

# Syntax natürlicher Sprachen

## Vorlesung 4: Konstituentenstruktur

A. Wisiorek

Centrum für Informations- und Sprachverarbeitung,  
Ludwig-Maximilians-Universität München

15.11.2022

# 1. Konstituentenstruktur

- 1 Konstituentenstruktur
- 2 Phrasenstrukturregeln des Deutschen
- 3 X-Bar-Phrasenstrukturschema
- 4 Adäquatheit einer CFG als Syntaxmodell
- 5 Konstituenten-Treebanks und -Grammatiken
- 6 *Exkurs: Diskontinuierliche Konstituenten*

## Konstituentenanalyse

auch IC-Analyse (Analyse der *immediate constituents*)

- 1 **Zerlegung** syntaktischer Einheit in ihre Teile (Konstituenten)
- 2 Bildung von **Konstituentenklassen** (lexikalische und syntaktische Kategorien)
  - Ermittlung über **Konstituententests**
  - Ergebnis ist eine **hierarchisch gegliederte Struktur**

## unmittelbare Konstituenten (*immediate constituents*)

**unmittelbare Konstituenten** sind die **maximalen Konstituenten** einer Einheit (aus denen sie unmittelbar zusammengesetzt ist)

# Konstituenz-Relation

## Konstituenz

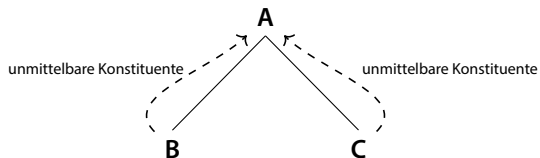
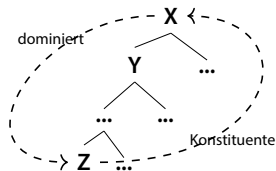
**Teil-Ganzes-Beziehung** zwischen sprachlichen Einheiten (Konstituenten)

## unmittelbare Dominanz

Beziehung der **unmittelbaren Dominanz** zwischen Einheit und ihren unmittelbaren Konstituenten

## Dominanz

Beziehung der **Dominanz** zwischen Einheit X und der unmittelbare Konstituente Y; sowie zwischen X und Z, wenn Y Z dominiert (transitive Relation)



# Konstituentenstruktur

- **Konstituentenstruktur:** Menge der durch die Relation der unmittelbaren Dominanz verbundenen Konstituenten
- durch Bezug auf **Konstituentenklassen** (lexikalische und syntaktische Kategorien als Knoten) und Abstraktion von der Wortebene ergeben sich **Konstituentenschemata**

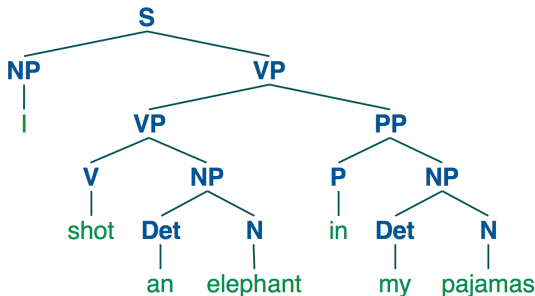


Abbildung: Syntaxbaum als Repräsentation der Konstituentenstruktur

- **Elemente der Struktur (Knoten):**
  - **Wörter** → *terminale Knoten*
  - **lexikalische Kategorien** → *präterminale Knoten*
  - **syntaktische Kategorien** → *nichtterminale Knoten*  
→ *Phrasen (Konstituente mit syntaktisch zusammengehörigen Subkonstituenten)*
- **Relationen der Struktur (Kanten)** → *Teil-Ganzes-Beziehung; unmittelbare Dominanz des Mutterknotens über Tochterknoten*
- **Strukturinformationen in Knoten des Syntaxbaums**

- jede Phrase hat einen **Kopf (auch: Phrasenkern)**, der die **Eigenschaften der Phrase bestimmt**
- alle anderen Wörter und Phrasen in der Phrase sind zum Kopf **dependent**
- Kopf vererbt **morphosyntaktische Merkmale** an Phrase
- Kopf steuert **syntaktisches Verhalten** der Konstituente im Satz
- Kopf bestimmt die **Phrasenkategorie** (Wortart X  $\rightarrow$  Phrasenkat. XP)

- Merkmale der Phrase werden getragen von Kopf
- Köpfe werden im Syntaxbaum nach oben weitergereicht
- **Kopf-Perkolations-Regeln:**
  - $head(NP) = head(N)$
  - $head(VP) = head(V)$
  - $head(S) = head(VP)$
- wichtig u.a. für **lexikalisierte Grammatiken** sowie die **Transformation einer Phrasenstruktur- in eine Dependenzgrammatik**



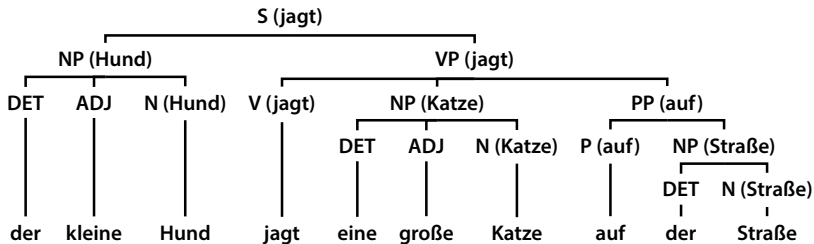


Abbildung: Phrasenstruktur Transitiver Satz des Deutschen mit Kopfregelelementen

## allgemeines Satzschema: $S = NP + VP$

- Ergebnis von Konstituententests (Reduktion auf Zweiwort-Satz)
- **Subjekt-NP und Verb interdependent**, also gegenseitig abhängig (sichtbar am **Verb-Agreement**)
- Subjekt (Satzgegenstand) - Prädikat (Satzaussage)
- abstrahiert von linearer Ordnung: **Wortstellung sprachabhängig**

