

# Formale Syntax: HPSG

## 01. Phrasenstruktur und Phrasenstrukturgrammatik

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>

Grundlage ist Stefans HPSG-Buch: <https://hpsg.hu-berlin.de/~stefan/Pub/hpsg-lehrbuch.html.de>

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>

Worum geht es heute?

- Vermittlung grundlegender Vorstellungen über deutsche Syntax
- Vorstellung für die Daten, Zusammenhänge und Komplexität
- Einführung in Grundannahmen in der HPSG
- Befähigung zum Schreiben formaler Grammatiken

Müller (2013b: Kapitel 1) bzw. Müller (2013a: Kapitel 1)  
Englische Version des Grammatiktheoriebuches: Müller (2020: Kapitel 1)

Wozu (formale) Syntax?

- **Zeichen** | Form-Bedeutungs-Paare (de Saussure 1916)
- Wörter, Wortgruppen, Sätze
- Sprache | **keine** (endliche) **Aufzählung** von Wortfolgen  
Endlichkeit von Sprache bei Annahme einer maximalen Satzlänge
  - (1) Dieser Satz geht weiter und weiter und weiter und weiter ...
  - (2) [Ein Satz ist ein Satz] ist ein Satz.
- Auf jeden Fall **sehr viele Sätze**, Unendlichkeitsproblem als Scheinfrage
- **Kompetenz** | (implizites) Wissen um grammatische Regularitäten
- **Performanz** | Nutzung des Wissens, Sprachproduktion
- **Kreativität** | Sätze bilden, die man nie zuvor gehört hat

Schon Kindern kann man ein Spiel um Kompetenz und Performanz zumuten:

*Und wir beeilten uns, den Jungen zu erzählen, wir hätten von Anfang an gewusst, dass es nur eine Erfindung von Lasse gewesen sei. Und da sagte Lasse, die Jungen hätten gewusst, dass wir gewußt hätten, es sei nur eine Erfindung von ihm. Das war natürlich gelogen, aber vorsichtshalber sagten wir, wir hätten gewusst, die Jungen hätten gewusst, dass wir gewusst hätten, es sei nur eine Erfindung von Lasse. Und da sagten die Jungen – ja – jetzt schaffe ich es nicht mehr aufzuzählen, aber es waren so viele „gewusst“, dass man ganz verwirrt davon werden konnte, wenn man es hörte.*

- **Grammatikalität** der Sätze | Einwandfrei feststellbar
- **Akzeptabilität** der Sätze | Vermindert durch **Performanzeffekte**

# Wozu Syntax? Bedeutung aus Bestandteilen ermitteln

Bedeutung einer Äußerung aus den Bedeutungen ihrer Teile bestimmen

- (3) Der Mann kennt den Kollegen.

**Syntax** | Art und Weise der Kombination, Strukturierung

- (4) a. Die Frau kennt die Kolleginnen.  
b. Die Frau kennen die Kolleginnen.

## Das Frege-Prinzip (Gottlob Frege, 1879)

Die Bedeutung eines Satzes ergibt sich aus der Bedeutung seiner Konstituenten und der Art ihrer Kombination.



# Warum formal?

*Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects. I think that some of those linguists who have questioned the value of precise and technical development of linguistic theory have failed to recognize the productive potential in the method of rigorously stating a proposed theory and applying it strictly to linguistic material with no attempt to avoid unacceptable conclusions by ad hoc adjustments or loose formulation. (Chomsky 1957: S. 5)*

*As is frequently pointed out but cannot be overemphasized, an important goal of formalization in linguistics is to enable subsequent researchers to see the defects of an analysis as clearly as its merits; only then can progress be made efficiently. (Dowty 1979: S. 322)*

Das bringt mir doch nichts für den Unterricht in der 5. oder 10. Klasse!

Erste Antwortmöglichkeit:

**Seien Sie froh!** Sie können jetzt im pessimistischsten Fall zum letzten Mal vor der Rente etwas machen, das Ihr Gehirn weiterbringt und nicht an die Zwecke der Arbeit gebunden ist.

Das ist aber in unserem Fall nicht die ganze Wahrheit ...

Sie möchten den **Bildungsspracherwerb** von Kindern/Jugendlichen fördern.  
Die Anforderungen an Sie ergeben sich aus den **Zielkompetenzen Ihrer Schüler**.

