

# Formale Syntax: HPSG

## 02. Merkmalstrukturen und Merkmalbeschreibungen

Roland Schäfer

Institut für Germanistische Sprachwissenschaft  
Friedrich-Schiller-Universität Jena

Stets aktuelle Fassungen: <https://github.com/rsling/VL-HPSG>

Basiert teilweise auf Folien von Stefan Müller: <https://hpsg.hu-berlin.de/~stefan/Lehre/S2021/hpsg.html>

Stefan trägt natürlich keinerlei Verantwortung für meine Fehler und Missverständnisse!

## Übersicht

- 1 Phrasenstruktur und Phrasenstrukturgrammatiken
- 2 Merkmalstrukturen und Merkmalbeschreibungen
- 3 Komplementation und Grammatikregeln
- 4 Verbsemantik und Linking (Semantik 1)
- 5 Adjunktion und Spezifikation
- 6 Lexikon und Lexikonregeln
- 7 Konstituentenreihenfolge und Verbbewegung
- 8 Nicht-lokale Abhängigkeiten und Vorfelddbesetzung
- 9 Quantorenspeicher (Semantik 2)
- 10 Unterspezifikationssemantik (Semantik 3)

<https://rolandschaefer.net/archives/2805>

<https://github.com/rsling/VL-HPSG/tree/main/output>

<https://hpsg.hu-berlin.de/~stefan/Pub/hpsg-lehrbuch.html>

# Einleitung

Worum geht es heute?

- Repräsentation von Merkmalen und ihren Werten in Grammatiken
- Strukturierte/hierarchische Merkmalstrukturen
- Unifikation von Merkmalstrukturen
- Merkmalstrukturen vs. Merkmalbeschreibungen

Müller (2013: Kapitel 2)

## Merken Sie sich die Strukturen von heute nicht als „korrekte Modellierung“ des Deutschen in HPSG!

Wir nehmen heute einige Vereinfachungen und Didaktisierungen vor,  
denn es geht darum, grundlegende Repräsentationen/Prinzipien einzuführen.

Völlig abwegig sind die Strukturen dieser Lektion aber auch nicht.

Generell haben Sie mehr davon, wenn Sie in jeder Woche zu verstehen versuchen,  
warum sich bestimmte Repräsentationen wieder ändern, als wenn Sie von Anfang an  
nur wissen wollen, wie das Endergebnis in den Prüfungen aussehen wird.

## Merkmalestrukturen

Problem mit einfachen **Phrasenstrukturgrammatiken**

- **Symbolinflation** | Selbst für einfachste Valenz-/Kongruenzphänomene
- Viele Regeln und viele Kategorien

**Merkmalstrukturen** wie in HPSG

- **Komplexe Symbole**, dadurch weniger Symbole
- Extrem einfache **Regeln** (Kombinatorik)



Merkmalstrukturen modellieren linguistische Objekte.

- Merkmal-Wert-Struktur
- Attribut-Wert-Struktur
- *Feature structure*

Wir nutzen Merkmalsbeschreibungen, um über Merkmalstrukturen zu sprechen.

- *Attribute-value matrix*
- *Feature matrix*

Shieber (1986), Pollard & Sag (1987), Johnson (1988), Carpenter (1992), King (1994), Richter (2004, 2021)

## Einfache Merkmalbeschreibung

$$\left[ \text{ATTRIBUT} \quad \text{wert} \right]$$

## Mehrere Attribut-Wert-Paare in einer Struktur

$$\begin{bmatrix} \text{ATTRIBUT}_1 & \text{wert} \\ \text{ATTRIBUT}_2 & \text{wert} \\ \dots & \dots \end{bmatrix}$$

## Komplexe Merkmale können Werte von Attributen sein!

$$\left[ \begin{array}{l} \text{ATTRIBUT}_1 \quad \text{wert} \\ \text{ATTRIBUT}_2 \quad \left[ \begin{array}{l} \text{ATTRIBUT}_{2-1} \quad \text{wert} \\ \text{ATTRIBUT}_{2-2} \quad \text{wert} \end{array} \right] \end{array} \right]$$

PHONE oder GRAPHEN | Aussprache bzw. Schreibung

[GRAPHEN *Tisch*]

Aber reicht diese Datenstruktur?

