

16. BEGLEITBLATT

13 Kapitel : The Structure of the Lexicon

13.1 Introduction

HPSG-Strategie: relativ viel syntaktische und semantische Information wird ins Lexikon geschoben. Zugleich liegt nach dem Bisherigen relativ viel Information mehrfach vor, weil jede Flexionsform (jede Vollform) als eigener lexikalischer Eintrag voll spezifiziert wurde.

----> Redundanz des Lexikons.

----> Verlust von Generalisierungen für ganze Klassen von Wörtern

----> Cognitive Unangemessenheit

ZIEL: Bessere Organisation der Lexikonkomponente

13.2 Lexemes

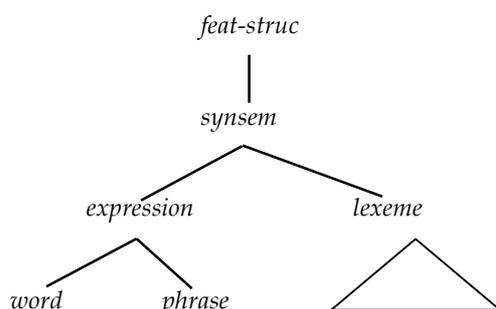
Bisher wurden nur Ausdrücke vom Typ *word* oder *phrase* unterschieden, wobei die Einheiten vom Typ *word* als Einträge des Lexikons dienten. Wortformen, die zum selben Paradigma gehörten, wurden nicht anders behandelt, als Wortformen, die nichts miteinander zu tun haben. Dies wird im 8. Kapitel von SWB geändert. Lexikalische Einträge sind jetzt Einheiten von einem neuen Typ *lexeme*. Lexeme dienen dann als Basis für die Ableitung der einzelnen Wortformen.

(184) Beispiel-Paradigma für das Lexem *viver* (Span.)

vivo 'I live'	vives 'you (sg.) live'	vive '(s)he/it lives'
vivimos 'we live'	vivís 'you (pl.) live'	viven 'they live'
vivía 'I lived'	vivías 'you (sg.) lived'	vivía '(s)he/it lived'
vivíamos 'we lived'	vivíais 'you (pl.) lived'	vivían 'they lived'
viviré 'I'll live'	vivirás 'you (sg.) 'll live'	vivirá '(s)he/it 'll live'
viviremos 'we'll live'	viviréis 'you (pl.) 'll live'	vivirán 'they'll live'

Diese Neuerung schlägt sich auch in der Typenhierarchie nieder:

(185)



Dadurch ist es nun möglich, Generalisierungen auch für ganze Wortklassen und sogenannte **lexikalische Regeln** zu formulieren.

Das Merkmal ARG-ST kennzeichnet nebenbei bemerkt nun sowohl *word-* als auch *lexeme-* Einheiten.

13.3 Default Constraint Inheritance

Wir haben bisher die verschiedensten Merkmalsstrukturepräsentationen kennengelernt, die sich voneinander dadurch unterschieden, dass in ihnen je unterschiedliche Merkmale vorkamen. Dementsprechend wurden diese AVMs unterschiedlich typisiert. Eine *syn-cat* AVM enthielt andere Merkmale, als eine *sem-cat* AVM oder eine *agr-cat* AVM. Zugleich wurden diese unterschiedlichen AVMs in eine Typenhierarchie eingeordnet. Zwischen Untertypen und Obertypen bestand eine Vererbungsbeziehung, indem die untergeordneten Typen alle Merkmalsdeklarationen ihrer Obertypen erben.

(186) Folie SWB p. 155 (55a) und p.156

Es galt folgende Regel:

(187)

Falls der Typ T2 ein Untertyp vom Typ T1 ist, dann

a. spezifiziert jedes Merkmal, das T1 spezifiziert, auch T2 und

b. erstreckt sich jedes Constraint, das T1 betrifft, auch auf sämtliche direkten und indirekten Untertypen von T1

Ein solcher Vererbungsmechanismus heißt **monoton**. Probleme mit monotoner Vererbung ergeben sich dann, wenn bestimmte Elemente, die man aus guten Gründen einem bestimmten Untertyp zuordnen möchte, nicht alle Merkmale und Beschränkungen ihrer Obertypen aufweisen, sie diesen also widersprechen.

