

Sprachorientierte Kl: Syntax und Parsing

Phrasenstrukturgrammatiken

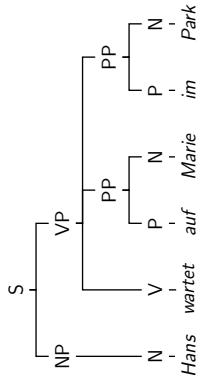
- Syntax als Untersuchungsgegenstand
- Wortartendisambiguierung
- Phrasenstrukturgrammatiken
 - Parsing mit Phrasenstrukturgrammatiken
 - Restringierte Phrasenstrukturgrammatiken
 - Unifikationsgrammatiken
 - Constraint-basierte Grammatiken
 - Robustes Parsing



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 1

Kontextfreie Grammatiken

- projektive Phrasenstrukturbäume (P-Marker) können durch kontextfreie Grammatiken erzeugt werden



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 3

Kontextfreie Grammatiken

- Phrasenstrukturbbaum ist Ableitungsprodukt der Grammatikanwendung
 - enger Zusammenhang zwischen
 - Regelstruktur
 - Strukturbeschreibung
 - Regельverwendung (Analyse/Generierung)
- Regeln lassen sich aus Phrasenstrukturbauem extrahieren



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 5

Kontextfreie Grammatiken

- strukturbildende Regeln, Grammatik
 - $S \rightarrow NP\ VP$
 - $VP \rightarrow V\ PP$
 - $VP \rightarrow V\ PP\ PP$
 - $PP \rightarrow P\ NP$
 - $NP \rightarrow N$
- erste Beschränkung für die Form der Regeln
 - Lexikon
 - $PT\text{-Symbol} \rightarrow T\text{-Symbol}$



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 7

Kontextfreie Grammatiken

- Kontextfreie Grammatiken
- Phrasale Kategorien
- ID/LP-Regeln
- Relationale Kategorien
- Subkategorisierung
- Rektion und Kongruenz
- Thematische Rollen
- Adjäquanzheit



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 2

Kontextfreie Grammatiken

- Regelstruktur
 - Nichtterminalsymbole: S, NP, VP, PP, ...
 - Terminalsymbole: *wartet, auf, im, Hans, Marie, Park*
 - NT-Symbol $\rightarrow \{T\text{-Symbol} \mid NT\text{-Symbol}\}^*$
- Regelverwendung
 - generativ
 - analytisch



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 4

Kontextfreie Grammatiken

- lexikalische Einsetzungsregeln, präterminale Regeln, Lexikon
 - $N \rightarrow Marie$
 - $N \rightarrow Hans$
 - $N \rightarrow Park$
 - $P \rightarrow auf$
 - $P \rightarrow im$
 - $V \rightarrow wartet$
- rekursive Regeln: potentiell unendlich viele Sätze können generiert werden
 - Kreativität der Sprachkompetenz
 - (noch) keine weiteren Beschränkungen für die Form der Regeln
 - Ziel linguistischer Modellierung!



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 6

Kontextfreie Grammatiken

- rekursive Regeln: potentiell unendlich viele Sätze können generiert werden
 - Kreativität der Sprachkompetenz
 - (noch) keine weiteren Beschränkungen für die Form der Regeln
 - Ziel linguistischer Modellierung!



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 8

Relationale Kategorien

Relationale Kategorien

- reine Dependenzrelationen
 $\text{depends_of}(X, Y)$
- verfeinert: funktionale Beschreibungen
(Unterordnungsrelationen)
 $\text{subject_of}(X, Y), \quad \text{object_of}(X, Y), \quad \dots$
- Subjekt: **Hans singt.**
 - direktes Objekt: **Hans trägt die Tasche.**
 - indirektes Objekt: **Hans gibt der Mutter die Tasche.**
 - Präpositionalobjekt: **Hans glaubt an das Gute im Menschen.**



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 17

Relationale Kategorien

Relationale Kategorien

- CHOMSKY 1965, 1981: funktionale Information ist redundant
 - Subjekt: NP, die unmittelbar vom Satzknöten dominiert wird
 - direktes Objekt: von den unmittelbar von einer VP dominierten NPs, die am weitesten links angeordnete
 - indirektes Objekt: die zweite, der unmittelbar von einer VP dominierten NPs
 - Präpositionalobjekt: die unmittelbar von einer PP dominierte NP



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 19

Relationale Kategorien

Relationale Kategorien

- Differenzierungsprinzip für Unterordnungsrelationen
 - Zwei Teilläume stehen in der gleichen Unterordnungsrelation zu einem Baumkontext wenn ihre wechselseitige Substitution keine ungrammatischen Bildungen zur Folge hat.
 - rein distributionelle Definition!