## Zielkompetenzen *Deutsch* 5.–11. Klasse (Thüringer RLP 2019; S. 7)

- 1 Texte rezipieren
- 2 Texte produzieren
- 3 Über Sprache, Sprachverwendung und Sprachenlernen reflektieren

## Das grammatikbezogene Aufgabenspektrum für Lehrpersonen

- Bildungssprache/Sprachbewusstheit unterrichten
- Sprachliche Leistungen fair bewerten
- Bewertungen und Lösungsstrategien erklären
- Deutsche Sprache vermitteln (falls nicht L1)
- Wie soll das ohne fundierte Grammatikkenntnisse funktionieren?
- Nach Morphologie, Syntax-Vorlesung und Syntax-Seminar geht es hier weiter!

## Konstituenz

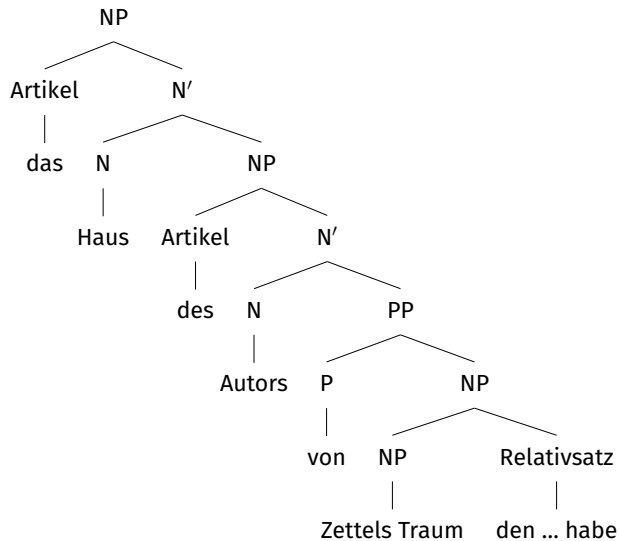
## Parataxe | Einbettung von ganzen Satzstrukturen

- (5) dass Max glaubt, [dass Julius weiß, [dass Otto behauptet, [dass Karl vermutet, [dass Richard bestätigt, [dass Friederike lacht]]]]]

## Parataxe als Spezialfall | Konstituenten in Konstituenten

- (6) [das Haus [des Autors [von Zettel Traum [den ich 1993 gelesen habe]]]]  
(7) [[den]][ich][1993][[gelesen]habe]]

# Naive Konstituenzanalyse



Welche **Konstituententests** kennen Sie?

- Substituierbarkeit/Pronominalisierungstest/Fragetest
- Weglasstest
- Verschiebetest (Umstelltest)/Vorfeldtest
- Koordinationstest



**Substituierbarkeit** Austauschbare Wortfolgen als potenzielle Konstituenten

- (8) Er kennt **den Mann**.
- (9) Er kennt **eine Frau**.

**Pronominalisierungstest** Dasselbe, aber spezifisch mit pronominalen Ein-Wort-Folgen

- (10) **Der Mann** schläft.
- (11) **Er** schläft.

**Fragetest** Erfragbarkeit von Konstituenten

(12) **Der Mann** arbeitet.

(13) **Wer** arbeitet?

**Verschiebetest** Umstellbarkeit von Konstituenten

(14) weil **keiner diese Frau** kennt.

(15) weil **diese Frau keiner** kennt.

**Koordinationstest** Konstituenten als koordinierbar

(16) **[[Der Mann] und [die Frau]]** arbeiten.

Köpfe

**Kopf** | Festlegung der syntaktisch relevanten **kategorialen Merkmale der Phrase**

(17) **Träumt** er?

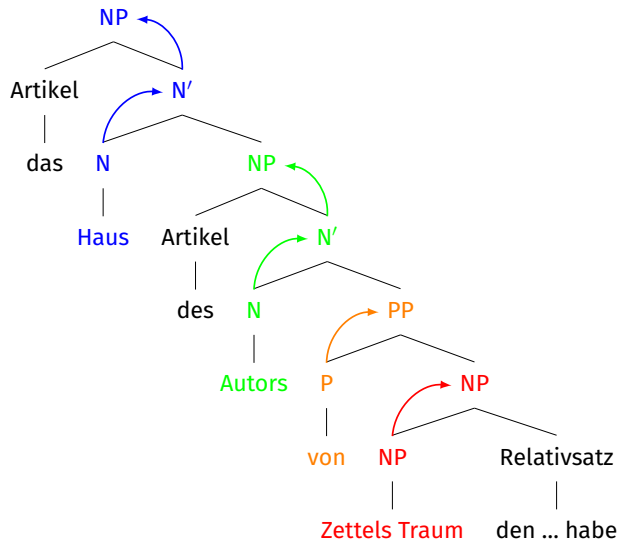
(18) **Erwartet** er einen dreiprozentigen Anstieg?

