

Syntax

Haotian Ye Ercong Nie Han-Ching Chen

Centrum für Informations- und Sprachverarbeitung

LMU München

09-02-2021

Outline

- 1 Syntaktische Grundlagen
- 2 Klassische syntaktische Theorien
 - Kategorialgrammatik
 - Dependenzgrammatik und Valenztheorie
 - Generative Transformationsgrammatik
- 3 Moderne syntaktische Theorien

Syntaktische Grundlagen

Syntaktische Grundlagen

- **Syntaktische Kategorisierung**
- **Syntaktische Tests**
- **Syntaktische Funktionen**

Syntaktische Kategorisierungen

- nach Wortarten
- nach Flexionskategorien
- nach Konstituenten-/Phrasenkategorien

Wortarten

- Wortartenanalyse bereits in den Antiken - Wörter zu einzelnen Wortklassen zugeordnet
- **Acht-Wortarten-Lehre** von Dionysios Thrax (1. Jh. v.Chr.)
 - Nomina, Verben, Partizipien, Artikel, Pronomina, Präpositionen, Adverbien, Konjunktionen
- Zahl der Klassen variieren in neueren Ansätzen
 - 4¹ bis mehreren Dutzend²

¹Busse (1997)

²Bergenholtz and Schaefer (1977)

Flexionsarten

■ Deklination

- Substantive
- Adjektive
- Artikel
- Pronomen
- Numerale

■ Konjugation

- Verben

■ Kriterien?

- Genus
- Kasus
- Numerus
- Person
- Tempora
- Modus
- Diathese

Phrasen & Konstituenten

- Phrase → Wortgruppe, die syntaktisch eine Einheit bilden
 - besteht minimal aus dem Kern
 - Kern → wonach die Phrase benannt wird
 - kann beliebig erweitert werden, z.B. mit anderen Phrasen

Phrasen & Konstituenten

■ Phrasen - Beispiele

- Nominalphrase (NP) → die große **Universität**
- Verbalphrase (VP) → **studiert** an der Universität
- Präpositionalphrase (PP) → **an** der Universität
- Adjektivphrase (AP) → sehr **klein**
- Adverbphrase (AdvP) → sehr **oft**

Phrasen & Konstituenten

- Konstituente → sprachliche Einheit, die Teil einer größeren Einheit ist
 - jedes Wort, jede Phrase → können Teil einer Phrase sein
 - möglicherweise Wortgruppen, die keine Phrasen sind → Evaluierung durch Konstituententests

Syntaktische Tests

- Identifizierung von syntaktischen Einheiten (Konstituenten)
- Syntagma → Kette von sprachlichen Zeichen (z.B. Satz)
- Paradigma → austauschbare Segmentierungen
- Bestehen von einem Konstituententest oft nicht genug → Kombination mehrerer Tests erhöht die Zuverlässigkeit

Syntaktische Tests

- Permutationstest → Teil kann zusammen umgestellt werden, ohne dass der Satz ungrammatisch wird
 - er lernt **in der Bibliothek**
 - **in der Bibliothek** lernt er
- Substitutionstest → Teil kann durch andere ausgetauscht werden, ohne dass der Satz ungrammatisch wird
 - er lernt **in der Bibliothek**
 - er lernt **dort**

Syntaktische Tests

- Eliminierungstest → Teil kann zusammen weggelassen werden
 - er lernt **in der Bibliothek**
 - er lernt
- Koordinationstest → Teil kann mit anderen koordiniert werden (z.B. durch Konjunktionen)
 - er lernt **in der Bibliothek** und **zu Hause**

Syntaktische Funktionen

- beschreiben die Verhältnisse der Konstituenten eines Satzes zueinander
- Subjekt
 - (Nom. Akk. Sprachen) Agens in aktiven, Patient in passiven Sätze
 - er hat das Buch gelesen
- Prädikat
 - verbaler Kern des Satzes
 - er **hat** das Buch **gelesen**
- Prädikativum
 - wird hinter einem Kopulaverb gebildet
 - er *ist* **Student**
 - er *wird* **krank**
 - kann auch auf das Objekt beziehen
 - sie *hält* ihn *für einen Student*