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 20

Subkategorisierung

Subkategorisierung

- Regeln einer Phrasenstrukturgrammatik ergeben sich direkt aus dem Phrasenstrukturbau, z.B.
 $S \rightarrow NP VP$
 $VP \rightarrow V NP PP$
 $PP \rightarrow P NP$
- Beseitigung von Redundanz
 $VP \rightarrow V NP$
 $VP \rightarrow V$
 - optionale Phrasenbestandteile $VP \rightarrow V (NP)$



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 21

Subkategorisierung

Subkategorisierung

- Stelligkeit der Verben
 - intransitiv: schlafen,
 - transitiv: tragen, beruhigen
 - bitransitiv: geben, verkaufen
 - dass-Satz: wissen, vermuten, glauben, denken
- Subkategorisierung: Unterscheidung bezüglich der Anzahl und der Kategorie von Ergänzungen (Komplementen) für eine lexikalische Kategorie



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 22

Subkategorisierung

Subkategorisierung

- feinere Unterteilung des Kategorieninventars
 - $V_i, V_t, V_{bt}, V_{dass}$
- Generalisierungen gehen verloren
 - z.B. Modifikation durch Adverbien
 - ... weil er viel schlält.
 - ... weil er den Koffer ungern trägt.
 - ... weil er ihr die Bücher günstig verkauft.
 - ... weil er sicher weiß, dass ...
- keine allgemeine Regel mehr



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 23



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 24

Relationale Kategorien

- andere syntaktische Perspektive: Valenzorientierte Beschreibung
- Abbildung auf semantische Funktionen ('Tiefenkasus'):
 - Agens, Rezipient, Patiens, Benefizient, ...



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 17

Relationale Kategorien

- andere syntaktische Perspektive: Valenzorientierte Beschreibung
- Abbildung auf semantische Funktionen ('Tiefenkasus'):
 - Agens, Rezipient, Patiens, Benefizient, ...



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 20

Relationale Kategorien

- andere syntaktische Perspektive: Valenzorientierte Beschreibung
- Abbildung auf semantische Funktionen ('Tiefenkasus'):
 - Agens, Rezipient, Patiens, Benefizient, ...



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 22

Subkategorisierung

Subkategorisierung

- Alternative: Subkategorisierungsrahmen
 - sehr allgemeine Phrasenstrukturregeln
 $VP \rightarrow V(NP)(NP)(PP)(S)$
 - Subkategorisierungsrahmen als zusätzliche Beschränkung für zulässige Phrasenstrukturbäume
 - schlafen: $__ NP$
 - tragen: $__ NP$
 - geben: $__ NP\ NP$
 - wissen: $__ S$
 - werden im Wörterbuch kodiert
 - lizensieren eine bestimmte syntaktische Umgebung



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 25

Rektion und Kongruenz

- Verben bestimmen nicht nur Anzahl und Art ihrer Komplemente zusätzlich auch deren Kasus (bzw. die Präposition von Präpositonalobjekten)
 - danken: Dativ
 - kennen: Akkusativ
 - gedenken: Genitiv
 - hoffen: auf (+ Akkusativ)
 - verkaufen: Akkusativ, Dativ



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 27

Subkategorisierung

- Rektion
 - Modellierung durch Kategoridifferenzierung in den Subkategorisierungsrahmen
 - tragen: $__ NP_{akk}$
 - drehen: $__ NP_{dat}$
 - gedenken: $__ NP_{gen}$
 - hoffen: $__ PP_{auf}$
 - verkaufen: $__ NP_{akk}\ NP_{dat}$



