



Computerlinguistische Grundlagen

Jürgen Hermes

Wintersemester 19/20

Sprachliche Informationsverarbeitung

Institut für Digital Humanities

Universität zu Köln



Syntax

- **Grammatik – Satz – Phrase**
- **Konstituenten – Ermittlung und Klassifikation**
- **Phrasenstrukturregeln**
- **Referat: Dependenzparser**
- **Das X-bar-Schema**
- **Kanonische Satzstruktur**



Syntax: Grundbegriffe

- Das Prinzip der doppelten Artikulation
 - Inventar sprachlicher Einheiten (phonologische Ebene)
 - Verknüpfungsregeln (syntaktische Ebene)
- Satzhaftigkeit: Vollständigkeit + Grammatikalität
- Grammatikalitätsurteile
- Strukturabhängige syntaktische Regularitäten
- Konstituenten (Wortarten – Phrasen)



Ermittlung von Konstituenten

- Ersetzungsprobe
- Pronominalisierungstest
- Fragetest
- Verschiebeprobe (Permutationstest)
- Koordinationstest
- Tilgungstest

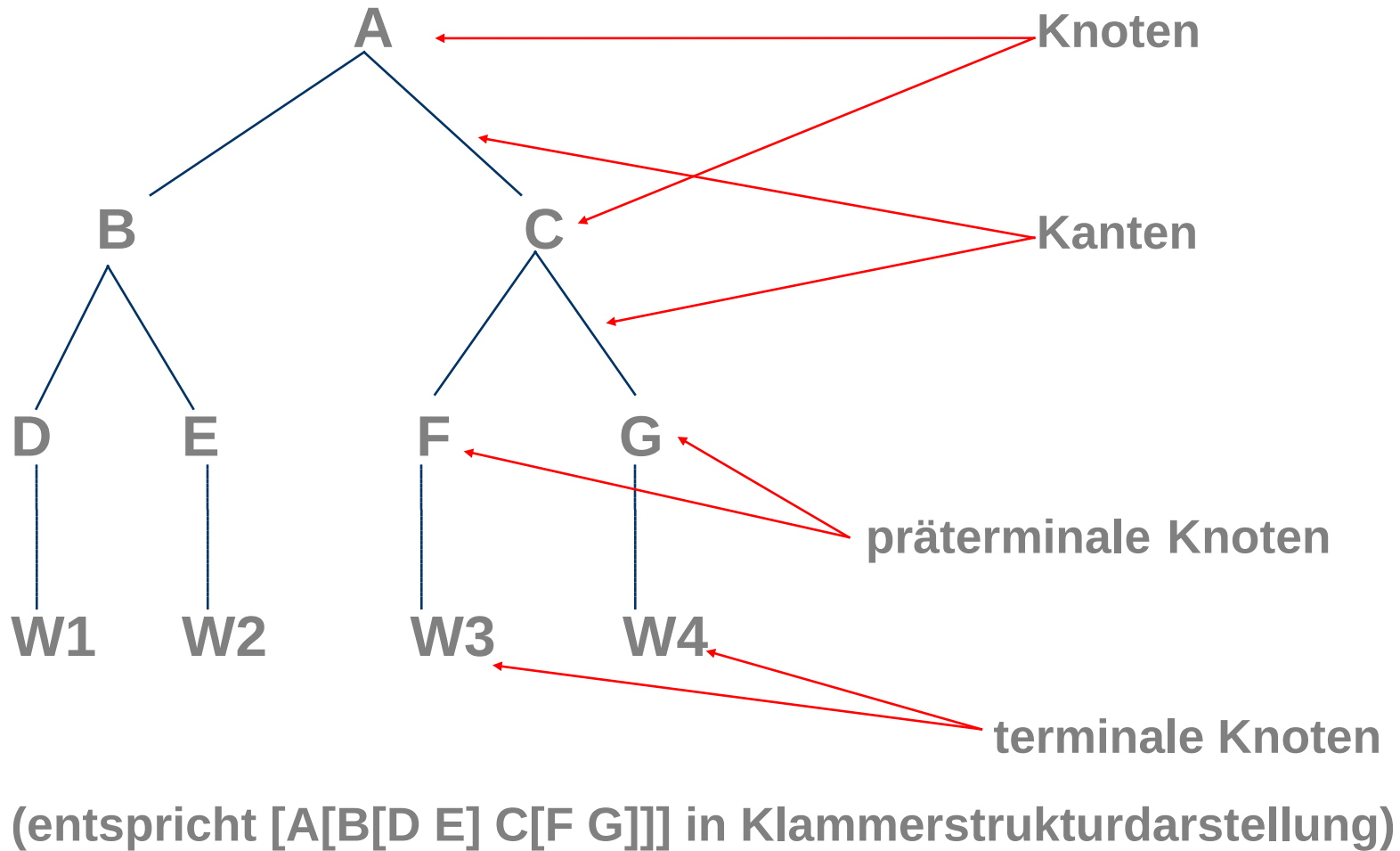


Klassifikation von Konstituenten

- Wortartenklassifikation
 - syntaktische Eigenschaften (Distribution)
 - morphosyntaktische Eigenschaften (Merkmale)
- Klassifikation von Phrasen
 - Nominalphrasen
 - Verbalphrasen
 - Adjektivphrasen
 - Präpositionalphrasen



Phrasenstrukturbäume





Phrasenstrukturregeln

- Syntaktische Regeln:
 - S → NP VP
 - NP → det n
 - VP → v NP
- Lexikalische Regeln:
 - det → die
 - n → katze
 - n → maus
 - v → frisst



Schema von Phrasenstrukturregeln

- **Links** steht genau ein Kategorialesymbol (kontextfreie Grammatik)
- **Rechts** steht mindestens ein kategorielles bzw. lexikalisches Symbol
- **Rekursion** ist möglich: $NP \rightarrow Det N NP$
- **Subkategorisierungsrahmen** für Verben muß eingehalten werden:
 - *Das Kind wohnt
 - *Die Mutter schenkt dem Kind