- *Tisch* | Sieht aus wie ein **Symbol** ohne Struktur
- Phonetik/Phonologie | **Ketten** von Phonen/Phonemen  
Bei Schäfer (2018) und anderen: Segmente
- Phonologische Grammatik | Zugriff auf einzelne Segmente  
Auslautverhärtung | Zugriff auf letztes Segment einer Silbe

## Lösung für GRAPH(EN) oder PHON(E) | Geordnete Listen

$\left[ \text{GRAPH} \langle T, i, s, c, h \rangle \right]$

Auf einer Liste stehen eigentlich auch Merkmalbeschreibungen.

$\left[ \text{PHON} \left\langle \left[ \begin{array}{ll} \text{MANNER} & \textit{plosive} \\ \text{PLACE} & \textit{alveolar} \end{array} \right], \left[ \begin{array}{ll} \text{MANNER} & \textit{vokal} \\ \text{BACKNESS} & \textit{front} \\ \text{HEIGHT} & \textit{high} \\ \text{ROUND} & \textit{no} \end{array} \right], \left[ \begin{array}{ll} \text{MANNER} & \textit{fricative} \\ \text{PLACE} & \textit{alveolar} \end{array} \right] \right\rangle \right]$

Strenggenommen falsche Kurzschreibweisen für PHON in typischer HPSG

$\left[ \text{PHON} \textit{Tisch} \right]$

$\left[ \text{PHON} \langle \textit{Tisch} \rangle \right]$

# Morphosyntaktische Merkmale

Lösung für Probleme mit Genus usw. in PSGs von letzter Woche

|                |                  |
|----------------|------------------|
| PHON           | <i>Tisch</i>     |
| PART-OF-SPEECH | <i>noun</i>      |
| GENDER         | <i>masculine</i> |
| NUMBER         | <i>sg</i>        |
| CASE           | <i>nom</i>       |

Andere Merkmalausstattungen = andere sprachliche Zeichen

|      |              |      |              |
|------|--------------|------|--------------|
| PHON | <i>Tisch</i> | PHON | <i>Tisch</i> |
| POS  | <i>n</i>     | POS  | <i>n</i>     |
| GEN  | <i>masc</i>  | GEN  | <i>masc</i>  |
| NUM  | <i>sg</i>    | NUM  | <i>sg</i>    |
| CAS  | <i>acc</i>   | CAS  | <i>dat</i>   |

Abgekürzte Schreibweise mit *oder* bzw.  $\vee$

|      |  |
|------|--|
| PHON | <i>Tisch</i>   |
| POS  | <i>n</i>   |
| GEN  | <i>masc</i>  |
| NUM  | <i>sg</i>  |
| CAS  | <i>nom <math>\vee</math> acc <math>\vee</math> dat</i> |

# Dasselbe für eine Verbform

Verben | Teilweise dieselben, teilweise andere Merkmale verglichen mit Nomina

|      |              |
|------|--------------|
| PHON | <i>sieht</i> |
| POS  | <i>v</i>     |
| PER  | <i>3</i>     |
| NUM  | <i>sg</i>    |

Syntaktisch relevant auch **Fintheit** bzw. **Status**

|              |              |               |                |
|--------------|--------------|---------------|----------------|
| PHON         | <i>sieht</i> | PHON          | <i>gesehen</i> |
| POS          | <i>v</i>     | POS           | <i>v</i>       |
| PER          | <i>3</i>     | <b>FINIT</b>  | <i>false</i>   |
| NUM          | <i>sg</i>    | <b>STATUS</b> | <i>3</i>       |
| <b>FINIT</b> | <i>true</i>  |               |                |

