

# Einführung in die Computerlinguistik

## Syntax II

Hinrich Schütze & Robert Zangenfeind

Centrum für Informations- und Sprachverarbeitung, LMU München

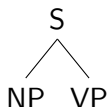
18.12.2015

# Take-away

- Phrasenstrukturgrammatik (PSG)
- Phrasenstrukturregeln
- Rekursion
- Lexikonregeln

# Phrasenstrukturgrammatik (PSG)

- Anwendung der Konstituentenstruktur-Syntax
- generative Grammatik (N. Chomsky)
- Beziehungen der unmittelbaren Dominanz als Regeln zur Erzeugung von Sätzen (Konstituentenstrukturen) , z.B.:



- Regel (rewriting rule, Phrasenstrukturregel) hierzu:
- $S \rightarrow NP \quad VP$
- zu lesen als: “expandiere S zu ... / ersetze S durch ... / S dominiert unmittelbar ... / schreibe S um zu ...”
- S: Startsymbol (“Satz”)

# Schema für Phrasenstrukturregeln

Allgemeines Schema einer PS-Regel:

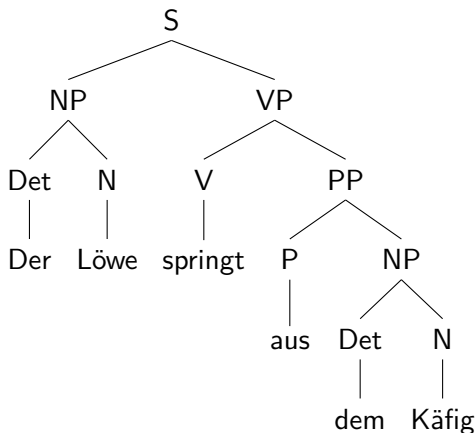
$$X \rightarrow (W) \quad Y \quad (Z)$$

(W, X, Y, Z: Kategoriale Symbole)

- X: Eingabesymbol (genau ein solches steht links des Pfeils)
- W, Y, Z: Ausgabesymbole
- Y: obligatorisches Element
- W, Z: fakultative Elemente

PS-Regeln sind kontextfrei, d.h. genau ein Nichtterminalsymbol (links) wird expandiert zu beliebiger Folge von Nichtterminal- und Terminalsymbolen

## Beispiel für Phrasenstrukturregeln (1)



$S \rightarrow NP \quad VP;$        $NP \rightarrow Det \quad N$   
 $VP \rightarrow V \quad PP;$        $PP \rightarrow P \quad NP$

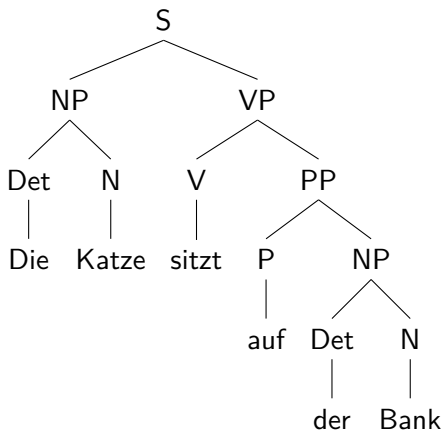
## Beispiel für Phrasenstrukturregeln (2)

$$\begin{array}{ll} S \rightarrow NP \quad VP; & NP \rightarrow \text{Det} \quad N \\ VP \rightarrow V \quad PP; & PP \rightarrow P \quad NP \end{array}$$

Mit diesen Regeln kann obige Baumstruktur bis zur Ebene der Wortkategorien (präterminale Kette) generiert werden.

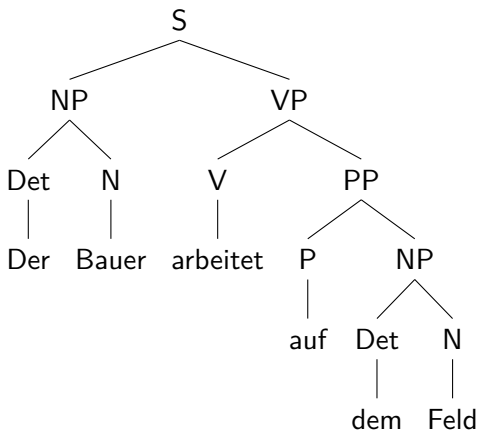
Außerdem sind z.B. möglich:

## Beispiel für Phrasenstrukturregeln (3)



$S \rightarrow NP \quad VP;$        $NP \rightarrow Det \quad N$   
 $VP \rightarrow V \quad PP;$        $PP \rightarrow P \quad NP$