## $VP = VERB + \text{Komplemente} + \text{Adjunkte}$

- Komplemente = obligatorische (valenzgeforderte) Erweiterungen
- Adjunkte = nicht-obligatorisch Erweiterungen, Anzahl nicht begrenzt

## $NP = NOUN + \text{nominale Adjunkte (Attribute)}$

- Links- und Rechtserweiterungen um Nomen (als NP-Kopf)

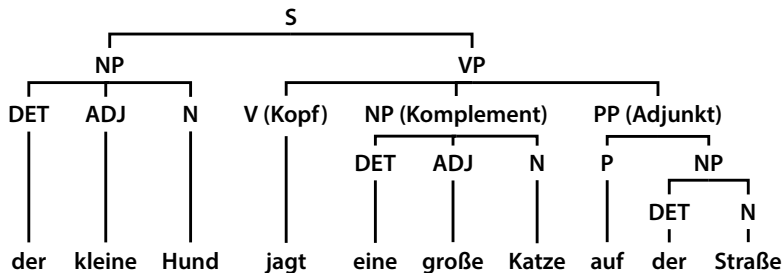


Abbildung: Phrasenstruktur Transitiver Satz des Deutschen

# NP-Struktur des Deutschen

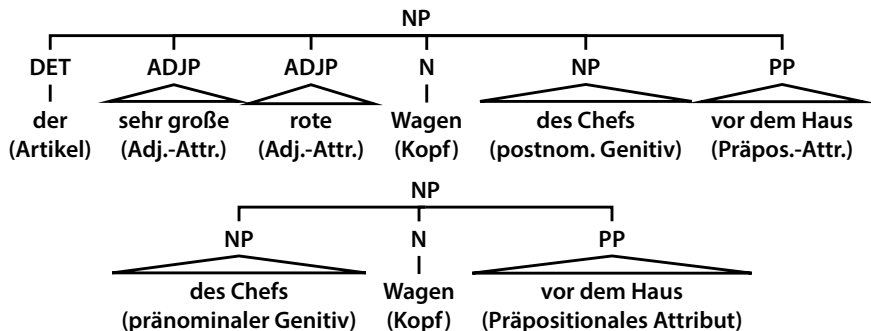


Abbildung: NP-Konstruktionen Deutsch (2: Genitivattribut als Determinativ-Vertreter)

# Weitere NP-Strukturen

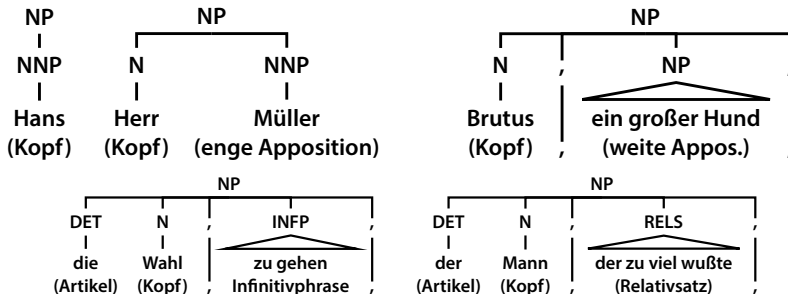
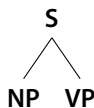


Abbildung: NP-Konstruktionen Deutsch

- **Syntax als Satzstrukturanalyse:**
  - Beschreibung der **syntaktischen Regeln**, die den beobachtbaren **Strukturen** zugrunde liegen (Grammatik)
- **Konstituenten-/Phrasenstruktur** natürlicher Sprache ist formal beschreibbar durch eine **kontextfreie Grammatik (CFG)**
  - *Phrasenstrukturgrammatik im engeren Sinne (PSG)*
  - *Regeln der Verkettung von lexikalischen und phrasalen Kategoriensymbolen (Nicht-Terminale)*
- **syntaktische Regel:** bestimmt, zu welchen **Klassen** die unmittelbaren **Konstituenten einer syntaktischen Kategorie** gehören:  
 $NP \rightarrow DET N PP$
- **lexikalische Regel:** bestimmt die **Zugehörigkeit einer elementaren Konstituente (Wort) zu einer lexikalischen Kategorie:**  $N \rightarrow Hund$

- PSG-Regeln können als **Konstruktionsanweisung für Syntaxbäume** interpretiert werden:

$S \rightarrow NP VP$  als *'expandiere S zu Folge NP + VP'*



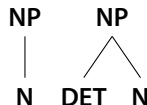
- PSG-Regel definiert **Relation der unmittelbaren Dominanz zwischen Mutterknoten und Tochterknoten**
  - *'S dominiert unmittelbar NP und VP'*
  - *'S dominiert vollständig die Folge NP + VP'*
  - und: *'NP und VP sind Ko-Konstituenten' (Geschwisterknoten)*
- PSG **erkennt durch Ableitung** Sätze als zur Sprache gehörig und weist ihnen die ihren Regeln entsprechende **Strukturbeschreibung** zu
  - *Strukturbeschreibung = die auf Kategorien bezugnehmende Konstituentenstruktur*
  - *'Die Folge NP + VP ist ein S'*