Wenn die Bedingungen an die Typenhierarchie nicht ganz so strikt formuliert werden und es im Prinzip auch zulässig ist, dass sich nicht jede Eigenschaft eines übergeordneten Typs auf einen untergeordneten Typ bzw. eine Instanz vererben muss, spricht man von **Hierarchien mit Default-Vererbung**. Unter anzugebenden Umständen ist es dann erlaubt, dass eine per Default ererbte Eigenschaft von einer entsprechenden Eigenschaft des jeweiligen Subtyps bzw. eines bestimmten Exemplars (einer „Instanz“) des Typs „überschrieben“ wird. Notationell werden Merkmalspezifikationen, die überschrieben werden können, fortan mit einem Schrägstrich „/“ gekennzeichnet.

(188) Beispiel für Defaultvererbung; Satz- bzw. VP-Modifikation

(i) T_i = Spezifikation für einen Default-Satz-Modifikator:

$$T_i : \left[\text{SYN} \left[\text{VAL} \left[\text{MOD} \left\langle \left[\text{SYN} / \left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{VAL} \left[\begin{array}{l} \text{SPR} \langle \rangle \\ \text{COMPS} \langle \rangle \end{array} \right] \end{array} \right] \right] \right] \right] \right] \right]$$

(ii) T_j : Spezifikation eines Untertyps zu T_i , die die Spezifikation von T_i in Teilen überschreibt:

$$T_j : \left[\text{SYN} \left[\text{VAL} \left[\text{MOD} \left\langle \left[\text{SYN} \left[\text{VAL} \left[\text{SPR} \langle \text{NP} \rangle \right] \right] \right] \right] \right] \right]$$

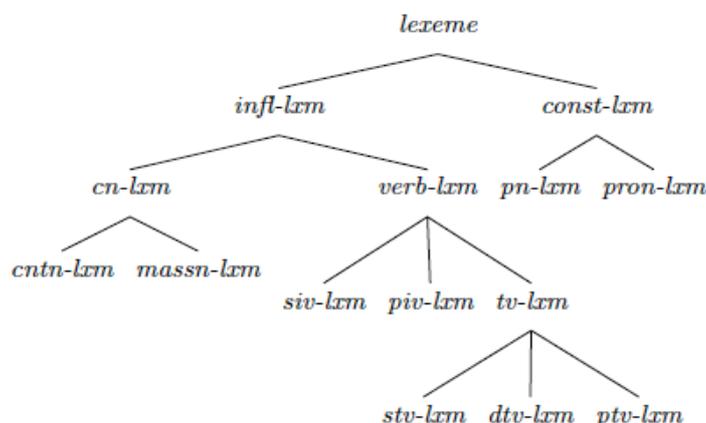
(iii) Spezifikation gemeinsam von T_i und T_j . Was T_i unberührt lässt, bleibt erhalten:

$$T_i \ \& \ T_j : \left[\text{SYN} \left[\text{VAL} \left[\text{MOD} \left\langle \left[\text{SYN} \left[\begin{array}{l} \text{HEAD } / \textit{verb} \\ \text{VAL} \left[\begin{array}{l} \text{SPR} \langle \text{NP} \rangle \\ \text{COMPS } / \langle \rangle \end{array} \right] \end{array} \right] \right] \right] \right] \right]$$

13.4 Some Lexemes of Our grammar

Die Lexikon besteht jetzt (u.a.) aus **lexikalischen Einträgen** vom Typ *lexeme*. Diese Einträge sind Paare aus phonologischer Form und Merkmalsstruktur. Diese Lexeme sind der Input für lexikalische Regeln, die aus Lexemen Einheiten vom Typ *word* ableiten (z.B. die einzelnen Flexionsformen).