Syntaktische Funktionen

■ Objekt

- Ziel des Geschehens, direkt (Patiens) oder indirekt (Rezipient)
- er liest **das Buch**
- er liest **dem Kind das Buch** vor

■ Adverbial

- kann als Adverbien/PPs/NPs/Nebensätze auftreten
- er lernt **viel**
- er lernt **in der Bibliothek**
- er lernt **die ganze Nacht**
- er lernt, **bis er die Prüfung besteht**

Syntaktische Funktionen

■ Freier Dativ

- wenn Dativ kein indirektes Objekt darstellt
- Dativus Iudicantis (Standpunkt) - er lernt **mir** zu viel
- Dativus Ethicus (emotionelle Anteilnahme) - Lerne **mir** nicht zu viel!

■ Attribut

- Beifügung zu Substantive/Adjektive, als APs/PPs/NPs/Nebensatz
- ein **fleißiger** Student
- ein Student **in der Bibliothek**
- ein Student **dieser Universität**
- ein Student, **der die ganze Nacht lernt**

Satzstrukturen des Deutschen

- bezogen auf die Stellung des finiten Verbs
- grammatische Merkmale → Person, Tempus, Numerus

(V/1)

- Lernt er heute?
- Lerne heute!
- Hat er heute gelernt?

(V/2)

- Er lernt in der Bibliothek.
- Wer lernt in der Bibliothek?

(V/E)

- Er lernt, weil er eine Prüfung hat.
- Er liest das Buch, das ihm gefällt.

Kategorialgrammatik

Grundidee

- Einzelnen Wörtern und größeren syntaktischen Einheiten werden 'Kategorien' zugeordnet. Pafel (2011)
(Ausdrücke haben Kategorien.)
- Durch die Kategorien ergibt sich, mit welchen anderen syntaktischen Einheiten Ausdrücke sich zu sinnvollen syntaktischen Einheiten, letztlich zu Sätzen, kombinieren lassen.
(Eine Kategorie gibt an, mit welchen anderen Kategorien ein Ausdruck kombiniert werden kann bzw. muss, um einen größeren Ausdruck zu erhalten.)

Arten von Kategorien

■ Grundkategorien

1. *S*: Kategorie eines Satzes
2. *NP*: Nominalphrase oder *N*: Name

■ Funktorkategorien

Gebildet nach einer allgemeinen Regel, z.B. $S \setminus NP$

Funktions-Argument-Unterscheidung

- Rechts vom Schrägstrich → Argument des Funktors
- Links vom Schrägstrich → Resultat der Anwendung des Funktors

Bemerkung: Die Ausrichtung des Schrägstrichs zeigt an, ob der Funktor das Argument rechts (/) oder links (\) *zusichnimmt*.

Dependenzgrammatik und Valenztheorie

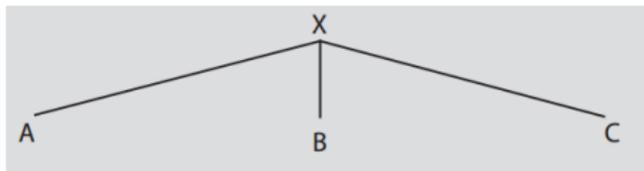
Relation von Dependenz und Valenz

- Miteinander wissenschaftsgeschichtlich **eng** verbunden, aber nicht gleichgesetzt
- Die Dependenzgrammatik modelliert die **strukturelle Organisation** der Grammatik mithilfe von verschiedenen **Abhängigkeitsrelationen**.
- Valenztheorie ist eine **konkrete Realisierung** von Abhängigkeitsbeziehungen für den Aufbau syntaktischer Struktur.
- Tesnière beschreibt seinen Valenzgedanken mittels der Dependenzgrammatik.

Dependenzgrammatik

Konstituentenstruktur vs. Dependenzstruktur

Konstituentenstruktur



- An der Spitze steht **X** als Variable für den ganzen Satz.
- Dieser Satz **X** setzt sich aus den Konstituenten **A**, **B** und **C** zusammen.

