

Einführung in die Grammatikentwicklung

Stefan Müller

Deutsche Grammatik
Institut für Deutsche und Niederländische Philologie
Fachbereich Philosophie und Geisteswissenschaften
FU Berlin

Stefan.Mueller@fu-berlin.de

15. April 2009

Kursziele

- Kennelernen eines Grammatikentwicklungssystems
- Lernen, wie man Grammatiken mit getypten Merkmalsbeschreibungen formalisiert
- Anpassen und entwickeln einer Reihe einfacher HPSG-Grammatiken in TRALE

Warum Grammatikentwicklung mit dem Computer?

- **Forschung** Formalisierung linguistischer Theorien mit komplexen Interaktionen zwischen einzelnen Phänomenen.
Identifizierung von sprachübergreifenden Generalisierungen
- **Lehre** Lehre von Frameworks oder Analysen in Morphologie, Syntax und Semantik; Förderung der experimentellen Arbeit von Studenten
- **Anwendung** Einbettung grammatikbasierter Analyse/Generierung von natürlichen Sprachen in Forschungsprototypen bzw. kommerzielle Anwendungen

Alte Weisheit

[Grammatik ist] das Tor zur Freiheit, die Medizin für die Krankheiten der Sprache, der Reiniger aller Wissenschaften; sie verbreitet ihr Licht über ihnen; . . . sie ist die erste Sprosse auf der Leiter, die zur Realisierung übernatürlicher Kräfte führt und der gerade, königliche Weg für diejenigen, die die Freiheit suchen. (Bhartrhari, Spruchdichter, gest. vor 650 n. Chr., aus *Vakyapadiya*, gefunden von Gabriele Knoll)

Wozu Syntax?

- Zeichen: Form-Bedeutungs-Paare (Saussure, 1916)
- Wörter, Wortgruppen, Sätze
- Sprache $\stackrel{?}{=}$ endliche Aufzählung von Wortfolgen
Sprache ist endlich, wenn man maximale Satzlänge annimmt
 - (1) a. Dieser Satz geht weiter und weiter und weiter und weiter ...
 - b. [Ein Satz ist ein Satz] ist ein Satz.

extrem viele Sätze, Beschränkung der Wiederholung willkürlich
- Unterscheidung zwischen **Kompetenz** (das Wissen darüber, was geht) und **Performanz** (der Benutzung des Wissens)

Die Kinder von Bullerbü

Und wir beeilten uns, den Jungen zu erzählen, wir hätten von Anfang an gewußt, daß es nur eine Erfindung von Lasse gewesen sei. Und da sagte Lasse, die Jungen hätten gewußt, daß wir gewußt hätten, es sei nur eine Erfindung von ihm. Das war natürlich gelogen, aber vorsichtshalber sagten wir, wir hätten gewußt, die Jungen hätten gewußt, daß wir gewußt hätten, es sei nur eine Erfindung von Lasse. Und da sagten die Jungen – ja – jetzt schaffe ich es nicht mehr aufzuzählen, aber es waren so viele „gewußt“, daß man ganz verwirrt davon werden konnte, wenn man es hörte. (S. 248)

Wir sind prinzipiell in der Lage, komplexere Sätze zu bilden (Kompetenz), aber irgendwann werden wir verwirrt, weil unsere Gehirne nicht mehr mitmachen (Performanz).

Kreativität

- Wir können Sätze bilden, die wir noch nie gehört haben →
muss Strukturierung, Muster geben

Direkte Evidenz für syntaktische Strukturen?

- Wir können feststellen, dass wir Regeln verwenden, indem wir Kinder beobachten.
Kinder wenden Regeln mitunter falsch an.
- Beispiel aus der Morphologie:
 - (2) a. * die Baggers
 - b. * die Ritters

Wozu Syntax? Bedeutung aus Bestandteilen ermitteln

- Bedeutung einer Äußerung aus den Bedeutungen ihrer Teile bestimmen
 - (3) Der Mann kennt diese Frau.
- Syntax: Art und Weise der Kombination, Strukturierung
 - (4)
 - a. Die Frau kennt die Mädchen.
 - b. Die Frau kennen die Mädchen.
 - c. Die Frau schläft.
 - d. Die Mädchen schlafen.

Subjekt-Verb-Kongruenz → Bedeutung von (4a,b) ist eindeutig

Warum formal?

Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects. I think that some of those linguists who have questioned the value of precise and technical development of linguistic theory have failed to recognize the productive potential in the method of rigorously stating a proposed theory and applying it strictly to linguistic material with no attempt to avoid unacceptable conclusions by ad hoc adjustments or loose formulation. (Chomsky, 1957, S. 5)

As is frequently pointed out but cannot be overemphasized, an important goal of formalization in linguistics is to enable subsequent researchers to see the defects of an analysis as clearly as its merits; only then can progress be made efficiently. (Dowty, 1979, S. 322)

- Was bedeutet eine Analyse genau?
- Welche Vorhersagen macht sie?
- Ausschluß anderer Analysen

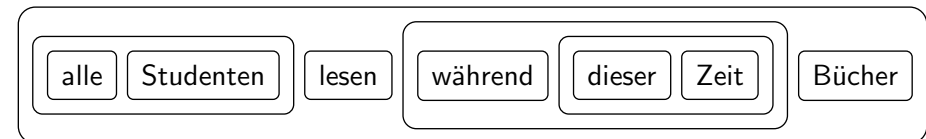
Einteilung in Einheiten

- Sätze können Sätze enthalten, die Sätze enthalten, die ... :
 - (5) dass Max glaubt, [dass Julius weiß, [dass Otto behauptet, [dass Karl vermutet, [dass Richard bestätigt, [dass Friederike lacht]]]]]]

Das funktioniert wie eine Matrjoschka bzw. wie eine Zwiebel.
- Genauso kann man in (6) Wörter zu Einheiten zusammenfassen:
 - (6) Alle Studenten lesen während dieser Zeit Bücher.

