

## 9. BEGLEITBLATT

Besprechung der Übungen aus Chapter 2: Problems 1, 2, 4, 6

### 8 Analyzing Features of Grammatical Categories

#### 8.1 Introduction

**Problem: Kookkurrenz-Restriktionen**, d.h. bestimmte Ausdrücke verlangen die Anwesenheit bestimmter anderer Ausdrücke in bestimmter Form, bzw. sind mit bestimmten Ausdrücken bestimmter Form nicht verträglich.

**CFG: Viele Regeln, wenig Verallgemeinerungen**

#### X-Bar-Syntax:

- Berücksichtigung der besonderen Rolle der Kopf-Konstituenten von Phrasen,
- allgemeine Prinzipien (Kopf-Prinzip, Phrasen-Prinzip, X-Bar-Schema...)

aber: die Berücksichtigung grober kategorialer Eigenschaften reicht nicht.

(94)

- (B309)
- (i) *Hein glaubt an Poseidons Offenbarungen.*
  - (ii) *Hein glaubt an Segeltuch aus Spitzbergen.*
  - (iii) *Hein glaubt Poseidons Offenbarungen.*
  - (iv) *\*Hein glaubt Segeltuch aus Spitzbergen.*

#### SWBs Strategie:

*„Our solution to the problem of redundancy is to make grammatical categories decomposable into component parts.“<sup>25</sup>*

#### 8.2 Feature Structures

##### Grundsätzliches vorweg:

Der Gegenstand der Grammatiker sind die Wörter und Sätze einer Sprache und die Regeln, nach denen diese gebildet werden. Diese „Gegenstände“ haben eine Vielzahl von Eigenschaften, von denen den Grammatiker letztlich nur eine relativ geringe, aber relevante Auswahl interessieren. So interessieren den Grammatiker an Wörtern beispielsweise deren morphologische Eigenschaften, deren Valenzeigenschaften, bestimmte Bedeutungseigenschaften etc. Wer ein Wort zuerst „erfunden“ hat, welchen Wandlungsprozessen es im Laufe seiner

---

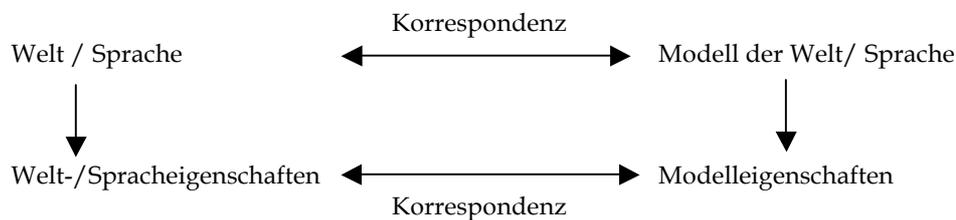
<sup>25</sup> Cf. S&W [1999:47].

Geschichte unterworfen war, ob es ein Lieblingswort von Heins Großmutter ist usw., das kümmert ihn in der Regel wenig.

**Anders gesagt:** die Wörter des Grammatikers sind nicht die Wörter selbst, die Totalität ihrer wesentlichen und unwesentlichen Eigenschaften, sondern eine Abstraktion davon.

In diesem Sinne ist das Lexikon als Teilkomponente eines Grammatiksystems auch nur ein Modell des Wortschatzes der untersuchten Sprache. Und analog handeln Syntax und Semantik mit Modellen für die Regeln der Satzbildung und Satzbedeutung. Allerdings: wenn die wissenschaftliche Beschäftigung mit Modellen einen Sinn haben soll, müssen Modelle und Wirklichkeit in den relevanten Eigenschaften einander ähnlich sein:

(95)



In der HPSG werden nun sprachliche Entitäten durch sog. Merkmalsstrukturen (engl. feature structures) modelliert<sup>26</sup>, die auf verschiedene Weise aufgefasst werden können, nämlich

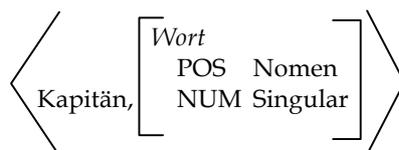
(i) als Funktion, (als Menge geordneter Paare, sodass es keine zwei Paare gibt, deren erste Elemente identisch sind), wobei jedes erste Element eines Paares ein Merkmal ist, jedes zweite Element ein Merkmalswert:

oder

(ii) als gerichteter Graph, bei dem die übergeordneten Knoten Merkmale darstellen, untergeordnete Knoten dagegen Merkmalswerte