Typen

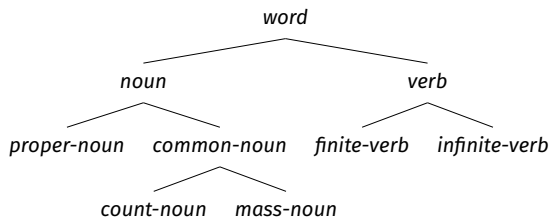
Nicht alle Wörter haben alle Merkmale. | **Typen** und **Beschränkungen** über Typen

|          |               |
|----------|---------------|
| <i>n</i> |               |
| PHON     | <i>Tischs</i> |
| GEN      | <i>masc</i>   |
| NUM      | <i>sg</i>     |
| CAS      | <i>gen</i>    |

|                    |              |                      |                |
|--------------------|--------------|----------------------|----------------|
| <i>finite-verb</i> |              | <i>infinite-verb</i> |                |
| PHON               | <i>sieht</i> | PHON                 | <i>gesehen</i> |
| PER                | <i>3</i>     | STATUS               | <i>3</i>       |
| NUM                | <i>sg</i>    |                      |                |
| TENSE              | <i>pres</i>  |                      |                |
| MOD                | <i>ind</i>   |                      |                |



Typen sind sehr wichtig in HPSG und bilden **Hierarchien**. Denkbare Beispiel:



- Typen sind die eigentlichen **Wortarten** in HPSG.
- **Monotonizität** | **Untertypen** erben alle Merkmale/Beschränkungen ihrer **Obertypen**.
- **Mehrfachvererbung** | Ein Typ kann **mehrere Obertypen** haben.
- Keine Sorge! Dazu kommen wir noch im Detail.

## Strukturteilung

Letzte Woche in PSGs | Valenz doppelt in **Kategorien** und **Regeln** kodiert

**Regel** für Satz mit intransitivem Verb

$S \rightarrow NP(\text{Per}, \text{Num}, \text{nom}) V_{\text{itr}}(\text{Per}, \text{Num})$

**Regel** für Satz mit transitivem Verb

$S \rightarrow NP(\text{Per}_1, \text{Num}_1, \text{nom}) NP(\text{Per}_2, \text{Num}_2, \text{akk}) V_{\text{tr}}(\text{Per}_1, \text{Num}_1)$

**Regel** für Satz mit ditransitivem Verb

$S \rightarrow NP(\text{Per}_1, \text{Num}_1, \text{nom}) NP(\text{Per}_2, \text{Num}_2, \text{dat}) NP(\text{Per}_3, \text{Num}_3, \text{akk}) V_{\text{dtr}}(\text{Per}_1, \text{Num}_1)$

**Typische Definition von Valenz allerdings**

Die **Liste** der Ergänzungen eines Worts.

## Valenz | Liste von Merkmalsbeschreibungen

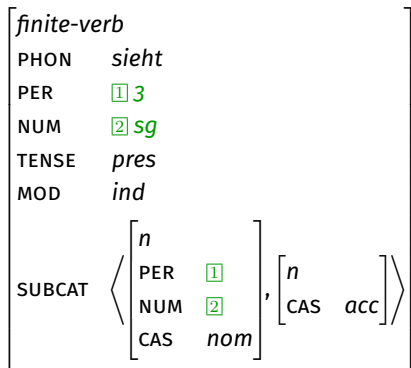
Bezeichnung für Valenz in HPSG: SUBCATEGORISATION, kurz SUBCAT

|                    |                            |
|--------------------|----------------------------|
| <i>finite-verb</i> |                            |
| PHON               | <i>sieht</i>               |
| PER                | 3                          |
| NUM                | <i>sg</i>                  |
| TENSE              | <i>pres</i>                |
| MOD                | <i>ind</i>                 |
| SUBCAT             | $\langle [n], [n] \rangle$ |