- **Startsymbol:**  $S$
- **Nichtterminalsymbole:**  $NP, VP, DET, N, V$
- **Terminalsymbole:** 'der', 'Hund', 'schläft'
- **Produktionsregeln:**  $S \rightarrow NP VP, \quad NP \rightarrow DET N, \quad VP \rightarrow V$ 
  - Ersetzungsregeln (linke mit rechter Seite)
  - CFG-Regel-Einschränkung (Chomsky-Hierarchie): links nur ein Nichtterminalsymbol  
→ *Ersetzung unabhängig von Kontext (Kontextfreiheit)*



- **syntaktische Regeln:**  $NP \rightarrow DET\ N\ (PP)$ 
  - links: syntaktische Kategorien (Phrasen/Satzknoten)
  - rechts: obligatorische und optionale Nichtterminale (syntaktische + lexikalische Kategorien)
- **lexikalische Regeln:**  $N \rightarrow 'Hund'$ 
  - links: lexikalische Kategorie/Wortart
  - rechts: Wörter (Terminale)
- **rekursive Regeln:**  $NP \rightarrow DET\ N\ NP$

- $NP \rightarrow (DET) N$
- äquivalent zu:  
 $NP \rightarrow DET N \mid N$
- äquivalent zu:  
 $NP \rightarrow DET N$   
 $NP \rightarrow N$



# Ableitung als top-down Erzeugung (Linksableitung)

$G = \langle \{S, NP, VP, DET, N, V\}, \{das, Tier, Ding, sieht\}, R, S \rangle$

$R = \{S \rightarrow NP VP, NP \rightarrow DET N, VP \rightarrow V NP, DET \rightarrow das, N \rightarrow Tier, N \rightarrow Ding, V \rightarrow sieht\}$

$S$	$\Rightarrow$	$NP VP$	$(S \rightarrow NP VP)$
	$\Rightarrow$	$DET N VP$	$(NP \rightarrow DET N)$
	$\Rightarrow$	$das N VP$	$(DET \rightarrow das)$
	$\Rightarrow$	$das Tier VP$	$(N \rightarrow Tier)$
	$\Rightarrow$	$das Tier V NP$	$(VP \rightarrow V NP)$
	$\Rightarrow$	$das Tier sieht NP$	$(V \rightarrow sieht)$
	$\Rightarrow$	$das Tier sieht DET N$	$(NP \rightarrow DET N)$
	$\Rightarrow$	$das Tier sieht das N$	$(DET \rightarrow das)$
	$\Rightarrow$	$das Tier sieht das Ding$	$(N \rightarrow Ding)$

- **Rekursiver Aufbau natürlicher Sprachen**

- erklärt **Nicht-Endlichkeit** natürlicher Sprachen

- endliche Anzahl an syntaktischen Regeln, aber unbegrenzte Anzahl an bildbaren Sätzen (z. B.: durch Hinzufügung von Adjunkten)

- modellierbar mit **rekursiven Phrasenstrukturregeln**

- Erzeugung **verschachtelter Strukturen**

- **einige rekursive Kategorien:**

- nominale Adjunkte (ADJ- oder PP-Attribute, Genitiv-NP-Attribute)
  - verbale Adjunkte (PP-Adverbiale)
  - Relativsätze (center-embedding)
  - Komplementsätze (*glauben, dass ...*)

## Direkte Rekursion

**NP** → DET N NP

## Beispiel direkte Rekursion

*der Schlüssel des Autos der Firma des Chefs ...*

## Indirekte Rekursion

**NP** → DET N PP

**PP** → P NP

## Beispiel indirekte Rekursion

*in der Frage nach den Problemen in seinen Ausführungen über die Freiheit auf dem Gebiete der Kunst während der 30er Jahre*  
(s. Dürscheid, Kap. 5, nach J. E. Schmidt, 1993:84)

# 2. Phrasenstrukturregeln des Deutschen

- 1 Konstituentenstruktur
- 2 Phrasenstrukturregeln des Deutschen**
- 3 X-Bar-Phrasenstrukturschema
- 4 Adäquatheit einer CFG als Syntaxmodell
- 5 Konstituenten-Treebanks und -Grammatiken
- 6 *Exkurs: Diskontinuierliche Konstituenten*

# ADVP-Strukturregeln (Adverbphrase)

## Phrasenschema ADVP

(ADV)

(PRT)

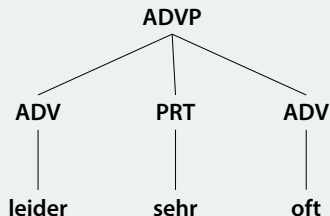
ADV

## ADVP-Produktionsregeln

*ADVP* → *ADV*

*ADVP* → *PRT ADV*

*ADVP* → *ADV PRT ADV*



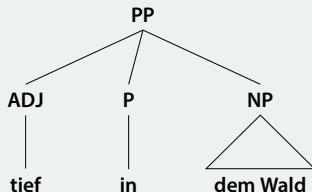
## Phrasenschema PP

(ADJ/ADV/PRT)    P    NP

## PP-Produktionsregeln

$PP \rightarrow P NP$

$PP \rightarrow PRT P NP$





# ADJP-Strukturregeln (nicht-rekursiv)

## Phrasenschema ADJP

(PP)\*

(PRT)