Merkmalsstrukturen werden in der HPSG durch Merkmalsstruktur-Beschreibungen repräsentiert, die sich durch die exzessive Verwendung großer, eckiger Klammern auszeichnen, nämlich in Gestalt sog. Merkmal-Wert-Matrices (AVMs, von engl. attribute-value-matrices). Beispiel einer einfachen AVM:

(96)



<sup>26</sup> Es wird hier und im folgenden darauf verzichtet, den Begriff der Merkmalsstruktur mit mathematischen Mitteln zu definieren. Wer sich für die Logik und Mathematik getypter Merkmalsstrukturen interessiert, findet in Carpenter [1992] die entsprechende Ausarbeitung und weitere Literatur.

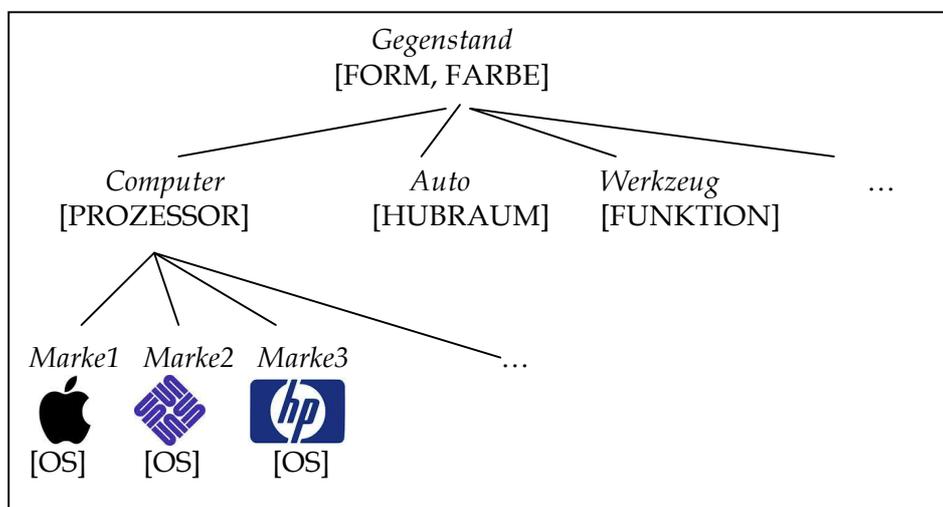
**Nebenbei:** in diesem Buch und auch leider sonst wird einige Verwirrung gestiftet, indem in Bezug auf diese Klammer-Gebilde manchmal von „Merkmalsstrukturbeschreibungen“, manchmal von „Merkmalsstrukturen“ selbst die Rede ist. Manchmal heißt es, diese AVMs *repräsentierten* eine Merkmalsstruktur, manchmal, diese *modellierten* Merkmalsstrukturen, manchmal sie *seien* Merkmalsstrukturen. Dies sind allesamt Ungenauigkeiten, denen auch wir nicht ganz entfliehen können. Daher sei hier festgestellt:

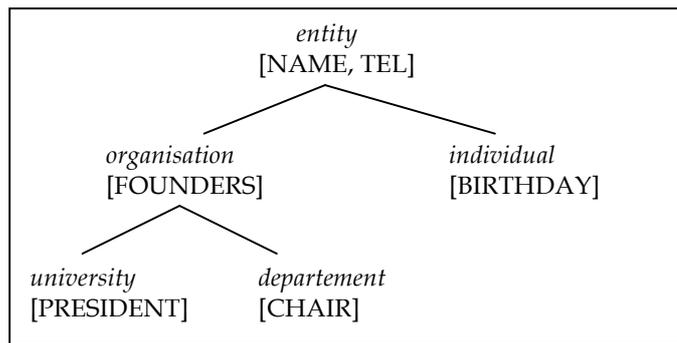
Diese Klammergebilde sind grundsätzlich **Merkmalsstrukturbeschreibungen**, auch dann, wenn zumeist der Einfachheit halber „Merkmalsstruktur“ gesagt wird. Diese Beschreibungen *repräsentieren* Merkmalsstrukturen, die ihrerseits Modelle von bestimmten Gegenständen der Welt sind (von sprachlichen Gegenständen nämlich).