(189) Verfeinerung der bisherigen Hierarchie (Ausschnitt)³⁵



³⁵ *infl-lxm*: inflectional-lexeme; *const-lxm*: constant-lexeme; *cn-lxm*: common-noun-lexeme; *pn-lxm*: proper-noun-lexeme (Eigennamen); *pron-lxm*: pronominal-lexeme; *cntn-lxm*: count-noun-lexeme; *massn-lxm*: mass-noun-lexeme; *siv-lxm*: strict-intransitiv-verb-lexeme; *piv-lxm*: prepositional-intransitive-verb-lexeme; *tv-lxm*: transitive-verb-lexeme; *stv-lxm*: strict-transitive-lexeme; *dtv-lxm*: ditransitive-lexeme; *ptv-lxm*: prepositional-transitive-lexeme.

(190) SHAC als Constraint für Lexeme vom Typ *infl-lxm*

Specifier-Head Agreement Constraint (SHAC)

$$infl\text{-}lxm : \left[\text{SYN} \left[\text{HEAD} \left[\text{AGR} \ \underline{1} \right] \right. \right. \\ \left. \left. \text{VAL} \left[\text{SPR} \ \langle \left[\text{AGR} \ \underline{1} \right] \rangle \right] \right] \right]$$

13.4.1 Nominal Lexemes**(191) Spezifikation für common nouns (Gattungsnomina)**

$$cn\text{-}lxm : \left[\text{SYN} \left[\text{HEAD} \left[\textit{noun} \right. \right. \right. \\ \left. \left. \text{AGR} \left[\text{PER} \ 3rd \right] \right] \right] \right] \\ \left[\text{SEM} \left[\text{MODE} \ / \ \textit{ref} \right. \right. \\ \left. \left. \text{INDEX} \ i \right] \right] \\ \left[\text{ARG-ST} \ \langle \text{DP}_i \rangle \oplus / \langle \ \rangle \right]$$

(192) Spezifikation für die Untertypen von *cn-noun*

- a. *cntn-lxm* : $\left[\text{ARG-ST} \ \langle \left[\text{COUNT} \ + \right], \dots \rangle \right]$
 b. *massn-lxm* : $\left[\text{ARG-ST} \ \langle \left[\text{COUNT} \ - \right], \dots \rangle \right]$

(193) typischer lexikalischer Eintrag

$$\left\langle \text{dog}, \left[\textit{cntn-lxm} \right. \right. \\ \left. \left. \text{SEM} \left[\text{INDEX} \ i \right. \right. \right. \\ \left. \left. \left. \text{RESTR} \ \left\langle \left[\text{RELN} \ \textit{dog} \right] \right. \right. \right. \\ \left. \left. \left. \left[\text{INST} \ i \right] \right] \right] \right] \right\rangle$$

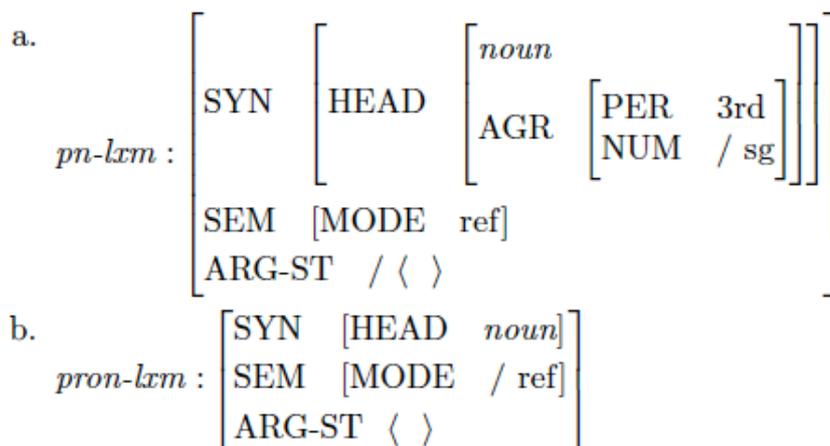
(194) AVM unter Berücksichtigung der bisherigen Constraints (190) bis (192):