Welche?

Schachteln



Wir tun alle Wörter, die zusammengehören, in eine Schachtel.

Diese Schachteln können wieder in andere Schachteln getan werden.

Im Beispiel ist intuitiv klar, was zusammengehört, aber gibt es Tests?

Konstituenz

Begriffe:

Wortfolge Eine beliebige linear zusammenhängende Folge von Wörtern, die nicht unbedingt syntaktisch oder semantisch zusammengehörig sein müssen.

Wortgruppe, Konstituente, Phrase Ein Wort oder mehrere Wörter, die eine strukturelle Einheit bilden.

Konstituententests

Welche kennen Sie?

- Substituierbarkeit/Pronominalisierungstest/Fragetest
- Weglaßtest
- Verschiebetest (Umstelltest)
- Koordinationstest

Konstituententests (I)

Substituierbarkeit Kann man eine Wortfolge einer bestimmten Kategorie in einem Satz gegen eine andere Wortfolge so austauschen, dass wieder ein akzeptabler Satz entsteht, so ist das ein Indiz dafür, dass die beiden Wortfolgen Konstituenten bilden.

- (7) a. Er kennt den Mann.
b. Er kennt eine Frau.

Pronominalisierungstest Alles, worauf man sich mit einem Pronomen beziehen kann, ist eine Konstituente.

- (8) a. Der Mann schläft.
b. Er schläft.

Konstituententests (II)

Fragetest Was sich erfragen läßt, ist eine Konstituente.

- (9) a. Der Mann arbeitet.
b. Wer arbeitet?

Verschiebetest Wortfolgen, die man ohne Beeinträchtigung der Korrektheit des Satzes verschieben bzw. umstellen kann, bilden eine Konstituente.

- (10) a. weil keiner diese Frau kennt.
b. weil diese Frau keiner kennt.

Koordinationstest Was sich koordinieren läßt, ist eine Konstituente.

- (11) Der Mann und die Frau arbeiten.

Bemerkungen zum Status der Tests: Expletiva (I)

Was ist mit *es* in (12)?

(12) Es regnet.

Substituierbarkeit und Fragetest schlagen fehl:

- (13) a. * Der Mann/er regnet.
b. * Wer/was regent?

Aus denselben Gründen schlägt der Koordinationstest fehl:

(14) * Es und der Mann regnet.

Bemerkungen zum Status der Tests: Expletiva (II)

Nur die (allerdings eingeschränkte) Umstellbarkeit ist gegeben:

- (15) a. Es regnet.
b. Regnet es?
c. weil es jetzt regnet
d. * weil jetzt es regnet

- (16) a. Er sah es regnen.
b. * Es sah er regnen.

Daraus folgt: Nicht alle Tests müssen positiv ausfallen, damit eine Wortfolge als Konstituente gelten kann, d. h., die Test stellen keine notwendige Bedingung dar.

Bemerkungen zum Status der Tests: Koordination

Was ist mit *der Mann einen Esel* und *die Frau ein Pferd* in (17)?

(17) Deshalb kaufte der Mann einen Esel und die Frau ein Pferd.

Diese Wörter kann man nur sehr bedingt gemeinsam umstellen:

(18) ?* Der Mann einen Esel kaufte deshalb.

Ein Ersetzung durch Pronomina ist nicht ohne Ellipse möglich:

- (19) a. # Deshalb kaufte er.
b. * Deshalb kaufte ihn.

Die Pronomina stehen nicht für beide logischen Argumente sondern nur für jeweils eins.

Daraus folgt: Auch wenn einige Tests erfüllt sind, muß es noch lange nicht sinnvoll sein, eine Wortfolge als Konstituente einzustufen, d. h., die Test stellen keine hinreichende Bedingung dar.

Bemerkungen zum Status der Tests: Voranstellung (I)

Normalerweise steht im Deutschen eine Konstituente vor dem Finitum.

- (20) a. [Alle Studenten] lesen während der vorlesungsfreien Zeit Bücher.
b. [Bücher] lesen alle Studenten während der vorlesungsfreien Zeit.
c. * [Alle Studenten] [Bücher] lesen während der vorlesungsfreien Zeit.
d. * [Bücher] [alle Studenten] lesen während der vorlesungsfreien Zeit.

Voranstellbarkeit vor das finite Verb wird in manchen Definitionen sogar zum ausschlaggebenden Kriterium für *Satzglied* (Duden, 2005, S. 783).

Bemerkungen zum Status der Tests: Voranstellung (II)

Satzgliedtest [Auch: Konsituententest]. Auf der → Topikalisierung beruhendes Verfahren zur Analyse komplexer Konstituenten. Da bei Topikalisierung jeweils nur eine Konstituente bzw. ein → Satzglied an den Anfang gerückt werden kann, lassen sich komplexe Abfolgen von Konstituenten (z. B. Adverbialphrasen) als ein oder mehrere Satzglieder ausweisen; in *Ein Taxi quält sich im Schrittempo durch den Verkehr* sind *im Schrittempo* und *durch den Verkehr* zwei Satzglieder, da sie beide unabhängig voneinander in Anfangsposition gerückt werden können. (Bußmann, 1983, S. 446)

nicht mehr enthalten in Bußmann, 1990

Bemerkungen zum Status der Tests: Voranstellung (III)

Nach Bußmann:

- Teile des Materials können einzeln vorangestellt werden. → Das Material bildet keine Konstituente.
- Material kann zusammen vorangestellt werden. → Das Material bildet eine Konstituente.

Beide Implikationen sind problematisch.