ADJ

## ADJP-Produktionsregeln (flacher Aufbau ohne Rekursion)

*ADJP* → *ADJ*

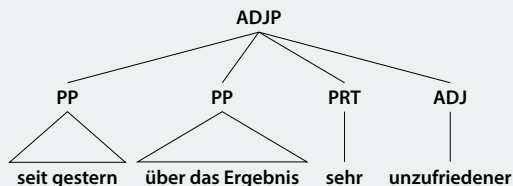
*ADJP* → *PRT ADJ*

*ADJP* → *PP ADJ*

*ADJP* → *PP PRT ADJ*

*ADJP* → *PP PP ADJ*

usw.



# ADJP-Strukturregeln (rekursiv)

## Phrasenschema ADJP

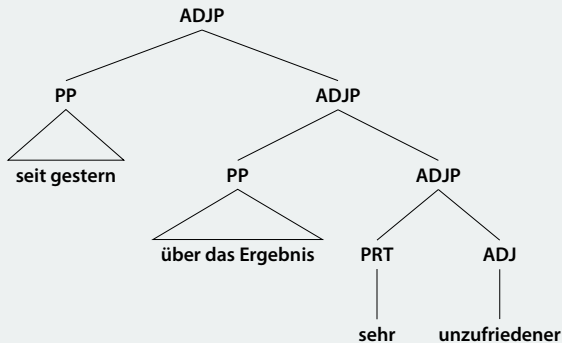
(PP)\*      (PRT)      ADJ

## ADJP-Produktionsregeln (rekursiver Aufbau)

*ADJP* → *PP ADJP*

*ADJP* → *PRT ADJ*

*ADJP* → *ADJ*



## Phrasenschema VP (vereinfacht, ohne AdvP)

V + Komplemente + (Adjunkte)\*

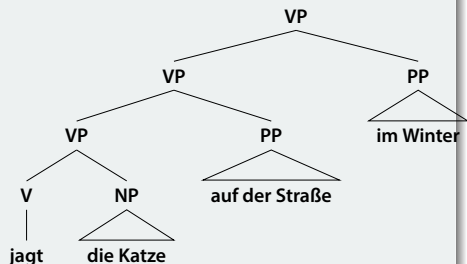
## VP-Produktionsregeln (rekursiv)

$VP \rightarrow VP PP$  (rekursive Adjunkte)

$VP \rightarrow V$  (ohne Kompl. = intransitiv)

$VP \rightarrow V NP$  (1 Kompl. = transitiv)

$VP \rightarrow V NP NP$  (2 Kompl. = ditrans.)



# NP-Strukturregeln (nicht-rekursiv)

## vereinfachtes Phrasenschema NP (ohne Rekursion)

(DET)      (ADJP)      N      (PP)

- (vereinfacht: nur Adjektiv-Phrase als Links- und PP als Rechtsattribut)  
→ (d.h. ohne innere Rechtsattribute oder Genitivattribute)

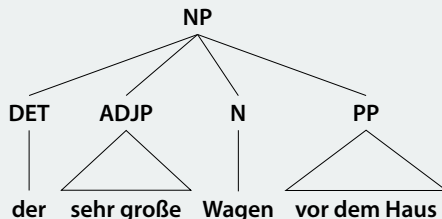
## NP-Produktionsregeln (flacher Aufbau ohne Rekursion)

$NP \rightarrow N \mid DET N$

$NP \rightarrow ADJP N \mid DET ADJP N$

$NP \rightarrow N PP \mid DET N PP$

$NP \rightarrow ADJP N PP \mid DET ADJP N PP$

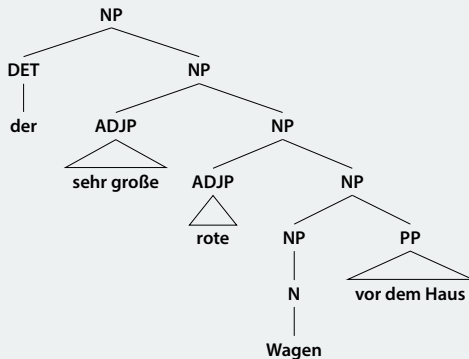


## vereinfachtes Phrasenschema NP mit rekursiven Attributen

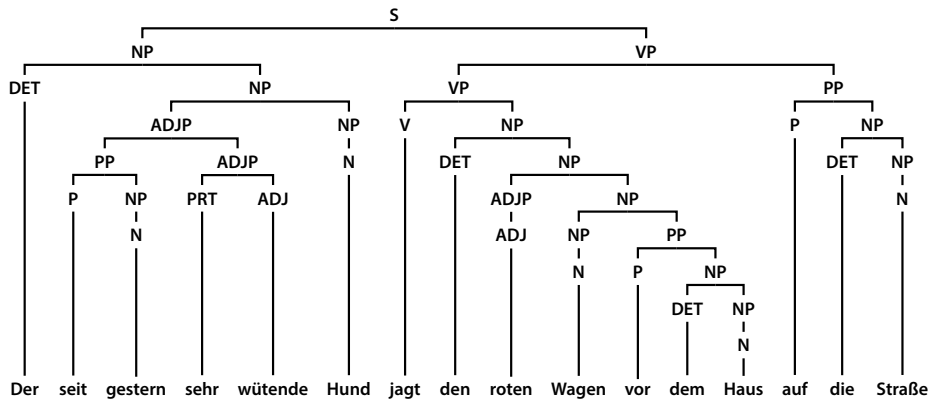
(DET)      (ADJP)\*      N      (PP)\*

## NP-Produktionsregeln mit Rekursion (vorläufig)

$NP \rightarrow N$   
 $NP \rightarrow ADJP\ NP$   
 $NP \rightarrow NP\ PP$   
 $NP \rightarrow DET\ NP\ (?)$