Dependenzgrammatik

Konstituentenstruktur vs. Dependenzstruktur

Dependenzstruktur

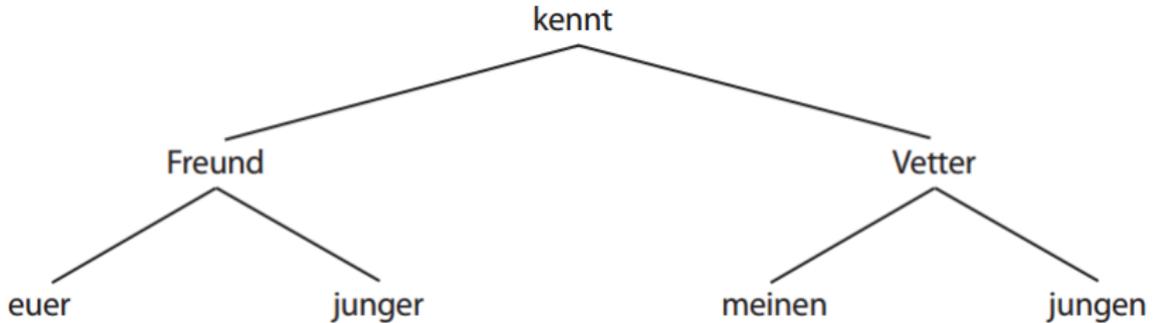


- Die Elemente **B** und **C** hängen von **A** ab.
- **A** ist hier das **Regens**.
- **B** und **C** sind die **Dependentien**.

Dependenzgrammatik und Valenztheorie

Dependenzgrammatik

Beispiel



Mustersatz: *Euer junger Freund kennt meinen jungen Vetter.*

Valenztheorie - Herkunft

Herkunft des Terminus Valenz

Der Terminus **Valenz** wurde ursprünglich in der Chemie gebraucht und von Tesnière in die Sprachwissenschaft eingeführt.³

Bei der Chemie

- Atom
- Bindet eine bestimmte Zahl von anderen **chemischen Elementen** an sich.

Bei der Sprache

- Verb
- Bindet eine bestimmte Zahl von anderen **sprachlichen Elementen**, d.h. **Aktanten** an sich

Beide verfügen über eine bestimmte **Wertigkeit**, d.h. **Valenz**.

³Dürscheid (2012)

Valenztheorie - Grundbegriffe

Valenz

die Verbindungsfähigkeit, die ein Wort hat, und die Realisierung der Fähigkeit in einem Satz.

Valenzgebundene Glieder

- **Ergänzungen (Komplemente)**: die Ausdrücke, die die Wertigkeiten füllen.
- **Angaben**: die Umstände, unter denen sich das Geschehen vollzieht.

Example

In der S-Bahn hat Jens¹ seinen Lehrer² getroffen.

Generative Transformationsgrammatik

Konzeptuelle Grundlagen

- **Sprachkompetenz:** Zentraler Bestandteil unserer kognitiven Fähigkeit (Eigentlicher Gegenstand linguistischer Untersuchungen)
- **Sprachperformanz:** der Gebrauch, den der Sprecher von dieser Fähigkeit macht (sekundär)
- **Universalgrammatik:** Der Mensch ist mit einem speziesspezifischen, genetisch vorprogrammierten Spracherwerbsmechanismus ausgestattet, der es ihm erlaubt, aus den Äußerungen seiner Umgebung beim (kindlichen) Spracherwerb zielsicher die richtigen Regeln der dahinter stehenden Sprache herauszufinden.
- **Prinzipien-und-Parameter-Modell**

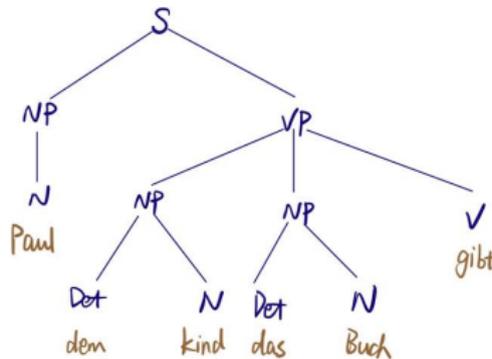
Generative Transformationsgrammatik

Phrasenstruktur

Phrasenstrukturregeln

PS-Regeln:

1. $S \rightarrow NP VP$
2. $NP \rightarrow (Det) N$
3. $VP \rightarrow V (NP) (NP)$



Transformationsgrammatik

- Der **Oberflächenstruktur (OS)** entspricht die konkrete Wort- und Satzgliedfolge.
- **Tiefenstruktur (TS)**: stellt die abstrakte Ebene dar, die der Oberflächenstruktur zugrunde liegt.
- **Transformationen** überführen die Tiefenstruktur in die jeweilige Oberflächenstruktur.

Die Idee der Transformationsgrammatik

Die Struktur, die Sätze an der Oberfläche aufweisen, muss durch Transformationen aus der Tiefenstruktur abgeleitet werden. Die Transformation von Sätzen wird durch **Bewegung** von Wörtern und Wortgruppen realisiert.

Moderne Syntaktische Theorien

Moderne syntaktische Theorien

Syntax ⁴

⁴Carstensen et al. (2009)

Moderne syntaktische Theorien

Syntax ⁴

- **Kontextfreie Grammatiken**

⁴Carstensen et al. (2009)

Moderne syntaktische Theorien

Syntax ⁴

- **Kontextfreie Grammatiken**
- **Unifikationsgrammatiken : $\langle PATR - II \rangle$**

⁴Carstensen et al. (2009)

Kontextfreie Grammatiken

Definition

Definition:

- Eine Grammatik $G = \langle \Phi, \Sigma, R, S \rangle$ heißt kontextfrei (oder auch Typ-2), falls alle Regeln von der Form $A \rightarrow \alpha$ mit $A \in \Phi$ und $\alpha \in \Gamma^* = (\Phi \cup \Sigma)^*$ sind.
- Die Sprache $L = \{a^i b a^i \mid i \geq 0\}$ wird von der kontextfreien Grammatik $G = \langle S, a, b, S \rightarrow b, S \rightarrow aSa \rangle$ erzeugt.
- Kontextfreie Sprachen lassen zwar Klammerstrukturen zu, nicht aber beliebig viele Verkettungen verschiedener Zeichen mit gleicher Anzahl.
e.g. $L = \{a^i b^i a^i \mid i > 0\} \Rightarrow$ nicht kontextfrei
- **Linksableitungen, Rechtsableitungen**

Kontextfreie Grammatik

Komponenten

Komponenten: $G = \langle \Phi, \Sigma, R, S \rangle$

- **Menge von Nichtterminalsymbolen Φ :**
 syntaktische Kategorien, präterminale Kategorien,
 Lexikonregeln
- **Menge von Terminalsymbolen Σ :** enthält sämtliche
 atomaren — zumindest syntaktisch — nicht weiter zerlegbaren
 Einheiten die Wörter der zu beschreibenden Sprachen.
- **Regelmengemenge R ,** die endlich viele Regeln der Form $A \rightarrow \alpha^5$
 enthält.
- **Startsymbol S :** Element des nichtterminalen Vokabulars

⁵ $A \rightarrow$ Nichtterminalsymbol, $\alpha \rightarrow$ eine Kette von Symbolen aus Φ und Σ 

Kontextfreie Grammatik

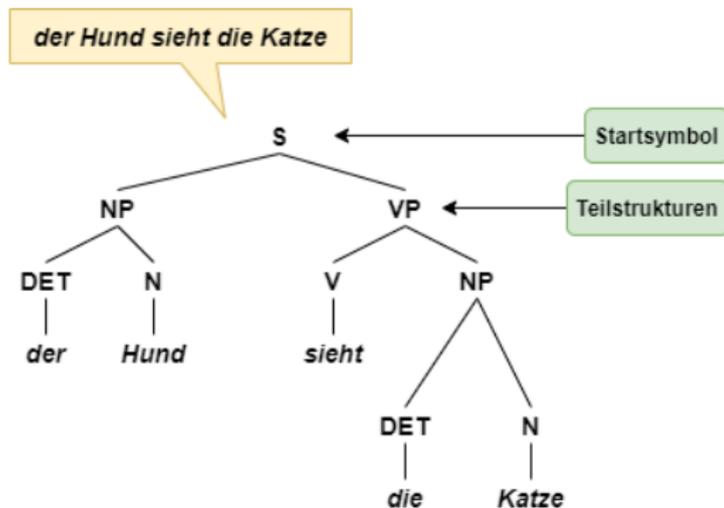
Komponenten

Beispiel:

Eine kontextfreie Grammatik für ein Fragment des Deutschen

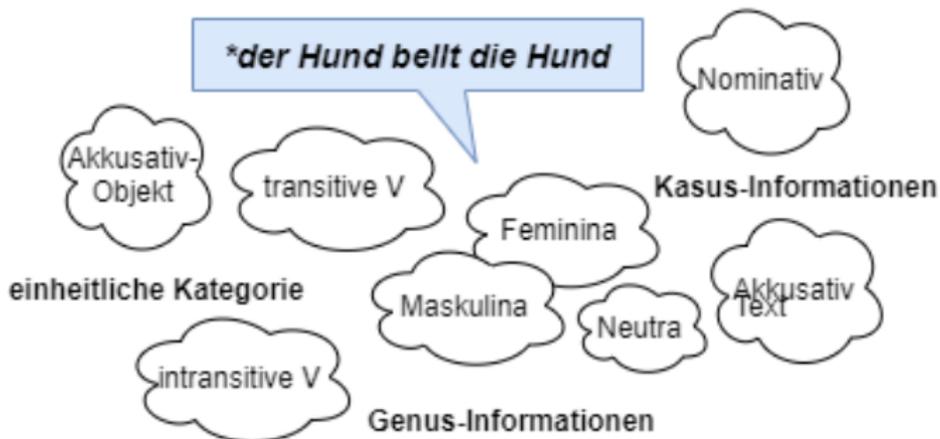
$$\begin{aligned}
 G = \langle \{ & S, NP, VP, DET, N, V \}, \\
 \{ & \text{der, Hund, bellt, sieht, die, Katze} \}, \\
 \{ & \begin{array}{ll}
 S & \rightarrow NP VP, & (1) \\
 NP & \rightarrow DET N, & (2) \\
 VP & \rightarrow V, & (3) \\
 VP & \rightarrow V NP, & (4) \\
 DET & \rightarrow \text{der}, & (5) \\
 DET & \rightarrow \text{die}, & (6) \\
 N & \rightarrow \text{Hund}, & (7) \\
 N & \rightarrow \text{Katze}, & (8) \\
 V & \rightarrow \text{bellt}, & (9) \\
 V & \rightarrow \text{sieht, } & (10) \\
 S & \rangle
 \end{array}
 \end{aligned}$$