Valenzliste | Hinreichend eingrenzende Beschreibung der Ergänzungen des Verbs

|   |              |
|---|--------------|
| <i>finite-verb</i>  |              |
| PHON  | <i>sieht</i> |
| PER   | 3            |
| NUM   | <i>sg</i>    |
| TENSE   | <i>pres</i>  |
| MOD   | <i>ind</i>   |
| SUBCAT $\left\langle \begin{bmatrix} n \\ \text{PER} & 3 \\ \text{NUM} & sg \\ \text{CAS} & nom \end{bmatrix}, \begin{bmatrix} n \\ \text{CAS} & acc \end{bmatrix} \right\rangle$ |              |

## Übereinstimmung von Merkmalen | Hart verdrahtet durch Strukturteilung

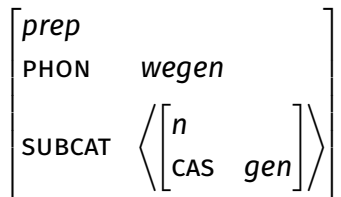


Strukturteilung bedeutet Token-Identität von Werten, nicht Kopie!

Man kann sich die Nummern als Zeiger auf dieselbe Datenstruktur vorstellen.

# Beispiel für Valenz einer Präposition

Valenz von Präpositionen | NP in einem bestimmten Kasus



- Was ist mit **argumentmarkierenden Präpositionen**/Präpositionalobjekten?  
*leiden unter, abhängen von, glauben an* usw.
- Was ist mit **Wechselpräpositionen** mit Akkusativ oder Dativ?  
*unter, neben, über* usw.

## Phrasen und Kopfmerkmale



Kongruenzmerkmale innerhalb der NP auch beim Determinierer erforderlich

|            |             |
|------------|-------------|
| <i>det</i> |             |
| PHON       | <i>des</i>  |
| GEN        | <i>masc</i> |
| NUM        | <i>sg</i>   |
| CAS        | <i>gen</i>  |

# Determinierer in der NP

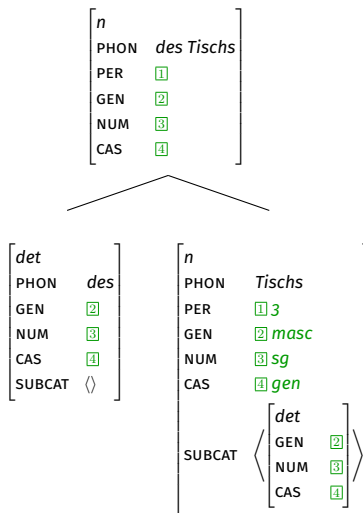
DP oder NP? | Für Deutsch ist eine NP-Analyse näherliegend.

|        |   |
|--------|---|
| $n$    |   |
| PHON   | <i>Tischs</i>   |
| PER    | 3   |
| GEN    | 1 <i>masc</i>   |
| NUM    | 2 <i>sg</i>   |
| CAS    | 3 <i>gen</i>  |
| SUBCAT | $\left\langle \begin{bmatrix} \textit{det} \\ \text{GEN} & 1 \\ \text{NUM} & 2 \\ \text{CAS} & 3 \end{bmatrix} \right\rangle$ |

Wie kann man Notwendigkeit von und Kongruenz mit Determinierern kodieren?