# Satzableitung aus den Grammatikregeln (mit $S \rightarrow NP VP$ )



# NLTK: Übergenerierung vorläufiger NP-Grammatik

```
1 grammar = nltk.CFG.fromstring("""
2 NP -> ADJP NP
3 NP -> NP PP
4 NP -> N
5 NP -> DET NP
6 N -> 'Nomen'
7 DET -> 'das'
8 ADJP -> 'schöne'
9 PP -> 'darin'
10 """)
11
12 for sentence in generate(grammar, depth=9):
13     print(' '.join(sentence))
14
15 #...
16 #das das schöne das schöne Nomen darin
17 #...
```

# Problem rekursiver DET-Regel

## rekursive DET-Regel: $NP \rightarrow DET\ NP$

Mit der Regel können zwar NPs gemäß des Phrasenschemas abgeleitet werden

## Problem: Übergenerierung

- die Grammatik ist aber stark **übergenerierend** und damit kein adäquates Modell des NP-internen Strukturaufbaus
- **Übergenerierung 1:** DET rekursiv wiederholbar an erster Position (richtige Strukturposition aber falsche Anzahl)
- **Übergenerierung 2:** DET wiederholbar an falscher Strukturposition (z. B.: ADJP DET N)

## Lösung: X-Bar-Struktur

- **Strukturbegrenzung notwendig!**
- möglich durch **X-Bar-Theorie** (Chomsky)



## 3. X-Bar-Phrasenstrukturschema

- 1 Konstituentenstruktur
- 2 Phrasenstrukturregeln des Deutschen
- 3 X-Bar-Phrasenstrukturschema**
- 4 Adäquatheit einer CFG als Syntaxmodell
- 5 Konstituenten-Treebanks und -Grammatiken
- 6 *Exkurs: Diskontinuierliche Konstituenten*

- Entwicklung durch Chomsky im Rahmen der *Government & Binding*-Theorie
- ursprüngliches Symbol: Balken (*bar*):  $\bar{X}$

## Beschränkung der Struktur von Phrasen

- 1 **binäre Verzweigung** ( $A \rightarrow B C$ )
- 2 Einführung **phrasaler Zwischenebene** ( $\bar{X}$  oder  $X'$ )
- 3 **gleicher Strukturaufbau für alle Phrasenarten** ( $X$  als Wortart-Variable)

**bisher: Anzahl und Art unmittelbarer Konstituenten einer Phrase nicht beschränkt**

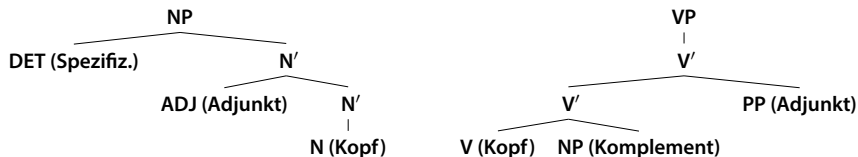
- Mischung aus nicht festgelegter Anzahl aus lexikalischen und phrasalen Kategorien
- keine festgelegte Ordnung zwischen Kopf und Erweiterungen

**Problem: Übergenerierung**

führt zu Problemen bei rekursiver Strukturdefinition (s. oben)

## Lösung: phrasale Zwischenebene: X'

- Einführung **phrasaler Zwischenebene** (X') zwischen **Gesamtphrase** (XP, **maximale Projektion**) und **Kopf** (X)
- Erlaubt die Differenzierung verschiedener Arten von Dependents in Phrase durch Strukturposition

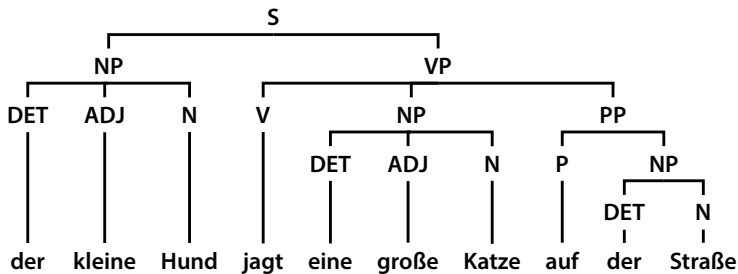


## Spezifizierer

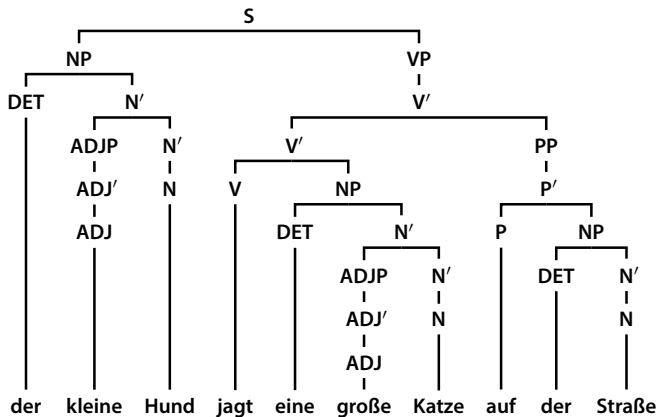
**phrasenabschließende** nicht-obligatorische Erweiterung, Anzahl begrenzt (nur ein Element)

