct* in which the HEAD-DTR value is a lexical sign.“<sup>34</sup>

Die Abbildung 9 ist daher gleichbedeutend mit folgender Notationsweise<sup>35</sup>:

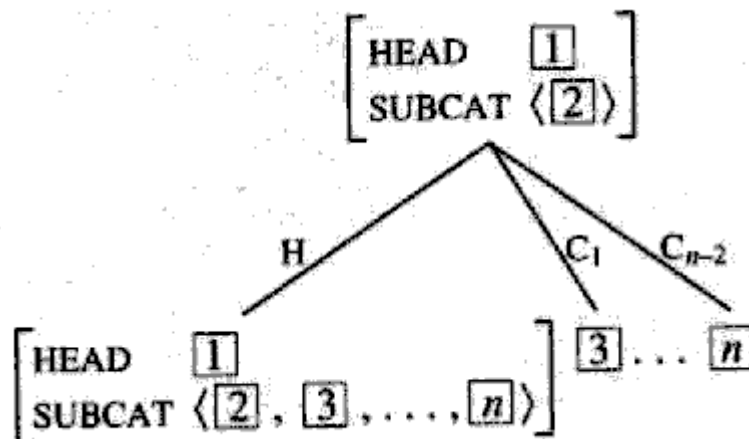


Abbildung 10

Das Regelschema 2 subsumiert folglich alle konventionellen Phrasenstrukturregeln, die eine Phrase als lexikalischen Kopf mit seinen Komplementen (ausser dem am wenigsten Obliquen, also dem Subjekt) expandieren; es sind dies sowohl alle VP-Regeln als auch Regeln für die Modifikation von Nomen mit Nondeterminier-

<sup>34</sup> Pollard / Sag (1994), S. 38.

<sup>35</sup> Pollard / Sag (1994), S. 39.

Komplementen.<sup>36</sup> Die entsprechende Regel aus der X-Bar-Theorie wird so ausgedrückt:  $X' \rightarrow X \text{ Comp}^*$ .

Die HPSG-Spezifikation lässt offen, ob alle Komplementsvalenzen auf einmal gesättigt werden oder ob sie schrittweise abgearbeitet werden. Das Regelschema 2 ist die vielleicht wichtigste syntaktische Grundregel.

### 3.2.3. Regelschema 3 (Head-Subject-Complement Schema)

Das Regelschema 3 beschreibt eine wohlgeformte Phrase als „a saturated ([SUBCAT <>]) phrase with DTRS value of sort *head-comp-struct* in which the HEAD-DTR value is a lexical sign“<sup>37</sup>.

Graphisch dargestellt:

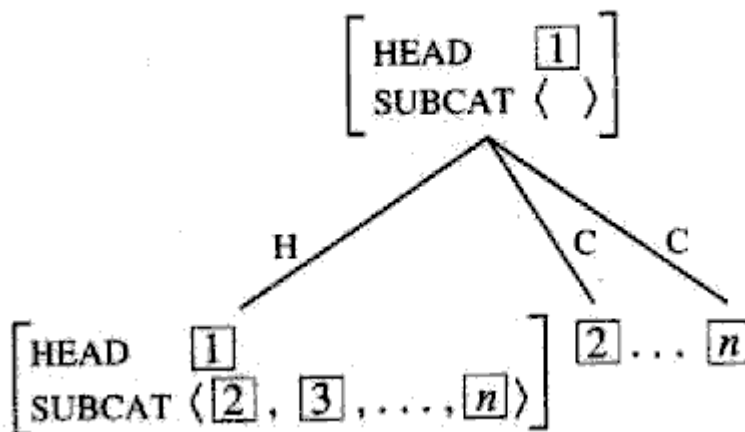


Abbildung 11

Als Konsequenz des Head-Feature Prinzips stellt die generelle Form von Regelschema 3 eine Phrase dar, in der alle Komplemente (einschliesslich dem Subjekt) als Schwestern des lexikalischen Kopfes realisiert werden.

<sup>36</sup> Nondeterminer-Komplemente sind solche, die nicht in Determinerstellung auftreten, also nicht Subjekt sind.

<sup>37</sup> Pollard / Sag, S. 40.