Die erste ist wegen Beispielen wie (21) problematisch:

- (21) a. Keine Einigung erreichten Schröder und Chirac über den Abbau der Agrarsubventionen. (tagesschau, 15.10.2002, 20:00)
- b. Über den Abbau der Agrarsubventionen erreichten Schröder und Chirac keine Einigung.

Bemerkungen zum Status der Tests: Voranstellung (IV)

Obwohl Teile der NP einzeln vorangestellt werden können, wollen wir die Wortfolge als eine NP analysieren, wenn sie nicht vorangestellt ist.

- (22) Schröder und Chirac erreichten keine Einigung über den Abbau der Agrarsubventionen.

Diese Wortgruppe kann auch gemeinsam vorangestellt werden:

- (23) Keine Einigung über den Abbau der Agrarsubventionen erreichten Schröder und Chirac.

Keine Einigung über den Abbau der Agrarsubventionen ist eine Konstituente, die unter gewissen Umständen aufgespalten werden kann.

Bei Aufspaltung können die einzelnen Teilkonstituenten unabhängig voneinander umgestellt werden.

Bemerkungen zum Status der Tests: Voranstellung (V)

Der zweite Teil des Konstituententests ist ebenfalls problematisch:

- (24) a. [Dauerhaft] [mehr Arbeitsplätze] gebe es erst, wenn sich eine Wachstumsrate von mindestens 2,5 Prozent über einen Zeitraum von drei oder vier Jahren halten lasse. (taz, 19.04.2000, S. 5)
- b. [Wenig] [mit Sprachgeschichte] hat der dritte Beitrag in dieser Rubrik zu tun, [...] (ZS für Dialektologie und Linguistik, LXIX, 3/2002, S. 339)

Mehr Daten in Müller, 2003.

Wörter vor Finitum stehen weder in semantischer noch in syntaktischer Beziehung zueinander → nicht sinnvoll, sie als eine Konstituente zu analysieren

Die Daten kann man mit einem leeren verbalen Kopf im Vorfeld analysieren, so dass letztendlich wieder V2-Strukturen vorliegen (Müller, 2005). Trotzdem sind die Daten für Konstituententests problematisch.

Voranstellbarkeit ist nicht hinreichend für Konstituentenstatus.

Bemerkungen zum Status der Tests: Voranstellung (VI)

- (25) a. Er bringt es bis zum Professor.
 b. # Es bringt er zum Professor.

es ist Konstituente, obwohl es nicht vorangestellt werden kann.

Genauso:

- (26) a. Karl hat sich nicht erholt.
 b. *Sich hat Karl nicht erholt.

- (27) a. Er hörte es regnen.
 b. *Es hörte er regnen.

→ Vorstellbarkeit ist nicht notwendig.

Also: Vorstellbarkeit ist weder hinreichend noch notwendig.

Köpfe

Kopf bestimmt die wichtigsten Eigenschaften einer Phrase

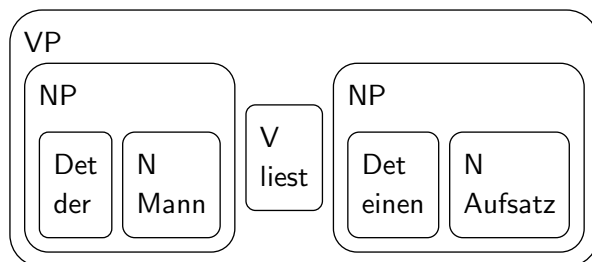
- (28) a. **Träumt** er?
 b. **Erwartet** er einen dreiprozentigen Anstieg?
 c. **in** diesem Haus
 d. ein **Mann**

Kombination eines Kopfes mit anderem Material wird **Projektion des Kopfes** genannt.

Eine vollständige Projektion ist eine **Maximalprojektion**

Ein Satz ist die Maximalprojektion eines finiten Verbs.

Beschriftete Schachteln



Wer schon einmal umgezogen ist, weiß, dass es sinnvoll ist, Schachteln zu beschriften.

Im obigen Bild steht auf jeder Schachtel etwas über das wichtigste Element in der Schachtel.

Schachteln sind austauschbar

- Der genaue Inhalt einer Schachtel ist egal:

- (29) a. er
 b. der Mann
 c. der Mann aus Stuttgart
 d. der Mann aus Stuttgart, den wir kennen

Wichtig ist, dass die Wörter bzw. Wortfolgen in (29) alle nominal sind und vollständig: NP.

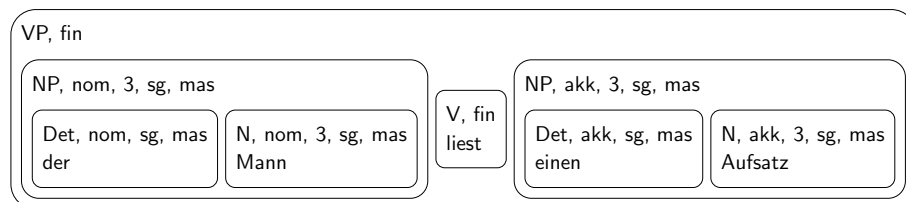
Man kann sie innerhalb größerer Schachtel gegeneinander vertauschen.

- Das geht aber nicht mit allen NPen:

- (30) a. Der Mann liest einen Aufsatz.
 b. *Die Männer liest einen Aufsatz.
 c. *Des Mannes liest einen Aufsatz.

- Es gibt Eigenschaften, die für die Verteilung (Distribution) von Phrasen wichtig sind.

Ausführlich beschriftete Schachteln



Alle Merkmale, die für die Distribution der gesamten Phrase wichtig sind, werden projiziert.

Diese Merkmale werden auch **Kopfmerkmale** genannt.