- Verbindung mit  $X'$ , bilden zusammen den Abschluß der XP-Phrase
- Deutsche NP: Determinierer, Possessivpronomen, Quantifizierer, sächlicher Genitiv (*seines Glückes Schmied*)
- verallgemeinert: als Strukturposition ( $XP \rightarrow (SPEC) X'$ ) im Schema für Elemente mit entsprechenden Eigenschaften
- z. B. AUX als VP-Spezifizierer (s. unten)

## zum Vergleich: nicht-beschränkte Phrasenstrukturanalyse



# Vollständige X-Bar-Analyse



# Verallgemeinerung verbaler Komplement-Adjunkt-Distinktion

## Komplement

**obligatorische** (valenzgeforderte) **Erweiterung**, Anzahl abhängig von Valenz/Subkategorisierungsrahmen

- *direkte Verbindung mit Phrasenkopf X, bilden zusammen X'-Phrase*
- *Modellierung der engen Verbindung Komplement mit Kopf*
- *Deutsche NP: Genitiv-NP-Erweiterung oder von-PP*

## Adjunkt

**nicht-obligatorische Erweiterung**, Anzahl nicht begrenzt

- *Verbindung mit X'-Konstituente, bilden zusammen wieder X'*
- *Deutsche NP: Adjektiv-Attribut, PP-Attribut*



# Allgemeines X-Bar-Schema

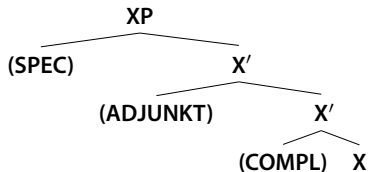


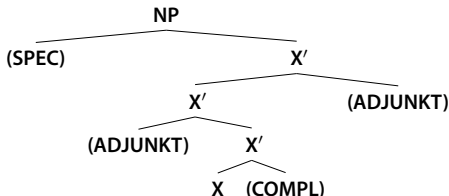
Abbildung: Allgemeines X-Bar-Schema

$XP \rightarrow (SPEC), X'$

$X' \rightarrow (ADJUNKT), X'$  (rekursive Regel)

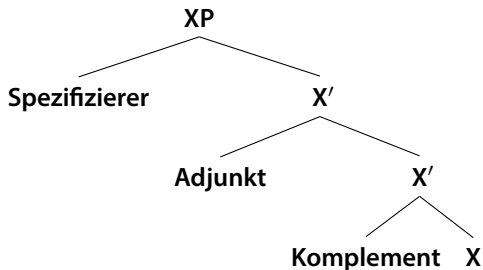
$X' \rightarrow (COMPL), X$

- X-Bar-Schema: **ohne implizierte lineare Struktur** (Wortstellung)  
→ *Einzelsprachliche Regeln*
- **Links- vs. Rechtsverzweigung** (*left- vs rightbranching*)  
→ *Linksverzweigung: head-final* (OV-Sprachen)  
→ *Rechtsverzweigung: head-initial* (VO-Sprachen)  
→ **Deutsche NP: links- und rechtsverzweigend**



# Strukturpositionen im X-Bar-PSG-Schema

	<b>X (Kopf)</b>	<b>X'</b>	<b>XP</b>
<b>Komplement</b>	Schwester	Tochter	
<b>Adjunkt</b>		Schwester und Tochter	
<b>Spezifizierer</b>		Schwester	Tochter