Diese Regel kommt bei Sprachen mit freier Konstituentenordnung zum Zuge (in der deutschen Sprache beispielsweise, wo die Komplemente mehr oder weniger frei angeordnet werden können, mit der Einschränkung, dass das Verb je nach Satzart satzinitial oder satzfinal stehen muss). Auch die Subjekt-Auxiliar-Inversion im Englischen wird so behandelt.

### 3.2.4. Weitere Regeln

Pollard / Sag führen drei weitere Grammatikregeln auf. An dieser Stelle seien nur kurz ihre jeweiligen Verwendungszwecke erwähnt.

#### 3.2.4.1. Regel 4 (Head-Marker Schema)

Regel 4 wird bei mit *that* und *for* eingeleiteten Komplementsätzen wie *I know [that John left]* angewendet.

#### 3.2.4.2. Regel 5 (Head-Adjunct Schema)

Pollard / Sag gehen davon aus, dass Adjunkte ihren zugehörigen Kopf selektieren. Schema 5 regelt die Art und Weise, in der ein Adjunkt und der Kopf, den es selektiert, kombiniert werden können.

#### 3.2.4.3. Regel 6 (Head-Filler Schema)

Regel 6 kommt bei ungebundenen Abhängigkeitsphänomenen zum Einsatz.

## 3.3. Lexikon

Das Lexikon beinhaltet in HPSG die meisten syntaktischen Informationen; es ist äusserst reich an Einträgen und sehr komplex. Trotzdem, so betont Borsley, ist das Lexikon in HPSG nicht überladen, denn es wird in Vererbungshierarchien strukturiert.<sup>38</sup>

---

<sup>38</sup> Borsley (1996), S. 38f.

Das heisst, das Lexikon wird von verschiedenen Vererbungsprinzipien organisiert, welche die Lexikon-Einträge um ein Vielfaches reduzieren. Das HPSG-Lexikon kann folglich nicht als unstrukturiertes Set von Wort-Merkmalstrukturpaaren angesehen werden.

Das Lexikon in HPSG besteht einerseits aus Hierarchien, die Merkmale vererben, und andererseits aus lexikalischen Regeln.

### 3.3.1. Vererbungsprinzipien

Die ursprünglich für Wissensrepräsentationen im Gebiet der Künstlichen Intelligenz verwendeten Vererbungshierarchien dienen in HPSG dazu, die lexikalischen Einträge einer gleichen Sorte unter einem einzigen Gattungseintrag zusammenzufassen. Als veranschaulichendes Beispiel nennt Borsley<sup>39</sup> das Wort *gives*: *gives* teilt seine Eigenschaft *3. pers. sg. präsens, ditransitiv* einerseits mit allen Verben der Form *3. pers. sg. präsens* sowie, andererseits, mit allen ditransitiven Verben. Somit können sämtliche Verben der Form *3. pers. sg. präsens, ditransitiv* in nunmehr zwei Gattungseinträgen erfasst werden. Der lexikalische Eintrag eines Wortes *give* besteht daher aus

1. einem Verweis auf diese beiden Gattungseinträge, und
2. den dem Wort eigenen phonologischen und semantischen Informationen.

Des Weiteren gilt, dass in HPSG die Gattungseinträge lexikalische Sorten sind und somit auch zur Vererbungsstruktur der der HPSG-Grammatik eigenen Vererbungshierarchie gehören. Jeder lexikalische Eintrag ist in allen relevanten Hierarchien (der Grammatik, des Lexikons) nach grammatischen Kriterien kreuzklassifiziert. Die Knoten oder Verzweigungsstellen einer jeden Hierarchie entsprechen typisierten Merkmalstrukturen, und die Abhängigkeiten - die Dominanzrelationen - innerhalb der Hierarchie sind mit Subsumptionsrelationen zu vergleichen. Jeder Mutterknoten gibt also an, welche Merkmale auf alle seine Töchter zutreffen, und jede Tochter kann weitere Merkmale hinzufügen.

---

<sup>39</sup> *ibid.*, S. 38.

Dieser Ansatz vermindert Redundanz im Lexikon. Die Organisation des HPSG-Lexikons erlaubt daher Verallgemeinerungen, welche die Anzahl der Einträge verringern.