(19) **in** diesem Haus

(20) ein **Mann**

- **Projektion** | Kombination eines Kopfes mit anderem Material
- **Maximalprojektion** | Vollständige Projektion
- **Satz** | Maximalprojektion eines finiten Verbs

# Naive Konstituenzanalyse mit Markierung der Köpfe



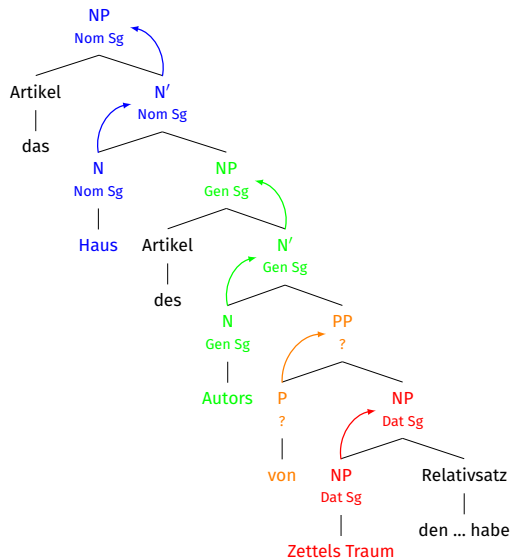
Der **interne Aufbau** einer Phrase ist für den Kontext **irrelevant**:

- (21) er
- (22) der Mann
- (23) der Mann aus Stuttgart
- (24) der Mann aus Stuttgart, den wir kennen

Bestimmte **Merkmale** des Kopfs sind aber **kontextrelevant**:

- (25) Der Kollege liest einen Aufsatz.
- (26) \* Die Kollegen liest einen Aufsatz.
- (27) \* Des Kollegen liest einen Aufsatz.

# Naive Konstituenzanalyse mit Projektion von Kopfmerkmalen



## Argumente und Adjunkte



Nicht alle Phrasen, die vom Verb abhängen, stehen in derselben Art Relation zu ihm.

- Konstituenten | Verschiedenartige Beziehungen zu ihrem Kopf
- Semantische Beteiligte – **Aktanten** – als **feste Teile der Verbbedeutung**
- Semantik von *sehen* | Immer ein **Sehender**, ein **Gesehenes**

(28) Dani sieht den Chaoten.

- **Logische Argumente von *sehen*** | Dani und der Chaot
- Valenz | Abbildung logischer Argumente auf grammatische Argumente

Semantische Argumente | Nicht immer syntaktisch erforderlich

(29) Er wartet auf den Installateur.

(30) Er wartet.

Bei Nominalisierung | Alle Argumente optional

(31) Arno liest diese Bücher.

(32) das Lesen dieser Bücher durch Arno

(33) das Lesen dieser Bücher

(34) das Lesen

Adjunkte | Keine verbgebundene, sondern selbst mitgebrachte Rolle

(35) Dani sieht den Chaoten bellend auf der Brücke.

Deutliche Unterschiede zwischen Argumenten und Adjunkten

- Sehende und Gesehener | Fester Teil einer *sehen*-Situation
- Ort | Teil so ziemlich jedes Geschehens, nicht *sehen*-spezifisch
- Verhalten des Beteiligten | Erst recht nicht *sehen*-spezifisch

## Üblicher Terminologie-Wildwuchs in der Linguistik

- Argument = Ergänzung
- Adjunkt = (freie) Angabe
- Argumente | Beim Verb aufgeteilt in Subjekte und Komplemente
- Aktant Subjekte und Objekte (nicht Prädikative und Adverbiale)
- Adverbial | Angabe beim Verb
  - ▶ Raum (Lage, Richtung/Ziel, Herkunft, Weg)
  - ▶ Zeit (Zeitpunkt, Anfang, Ende, Dauer)
  - ▶ Grund (inkl. Gegengrund, Bedingung)
  - ▶ Art und Weise

## Grammatische Funktionen

Grammatische Funktionen/Relationen sind oft nicht unabhängig definierbar!