## Beispiel für Phrasenstrukturregeln (4)



$S \rightarrow NP \quad VP;$        $NP \rightarrow Det \quad N$   
 $VP \rightarrow V \quad PP;$        $PP \rightarrow P \quad NP$



# Rekursion (1)

Rekursion ist möglich, d.h. phrasale Gruppen können Konstituenten derselben phrasalen Kategorie enthalten:

(i) Ausgabe einer Regel dient als Eingabe der gleichen Regel:

z.B.  $NP \rightarrow Det\ N\ NP$

→ komplexe NP:

z.B. *Die Farbe der Augen der Tochter des Pianisten ...*

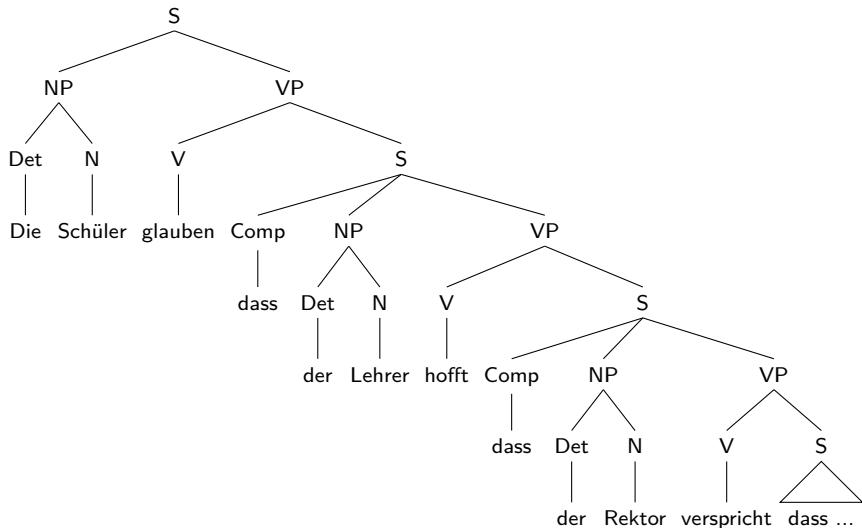
(ii) Ausgabe einer Regel dient als Eingabe einer anderen Regel:

$VP \rightarrow V\ S$

$S \rightarrow Comp\ NP\ VP$

z.B. *Die Schüler glauben, dass der Lehrer hofft, dass der Rektor verspricht, dass ...*

## Rekursion (2)



$S \rightarrow NP VP$ ;  $VP \rightarrow V S$ ;  $S \rightarrow Comp NP VP$ ; ...

# Lexikonregeln (1)

Zur Generierung der terminalen Kette (z.B. *Der Löwe springt aus dem Käfig*) sind noch weitere Regeln nötig:

Det → der

N → Löwe

V → springt

P → aus

Det → dem

N → Käfig

- zu lesen als: *Käfig* ist ein Exemplar der Menge der Nomen
- Regeln entsprechen Lexikon
- mittels Lexikonregeln werden Wörter in die durch PS-Regeln erzeugten Strukturbäume eingesetzt (lexikalische Einsetzung)

## Lexikonregeln (2)

mehr Information nötig, weil:

z.B. *\*Der Löwe schläft aus dem Käfig.*

→ Lexikoneintrag muss Informationen über den zugelassenen syntaktischen Rahmen des Verbs enthalten → Subkategorisierung (vgl. Rektionsmodell):

z.B.:

- a. schlafen V, [ \_\_\_ ]
- b. helfen V, [ \_\_\_ NP]
- c. schenken V, [ \_\_\_ NP, NP]

## Zusammenfassung: Zwei Typen von Kategorien (1)

Wortkategorien (Wortarten):

- z.B. N, V, Det
- terminale Kategorien
- dargestellt in präterminaler Kette
- bezieht sich auf terminale Kette: Wortformen (Elemente, die auf Satzebene nicht weiter zerteilt werden)
- setzt Klassifizierung der Wörter voraus (Klassifikationsbaum, Wörterbuch)

## Zusammenfassung: Zwei Typen von Kategorien (2)

Phrasenkategorien:

- Konstituenten, die mehr als ein Wort enthalten: Phrasen (grammatische Klassen)
- Phrase besteht aus Kopf (obligatorischer Kern) und optional aus weiteren (dominierten) Wörtern
- Klassifizierung nach der Wortart, zu der der Kopf einer Phrase gehört
- z.B. *großes Haus*: grammatikalische Eigenschaften wie *Haus*  
→ *Haus* ist Kopf

→ Kategoriensymbole ersetzen die Knoten im Baum →  
Konstituentenstruktur

# Take-away

- Phrasenstrukturgrammatik (PSG)
- Phrasenstrukturregeln
- Rekursion
- Lexikonregeln