### 3.3.2. Lexikalische Regeln

Weitere Generalisierungen, nämlich solche zwischen einzelnen Formen eines Wortes, werden durch lexikalische Regeln ausgedrückt. Eine lexikalische Regel für das Passiv beispielsweise verbindet Merkmalstrukturbeschreibungen von verbalen Grundformen mit Beschreibungen ihrer Passivpartizipien

Beispiel<sup>40</sup>:

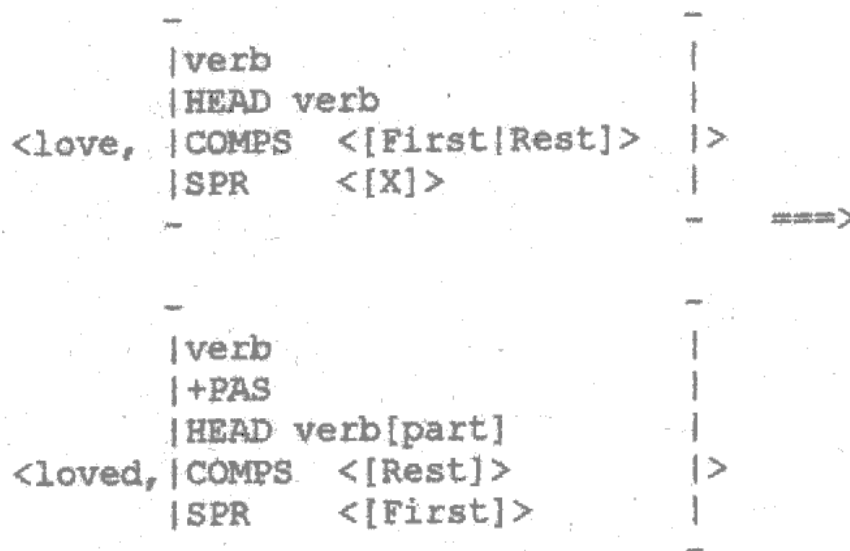


Abbildung 12

Gewisse lexikalische Regeln sind für die Bildung von flektierten Verbformen wie 3. pers. sg. aus Infinitiven zuständig:

Beispiel<sup>41</sup>: *walks* aus *walk*

<sup>40</sup> Schneider (1999), Skript zum Passiv.

<sup>41</sup> Schneider (1999), Skript zum Passiv.