$$\left\langle \text{dog}, \left[\textit{cntn-lxm} \right. \right. \\ \left. \left. \text{SYN} \left[\text{HEAD} \left[\textit{noun} \right. \right. \right. \right. \\ \left. \left. \left. \text{AGR} \ \underline{1} \left[\text{PER} \ 3rd \right] \right] \right] \right. \right. \\ \left. \left. \left. \text{VAL} \left[\text{SPR} \ \langle \left[\text{AGR} \ \underline{1} \right] \rangle \right] \right] \right] \right. \\ \left. \left. \text{SEM} \left[\text{MODE} \ \textit{ref} \right. \right. \right. \\ \left. \left. \left. \text{INDEX} \ i \right] \right. \right. \\ \left. \left. \left. \text{RESTR} \ \left\langle \left[\text{RELN} \ \textit{dog} \right] \right. \right. \right. \\ \left. \left. \left. \left[\text{INST} \ i \right] \right] \right] \right] \right. \\ \left. \left. \text{ARG-ST} \ \left\langle \text{DP} \left[\text{COUNT} \ + \right] \right\rangle \right] \right\rangle$$

(195) was zu tun übrig bleibt...(p. 240)

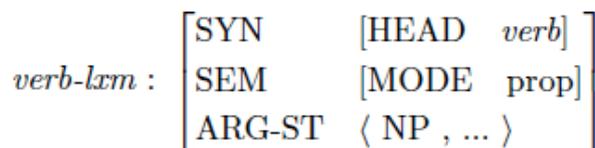
Essentially, all that remains to be stipulated in a given lexical entry is its phonological form, the particular predication in its semantic restriction, and any exceptional properties it may have. The rest follows from ‘the logic of the lexicon’. This is precisely what lexical stipulation should be reduced to.