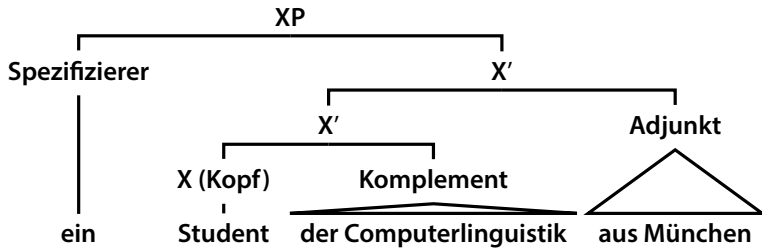


Abbildung: allgemeine X-Bar-Analyse am Beispiel NP

# X-Bar-NP-Regeln

## Spezifizierer-Regel (Phrasenabschluss)

$NP \rightarrow DET N'$

## Adjektiv-Adjunkt-Regel (rechtsrekursiv)

$N' \rightarrow ADJP N'$

## PP-Adjunkt-Regel (linksrekursiv)

$N' \rightarrow N' PP$

## Genitiv-NP-Komplement-Regel (Ko-Konstituente/Schwester Kopf)

$N' \rightarrow N NP$

## Kopf-Regel

$N' \rightarrow N$

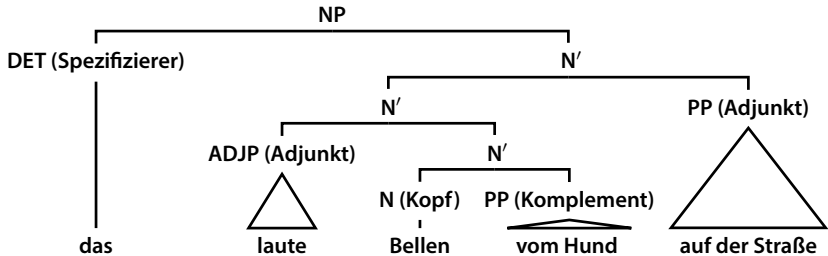


Abbildung: X-Bar-Analyse NP mit Komplement-Adjunkt-Spezifizierer-Distinktion

## Phrasenabschluss

$$VP \rightarrow V'$$

## Adjunkt-Regel

$$V' \rightarrow V' PP$$

## Komplement-Regel (transitives Verb)

$$V' \rightarrow V NP$$

## Kopf-Regel (intransitives Verb: kein Komplement)

$$V' \rightarrow V$$

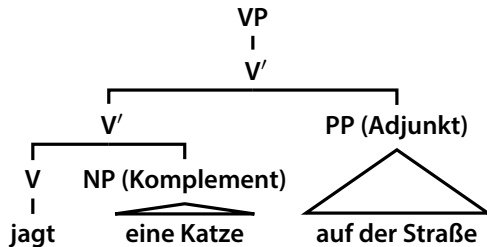


Abbildung: X-Bar-Analyse VP mit Komplement-Adjunkt-Distinktion



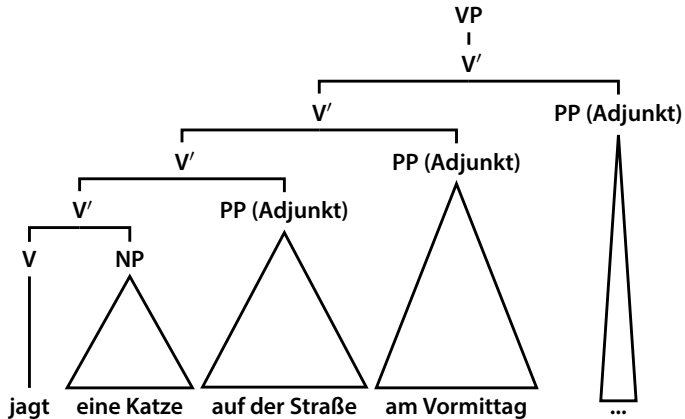


Abbildung: X-Bar-Analyse VP mit rekursiver Adjunktionsregel

# Parallele X-Bar-Strukturanalyse für NP und VP: Auxiliar als Spezifizierer

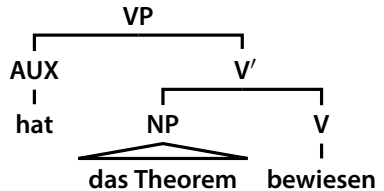
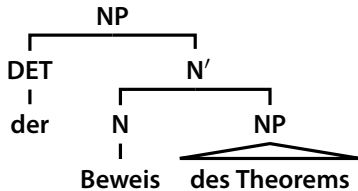


Abbildung: Parallele X-Bar-Strukturanalyse für NP und VP (Spezifizierer und Komplement)

- **Auxiliare = Hilfsverben**

- *begleiten Verb (Träger lexikalischer Bedeutung)*

- Ausdruck von **grammatischen Merkmalen** (Tempus, Modus; Diathese; Flexionsmerkmale)

- Analyse ist stark **theorieabhängig** (Strukturposition)

- u. a. als Spezifizierer (s. o.)

- ebenso: Analyse als Verbgruppe (Verb + Auxiliare)

- ebenso: Eintrag in Subkategorisierungslexikon

## 4. Adäquatheit einer CFG als Syntaxmodell

- 1 Konstituentenstruktur
- 2 Phrasenstrukturregeln des Deutschen
- 3 X-Bar-Phrasenstrukturschema
- 4 Adäquatheit einer CFG als Syntaxmodell**
- 5 Konstituenten-Treebanks und -Grammatiken
- 6 *Exkurs: Diskontinuierliche Konstituenten*

## X-Bar-Phrasenstrukturgrammatiken

- schränken zwar durch ihre Strukturrestriktionen die Übergenerierung ein

## Problem

- bestimmte Formen der Übergenerierung bleiben aber weiter bestehen

## Nichtberücksichtigung von Morphologie

- Rektion (Kasus):

*\*der Mann sieht des Kindes*

- Kongruenz (Agreement in Merkmalen):

*\*der Mann sieht das Kinder*

## Nichtberücksichtigung von Subkategorisierung

- Art und/oder Anzahl von Komplementen

*\*der Hund starb die Katze*

## Splitting atomarer Kategorien

- z.B. Numerus-Kongruenz NP:
- NP gesplittet in SgNP und PlNP:
  - SgN, PlN, SgDET, PlDET, SgNP, PlNP
  - SgNP  $\rightarrow$  SgDET SgN, PlNP  $\rightarrow$  PlDET PlN
- Problem: Regelvervielfachung

## Merkmale in Lexikon

- Merkmalsstrukturen und Unifikationsconstraints erlauben Feststellung Merkmalskongruenz

## Auswahl durch probabilistisches Modell (PCFG)

- Übergenerierung erlauben
- ungrammatische Sätze als unwahrscheinliche ausschließen