 - *Der Mann ruft die Frau den Hund



Literatur / Hausaufgabe

Zur Nachbereitung:

Lesen Sie: Ramers (2000): Kapitel 1 (S. 11-32)

Lösen Sie bitte die Aufgaben 1a, 1b, 2a, 2b, 3a und 3b
(Seite 33)

Zur Vorbereitung:

Lesen Sie: Ramers (2000): Kapitel 2 & 3



Probleme von Phrasenstrukturregeln

- Lernbarkeit von teilweise extrem komplexen Regeln ist nicht nachgewiesen / nicht nachweisbar
- Redundante Kodierung durch PS-Regeln und Subkategorisierungsrahmen
- Keine allgemeinen Restriktionen, die von vornherein unmögliche Strukturen (VP → AP NP) verhindern



Zielsetzungen der generativen Grammatik

- **Universalität:** Alle Phrasentypen aller natürlichen Sprachen müssen erfasst werden
- **Lernbarkeit:** Die postulierten Phrasenstrukturen können im Spracherwerb gelernt werden



Das X-bar Modell der Phrasenstruktur

- **Kopfprinzip** (Prinzip der Endozentrität): Alle Phrasen sind Expansionen eines Kerns mit gleicher kategorialer Prägung: $(XP \rightarrow \dots X \dots)$
- **Ebenenprinzip**: Phrasen sind auf mehr als zwei Ebenen gegliedert (X' - Ebene)
- **Phrasenprinzip**: Ergänzungen zu syntaktischen Wörtern oder ihren Projektionen sind immer Phrasen
- **Verzweigungsprinzip**: Knoten verzweigen entweder unär oder binär



Schema des X-bar-Modells

- **Spezifiziererregel:**

$X'' \rightarrow$ Spezifizierer X'

- **Komplementregel:**

$X' \rightarrow$ Komplement X°

- **Adjunktregeln (rekursiv):**

$X' \rightarrow$ Adjunkt X'

$X'' \rightarrow$ Adjunkt X''



Auxiliare, Finitheit und der Nominativ (I)

- Der Frosch sitzt auf der Seerose.
- Der Frosch hat auf der Seerose gegessen.
- Der Storch sieht *der/den Frosch auf der Seerose sitzen.
- Es ist schön, (*der Frosch) auf der Seerose zu sitzen.