Kontextfreie Grammatik



Baumgraph von einem Fragment des Deutschen

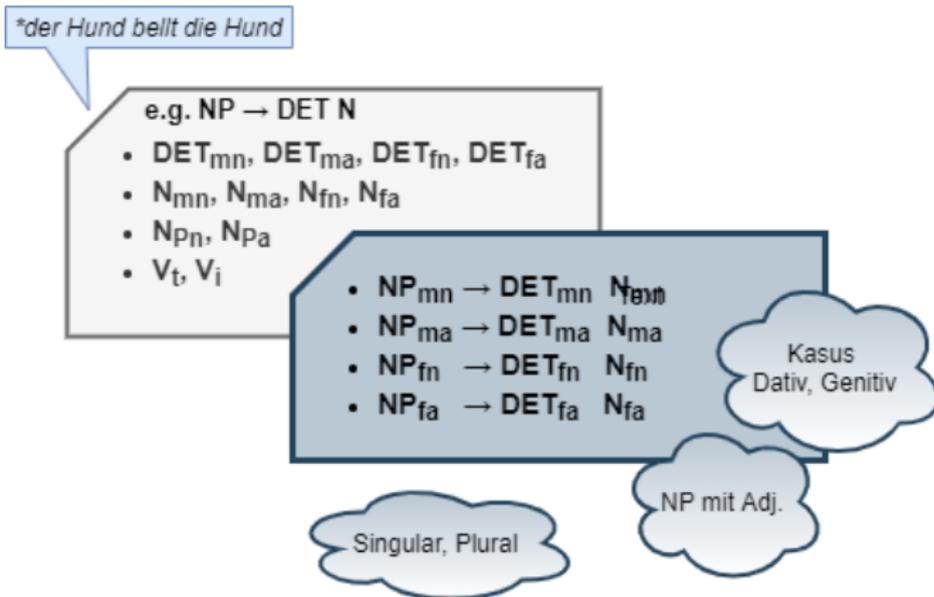
Kontextfreie Grammatik



Ungrammatischer Ausdruck

Kontextfreie Grammatiken

Kategorieninventars



Kontextfreie Grammatiken

Fazit

- Mit kontextfreien Grammatiken lassen sich Grundelemente der syntaktischen Struktur von natürlichsprachlichen Ausdrücken beschreiben.
- Wenn es um die explizite Repräsentation von bestimmten syntaktischen Phänomenen geht, zu denen unter anderem die folgenden zählen:
 - Kongruenz
 - Subkategorisierung
- **Lösungsansätze** für die genannten Probleme wird von **Unifikationsgrammatiken** angeboten.

Unifikationsgrammatik

PATR-II

- ist der einfachste der unifikationsbasierten Grammatikformalismen.
Einfachheit bezieht sich auf die **Anzahl** und **Komplexität** der zugrunde liegenden linguistischen Annahmen.
- **Werkzeugformalismen**⁶ (engl. *tool formalisms*)
 - Kongruenz
 - Subkategorisierung

⁶Stuart Shieber hat Systeme wie PATR-II oder DCG (Definite Clause Grammar) als **Werkzeugformalismen** charakterisiert. Grammatikmodelle, die auf mehr oder weniger expliziten linguistischen Annahmen beruhen (z.B. GPSG, LFG und HPSG) bezeichnet er als **Theorieformalismen** (engl. *theory formalisms*)

Unifikationsgrammatik

Vergleich der Darstellungen von PATR-II und Kontextfreie Grammatik

$X_0 \rightarrow X_1 X_2$
 $\langle X_0 \text{ CAT} \rangle = \text{NP}$
 $\langle X_1 \text{ CAT} \rangle = \text{DET}$
 $\langle X_2 \text{ CAT} \rangle = \text{N}$

PATR-II-Regel
 mit Hilfe der
 Pfadgleichungen



Kontextfreie Grammatik
 mit Hilfe der Grammatik
 ableitbaren Ketten

Unifikationsgrammatik

Lösungsansatz vom PATR-II zur Kontextfreien Grammatik Defizite - Kongruenz

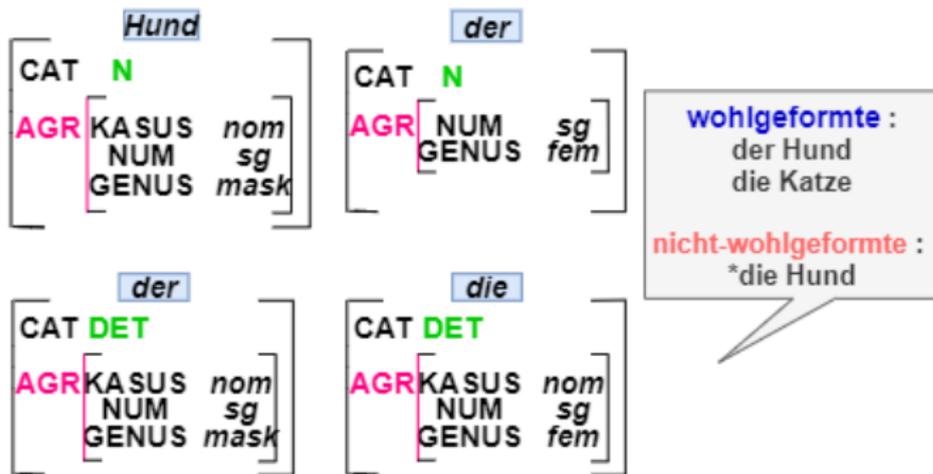
- Beispiel: die Übereinstimmung hinsichtlich Kasus, Numerus und Gernus in der deutschen Nominalphrase