Eine AVM charakterisiert einen sprachlichen Ausdruck im Rahmen des Grammatikmodells durch versal gesetzte Merkmalsbezeichnungen (in (96) POS, für engl. „**p**art **o**f **s**peech“, dt. „Wortart“, und NUM, für engl. „**n**umber“, dt. „Numerus“) und diesen Merkmalsbezeichnungen jeweils zugeordnete Wertbezeichnungen (in (96) resp. „Nomen“ und „Singular“). Die AVMs werden ihrerseits durch eine Bezeichnung gekennzeichnet, die den **Typ** der jeweiligen AVM angibt (im gegebenen Falle (96) durch die kursiv gesetzte Typen-Bezeichnung *Wort*).

Ein Typ ist eine Klasse von Objekten, die sich in mindestens einer Hinsicht ähneln. Typen lassen sich gemäß ihrer Ähnlichkeiten und Unterschiede in eine Typenhierarchie mit Ober- und Untertypen einordnen. Teil der Typenhierarchie sind **Beschränkungen** (constraints), die bestimmen, welche Merkmale für die jeweiligen Typen charakteristisch sind und welche Arten von Werten diesen zuzuordnen sind.

### (97) 1. Beispiel einer Typenhierarchie



**(98) 2. Beispiel einer Typenhierarchie**

Der Vorteil von Typenhierarchien besteht darin, dass die Merkmale, die ein bestimmtes Objekt kennzeichnen, sich an genau der Stelle im Baum ansiedeln lassen, wo sie wirklich einschlägig sind. Die Hierarchie selbst sorgt dafür, dass jeder Merkmalsstrukturtyp die Merkmale aller übergeordneten Merkmalsstrukturtypen **erbt**. Das führt letztlich auch zu deutlich knapperen Merkmalsstrukturepräsentationen. Wer etwa von einer Universität spricht, braucht nicht explizit anzugeben, dass die Universität von irgendwem gegründet wurde und dass sie eine Telefonnummer hat. Das ergibt sich automatisch aus der Hierarchie.

**(99) SWB 54-7**

<i>university</i>	
NAME	Stanford University
FOUNDERS	$\left\langle \begin{array}{l} \textit{individual} \\ \text{NAME} \quad \text{Leland Stanford} \end{array} \right\rangle ,$ $\left. \begin{array}{l} \textit{individual} \\ \text{NAME} \quad \text{Jane Stanford} \end{array} \right\rangle$
PRESIDENT	$\left[ \begin{array}{l} \textit{individual} \\ \text{NAME} \quad \text{John Hennessy} \\ \text{BIRTHDAY} \quad 9-22-52 \\ \text{TEL} \quad 650-723-2481 \end{array} \right]$
TEL	650-723-2300

<i>department</i>	
NAME	Stanford Linguistics
FOUNDERS	$\left\langle \begin{array}{l} \textit{individual} \\ \text{NAME} \quad \text{Joseph Greenberg} \\ \text{BIRTHDAY} \quad 5-28-15 \end{array} \right\rangle ,$ $\left. \begin{array}{l} \textit{individual} \\ \text{NAME} \quad \text{Charles Ferguson} \\ \text{BIRTHDAY} \quad 7-6-21 \end{array} \right\rangle$
CHAIR	$\left[ \begin{array}{l} \textit{individual} \\ \text{NAME} \quad \text{Eve Clark} \\ \text{BIRTHDAY} \quad 7-26-42 \\ \text{TEL} \quad 650-723-4284 \end{array} \right]$
TEL	650-723-4284

(100) SWB 55-8<sup>27</sup>

TYPE	FEATURES/VALUES	IST
<i>entity</i>	[NAME <i>string</i> TEL <i>number</i> ]	
<i>organization</i>	[FOUNDERS <i>list(individual)</i> ]	<i>entity</i>
<i>university</i>	[PRESIDENT <i>individual</i> ]	<i>organization</i>
<i>department</i>	[CHAIR <i>individual</i> ]	<i>organization</i>
<i>individual</i>	[BIRTHDAY <i>date</i> ]	<i>entity</i>

Beschränkungen lassen sich, sofern sich ihre jeweiligen Merkmalspezifikationen nicht widersprechen, miteinander kombinieren (vereinigen, „**unifizieren**“). Nehmen wir an, wir hätten zwei miteinander verträgliche Beschränkungen A und B, die jeweils von einer Menge von Merkmalsstrukturen M und N **erfüllt** werden, dann ist die Menge der Merkmalsstrukturen, die  $A \cup B$  erfüllen, der Durchschnitt aus M und N, also  $M \cap N$ .