(196) Spezifikationen für *pn-lxm* und *pron-lxm*

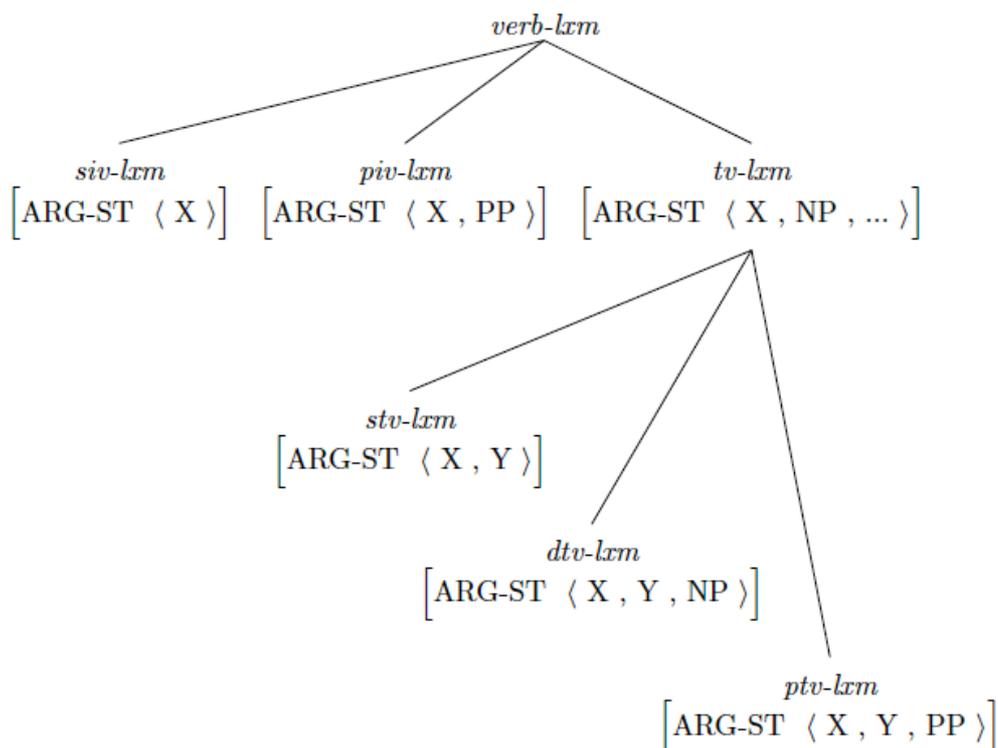


13.4.2 Verbal Lexemes

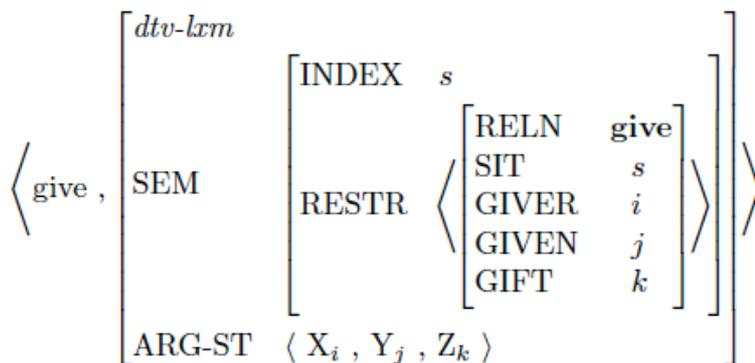
(197)



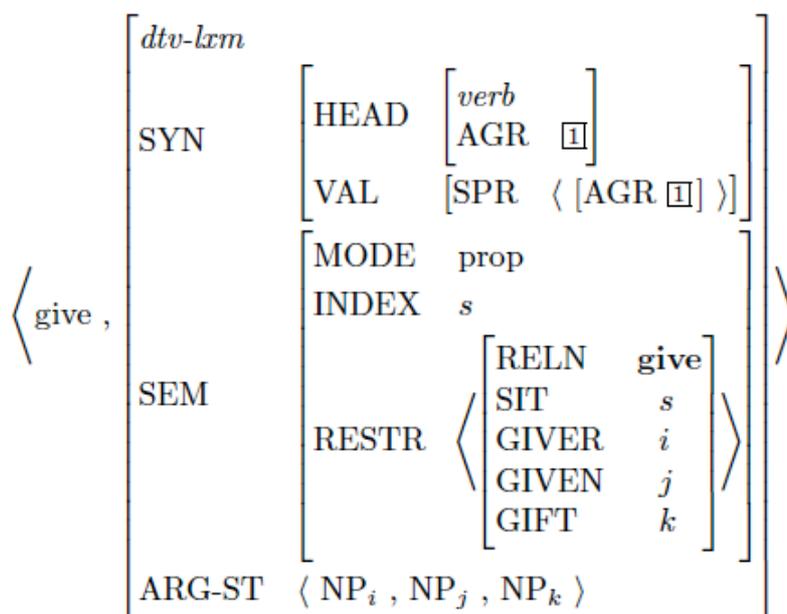
(198) Typenhierarchie für Verben (Ausschnitt)



(199) lexikalischer Eintrag für *give*

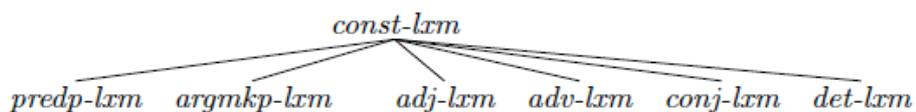


(200) AVM unter Berücksichtigung der bisherigen Constraints

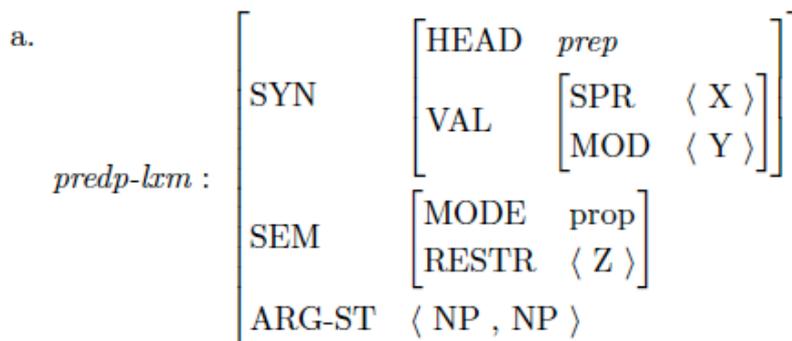


13.4.3 Constant Lexemes

(201)