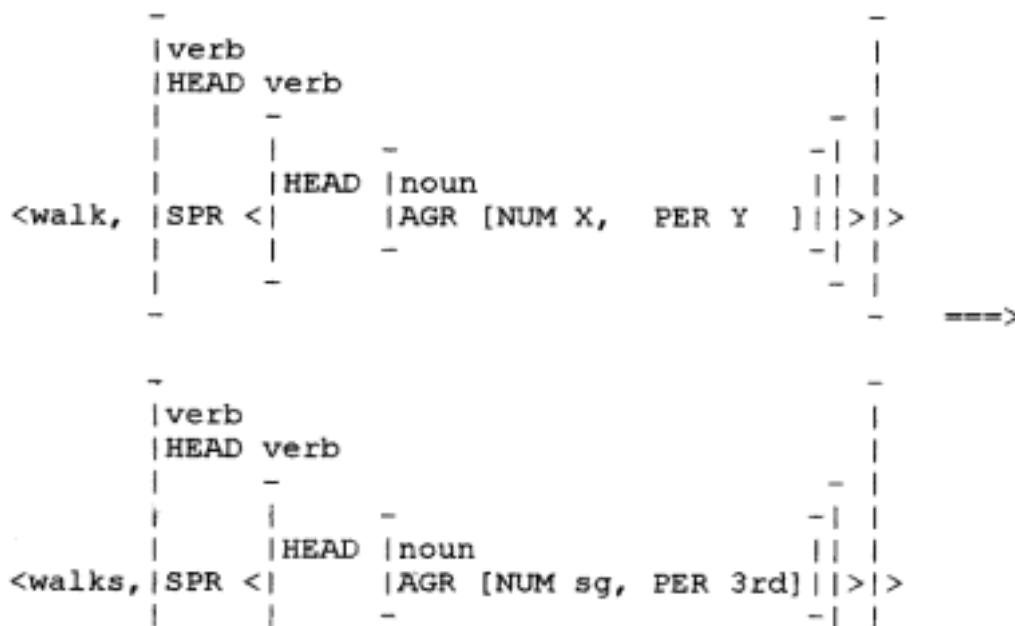


Abbildung 13

Andere lexikalische Regeln stellen Relationen zwischen allen Verbformen, verschiedenen Adjektivformen und nominalen Pluralformen her. Lexikalische Regeln in HPSG sollen also in erster Linie aus Basisformen flektierte Formen bilden. Mit Hilfe solcher Regeln kann nun das Lexikon ausschliesslich in Basisformen gehalten werden. Die Abarbeitung kann vor dem Parsing (offline) oder auch während dem Parsing (online) erfolgen. Eine offline-Abarbeitung bedeutet, dass der Parser schon von Anfang an mit einem Vollformenlexikon arbeitet und keine morphologischen Schritte mehr durchführt. Dem Geschwindigkeitsvorteil steht ein exponentiell gewachsenes Lexikon gegenüber. Gewisse derivationale Morphologieprozesse wie Nominalkomposition lassen sich für umfassende Grammatiken jedoch nur online durchführen.

### 3.4. Zusammenfassung

Die Grammatik in HPSG fungiert als Beschränkung für die linguistische Wohlgeformtheit der vorkommenden Zeichen. Das heisst, dass jedes wohlgeformte Zeichen mit der Grammatik unifizierbar sein muss: Unifikation operiert über

Merkmalstrukturen. Die Gesamtheit der korrekten Sätze in einer HPSG-Grammatik lässt sich mit der folgenden Formel erfassen:<sup>42</sup>

$$P_1 \text{ unifiziert mit } \dots \text{ unifiziert mit } P_i \text{ unifiziert mit } (R_m \vee \dots \vee R_n \vee L_1 \vee \dots \vee L_n)$$

wobei gilt:

$P_i$  sind die Grammatik-Prinzipien

$R_i$  sind die Grammatik-Regeln

$L_i$  sind die lexikalischen Einträge

$\vee$  Disjunktion

Die Grammatik in HPSG ist folglich die Disjunktion aller Regeln und aller lexikalischen Einträge, in Konjunktion gesetzt (unifiziert) mit den grammatischen Prinzipien.

## 4. Behandlung einiger Phänomene in HPSG

Im Hinblick auf die später (im zweiten Teil unserer Arbeit in Zusammenhang mit der Grammatik LinGO) behandelten Beispielsätze erscheint es angemessen, hier einige HPSG-Lösungsansätze theoretisch vorzustellen.<sup>43</sup>

### 4.1. Agreement

Pollard / Sag sehen Strukturen, die infolge von Kongruenzphänomenen identisch sein müssen als koindexiert an. Es wird also ein Merkmal INDEX benötigt. Die Form des Verbes zum Beispiel wird eingeschränkt, wenn die Grammatik Koreferenz zwischen dem Wert seines Merkmals INDEX mit einem von einem anderen Ausdruck festgelegten Index verlangt.

<sup>42</sup> Cooper (1996), S. 192.

<sup>43</sup> Wir stützen uns dabei auf Borsley (1997).

## 4.2. Komplemente und Adjunkte

Die Unterscheidung zwischen Komplementen und Adjunkten ist für alle Kategorien relevant; im Folgenden wird sie jedoch nur im Zusammenhang mit Verben erläutert werden.

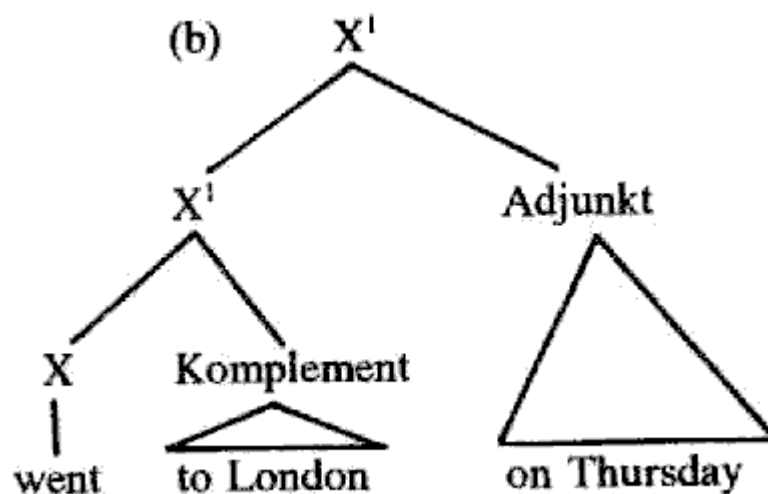
Komplemente sind mit spezifischen lexikalischen Köpfen auf eine Weise verknüpft, wie dies für Adjunkte nicht gilt: Bestimmte lexikalische Köpfe erscheinen gleichzeitig mit bestimmten Komplementen, während Adjunkte eines bestimmten Typs generell in jeder Phase möglich sind - unabhängig davon, was deren Kopf ist. Ausserdem tendieren Komplemente dazu, obligatorisch zu sein, während Adjunkte immer optional sind.