Argumente

- Konstituenten stehen in verschiedenartigen Beziehungen zu ihrem Kopf.
- Man unterscheidet zwischen **Argumenten** und **Adjunkten**.
- Bestimmte Mitspieler (Aktanten) gehören zur Bedeutung eines Verbs. Z. B. gibt es in Situationen, die durch *lieben* beschrieben werden, immer einen *liebenden* und einen *geliebten* / etwas *geliebtes*.

- (31) a. Peter liest Maria.
b. *lieben'*(Peter', Maria')

(31b) ist eine logische Repräsentation für (31a).

Peter' und Maria' sind **logische Argumente** von *lieben'*.

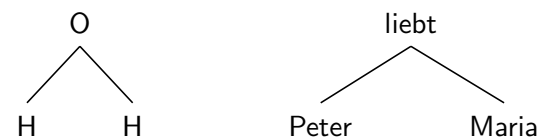
- Syntaktische Argumente entsprechen meistens den logischen (später mehr).
- Solche Beziehungen zwischen Kopf und Argumenten werden mit dem Begriff **Selektion** bzw. **Valenz** erfasst.
- Tesnière (1959) überträgt Valenzbegriff aus der Chemie auf die Linguistik.

Valenz in der Chemie

- Atome können sich mit anderen Atomen zu mehr oder weniger stabilen Molekülen verbinden.
- Wichtig für die Stabilität ist, wie Elektronenschalen besetzt sind.
- Eine Verbindung mit anderen Atomen kann dazu führen, dass eine Elektronenschale voll besetzt ist, was dann zu einer stabilen Verbindung führt.
- Die Valenz sagt etwas über die Anzahl der Wasserstoffatome aus, die mit einem Atom eines Elements verbunden werden können.
- In der Verbindung H₂O hat Sauerstoff die Valenz 2.
- Man kann nun die Elemente in Valenzklassen einteilen. Elemente mit einer bestimmten Valenz werden im Periodensystem von Mendeleev in einer Spalte repräsentiert.

Valenz in der Linguistik

- Ein Kopf braucht bestimmte Argumente, um eine stabile Verbindung einzugehen.
- Wörter mit der gleichen Valenz (mit gleicher Anzahl und Art von Argumenten) werden in Valenzklassen eingeordnet, da sie sich in bezug auf die Verbindungen, die sie eingehen, gleich verhalten.



Verbindung von Sauerstoff mit Wasserstoff und Verbindung eines Verbs mit seinen Argumenten

Optionale Argumente

- Argumente müssen nicht immer realisiert werden:

- (32) a. Er wartet auf den Installateur.
b. Er wartet.

Das Präpositionalobjekt von *warten* ist ein **fakultatives Argument**.

- In nominalen Umgebungen sind Argumente immer optional!

- (33) a. Jemand liest diese Bücher.
b. das Lesen dieser Bücher
c. das Lesen

Syntaktische Argumente, die keine logischen sind

- In unserem bisherigen Beispiel entsprechen die syntaktischen den logischen Argumenten:

- (34) a. Peter liebt Maria.
b. *lieben'*(*Peter'*, *Maria'*)

- Allerdings gibt es auch Argumente, die keinen semantischen Beitrag leisten:

- (35) a. Es regnet.
b. Peter erholt sich.

es und *sich* sind **syntaktische Argumente**,
aber keine **logischen Argumente**.

Argumente und Adjunkte

- Adjunkte füllen keine semantische Rolle
- Adjunkte sind optional
- Adjunkte sind iterierbar

Adjunkte füllen keine semantische Rolle

- In einer *lieben*-Situation gibt es einen Liebenden und etwas Geliebtes.
seit der Schulzeit in (36) ist von anderer Art:

- (36) Peter liebt Maria seit der Schulzeit.

Es sagt zusätzlich etwas über die Dauer der Relation aus,
in der Peter und Maria zueinander stehen.

Adjunkte sind optional

- Adjunkte sind optional:
 - (37) a. Peter liebt Maria.
 - b. Peter liebt Maria seit der Schulzeit.
 - c. Peter liebt Maria aufrichtig.
- Vorsicht! Das ist auch bei Argumenten mitunter der Fall:
 - (38) a. Er gibt den Armen Geld.
 - b. Er gibt den Armen.
 - c. Er gibt Geld.
 - d. Er gibt gerne.
 - e. Du gibst. (beim Skat)
 - f. Gib!

Adjunkte sind iterierbar

- Argumente können nur einmal mit dem Kopf kombiniert werden:
 - (39) * Der Mann der Mann schläft.

Die entsprechende Andockstelle des Kopfes (*schläft*) ist besetzt.
- Bei Adjunkten ist das anders:
 - (40) A: Alle klugen Frauen sind unglücklich.
 - B: Nein, ich kenne eine glückliche kluge Frau.
 - A: Aber alle glücklichen klugen Frauen sind schön.
 - B: Nein, ich kenne eine hässliche glückliche kluge Frau.
 - ...

Weiter Beispiele für Adjunkte

Adverbial gebrauchtes Adjektiv (nicht alle):

(41) Karl schnarcht *laut*.

Relativsätze (nicht alle):

(42) a. der Mann, *den Maria liebt*
 b. der Mann, *der Maria liebt*

Präpositionalphrasen (nicht alle):

(43) a. Die Frau arbeitet *in Berlin*.
 b. die Frau *aus Berlin*

Andere Bezeichnungen

- Argument: Ergänzung
- Adjunkt: (freie) Angabe
- Argumente werden mitunter in Subjekt und Komplemente aufgeteilt.
- auch Aktant für Subjekte und Objekte (aber nicht Prädikative und Adverbialia)
- Zirkumstant für Adverbialia
 - Adverbiale des Raumes (Lage, Richtung/Ziel, Herkunft, Weg)
 - Adverbiale der Zeit (Zeitpunkt, Anfang, Ende, Dauer)
 - Adverbiale des Grundes.
Hierher werden traditionellerweise auch Adverbialien gestellt, die einen Gegengrund oder eine Bedingung ausdrücken.
 - Adverbiale der Art und Weise.