- Typen von Argumenten/Adjunkten mit spezifischen Eigenschaften
- **Subjekt** | Siehe nächste Folien
- **Objekt/Komplement** | Nicht-Nominativ-Argumente
- **Adverb/Adverbiale Bestimmung** | Angabe des Verbs

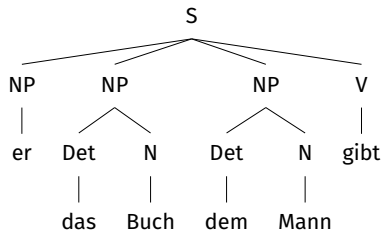
Für **deutsche Subjekte** benannte definitorische Kriterien:

- 1 **Kongruenz** mit dem finiten Verb
- 2 **Nominativ** in nicht-kopulativen Sätzen
- 3 Weglassbarkeit in **Infinitivkonstruktionen** (Kontrolle)
- 4 Weglassbarkeit in **Imperativsätzen**

Reis (1982) | Nur (2) relevant!

## Phrasenstrukturgrammatiken

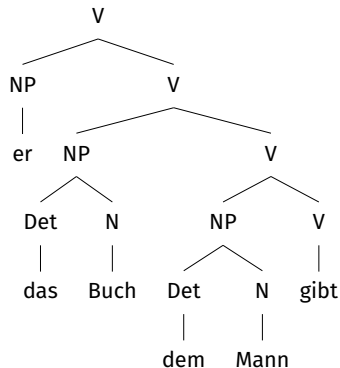




Grammatik

$NP \rightarrow Det\ N$

$S \rightarrow NP\ NP\ NP\ V$



Grammatik

$NP \rightarrow Det\ N$

$V \rightarrow NP\ V$

# Wie PSG-Regeln als Ersetzungsregeln funktionieren

## Ersetzungsregeln und Bäume als Protokoll der Ersetzung

Grammatik

NP → Det N

S → NP NP NP V

Lexikon (gleiches Format)

NP → er

Det → das

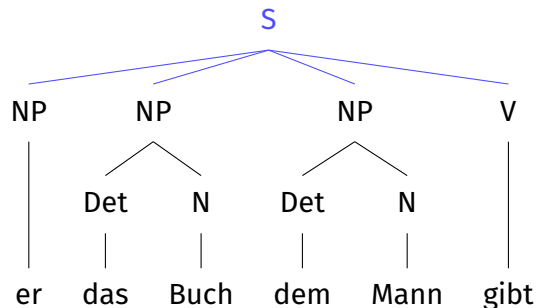
Det → dem

N → Buch

N → Mann

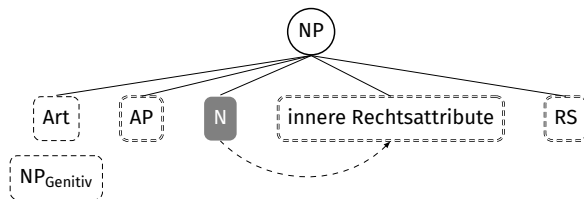
V → gibt

er	das	Buch	dem	Mann	gibt
NP	das	Buch	dem	Mann	gibt
NP	Det	Buch	dem	Mann	gibt
NP	Det	N	dem	Mann	gibt
NP		NP	dem	Mann	gibt
NP		NP	Det	Mann	gibt
NP		NP	Det	N	gibt
NP		NP		NP	gibt
NP		NP		NP	V
					S



# Phrasenstrukturschemata

Manche kennen die **Phrasenschemata** aus Schäfer (2018).



Es handelt sich um **abgekürzte Phrasenstrukturregeln**.

$NP \rightarrow N$

*Bücher*

$NP \rightarrow N \text{ Rechtsattribut}^n$

*Bücher über Poe*

$NP \rightarrow N \text{ RS}^n$

*Bücher, die gefallen*

$NP \rightarrow \text{Art } N$

*das Buch*

$NP \rightarrow \text{Art } N \text{ Rechtsattribut}^n$

*das Buch über Poe*

$NP \rightarrow \text{Art } N \text{ RS}^n$

*das Buch, das gefällt*

$NP \rightarrow \text{NP}_{\text{Gen}} N$

*Arnos Buch*

$NP \rightarrow \text{NP}_{\text{Gen}} N \text{ Rechtsattribut}^n$

*Arnos Buch über Poe*

$NP \rightarrow \text{NP}_{\text{Gen}} N \text{ RS}^n$

*Arnos Buch, das gefällt*

usw.