Beispiel: Optionales Komplement

Stefan was eating [the carrot].

Es wird allgemein angenommen, dass es einen strukturellen Kontrast zwischen Komplementen und Adjunkten gibt. Speziell geht man davon aus, dass Komplemente mit einer lexikalischen Kategorie verbunden sind und mit ihr eine verwandte intermediäre Kategorie bilden. Adjunkte hingegen verbinden sich mit einer beliebigen Kategorie und formen nur wieder eine Kategorie des gleichen Typs.

Beispiel:<sup>44</sup>



<sup>44</sup> Borsley (1997), S. 123.

Abbildung 14

HPSG benutzt zur Behandlung solcher Phänomene das Merkmal SUBCAT, das eine Liste von Kategorien als Wert annimmt.<sup>45</sup> Die SUBCAT-Liste soll genau jene Komplemente einschliessen, die der lexikalische Kopf fordert. Ist der Wert von SUBCAT eine leere Liste, so werden keine Komplemente gefordert bzw. zugelassen. Abgearbeitet wird die SUBCAT-Liste in Interaktion mit der Head-Complement-Regel<sup>46</sup> und dem Subkategorisierungsprinzip<sup>47</sup>.

Für Adjunkte hingegen wird ein Head-Feature MODIFIED (MOD) eingeführt, worin die Werte für die Merkmale INDEX und REST mit denen der Adjunkttochter übereinstimmen müssen. Den verschiedenen Methoden zur Behandlung von Adjunkten wird in HPSG grosse Beachtung geschenkt; es ist jedoch auch heute noch unklar, welche Analyse für Adjunkte korrekt ist.<sup>48</sup>

### 4.3. Subjekte und Prädikate

Im Englischen, Deutschen sowie in vielen anderen Sprachen besteht ein typischer Satz aus einem Subjekt und einem Prädikat. In der gleichen Weise, in der lexikalische Köpfe bestimmte Komplemente nehmen, verbinden sich Prädikate mit bestimmten Subjekten.

Pollard / Sag behandeln Subjekte als die Realisierungen einer weiteren Kategorie in der SUBCAT-Liste. Dies ist der Grund, weshalb die Head-Complement-Regel ausschliesst, dass ein Kopf sich mit allen Kategorien seiner SUBCAT-Liste verbindet: Die letzte Kategorie der SUBCAT-Liste, das Subjekt, ist von der Verbindung mit dem Kopf ausgenommen. Diese letzte Kategorie erscheint als die SUBCAT-Liste des Mutterknotens. Für die Verbindung von Subjekt und Prädikat wird nun eine Regel

---

<sup>45</sup> Vergleiche dazu Kapitel 2.2.1.

<sup>46</sup> Siehe Kapitel 3.2.1.

<sup>47</sup> Siehe Kapitel 3.1.2.

<sup>48</sup> Neben dem eben vorgestellten Lösungsansatz existieren u.a. auch Methoden, welche Adjunkte als Schwestern von Komplementen behandeln. In diesem Ansatz wird man den im Englischen selten auftretenden Fällen von Adjunkten, welche Komplementen vorangestellt sind (Bsp.: *Sandy thought until she met him that Kim was too old*), gerecht. Dazu Borsley (1996), S. 110-113.

benötigt, die eine VP-Kategorie, welche ein einziges Element, das Subjekt, in ihrer SUBCAT-Liste fordert, mit einem entsprechenden Element des Kontextes verbindet, das Head-Subject Schema.