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 25

Rektion und Kongruenz

- Rektion (2)
 - Problem: Kombinatorik bei der Beschreibung komplexer Nominalgruppen
 - Nominalgruppen
 - NP_{akk} → NP_{gen} NP_{akk} Er trägt Marias Taschen.
 - NP_{akk} → NP_{akk} PP Er trägt die Taschen mit den Streifen.
 - NP_{akk} → NP_{akk} NP_{gen} Er trägt die Taschen der Kinder.
 - NP_{akk} → NP_{akk} S Er trägt die Taschen, die hier standen.
 - ...
■ NP_{dat} → NP_{gen} NP_{dat} Er droht Marias Töchtern.
 - NP_{dat} → NP_{dat} PP Er droht dem Mädchen mit den Locken.
 - NP_{dat} → NP_{dat} NP_{gen} Er droht den Töchtern der Nachbarin.
 - NP_{dat} → NP_{dat} S Er droht den Mädchen, die dort spielen.
 - ...



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 29

Rektion und Kongruenz

- Kongruenz
 - Subjekt-Prädikat: Numerus, Person
 - Modellierung durch Kategoridifferenzierung in den Regeln
 - S → NP_{nom,sg,1} VP_{sg,1}
 - S → NP_{nom,sg,2} VP_{sg,2}
 - S → NP_{nom,sg,3} VP_{sg,3}
 - S → NP_{nom,pl,1} VP_{pl,1}
 - S → NP_{nom,pl,2} VP_{pl,2}
 - S → NP_{nom,pl,3} VP_{pl,3}
 - ...



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 28

Rektion und Kongruenz

- Kombinatorik
 - zusätzlich: Fortpflanzung der Differenzierung auf weitere Regeln
 - NP_{nom,sg,3} → D_{nom,sg} N_{nom,sg}
 - ...
 - VP_{sg,1} → V_{sg,1} (NP) (NP) (PP) (S)
 - ...
 - Explosionsartige Vermehrung der Regeln
- Verlust von Generalisierungen
 - systematische Beziehungen gehen verloren z.B. zwischen NP_{nom,sg,1} und NP_{nom,pl,3}
 - Kategorien sind atomare Symbole, innere Struktur hat keinerlei Bedeutung innerhalb der Grammatik
 - kontextfreie Grammatiken sind deskriptiv inadäquat
 - flexionsmorphologische Fakten aus der Grammatik ausschließen
 - → Merkmalsbasierte Unifikationsgrammatiken



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 31

Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 25

Thematische Rollen

Thematische Rollen

- Interface zur Semantik
- Verben ~ Prädikate
 - jemand schläft \Rightarrow schlafen(X)
 - Valzen ~ Argumentstellen
 - Argumentstellen werden durch thematische Rollen (Θ -Rollen) beschrieben
 - Stelligkeit eines semantischen Prädikats: Θ -Raster



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 33

Thematische Rollen

Thematische Rollen

- Thematische Rollen
 - THEME (patients) Entität, auf die der Effekt einer Aktion ausgeübt wird
Marie fiel hin.
 - AGENT (actor) Verursacher einer Aktion
Hans traf das Ziel.
 - EXPERIENCER Entität, die einen bestimmten (psychologischen) Zustand einnimmt
Marie ist glücklich.
 - BENEFACTIVE Entität, die eine Auswirkung einer Aktion erfährt
Marie umarmte Hans.
 - LOCATIVE Ort an dem sich etwas befindet, oder an dem eine Aktion stattfindet
Hans kommt aus der Schule.



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 35

Thematische Rollen

Thematische Rollen

- selektionale Einschränkungen für die Zuweisung der Θ -Rollen
 - Der Vater singt ein Lied.*
 - ? *Der Stuhl singt ein Lied.*
 - ? *Der Vater singt eine Tasse.*
- verkaufen[agent (human) , theme , goal (human)]
diskutieren[agent (human) , theme (abstract)]
schlafen[agent (animated)]
- aber:
 - übertragene Bedeutungen
 - Granularität der Beschreibungen
 - z.B. THEME bei essen, singen, auffädeln, ...