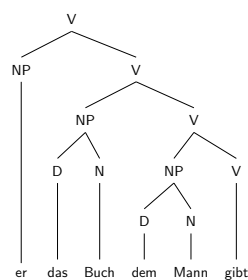
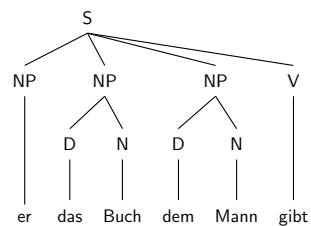
Verschiedene Grammatikmodelle (I)

- **Dependenzgrammatik (DG)**
(Tesnière, 1980; Kunze, 1975; Weber, 1992; Heringer, 1996; Eroms, 2000)
- **Kategorialgrammatik (CG)**
(Ajdukiewicz, 1935; Steedman, 2000)
- **Phrasenstrukturgrammatik (PSG)**
- **Transformationsgrammatik und deren Nachfolger**
 - **Transformationsgrammatik**
(Chomsky, 1957; Bierwisch, 1963)
 - **Government & Binding**
(Chomsky, 1981; von Stechow und Sternefeld, 1988; Grewendorf, 1988)
 - **Minimalismus**
(Chomsky, 1995; Grewendorf, 2002)

Verschiedene Grammatikmodelle (II)

- **Tree Adjoining Grammar**
(Joshi, Levy und Takahashi, 1975; Joshi, 1987; Kroch und Joshi, 1985)
- **Generalisierte Phrasenstrukturgrammatik (GPSG)**
(Gazdar, Klein, Pullum und Sag, 1985; Uszkoreit, 1987)
- **Lexikalisch Funktionale Grammatik (LFG)**
(Bresnan, 1982, 2001; Berman und Frank, 1996; Berman, 2003)
- **Head-Driven Phrase Structure Grammar (HPSG)**
(Pollard und Sag, 1987, 1994; Müller, 1999, 2002, 2008)
- **Construction Grammar (CxG)**
(Fillmore, Kay und O'Connor, 1988; Goldberg, 1995, 2006; Fischer und Stefanowitsch, 2006)

Phrasenstrukturen



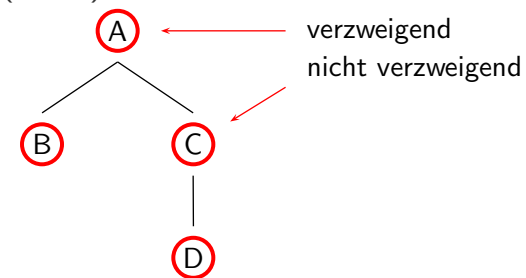
NP → D, N
S → NP, NP, NP, V

NP → D, N
V → NP, V

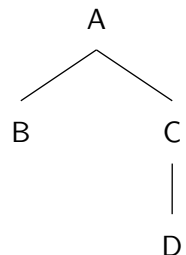
Das Eigentliche sind die Ersetzungsregeln! Die Bäume sind nur die Visualisierung.
Aus Platzgründen auch Klammerschreibweise:

[S [NP er] [NP [D das] [N Buch]] [NP [D dem] [N Mann]] [V gibt]]

Knoten (node)



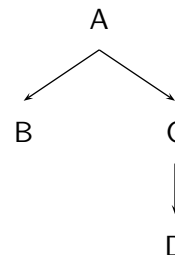
Mutter, Tochter und Schwester



A ist die Mutter von B und C
C ist die Mutter von D
B ist die Schwester von C

Verhältnisse wie in Stammbäumen

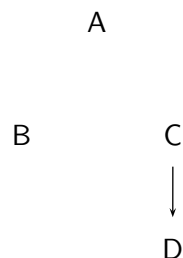
Dominanz (*dominance*)



A dominiert B, C und D
C dominiert D

A dominiert B genau dann, wenn A höher im Baum steht und wenn es eine ausschließlich abwärts führende Linie von A nach B gibt.

Unmittelbare Dominanz (*immediate dominance*)



A dominiert unmittelbar B und C
C dominiert unmittelbar D

A dominiert unmittelbar B genau dann, wenn A B dominiert und es keinen Knoten C zwischen A und B gibt.

Begriffe

Präzedenz (*precedence*)

A geht B voran, wenn A in einer Baumgrafik vor B steht und keiner der beiden Knoten den anderen dominiert.

Unmittelbare Präzedenz (*immediate precedence*)

Kein Element C zwischen A und B.

Beispielableitung bei Annahme flacher Strukturen

NP → D, N	NP → er	N → Buch
S → NP, NP, NP, V	D → das	N → Mann
	D → dem	V → gibt
er das Buch dem Mann gibt		
NP das Buch dem Mann gibt	NP → er	
NP D Buch dem Mann gibt	D → das	
NP D N dem Mann gibt	N → Buch	
NP NP dem Mann gibt	NP → D, N	
NP NP D Mann gibt	D → dem	
NP NP D N gibt	N → Mann	
NP NP NP gibt	NP → D, N	
NP NP NP V	V → gibt	
	S → NP, NP, NP, V	

Von der Grammatik beschriebene Sätze

- die Grammatik ist zu ungenau:
 - NP → D, N
 - S → NP, NP, NP, V
- (44)
- er das Buch dem Mann gibt.
 - * ich das Buch dem Mann gibt.
(Subjekt-Verb-Kongruenz *ich, gibt*)
 - * er das Buch den Mann gibt.
(Kasusanforderungen des Verbs *gibt* verlangt Dativ)
 - * er den Buch dem Mann gibt.
(Determinator-Nomen-Kongruenz *den, Buch*)

Subjekt-Verb-Kongruenz (I)

- Übereinstimmung in Person (1, 2, 3) und Numerus (sg, pl)
- (45)
- Ich schlafe. (1, sg)
 - Du schläfst. (2, sg)
 - Er schläft. (3, sg)
 - Wir schlafen. (1, pl)
 - Ihr schlaft. (2, pl)
 - Sie schlafen. (3,pl)
- Wie drückt man das in Regeln aus?