Expletiva (dummy it/es) werden von PS-Grammatiken nicht gesondert behandelt, sondern generell als Dummies für Subjekte behandelt und von normalen NPs durch ein Merkmal NFORM unterschieden, das als Wert NORM, THERE/DA und IT/ES hat.

#### 4.4. Grammatische Funktionen

Mit grammatischen Funktionen/Relationen sind Begriffe wie Subjekt und Objekt gemeint. In der von Pollard / Sag (1994) beschriebenen HPSG-Version wird das Subjekt als Realisierung einer zusätzlichen Kategorie der SUBCAT-Liste verstanden, in anderen Versionen als Realisierungen eines separaten SUBJ-Merkmals. In ersterer Version könnte ein Subjekt wie folgt definiert werden:

„Ein Subjekt ist ein Ausdruck, der das letzte Element in einer SUBCAT-Liste realisiert.“<sup>49</sup>

Wenn nun Subjekte die Realisierungen einer zusätzlichen Kategorie in der SUBCAT-Liste sind, dann können nominale Objekte wie folgt definiert werden:

„Ein Objekt ist eine NP, die die vorletzte Kategorie in der SUBCAT-Liste eines Verbs realisiert.“<sup>50</sup>

#### 4.5. Wh-Fragen

*Wh*-Fragen sind Satztypen, in denen ein Komplement oder ein Subjekt fehlt. In den PS-Grammatiken wird generell ausgesagt, dass sie leere Kategorien wie die Spuren (traces) in der GB-Theorie aufweisen.

---

<sup>49</sup> Borsley (1997), S. 203.

<sup>50</sup> *ibid.*, S. 204.



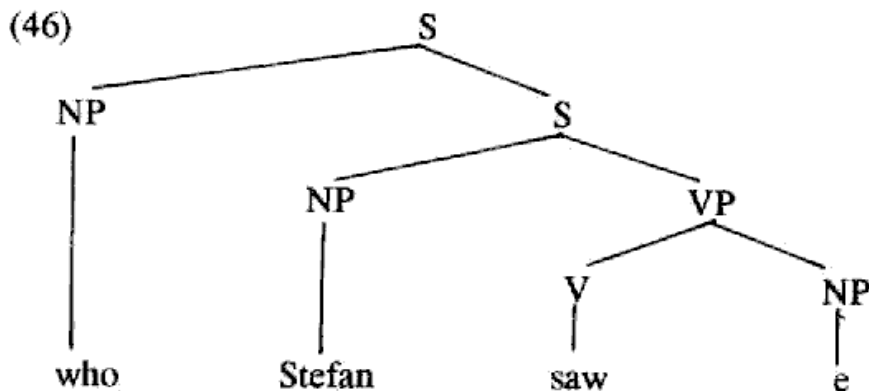
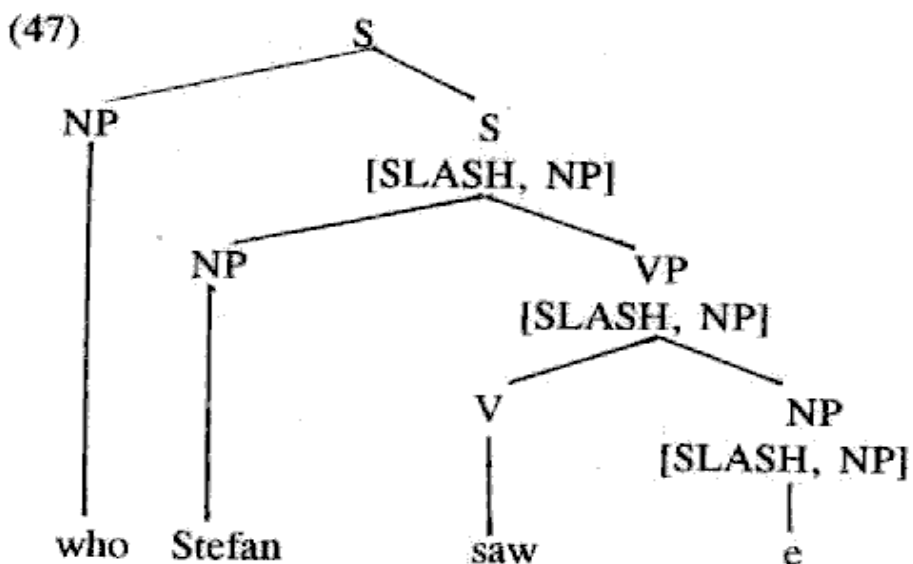