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 37

Thematische Rollen

- Zuweisung der thematischen Rollen ist unabhängig von der syntaktischen Strukturierung
 - Die Vase zerbrach die Fensterscheibe.*
 - Die Vase zerbrach.*
- gleiche syntaktische Struktur, aber INSTRUMENT in der transitiven Verwendung gegenüber THEME in der ergativen Verwendung
 - unterschiedlicher thematischer Status führt zu unterschiedlichen selektionalen Beschränkungen
 - Die Vase / der Fußball / der Knall zerbrach die Fensterscheibe.*
 - Die Vase /² Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 39*



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 40

Thematische Rollen

- Theta-Kriterium: eineindeutige Beziehungen zwischen der syntaktischen Struktur und den thematischen Rollen
- prinzipielle Zuordnung
 - ein intransitives Verb weist nur seiner Subjektposition eine Θ Rolle zu
 - * *Hier schläft Hans.*



Hier schläft.

* *Hier wird geschlafen.*



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 34

Thematische Rollen

Thematische Rollen

- Thematische Rollen
 - INSTRUMENT Hilfsmittel für das Ausführen einer Aktion
Hans aß den Salat mit der Gabel.
 - LOCATIVE Ort an dem sich etwas befindet, oder an dem eine Aktion stattfindet
Marie suchte die Pralinen im Schrank.
 - GOAL Entitäts zu der sich etwas hinbewegt
Er legte das Schreiben in die Ablage.
 - SOURCE Entität von der sich etwas weg bewegt
Hans kommt aus der Schule.



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 36

Thematische Rollen

- Modellierung der Ähnlichkeit zwischen verschiedenen syntaktischen Verwendungen eines Wortes
- transitiv: *Hans rollte den Ball den Berg hinunter.*
- ergativ: *Der Ball rollte den Berg hinunter.*
- Objekt wird bei intrans. (ergativen) Verwendung zum Subjekt
- Zuweisung der gleichen thematischen Rolle: THEME
 - gestützt durch gleiche selektionale Distribution
 - Hans rollte den Ball / den Stein / den Käse / ?die Ansprache / ?die Dämmerung den Berg hinunter.*
 - Der Ball / der Stein / der Käse / ?die Ansprache / ?die Dämmerung rollte den Berg hinunter.*



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 38

Adjäquatheit

- Erzeugungsadjäquatheit
- Beschreibungsadjäquatheit
- Erklärungsadjäquatheit
- Verarbeitungsadjäquatheit



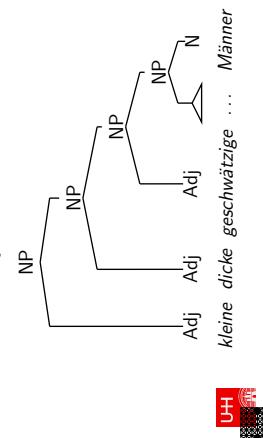
Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 40

Erzeugungssadäquatheit

- sprachliche Abdeckung / Übergeneralisierung
- prinzipiell ausreichend
- nur ganz seltene Phänomene mit "milder" Kontextsensitivität
- Schweizerdeutscher Dialekt der Zürcher Region (Huybrechts 1985, Shieber 1985)
- *Claudia beobachtete Helmut ließ Eva helfen Hans mache Ulrike arbeitete.*
- *Claudia Helmut Eva Hans Ulrike beobachtete ließ helfen machte arbeite.*
- überlagert mit Kasusrektion:
- $a^m b^n c^m d^n$ mit $n > 0$

Beschreibungssadäquatheit

- rekursive Registriertuktur mit Hilfsknoten
- $NP \rightarrow N$
 - $NP \rightarrow Adj\ NP$



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 43

Beschreibungssadäquatheit

- Repräsentationsprobleme
 - (unendliche) Adjektivreihungen
-
- NP → N
NP → Adj N
NP → Adj Adj N
NP → Adj Adj Adj N
...
Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 42

Beschreibungssadäquatheit

- suggeriert ein Dromianzverhältnis, das nicht gegeben ist
- ähnliche Probleme bei der Behandlung freier Adjunkte
- Problem tritt bei der Dependenzgrammatik in analoger Weise auf



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 44

Beschreibungssadäquatheit

- nichtprojektive Strukturen
- abtrennbare Verbzusätze
- elliptische Konstruktionen bei der Koordination
- diskontinuierliche Konstituenten
- Extrapolation
- Fragestanzbildung
- Topikalisierung