Subjekt-Verb-Kongruenz (II)

- Verfeinerung der verwendeten Symbole aus S → NP, NP, NP, V wird
- S → NP_{1_sg}, NP, NP, V_{1_sg}
 S → NP_{2_sg}, NP, NP, V_{2_sg}
 S → NP_{3_sg}, NP, NP, V_{3_sg}
 S → NP_{1_pl}, NP, NP, V_{1_pl}
 S → NP_{2_pl}, NP, NP, V_{2_pl}
 S → NP_{3_pl}, NP, NP, V_{3_pl}
- sechs Symbole für Nominalphrasen, sechs für Verben
 - sechs Regeln statt einer

Kasuzuweisung durch das Verb

- Kasus muß repräsentiert sein:
 $S \rightarrow NP_{1_sg_nom}, NP_{dat}, NP_{acc}, V_{1_sg_ditransitiv}$
 $S \rightarrow NP_{2_sg_nom}, NP_{dat}, NP_{acc}, V_{2_sg_ditransitiv}$
 $S \rightarrow NP_{3_sg_nom}, NP_{dat}, NP_{acc}, V_{3_sg_ditransitiv}$
 $S \rightarrow NP_{1_pl_nom}, NP_{dat}, NP_{acc}, V_{1_pl_ditransitiv}$
 $S \rightarrow NP_{2_pl_nom}, NP_{dat}, NP_{acc}, V_{2_pl_ditransitiv}$
 $S \rightarrow NP_{3_pl_nom}, NP_{dat}, NP_{acc}, V_{3_pl_ditransitiv}$
- insgesamt $3 * 2 * 4 = 24$ neue Kategorien für NP
- $3 * 2 * x$ Kategorien für V ($x =$ Anzahl der Valenzmuster)

Determinator-Nomen-Kongruenz

- Übereinstimmung in Genus (fem, mas, neu), Numerus (sg, pl) und Kasus (nom, gen, dat, acc)

- (46)
- a. der Mann, die Frau, das Buch (Genus)
 - b. das Buch, die Bücher (Numerus)
 - c. des Buches, dem Buch (Kasus)

- aus NP \rightarrow D, N wird

$NP_{3_sg_nom} \rightarrow D_{fem_sg_nom}, N_{fem_sg_nom}$	$NP_{gen} \rightarrow D_{fem_sg_gen}, N_{fem_sg_gen}$
$NP_{3_sg_nom} \rightarrow D_{mas_sg_nom}, N_{mas_sg_nom}$	$NP_{gen} \rightarrow D_{mas_sg_gen}, N_{mas_sg_gen}$
$NP_{3_sg_nom} \rightarrow D_{neu_sg_nom}, N_{neu_sg_nom}$	$NP_{gen} \rightarrow D_{neu_sg_gen}, N_{neu_sg_gen}$
$NP_{3_pl_nom} \rightarrow D_{fem_pl_nom}, N_{fem_pl_nom}$	$NP_{gen} \rightarrow D_{fem_pl_gen}, N_{fem_pl_gen}$
$NP_{3_pl_nom} \rightarrow D_{mas_pl_nom}, N_{mas_pl_nom}$	$NP_{gen} \rightarrow D_{mas_pl_gen}, N_{mas_pl_gen}$
$NP_{3_pl_nom} \rightarrow D_{neu_pl_nom}, N_{neu_pl_nom}$	$NP_{gen} \rightarrow D_{neu_pl_gen}, N_{neu_pl_gen}$
...	...
Dativ	Akkusativ

- 24 Symbole für Determinatoren, 24 Symbole für Nomen
- 24 Regeln statt einer

Probleme dieses Ansatzes

- Gernalisierungen werden nicht erfaßt.
- weder in Regeln noch in Categoriesymbolen
 - Wo kann eine NP oder NP_{nom} stehen?
Nicht wo kann eine NP_{3sgnom} stehen?
 - Gemeinsamkeiten der Regeln sind nicht offensichtlich.
- Lösung: Merkmale mit Werten und Identität von Werten
 Categoriesymbol: NP Merkmal: Per, Num, Kas, ...
 Wir erhalten z. B. die Regeln:
 $NP(3,sg,nom) \rightarrow D(fem,sg,nom), N(fem,sg,nom)$
 $NP(3,sg,nom) \rightarrow D(mas,sg,nom), N(mas,sg,nom)$

Merkmale und Regelschemata (I)

- Regeln mit speziellen Werten zu Regelschemata verallgemeinern:
 $NP(3,Num,Kas) \rightarrow D(Gen,Num,Kas), N(Gen,Num,Kas)$
- Gen-, Num- und Kas-Werte sind egal,
Hauptsache sie stimmen überein (identische Werte)
- Der Wert des Personenmerkmals (erste Stelle in NP(3,Num,Kas)) ist durch die Regel festgelegt: 3.