Abbildung 15

Wenn man eine solche Struktur annimmt, muss man die leere Kategorie sensitiv machen für die Anwesenheit der *wh*-Phrase weiter oben im Baum. Wie in GPSG wird zu diesem Zwecke auch in HPSG ein Merkmal SLASH eingeführt. Dieses Merkmal nimmt eine Kategorie als Wert und zeigt an, dass eine Konstituente eine leere Kategorie bestimmter Art enthält. Der vorher eingeführte Baum kann also wie folgt erweitert werden:<sup>51</sup>



<sup>51</sup> Borsley (1997), S. 306.

Abbildung 16

Es muss nun noch eine Regel eingefügt werden, die solche Strukturen lizenziert.

#### 4.6. Satzfragen

Die Behandlung von Satzfragen lässt sich in zwei Teilgebiete gliedern: Die Behandlung von Hilfsverben und der Subjekt-Verb-Inversion.

##### 4.6.1 Hilfsverben

Es lassen sich einige Generalisierungen über Hilfsverben zusammenfassen:<sup>52</sup>

1. Sie sind optional.
2. Sie determinieren die Form des folgenden Verbs.
3. Mehrere Auxiliare können gemeinsam stehen, allerdings nur in einer bestimmten Reihenfolge.
4. Hilfsverben eines bestimmten Typs können sich nicht wiederholen.

Während sich Borsley<sup>53</sup> über die allgemeine Betrachtung von Auxiliaren weitgehend ausschweigt, bieten Sag / Wasow folgende Analyse an: Sie betrachten Hilfsverben generell als Verben, die VP-Komplemente verlangen, ein Ansatz, der das Problem der zweiten Generalisierung sehr elegant löst, während (beinahe) die gesamte restliche Information im Lexikoneintrag des jeweiligen Hilfsverbs enthalten ist.

Satzfragen, in denen kein Auxiliar vorhanden ist, werden bekanntlich mit *do* konstruiert. Der von Sag / Wasow vorgeschlagene Lexikoneintrag für *do* sieht so aus:

---

<sup>52</sup> Nach Sag / Wasow (1999).

<sup>53</sup> Borsley (1997).

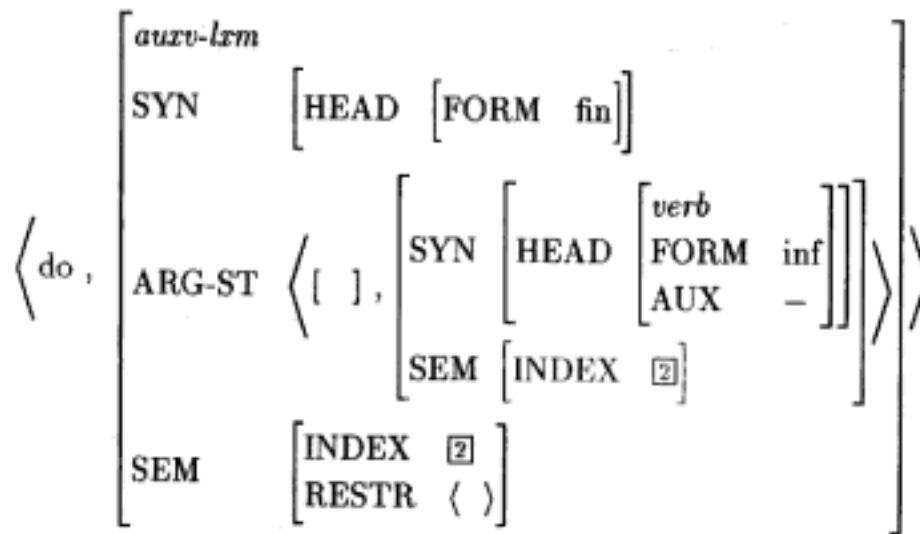


Abbildung 17

Diese Merkmalstruktur sagt unter anderem aus, dass der Kopf des Komplements von *do* ein Infinitiv sein muss, aber kein Hilfsverb sein kann, und dass das Auxiliär *do* selbst nur in finiten Formen auftritt. Der Semantikwert ausserdem garantiert, dass *do* keinen Beitrag an die Semantik des Satzes leistet.