$NP \rightarrow (\text{Art} \mid \text{NP}_{\text{Gen}}) (\text{AP}^n) N (\text{Rechtsattribut}^n) (\text{RS}^n)$

Rechtsattribut  $\rightarrow \text{PP}, \text{NP}_{\text{Gen}}, \text{CP}, \text{IP}, \dots$

Die folgende Grammatik **übergeneriert!**

NP → Det N

S → NP NP NP V

(36) er das Buch dem Mann gibt

(37) \* ich das Buch dem Mann gibt  
Subjekt-Verb-Kongruenz | *ich – gibt*

(38) \* er das Buch den Mann gibt  
Valenz/Rektion | *gibt + Dativ*

(39) \* er den Buch dem Mann gibt  
Determinator-Nomen-Kongruenz | *den – Buch*

Übereinstimmung in **Person (1, 2, 3)** und **Numerus (sg, pl)**

- (40) Ich schlafe. (1, sg)
- (41) Du schläfst. (2, sg)
- (42) Er schläft. (3, sg)
- (43) Wir schlafen. (1, pl)
- (44) Ihr schlaft. (2, pl)
- (45) Sie schlafen. (3,pl)

Wie drückt man das in Regeln aus?

Verfeinerung der verwendeten Symbole | Statt  $S \rightarrow NP\ NP\ NP\ V$

$S \rightarrow NP\_1\_sg\ NP\ NP\ V\_1\_sg$

$S \rightarrow NP\_2\_sg\ NP\ NP\ V\_2\_sg$

$S \rightarrow NP\_3\_sg\ NP\ NP\ V\_3\_sg$

$S \rightarrow NP\_1\_pl\ NP\ NP\ V\_1\_pl$

$S \rightarrow NP\_2\_pl\ NP\ NP\ V\_2\_pl$

$S \rightarrow NP\_3\_pl\ NP\ NP\ V\_3\_pl$

**Sechs Regeln** ( $3 \times 2$ ) statt einer!

Hier für ein Valenzmuster (**ditransitiv**) die Kongruenzkodierung.

S → NP\_1\_sg\_nom NP\_dat NP\_acc V\_1\_sg\_ditransitiv  
S → NP\_2\_sg\_nom NP\_dat NP\_acc V\_2\_sg\_ditransitiv  
S → NP\_3\_sg\_nom NP\_dat NP\_acc V\_3\_sg\_ditransitiv  
S → NP\_1\_pl\_nom NP\_dat NP\_acc V\_1\_pl\_ditransitiv  
S → NP\_2\_pl\_nom NP\_dat NP\_acc V\_2\_pl\_ditransitiv  
S → NP\_3\_pl\_nom NP\_dat NP\_acc V\_3\_pl\_ditransitiv

NP |  $3 \times 2 \times 4 = 24$  neue Kategorien

V | Für  $n$  Valenzmuster  $3 \times 2 \times n$  Kategorien

Übereinstimmung in **drei Genera**, **zwei Numeri** und **vier Kasus**!

(46) der Mann, die Frau, das Buch (Genus)

(47) das Buch, die Bücher (Numerus)

(48) des Buches, dem Buch (Kasus)