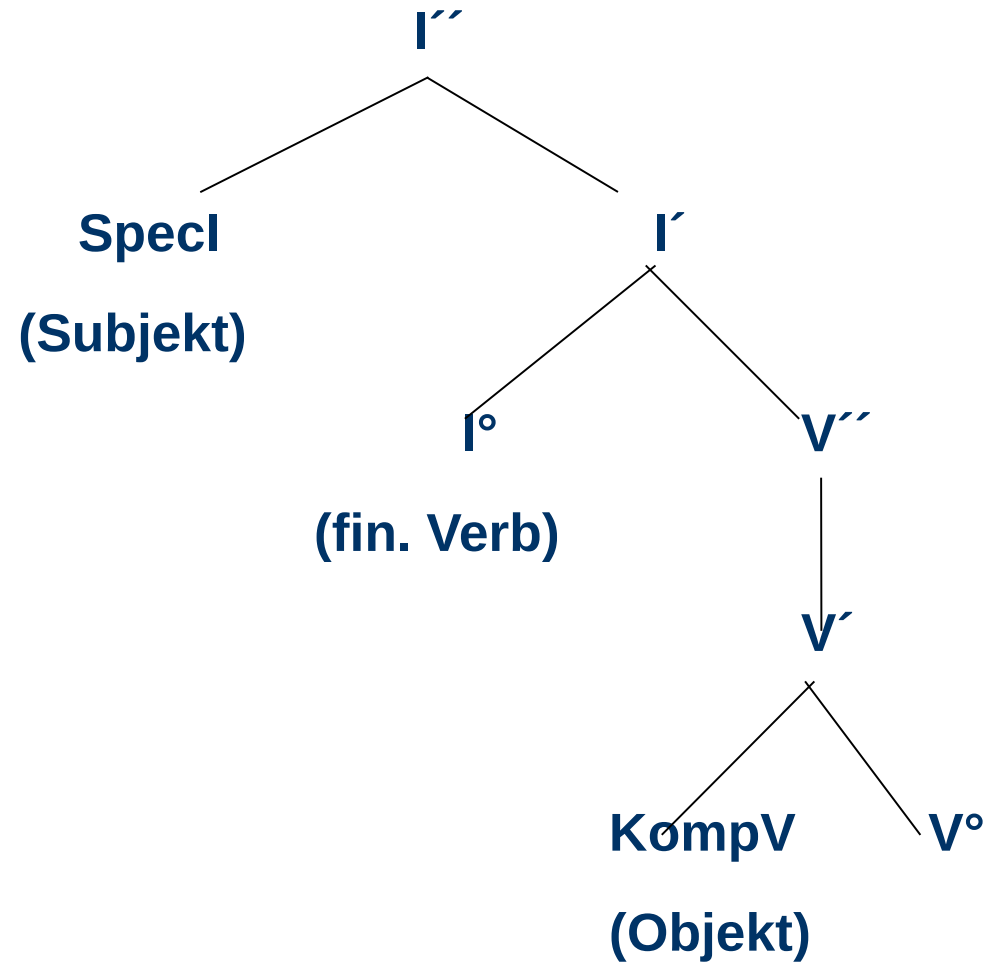


Auxiliare, Finitheit und der Nominativ (II)

- Der Frosch **sitzt** auf der Seerose.
 - Der Frosch **hat** auf der Seerose gegessen.
 - Der Storch **sieht** *der/den Frosch auf der Seerose sitzen.
 - Es **ist** schön, (*der Frosch) auf der Seerose zu sitzen.
- Nominativ tritt nur auf, wenn sich im Satz ein finiter Bestandteil (Auxiliar oder Vollverb) findet.



Inflektionsphrase, die erste





Das topologische Feldermodell

Hauptsätze

Der Frosch **hat** auf der Seerose **gesessen**.

Nebensätze

... **weil** der Frosch auf der Seerose **gesessen hat**.