# NP mit Kongruenz als Baum

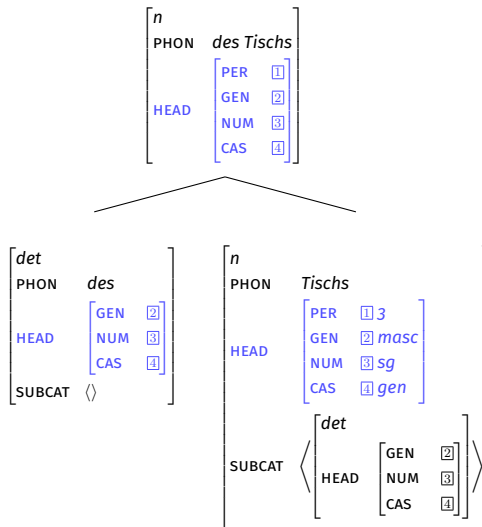
In HPSG gibt es eigentlich keine Bäume. Zur Illustration aber hilfreich:



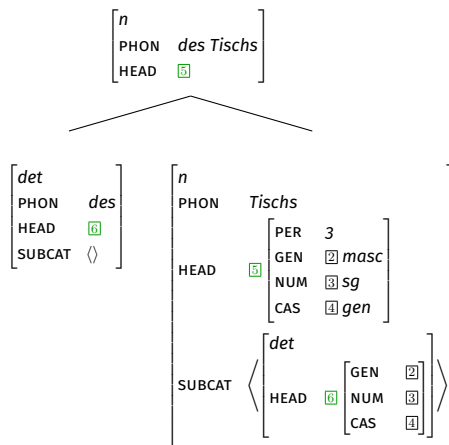
Wir haben jetzt so getan, **als hätten wir schon eine Syntax!**

- Eigentlich **nur Lexikoneinträge**
- Fehlende **Regeln für Kombinationsmechanismus**
- NP auf der letzten Folie | Nur eine grobe Idee, wo wir hin wollen
- **Projektionsebenen** (N vs. NP) nicht unterscheidbar
- Also auch keine Identifikation von **Köpfen**
- Identifikation der **Merkmale, die vom Kopf zur Phrase projizieren**
- Zusammenbau von ***des Tischs*** aus ***des*** und ***Tischs***

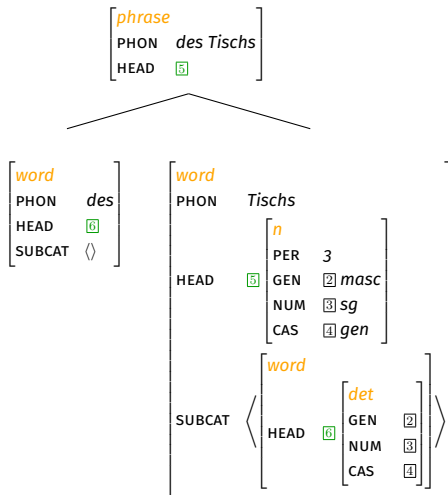
Head features | Bündel der Merkmale, die vom Kopf zur Phrase projizieren



Durch Merkmalbündel | Optimale Struktur finden/Generalisierungen abbilden



Wortartenspezifisch sind die **HEAD-Bündel**, nicht die **Wörter/Phrasen**.



Unifikation



# Zusammenlegen von Informationen

## Beispiel | Lexikalische Spezifikation der Valenz einer Präposition

|        |   |
|--------|---|
| word   |   |
| PHON   | wegen   |
| HEAD   | [prep]  |
| SUBCAT | $\left\langle \left[ \text{HEAD } \boxed{1} \left[ \begin{array}{l} n \\ \text{CAS } gen \end{array} \right] \right] \right\rangle$ |

Die NP kommt mit viel mehr Information daher.

|        |   |
|--------|---|
| phrase |   |
| PHON   | des Tischs  |
| HEAD   | $\boxed{1} \left[ \begin{array}{l} n \\ \text{PER } 3 \\ \text{GEN } masc \\ \text{NUM } sg \\ \text{CAS } gen \end{array} \right]$ |
| SUBCAT | $\langle \rangle$   |

Die Informationen unter  $\boxed{1}$  sind aber kompatibel und **unifizieren** daher.

## Unifikation | Mehrere Merkmalstrukturen zu einer machen

Bedingungen für Unifikation von zwei Merkmalstrukturen A und B:

- A und B enthalten **keine widersprüchlichen Informationen**.