NP\_3\_sg\_nom → Det\_fem\_sg\_nom N\_fem\_sg\_nom

NP\_3\_sg\_nom → Det\_mas\_sg\_nom N\_mas\_sg\_nom

NP\_3\_sg\_nom → Det\_neu\_sg\_nom N\_neu\_sg\_nom

NP\_3\_pl\_nom → Det\_fem\_pl\_nom N\_fem\_pl\_nom

NP\_3\_pl\_nom → Det\_mas\_pl\_nom N\_mas\_pl\_nom

NP\_3\_pl\_nom → Det\_neu\_pl\_nom N\_neu\_pl\_nom

... → Dativ

NP\_gen → Det\_fem\_sg\_gen N\_fem\_sg\_gen

NP\_gen → Det\_mas\_sg\_gen N\_mas\_sg\_gen

NP\_gen → Det\_neu\_sg\_gen N\_neu\_sg\_gen

NP\_gen → Det\_fem\_pl\_gen N\_fem\_pl\_gen

NP\_gen → Det\_mas\_pl\_gen N\_mas\_pl\_gen

NP\_gen → Det\_neu\_pl\_gen N\_neu\_pl\_gen

... → Akkusativ

**Je 24 Symbole** für Determinatoren und Substantive, **24 Regeln**



# Das Problem sind nicht die vielen Regeln!

Syntaktische **Generalisierungen werden nicht erfasst.**

- Beispiel Generalisierung | **Wo kann eine NP oder NP\_nom stehen?**
- Nicht: **Wo kann eine NP\_3\_sg\_nom stehen?**

Lösung | **Komplexe Kategorien** mit Merkmalen, Werten und Identität von Werten

$NP(3,sg,nom) \rightarrow Det(fem,sg,nom) N(fem,sg,nom)$

$NP(3,sg,nom) \rightarrow Det(mask,sg,nom) N(mask,sg,nom)$

Regelschemata mit variablen Werten und ggf. festen Werten

$NP(3, Num, Kas) \rightarrow Det(Gen, Num, Kas) N(Gen, Num, Kas)$

- Genus | Festgelegt durch Regel (NP mit Appellativum)
- Numerus und Kasus | Müssen übereinstimmen, sind an Projektion sichtbar
- Genus | Muss übereinstimmen, an Projektion sichtbar
- Wohlgeformte und nicht wohlgeformte NP nach dieser Regel:
  - ▶ des Baums  
 $NP(3, sg, gen) \rightarrow Det(mask, sg, gen) N(mask, sg, gen)$
  - ▶ des Bäumen  
 $NP(3, ?, ?) \rightarrow Det(mask, sg, gen) N(mask, pl, dat)$

## Grammatik mit Kongruenz und rudimentärer Valenz

$NP(3, Num, Kas) \rightarrow Det(Gen, Num, Kas) N(Gen, Num, Kas)$

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

$S \rightarrow NP(Per_1, Num_1, nom) NP(Per_2, Num_2, akk) V_{tr}(Per_1, Num_1)$

$S \rightarrow NP(Per_1, Num_1, nom) NP(Per_2, Num_2, dat) NP(Per_3, Num_3, akk) V_{dtr}(Per_1, Num_1)$

- Kongruenzmerkmale
- Valenz noch in der Regel und einem Verbsymbol kodiert

Merkmalsmengen in den obigen Regeln müssen geordnet sein!

- N(mask, sg, nom) | Werte in bestimmter Reihenfolge: Genus, Numerus, Kasus
- N(Gen, Num, Kas) | Variablen für Werte in dieser Reihenfolge
- N(Bim, Bam, Bum) | Genau so gute Variablennamen (gleiche Reihenfolge!)
- N(V1, V2, V3) | Indizierte Variablennamen (gleiche Reihenfolge!)
- N( \_, \_ , \_ ) | Irrelevante Werte für Genus, Numerus, Kasus in dieser Reihenfolge

Andere Möglichkeit | Trennung von Merkmal und Wert

- N{Gen=mask, Num=sg, Kas=nom} | Benennung von Merkmal, Wert
- N{Kas=nom, Gen=mask, Num=sg} | Reihenfolge egal

- Chomsky, Noam. 1957. *Syntactic Structures*. (Janua Linguarum / Series Minor 4).
- Dowty, David R. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and Montague's PTQ*. (Synthese Language Library 7). D. Reidel Publishing Company.
- Müller, Stefan. 2013a. *Grammatiktheorie*. 2. Aufl. (Stauffenburg Einführungen 20). Tübingen: Stauffenburg Verlag.
- Müller, Stefan. 2013b. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 3. Aufl. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.
- Müller, Stefan. 2020. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 4. Aufl. (Textbooks in Language Sciences 1). Berlin: Language Science Press.
- Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (Hrsg.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15), 171–211. Tübingen.
- de Saussure, Ferdinand. 1916. *Cours de linguistique générale*. (Bibliothèque Scientifique Payot none). Publié par Charles Bally and Albert Sechehaye. Paris: Payot.
- Schäfer, Roland. 2018. *Einführung in die grammatische Beschreibung des Deutschen*. 3. Aufl. (Textbooks in Language Sciences 2). Berlin.

## 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.