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 45

Beschreibungssadäquatheit

- Hauptproblem bei diskontinuierlichen Konstituenten: fehlende Parallelität bei strukturell verwandten Konstruktionen
- *Uns hat der Wein, den du so toll findest, nicht geschmeckt.*
- *Uns hat der Wein nicht geschmeckt, den du so toll findest.*
- Lösungsmöglichkeiten
 - Transformationen (derivationeller Ansatz)
 - Slash-Merkmal in Unifikationsgrammatiken (nichtderivationeller Ansatz)
- Kombinatorik und Generalisierungen bei



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 46

Erklärungssadäquatheit

- Lernbarkeit
- Unterbestimmtheit des sprachlichen Wissens durch die verfügbare sprachliche Evidenz
- poverty of stimulus (Fodor 1983)
- quantitative Unterbestimmtheit
- Sprachausschnitt, aus dem das Kind Hypothesen über Sprachstrukturen gewinnen muss ist sehr begrenzt
- nur ein geringer Teil der möglichen Strukturen wird dem Kind angeboten



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 47

Erklärungssadäquatheit

- qualitative Unterbestimmtheit
- Erlernen von Regeln erfordert ausreichend positive und negative Evidenz
- Erlerte Regeln müssen auf ihre Adäquatheit geprüft werden
- Erlernen von Regeln ist aufwendig und lässt die Kommunikation in den Hintergrund treten:
 - "Es ist jedoch hinreichend bekannt, dass Kinder Äußerungen produzieren, um ihre Gedanken und Wünsche mitzuteilen und nicht, um syntaktische Regeln auszuprobieren."



Wolfgang Menzel: Sprachorientierte Kl: Syntax und Parsing – p. 48

Erklärungsadäquatheit

Erklärungsadäquatheit

- insbesondere negative Evidenz ist unzureichend
- Nichtauftreten einer Konstruktion \neq Ungrammatikalität
- Eltern korrigieren ihre Kinder selten und unsystematisch wenn, dann vorrangig bezüglich inhaltlicher Aspekte
- systematische Korrekturversuche bringen nur wenig Vorteile für den Grammatikerwerb



Wolfgang Menzel: Sprachorientierte KI: Syntax und Parsing – p. 49

- Sprache wird nicht gelernt, sondern erworben

- Language acquisition device als mentales Organ
- Universalgrammatik mit universalgrammatischen Prinzipien
- beim Spracherwerb werden nur einzelsprachliche Varianten (Parameter) gelernt



Wolfgang Menzel: Sprachorientierte KI: Syntax und Parsing – p. 50

Erklärungsadäquatheit

Erklärungsadäquatheit

- Strukturelle Eigenschaften der Phrasenstrukturgrammatik erklären nicht, was eine natürliche Sprache ist
- wenn es möglich sein sollte, beliebige rekursive Regelsysteme zu erlernen, dann müsste, dies auf beliebige sequentielle Zeichenanordnungen übertragbar sein
- Phrasenstrukturgrammatiken sind für sich alleine noch keine Theorie der Sprachstruktur keinerlei Bezug zum Sprachvermögen des Menschen



Wolfgang Menzel: Sprachorientierte KI: Syntax und Parsing – p. 51

- einfacher Formalismus

- deklarativer Formalismus
- relativ effiziente Analysealgorithmen



Wolfgang Menzel: Sprachorientierte KI: Syntax und Parsing – p. 52

Verarbeitungsadäquatheit

Verarbeitungsadäquatheit

- einfacher Formalismus
- deklarativer Formalismus
- relativ effiziente Analysealgorithmen



Wolfgang Menzel: Sprachorientierte KI: Syntax und Parsing – p. 52

Erklärungsadäquatheit

Erklärungsadäquatheit

- Sprache wird nicht gelernt, sondern erworben
- Language acquisition device als mentales Organ
- Universalgrammatik mit universalgrammatischen Prinzipien
- beim Spracherwerb werden nur einzelsprachliche Varianten (Parameter) gelernt



Wolfgang Menzel: Sprachorientierte KI: Syntax und Parsing – p. 50