Merkmale und Regelschemata (II)

- Regeln mit speziellen Werten zu Regelschemata verallgemeinern:
 $NP(3, Num, Kas) \rightarrow D(Gen, Num, Kas), N(Gen, Num, Kas)$
 $S \rightarrow NP(Per1, Num1, nom),$
 $NP(Per2, Num2, dat),$
 $NP(Per3, Num3, akk),$
 $V(Per1, Num1)$
- Per1 und Num1 sind beim Verb und Subjekt gleich.
- Bei anderen NPen sind die Werte egal.
(Schreibweise für irrelevante Werte: '└')
- Die Kasus der NPen sind in der zweiten Regel festgelegt.

Übungsaufgaben

1. Schreiben Sie eine Phrasenstrukturgrammatik, mit der man u. a. die folgenden Sätze analysieren kann:
 - (47) a. Der Mann hilft der Frau.
 - b. Er gibt ihr das Buch.
 - c. Er wartet auf ein Wunder.
 - d. Er wartet neben dem Bushäuschen auf ein Wunder.

Head-Driven Phrase Structure Grammar (Hintergrund)

- in den 80ern als Nachfolger der Generalized Phrase Structure Grammar (GPSG) entwickelt
- wichtigste Veröffentlichungen Pollard und Sag (1987, 1994), über 1000 weitere Veröffentlichungen
 - Syntaxtheorie
 - Typologie
 - Computerlinguistik, Grammatikentwicklung (Deutsch, Englisch, Französisch, Norwegisch, Japanisch, Spanisch, Polnisch, Chinesisch, Persisch, Maltesisch, ...)
- seit 1994 jährliche Konferenzen:
Konferenzbände werden von CSLI online publications veröffentlicht

HPSG-Ressourcen

- Web pages:
<http://hpsg.stanford.edu/> und
<http://hpsg.fu-berlin.de/HPSG-Bib/> (Literatur)
- aktuelle HPSG-Lehrbücher:
Borsley, 1999; Sag, Wasow und Bender, 2003; Müller, 2008

TRALE

- TRALE = Controll + ALE
(Meurers, Penn und Richter, 2002)
- entwickelt in
 - Tübingen (Frank Richter)
 - Columbus/Ohio (Detmar Meurers)
 - Toronto (Gerald Penn)
 mit Hilfe von Holger Wunsch, Martin Lazarov (Tübingen), Stefan Müller (Berlin), Kordula De Kuthy (Columbus), Mohammad Haji-Abdolhosseini (Toronto), Vanessa Metcalf (Columbus)
- System kann Sätze analysieren und auch aus Bedeutungsrepräsentationen Sätze erzeugen.
- System kann semantische Repräsentationen nehmen und Äußerungen generieren.

Wieso TRALE?

Das einzige gegenwärtig unterstützte System zur Implementierung von HPSG-Analysen aus Publikationen:

- relationale Beschränkungen (append, shuffle, last)
- leere Elemente (Sag, Wasow und Bender, 2003, S. 464)
- komplexe Antecedentien in Typimplikationen
- zyklische Strukturen (Pollard und Sag, 1994; Soehn und Sailer, 2003)

Grammix

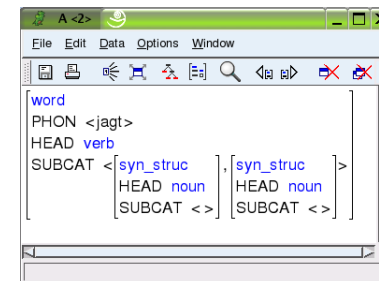
- bootbare CD-Rom
- Keine Installation nötig, Grammix-CD enthält alles, was man braucht:
 - vernünftiges Betriebssystem mit Unicode-Unterstützung, Tcl/TK, Java, PVM, ...
 - Trale (saved states, keine Prolog-Lizenz nötig)
 - DFKI Chart Display (Lizenz: Schreiben Sie Stephan Busemann)
 - [incr TSDB()] (+ LKB)
 - Utool
 - Babel
 - viele Beispielgrammatiken (Deutsch, Chinesisch, Maltesisch)
 - das Lehrbuch + Folien und Handouts
- hoffentlich plug & play
(Wenn der PC nicht zu neu ist ... ,
wenn der Mac nicht zu alt ist (Intel))

AVMs in Trale

```

jagt ~->
(word,
 head:verb,
 subcat:[ (head:noun, subcat:[]),
          (head:noun, subcat:[] ) ] ).

```



Repräsentation von Grammatikregeln in TRALE

- $\begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBCAT} & \boxed{2} \end{bmatrix} \rightarrow \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBCAT} & \langle \boxed{3} \rangle \oplus \boxed{2} \end{bmatrix}, \boxed{3} [\dots]$
- Beispiel:
 $V \rightarrow V, NP$

Information über Töchter

- Tochter-Information wird in Mutterstruktur mit repräsentiert:

$$\begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBCAT} & \boxed{2} \\ \text{DTRS} & \langle \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBCAT} & \langle \boxed{3} \rangle \oplus \boxed{2} \end{bmatrix}, \boxed{3} [\dots] \rangle \end{bmatrix}$$

Das ist für die Verarbeitung nicht nötig, wird aber zur Darstellung als Baum benötigt.

TRALE-Schreibweise

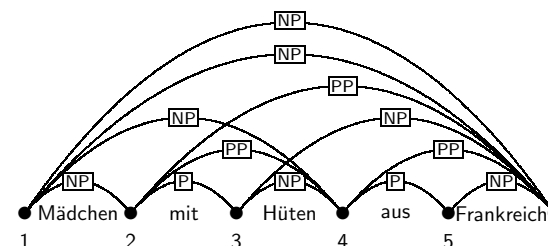
```
h_arg ## (head:Head,
          subcat:Subcat)
====>
cat> (head:Head,
      subcat:[NonheadDtr|Subcat]),
cat> NonheadDtr.
```

h_arg ist ein Bezeichner, der zur Identifikation der Regel dient.
 Er taucht z. B. auch im Chart-Display wieder auf.