Entscheidungsfragesätze

Hat der Frosch auf der Seerose **gesessen**?

W-Fragesätze

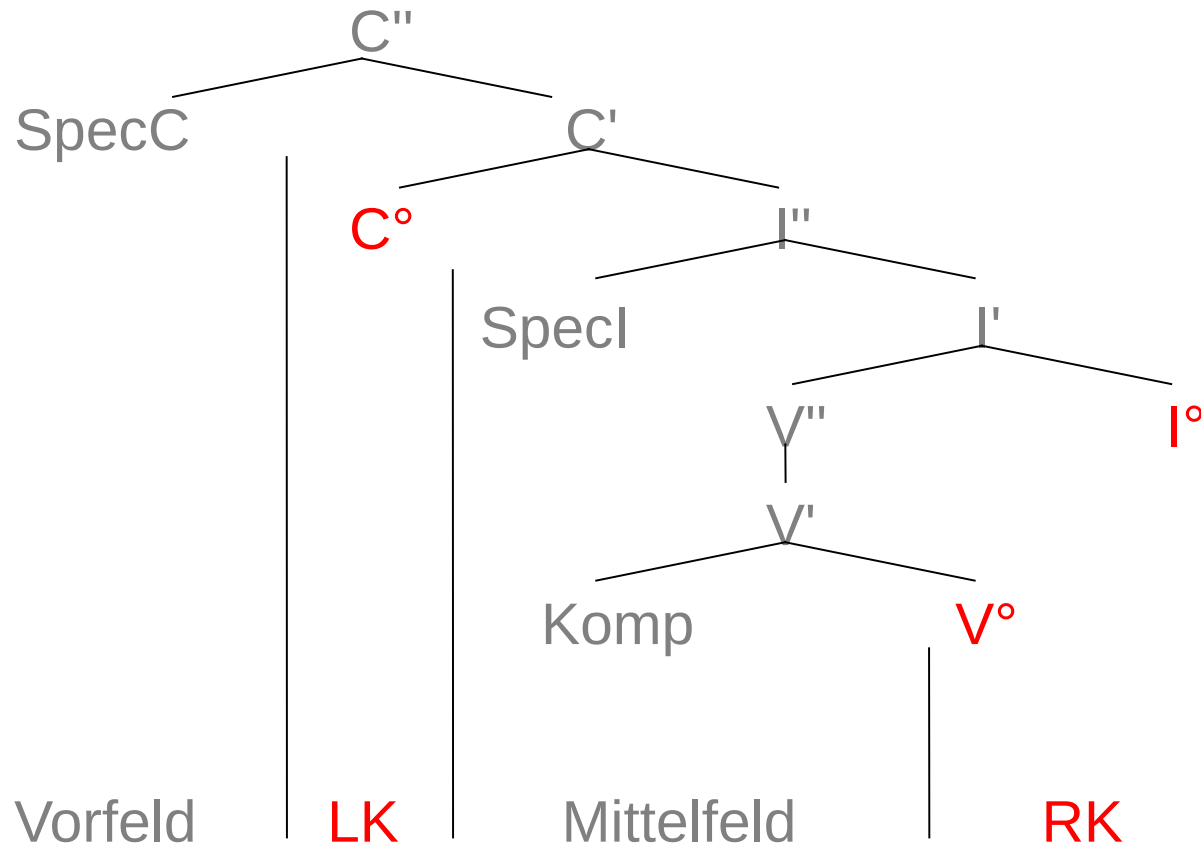
Worauf **hat** der Frosch **gesessen**?

Wer **hat** auf der Seerose **gesessen**?

→ Vorfeld – **linke SK** – Mittelfeld – **rechte SK** – Nachfeld



Das CP-IP-Modell





Literatur / Hausaufgabe

Zur Nachbereitung:

Lesen Sie: Ramers (2000): Kapitel 2 und 3 (S. 35-76)
bearbeiten Sie auch die Übungsaufgaben 2 und 3
jeweils a) bis c) und ß) (S. 55)

Zur Vorbereitung:

Lesen Sie: Schwarz/Chur (1996): Kapitel 1 (S. 13-36)