$\begin{bmatrix} \text{CAS} & \text{nom} \\ \text{PER} & 3 \end{bmatrix}$  und  $\begin{bmatrix} \text{CAS} & \text{acc} \end{bmatrix}$  unifizieren nicht.

- Aus nicht widersprüchlichen Informationen wird die **Vereinigungsmenge** gebildet.
- A kann mehr Informationen enthalten als B oder umgekehrt.

$\begin{bmatrix} \text{CAS} & \text{nom} \\ \text{PER} & 3 \end{bmatrix}$  und  $\begin{bmatrix} \text{CAS} & \text{nom} \end{bmatrix}$  unifizieren zu  $\begin{bmatrix} \text{CAS} & \text{nom} \\ \text{PER} & 3 \end{bmatrix}$

- A und B können beide mehr Informationen enthalten als die jeweils andere.

$\begin{bmatrix} \text{CAS} & \text{nom} \\ \text{PER} & 3 \end{bmatrix}$  und  $\begin{bmatrix} \text{CAS} & \text{nom} \\ \text{NUM} & \text{sg} \end{bmatrix}$  unifizieren zu  $\begin{bmatrix} \text{CAS} & \text{nom} \\ \text{PER} & 3 \\ \text{NUM} & \text{sg} \end{bmatrix}$

Nächste Woche

Nächste Woche geht es um Valenz und Valenzabbildung.

Sie sollten dringend vorher aus dem HPSG-Buch  
Abschnitt 3.1 und Kapitel 4 lesen!

Das sind gerade mal 15 Seiten.

- Carpenter, Bob. 1992. *The Logic of Typed Feature Structures*. (Cambridge Tracts in Theoretical Computer Science 32). Cambridge: Cambridge University Press.
- Johnson, Mark. 1988. *Attribute-Value Logic and the Theory of Grammar*. (CSLI Lecture Notes 16).
- King, Paul. 1994. *An Expanded Logical Formalism for Head-Driven Phrase Structure Grammar*. Arbeitspapiere des SFB 340 Nr 59. Tübingen: Universität.
- Müller, Stefan. 2013. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 3. Aufl. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.
- Pollard, Carl & Ivan A. Sag. 1987. *Information-Based Syntax and Semantics*. (CSLI Lecture Notes 13).
- Richter, Frank. 2004. *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar*. Universität Tübingen Phil. Dissertation (2000).
- Richter, Frank. 2021. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (Hrsg.), *Head-Driven Phrase Structure Grammar: The Handbook*, 89–124. Berlin.
- Schäfer, Roland. 2018. *Einführung in die grammatische Beschreibung des Deutschen*. 3. Aufl. (Textbooks in Language Sciences 2). Berlin.
- Shieber, Stuart M. 1986. *An Introduction to Unification-Based Approaches to Grammar*. (CSLI Lecture Notes 4). republished as 2003. *An Introduction to Unification-Based Approaches to Grammar*. Brookline, MA: Microtome Publishing, 2003.

## Kontakt

Prof. Dr. Roland Schäfer  
Institut für Germanistische Sprachwissenschaft  
Friedrich-Schiller-Universität Jena  
Fürstengraben 30  
07743 Jena

<https://rolandschaefer.net>  
[roland.schaefer@uni-jena.de](mailto:roland.schaefer@uni-jena.de)

## Creative Commons BY-SA-3.0-DE

Dieses Werk ist unter einer Creative Commons Lizenz vom Typ *Namensnennung - Weitergabe unter gleichen Bedingungen 3.0 Deutschland* zugänglich. Um eine Kopie dieser Lizenz einzusehen, konsultieren Sie

<http://creativecommons.org/licenses/by-sa/3.0/de/> oder wenden Sie sich brieflich an Creative Commons, Postfach 1866, Mountain View, California, 94042, USA.