(202) Zwei Arten Präpositionen: *predicational-preposition-lexeme* (a.) und *argument-marking-preposition-lexeme* (b.)



$$\text{b.} \quad \text{argmkp-lexm} : \left[\begin{array}{l} \text{SYN} \quad \left[\begin{array}{l} \text{HEAD} \quad \textit{prep} \\ \text{VAL} \quad \left[\text{SPR} \quad \langle \rangle \right] \end{array} \right] \\ \text{SEM} \quad \left[\begin{array}{l} \text{MODE} \quad \boxed{1} \\ \text{INDEX} \quad \boxed{2} \\ \text{RESTR} \quad \langle \rangle \end{array} \right] \\ \text{ARG-ST} \quad \left\langle \begin{array}{l} \text{NP} \\ \left[\begin{array}{l} \text{MODE} \quad \boxed{1} \\ \text{INDEX} \quad \boxed{2} \end{array} \right] \end{array} \right\rangle \end{array} \right]$$

(B396) *I wrapped a blanket around me.*

(203) Constraints für weitere nicht-flektierende Wortarten

$$\begin{array}{l}
 \text{a.} \quad \text{adj-lexm} : \left[\begin{array}{l} \text{SYN} \quad \left[\begin{array}{l} \text{HEAD} \quad \textit{adj} \\ \text{VAL} \quad \left[\begin{array}{l} \text{SPR} \quad \langle X \rangle \\ \text{MOD} \quad \langle [\text{HEAD} \quad \textit{noun}] \rangle \end{array} \right] \end{array} \right] \\ \text{ARG-ST} \quad \langle \text{NP}, \dots \rangle \\ \text{SEM} \quad [\text{MODE} \quad \textit{prop}] \end{array} \right] \\
 \text{b.} \quad \text{adv-lexm} : \left[\begin{array}{l} \text{SYN} \quad \left[\begin{array}{l} \text{HEAD} \quad \textit{adv} \\ \text{VAL} \quad [\text{MOD} \quad \langle [\text{HEAD} \quad \textit{verb}] \rangle] \end{array} \right] \\ \text{SEM} \quad [\text{MODE} \quad \textit{none}] \end{array} \right] \\
 \text{c.} \quad \text{conj-lexm} : \left[\begin{array}{l} \text{SYN} \quad [\text{HEAD} \quad \textit{conj}] \\ \text{SEM} \quad [\text{MODE} \quad \textit{none}] \\ \text{ARG-ST} \quad \langle \rangle \end{array} \right] \\
 \text{d.} \quad \text{det-lexm} : \left[\begin{array}{l} \text{SYN} \quad \left[\begin{array}{l} \text{HEAD} \quad \textit{det} \\ \text{VAL} \quad \left[\begin{array}{l} \text{SPR} \quad / \langle \rangle \\ \text{COMPS} \quad \langle \rangle \end{array} \right] \end{array} \right] \\ \text{SEM} \quad [\text{MODE} \quad \textit{none}] \end{array} \right]
 \end{array}$$

13.4.4 Lexemes vs. Parts of Speech (pos)

The subtypes of pos specify which features are appropriate for particular categories of words and phrases. They thus serve to organize the various parts of speech that our grammar has to recognize. The subtypes of lexeme, on the other hand introduce constraints on what combinations of feature values are possible, for example, the SHAC or the constraint that verbs require propositional mode.

(SWB p.245)

13.4.5 The Case Constraint

Bezogen auf *words*

(204) Case Constraint (CC)

An outranked NP is [CASE acc]

-----> Nur in seltenen Fällen der Kasus von Komplementen spezifiziert werden.

Aus CC folgt freilich nicht, dass eine NP im Akkusativ einen niedrigeren Rang als irgendeine andere haben müsste.