**Ein Beispiel:** Nehmen wir an, es gäbe rein zufällig drei Personen, auf die die AVM in (101) zutrifft, eine Person aus Hamburg, eine aus Lübeck und eine aus Kopenhagen. Nehmen wir ferner an, es gäbe zwei Schiffe auf dieser Welt, die die (partiellen) AVM in (102) erfüllen.

(101)  $\left[ \begin{array}{l} \text{Person} \\ \text{NAME} \quad \text{Hein Petersen} \\ \text{GEBURTSTAG} \quad 01.01.1950 \end{array} \right]$

(102)  $\left[ \begin{array}{l} \text{Schiff} \\ \text{NAME} \quad \text{Santa Maria} \\ \text{KAPITÄN} \quad [\text{Hein Petersen}] \end{array} \right]$

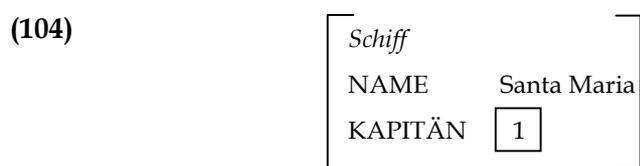
Da die beiden AVM (101) und (102) miteinander verträglich sind, sähe eine Unifizierung beider AVMs wie folgt aus:

(103)  $\left[ \begin{array}{l} \text{Schiff} \\ \text{NAME} \quad \text{Santa Maria} \\ \text{KAPITÄN} \quad \left[ \begin{array}{l} \text{Person} \\ \text{NAME} \quad \text{Hein Petersen} \\ \text{GEBURTSTAG} \quad 01.01.1950 \end{array} \right] \end{array} \right]$

<sup>27</sup> „IST“ in der Kopfzeile der Tabelle bedeutet *immediate supertype*, also wörtlich „direkter Obertyp“.

Die AVM in (103) sollte jedoch nur noch von einem Objekt erfüllt werden, nämlich von demjenigen Schiff mit dem Namen Santa Maria, das einen Hein Petersen als Kapitän hat, der am 01.01.1950 geboren wurde. Eine AVM, die sämtliche Merkmal-Wert-Paare genau einer Merkmalsstruktur innerhalb eines gegebenen Modells repräsentiert, heißt **totale AVM**. Merkmalsstrukturen selbst sind niemals partiell oder total (und wenn, dann immer total). Partiiell oder Total können nur Merkmalsstrukturbeschreibungen sein.

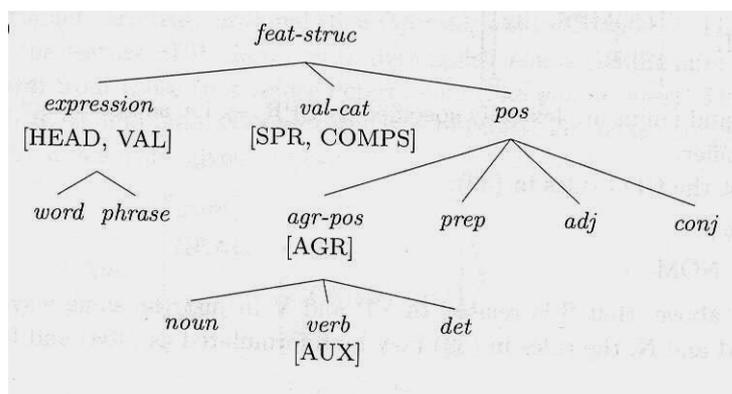
Um die Notation der AVMs zu vereinfachen, werden in der HPSG sog. **Tags** verwendet. Ein Tag ist eine kleine Box mit einer Ziffer darin, die stellvertretend für eine AVM steht. Die Unifikation aus (101) und (102) kann abkürzend daher wie folgt dargestellt werden, wobei sich das Tag mit der „1“ auf die AVM in (101) bezieht:



### 8.3 The Linguistic Application of Feature Structures

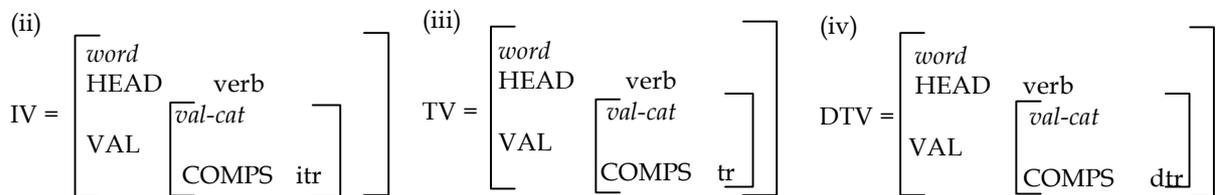
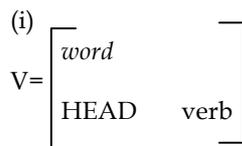
Die Gegenstände der Grammatik sind nicht Universitäten, Schiffe oder Kapitäne, sondern Wörter, Phrasen und Sätze. Die Wörter einer Sprache stehen im Lexikon, und wie Phrasen und Sätze gebildet werden, wird durch Regeln angegeben. In der HPSG werden Wörter und Phrasen mit ihren Eigenschaften als AVMs dargestellt. Wie Phrasen und Sätze zusammengefügt werden, wird in Prinzipien und Regeln erfasst. Wie später deutlicher werden wird, spielt in der HPSG, und anders aber als in allen Vorgängertheorien der generativen Grammatik, die lexikalischen Eigenschaften eine zentrale Rolle. Der Regel- und Prinzipienapparat wird dagegen verhältnismäßig kompakt gehalten.

#### (105) SWB-65-36: Typenhierarchie (vorläufige Version)

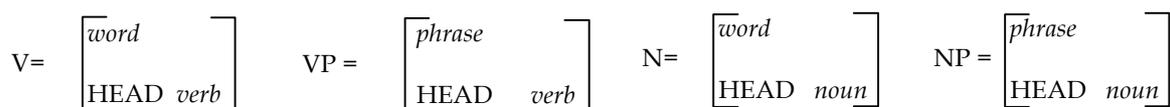


**(106) SWB-61-23 / 24: lexikalische Einträge und das HEAD-Merkmal**

(i) AVM für die Kategorie Nomen (ii) AVM für den lexikalischen Eintrag eines Nomens

**8.3.1 Valenz-Merkmale****(107) Verben**

Die AVM (i) in (107) verzichtet auf die Angabe des Merkmals VAL und gibt für dieses Merkmal auch keinen Wert an. Damit repräsentiert diese AVM sämtliche Verben, allerdings ohne Angaben zu ihren jeweiligen Valenzeigenschaften zu machen. In solchen Fällen wird von **Unterspezifizierung** gesprochen.

**(108) SWB 63-28,29,30,31: Reanalyse der PSG-Kategorien**

Der Unterschied zwischen Wörtern und Phrasen ist grob gesagt der, dass Phrasen als abgeschlossene Einheiten gelten, während Wörter eventuell noch mit anderen begleitenden Elementen angereichert werden müssen, um gesättigt zu sein. Verben etwa werden durch die Anbindung mit Komplementen zu Verbalphrasen. Auf die Nomina übertragen kann man argumentieren, dass erst die Anbindung eines Determinators aus einem Nomen eine Nominalphrase macht. Wenn man den Spezialfall der Eigennamen und der Pronomina einmal beiseite lässt, kann man den Standpunkt vertreten, dass der Determinator von seinem Nomen valenzabhängig ist. Diesen Gedanken setzten SWB so um, dass sie neben dem Valenzmerkmal COMPS, das als Wert Verbkomplemente fordert, es ein Valenzmerkmal SPECIFIER (SPR) geben muss, das (u.a.) als Wert einen passenden Determinator verlangt. Dasselbe

Merkmal wird - ganz im Einklang mit der generativen Tradition - verwendet, um die Valenzabhängigkeit des „Subjekts“ bzw. der ersten NP von der VP zu modellieren. Der Gedanke ist hier der, dass die VP einen valenzabhängigen Spezifikator in Gestalt einer NP benötigt, um einen (vollständigen und valenzmäßig gesättigten) Satz zu ergeben.

**(109) SWB 64- 32: Specifier-Merkmal**

$$NP = \begin{bmatrix} \textit{phrase} \\ \text{HEAD} \textit{ noun} \\ \text{VAL} \begin{bmatrix} \textit{val-cat} \\ \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ +} \end{bmatrix} \end{bmatrix} \quad \text{NOM} = \begin{bmatrix} \textit{phrase} \\ \text{HEAD} \textit{ noun} \\ \text{VAL} \begin{bmatrix} \textit{val-cat} \\ \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix}$$