## 5. Konstituenten-Treebanks und -Grammatiken

- 1 Konstituentenstruktur
- 2 Phrasenstrukturregeln des Deutschen
- 3 X-Bar-Phrasenstrukturschema
- 4 Adäquatheit einer CFG als Syntaxmodell
- 5 Konstituenten-Treebanks und -Grammatiken**
- 6 *Exkurs: Diskontinuierliche Konstituenten*



## Treebank

Sammlung von per Hand annotierten Syntaxbäumen in bestimmtem Annotationsschema

## *grammar induction* zur Erstellung von Syntaxmodellen

- Treebanks als **implizite Grammatik**
- CFG-Regeln können aus Treebank-Korpus gewonnen werden

- englisches Zeitungskorpus (bekannteste Treebank)
- relativ flache Struktur (z.B. NP → DT JJ NN)
- arbeitet mit *traces* um *long-distance-dependencies* zu markieren  
→ siehe unten: *Diskontinuierliche Konstituenten*
- Schema verwendet im englischen Stanford-Parser-Modell

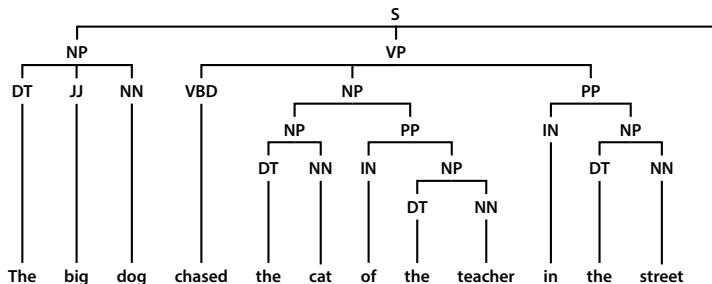


Abbildung: Beispiel-Parse Stanford-Parser en. (trainiert auf Penn-Treebank)

- wichtigste deutsche Treebank
- noch flacher als Penn-Treebank (z.B PP → APPR ART NN)
- Grundlage deutsches Stanford-Parser-Modell

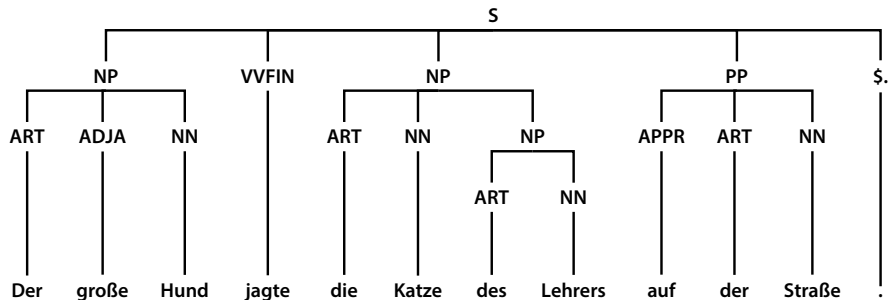


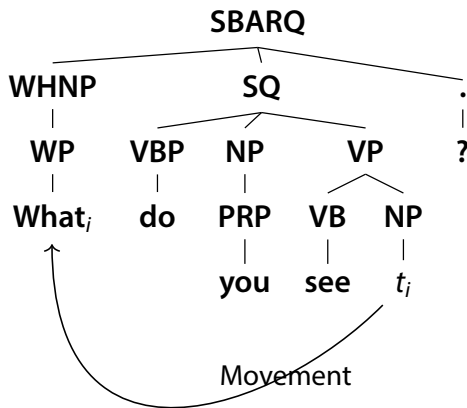
Abbildung: Beispiel-Parse Stanford-Parser dt. (trainiert auf TIGER-Korpus)

## 6. Exkurs: Diskontinuierliche Konstituenten

- 1 Konstituentenstruktur
- 2 Phrasenstrukturregeln des Deutschen
- 3 X-Bar-Phrasenstrukturschema
- 4 Adäquatheit einer CFG als Syntaxmodell
- 5 Konstituenten-Treebanks und -Grammatiken
- 6 *Exkurs: Diskontinuierliche Konstituenten*

- auch: *long distance dependencies*
- Heraustrennung von Teilkonstituenten einer Konstituente
- Problem für Baumdarstellung:  
→ *Überkreuzung* = nicht-projektiv
- Lösung: **leere Knoten** (*empty nodes*: 0,  $\epsilon$ ,  $\tau$ , NONE)  
→ *trace (Spur)*: Konzept der Transformationsgrammatik
- **Transformationsgrammatik**:  
→ Annahme: **Tiefen- und Oberflächenstruktur**  
→ *abstrakte vs. beobachtbare Form von Sätzen*  
→ z. B.: Annahme deutsche Tiefenstruktur der VP: OV (den Hund sehen)  
→ **Transformationsregelanwendung** zur Erzeugung der Oberflächenstruktur: *läßt Spur zurück*
- im Englischen relativ begrenzt: z.B. Topikalisierung, Extraposition, **Wh-fronting**

## Analyse *long distance dependencies* mit Spur (t)



# Rückblick auf heutige Themen

- 1 Konstituentenstruktur
- 2 Phrasenstrukturregeln des Deutschen
- 3 X-Bar-Phrasenstrukturschema
- 4 Adäquatheit einer CFG als Syntaxmodell
- 5 Konstituenten-Treebanks und -Grammatiken
- 6 *Exkurs: Diskontinuierliche Konstituenten*