$$\begin{aligned}
 X_0 &\longrightarrow X_1 X_2 \\
 \langle X_0 \text{ CAT} \rangle &= \text{NP} \\
 \langle X_1 \text{ CAT} \rangle &= \text{DET} \\
 \langle X_2 \text{ CAT} \rangle &= \text{N} \\
 \langle X_1 \text{ AGR} \rangle &= \langle X_2 \text{ AGR} \rangle
 \end{aligned}$$

PATR-II-Grammatik Kongruenz

Unifikationsgrammatik

Lösungsansatz vom PATR-II zur Kontextfreien Grammatik Defizite - Kongruenz

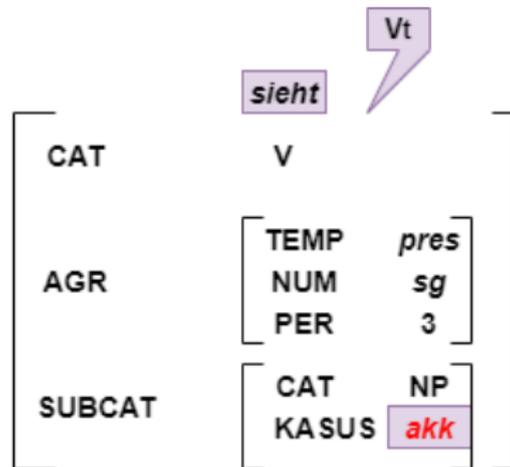
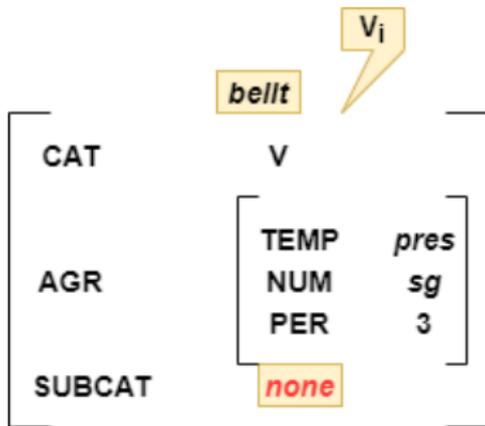


PATR-II-Grammatik Lexikoneinträge von Wörtern aus deklinierbaren Wortarten

Unifikationsgrammatik

Lösungsansatz vom PATR-II zur Kontextfreien Grammatik Defizite - Subkategorisierung

- Beispiel: die Unterscheidung zwischen transitiven und intransitiven Verben in der deutschen Verbalphrasen



Unifikationsgrammatik

Fazit

- Insgesamt sind also die Regeln einer **PATR-II-Grammatik** im Vergleich zu einer kontextfreien genereller und unterspezifizierter, die Lexikoneinträge sind hingegen **detaillierter** und **informativer**.

Literaturverzeichnis I

- Bergenholtz, H. and Schaeder, B. (1977). *Die Wortarten des Deutschen: Versuch e. syntakt. orientierten Klassifikation*. Klett.
- Busse, D. (1997). Wortarten und semantische typen.
Überlegungen zu den Grundlagen der lexikalisch-syntaktischen Wortarten-Klassifikation. In: Christa Dürscheid/Karl Heinz Ramers/Monika Schwarz (Hrsg.): *Sprache im Fokus*.(Festschrift für Heinz Vater) Tübingen: Niemeyer, pages 219–240.
- Carstensen, K.-U., Ebert, C., Ebert, C., Jekat, S., Klabunde, R., and Langer, H., editors (2009). *Computerlinguistik und Sprachtechnologie: Eine Einführung*. Spektrum, Heidelberg, 3 edition.

Literaturverzeichnis II

Dürscheid, C. (2012). *Syntax. Grundlagen und Theorien*. 6., aktualisierte Aufl.

Pafel, J. (2011). *Einführung in die Syntax*. JB Metzler.