**(110) SWB 64- 33: Specifier-Merkmal**

$$S = \begin{bmatrix} \textit{phrase} \\ \text{HEAD} \textit{ verb} \\ \text{VAL} \begin{bmatrix} \textit{val-cat} \\ \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ +} \end{bmatrix} \end{bmatrix} \quad \text{VP} = \begin{bmatrix} \textit{phrase} \\ \text{HEAD} \textit{ verb} \\ \text{VAL} \begin{bmatrix} \textit{val-cat} \\ \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix}$$

**(111) SWB 65-37: Umformulierung der PSG-VP-Regeln (vorläufig):**

$$\begin{array}{l} \text{a.} \begin{bmatrix} \textit{phrase} \\ \text{HEAD} \textit{ [1]} \\ \text{VAL} \begin{bmatrix} \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \textit{word} \\ \text{HEAD} \textit{ [1]} \\ \text{VAL} \begin{bmatrix} \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix} \\ \text{b.} \begin{bmatrix} \textit{phrase} \\ \text{HEAD} \textit{ [1]} \\ \text{VAL} \begin{bmatrix} \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \textit{word} \\ \text{HEAD} \textit{ [1]} \\ \text{VAL} \begin{bmatrix} \text{COMPS} \textit{ str} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix} \text{ NP} \\ \text{c.} \begin{bmatrix} \textit{phrase} \\ \text{HEAD} \textit{ [1]} \\ \text{VAL} \begin{bmatrix} \text{COMPS} \textit{ itr} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \textit{word} \\ \text{HEAD} \textit{ [1]} \\ \text{VAL} \begin{bmatrix} \text{COMPS} \textit{ dtr} \\ \text{SPR} \textit{ -} \end{bmatrix} \end{bmatrix} \text{ NP NP} \end{array}$$

Diese Regeln lassen sich, da der Wert des POS-Merkmals nicht angegeben wurde, im Prinzip auch auf passende Nomen, Präpositionen etc. anwenden.

**(112) SWB66-38,40,41 Reformulierung der PSG-Regeln mit Spezifikator**

$$\left\langle \text{bird}, \begin{bmatrix} \text{word} \\ \text{HEAD } \textit{noun} \\ \text{VAL} \begin{bmatrix} \text{COMPS } \textit{itr} \\ \text{SPR} \quad - \end{bmatrix} \end{bmatrix} \right\rangle$$

$$\begin{array}{l} \text{a.} \begin{bmatrix} \textit{phrase} \\ \text{HEAD } \boxed{1} \textit{verb} \\ \text{VAL} \begin{bmatrix} \text{COMPS } \textit{itr} \\ \text{SPR} \quad + \end{bmatrix} \end{bmatrix} \rightarrow \text{NP} \begin{bmatrix} \textit{phrase} \\ \text{HEAD } \boxed{1} \\ \text{VAL} \begin{bmatrix} \text{SPR} \quad - \end{bmatrix} \end{bmatrix} \\ \text{b.} \begin{bmatrix} \textit{phrase} \\ \text{HEAD } \boxed{1} \textit{noun} \\ \text{VAL} \begin{bmatrix} \text{COMPS } \textit{itr} \\ \text{SPR} \quad + \end{bmatrix} \end{bmatrix} \rightarrow \text{D} \begin{bmatrix} \textit{phrase} \\ \text{HEAD } \boxed{1} \\ \text{VAL} \begin{bmatrix} \text{SPR} \quad - \end{bmatrix} \end{bmatrix} \end{array}$$

$$\text{D} = \begin{bmatrix} \text{word} \\ \text{HEAD } \textit{det} \\ \text{VAL} \begin{bmatrix} \text{COMPS } \textit{itr} \\ \text{SPR} \quad + \end{bmatrix} \end{bmatrix}$$

**(113) SWB67- 43,44: Regel für Nomen ohne Determiner (z.B. Eigennamen)**

$$\left\langle \text{Alex}, \begin{bmatrix} \text{word} \\ \text{HEAD } \textit{noun} \\ \text{VAL} \begin{bmatrix} \text{COMPS } \textit{itr} \\ \text{SPR} \quad + \end{bmatrix} \end{bmatrix} \right\rangle$$

$$\begin{bmatrix} \textit{phrase} \\ \text{HEAD } \boxed{1} \textit{noun} \\ \text{VAL} \begin{bmatrix} \text{COMPS } \textit{itr} \\ \text{SPR} \quad + \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \text{word} \\ \text{HEAD } \boxed{1} \\ \text{VAL} \begin{bmatrix} \text{SPR} \quad + \end{bmatrix} \end{bmatrix}$$

(114) SWB68-45: Ein Beispielsatz