#### 4.6.2. Inversion

In Satzfragen steht das Auxiliär vor dem Subjekt. Trotzdem soll die Beziehung zwischen dem Hilfsverb und seinem Komplement, dem Hauptverb, erhalten bleiben. Das Problem besteht also darin, das Subjekt hinter dem Auxiliär zu positionieren. Sag / Wasow schlagen dazu vor, die postauxiliäre NP (also das Subjekt) nicht als Specifier, sondern vielmehr als erstes Komplement des Hilfsverbs zu behandeln. Da das Head-Complement Schemata Komplemente in der Reihenfolge, in der sie in der SUBCAT-Liste des Kopfes stehen, realisiert, erscheint die Subjekts-NP als erstes Komplement im Lexikoneintrag für das invertierte Auxiliär im Fragesatz an der richtigen Stelle. Es muss also nur noch eine lexikalische Regel für YES-NO-Fragen formuliert werden, die die Subjekt-Verb-Inversion erst bewirkt.

## 5. Schlusswort

In den letzten Jahren hat sich HPSG in der Computerlinguistik zum einflussreichsten Framework für die deklarative Spezifikation von Grammatik entwickelt. Es arbeiten mehr Leute an und mit implementierten HPSG-Grammatiken als mit jedem anderen linguistischen Grammatikmodell.

Welches sind denn die entscheidenden Vorteile von HPSG? Besser als in anderen grammatischen Frameworks können linguistische Generalisierungen isoliert und doch uniform in Prinzipien repräsentiert werden. Der HPSG-Formalismus lässt für eine konkrete Implementation relativ viel Freiheit, weshalb auch zahlreiche Versionen (von denen wir uns auf die von Pollard / Sag in 1994 und von Borsley in 1996 Präsentierten gestützt haben) in Gebrauch sind. Ebendiese Freiheit jedoch erschwert es, die essentiellen Eigenschaften von HPSG-Grammatikmodellen, Implementationen von HPSG-Formalisten oder konkret sprachspezifischen HPSG-Grammatiken zu ermitteln. Dies muss allerdings nicht unbedingt als Nachteil betrachtet werden, da gerade diese Konstellation Raum für weitere, innovative Entwicklungen in verschiedene Richtungen schafft. Ausserdem wurde HPSG von Anfang an bewusst darauf ausgerichtet, ein präzises Framework zur Artikulation einer Synthese von Ideen aus Grammatiken unterschiedlichster Traditionen zu bilden.

## 6. Bibliografie

### Literatur

Borsley, Robert D.: *Modern Phrase Structure Grammar*. Oxford: Blackwell Publishers, 1996.

Borsley, Robert D.: *Syntax-Theorie. Ein zusammengefasster Zugang*. Dt. Bearb. v. P. Suchsland. Tübingen: Niemeyer, 1997. (Konzepte der Sprach- und Literaturwissenschaft; 55)

Bussmann, Hadumod: *Lexikon der Sprachwissenschaft*. Stuttgart: Kröner, 1990.

Cooper, Richard P.: Head-driven Phrase Structure Grammar. In: Brown, Edward Keith: *Concise encyclopedia of syntactic theories*. Oxford: Pergamon, 1996.

Pollard, Carl und Sag, Ivan A.: *Head-driven Phrase Structure Grammar*. Chicago: The University of Chicago Press, 1994.

Sag, Ivan A. und Wasow, Thomas: *Syntactic Theory: A Formal Introduction*. Chicago: The University of Chicago Press, 1999.

### Vorlesungen / Skripts

Schneider, Gerold: Formale Grammatiken und Syntaxanalyse. Wintersemester 1999/2000.

Volk, Martin: Formale Grammatiken und Syntaxanalyse. Wintersemester 1998/1999.

**homepages**

<http://www.coli.uni-sb.de/~hansu/hpsg1.html>

<http://www.coli.uni-sb.de/~mineur/ltv/Carpenter.html>

<http://www.ling.ohio-state.edu/hpsg/>

(mit Beiträgen von Bob Carpenter, Ivan Sag und Hans Uszkoreit)