Nach den ## folgt eine Beschreibung der Ergebniskategorie (linke Regelseite). Nach dem ==> folgen Paare der Form cat> Category, wobei Category eine Beschreibung eines Elements auf der rechten Regelseite ist.

Bottom-Up Chart Parsing

- Initialisierung der Chart: Nachschlagen aller Lexikoneinträge für die Wörter in der Eingabekette
- Parsen: Anwenden aller Regeln auf alle Tupel adjazenter Kanten in der Chart
- Hinzufügen einer neuen Kante für jede erfolgreiche Instantiierung einer Grammatikregel.



Einführung in die Grammatikentwicklung

└ Literatur

- Ajdkiewicz, Kasimir. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1, 1–27.
- Berman, Judith. 2003. *Clausal Syntax of German*. Studies in Constraint-Based Lexicalism, Stanford: CSLI Publications.
- Berman, Judith und Frank, Anette (Hrsg.). 1996. *Deutsche und französische Syntax im Formalismus der LFG*. Linguistische Arbeiten, Nr. 344, Tübingen: Max Niemeyer Verlag.
- Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs*. studia grammatica II, Berlin: Akademie Verlag.
- Bloomfield, Leonard. 1933. *Language*. London: George Allen and Unwin.
- Borsley, Robert D. 1999. *Syntactic Theory: A Unified Approach*. London: Edward Arnold, zweite Auflage.
- Bresnan, Joan (Hrsg.). 1982. *The Mental Representation of Grammatical Relations*. MIT Press Series on Cognitive Theory and Mental Representation, Cambridge: Massachusetts, London: England: The MIT Press.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford, UK/Cambridge, USA: Blackwell.
- Bußmann, Hadumod. 1983. *Lexikon der Sprachwissenschaft*. Kröners Taschenausgabe, Nr. 452, Stuttgart: Alfred Kröner Verlag.
- Bußmann, Hadumod. 1990. *Lexikon der Sprachwissenschaft*. Stuttgart: Alfred Kröner Verlag.
- Chomsky, Noam. 1957. *Syntactic Structures*. Janua Linguarum / Series Minor, Nr. 4, The Hague/Paris: Mouton.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.
- Chomsky, Noam. 1995. *The Minimalist Program*. Current Studies in Linguistics, Nr. 28, Cambridge: Massachusetts, London: England: The MIT Press.
- Dowty, David R. 1979. *Word Meaning and Montague Grammar*. Synthese Language Library, Nr. 7, Dordrecht, Boston, London: D. Reidel Publishing Company.
- Duden. 2005. *Grammatik der deutschen Gegenwartssprache*, Band 4. Mannheim, Leipzig, Wien, Zürich: Dudenverlag, 7. Auflage.
- Erms, Hans-Werner. 2000. *Syntax der deutschen Sprache*. de Gruyter Studienbuch, Berlin: Walter de Gruyter Verlag.
- Fillmore, Charles J., Kay, Paul und O'Connor, Mary Catherine. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*. *Language* 64(3), 501–538.
- Fischer, Kerstin und Stefanowitsch, Anatol (Hrsg.). 2006. *Konstruktionsgrammatik. Von der Anwendung zur Theorie*. Stauffenburg Linguistik, Nr. 40, Tübingen: Stauffenburg Verlag.
- Gazdar, Gerald, Klein, Ewan, Pullum, Geoffrey K. und Sag, Ivan A. 1985. *Generalized Phrase Structure Grammar*. Cambridge, Massachusetts: Harvard University Press.
- Goldberg, Adele E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Cognitive

Einführung in die Grammatikentwicklung

└ Literatur

- Theory of Language and Culture, Chicago/London: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at Work. The Nature of Generalization in Language*. Oxford Linguistics, Oxford, New York: Oxford University Press.
- Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax. Eine Rektions-Bindungs-Analyse*. Studien zur deutschen Grammatik, Nr. 33, Tübingen: original Gunter Narr Verlag jetzt Stauffenburg Verlag.
- Grewendorf, Günther. 2002. *Minimalistische Syntax*. UTB für Wissenschaft: Uni-Taschenbücher, Nr. 2313, Tübingen, Basel: A. Francke Verlag GmbH.
- Heringer, Hans-Jürgen. 1996. *Deutsche Syntax dependentiell*. Stauffenburg Linguistik, Tübingen: Stauffenburg Verlag.
- Joshi, Aravind K. 1987. Introduction to Tree Adjoining Grammar. In Alexis Manaster-Ramer (Hrsg.), *The Mathematics of Language*, Seiten 87–114, Amsterdam, Philadelphia: John Benjamins Publishing Co.
- Joshi, Aravind K, Levy, Leon S. und Takahashi, Masako. 1975. Tree Adjunct Grammar. *Journal of Computer and System Science* 10(2), 136–163.
- Kroch, Anthony und Joshi, Aravind. 1985. The Linguistic Relevance of Tree Adjoining Grammar. Technischer Bericht MS-CIS-85-16, University of Pennsylvania. <http://babel.ling.upenn.edu/papers/faculty/tony.kroch/papers/relevance3.pdf>, 21.03.2008.
- Kunze, Jürgen. 1975. *Abhängigkeitsgrammatik*. Studia Grammatica XII, Berlin: Akademie Verlag.
- Meurers, Walt Detmar, Penn, Gerald und Richter, Frank. 2002. A Web-Based Instructional Platform for Constraint-Based Grammar Formalisms and Parsing. In Dragomir Radev und Chris Brew (Hrsg.), *Effective Tools and Methodologies for Teaching NLP and CL*, Seiten 18–25, proceedings of the Workshop held at 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, PA. <http://www.ling.ohio-state.edu/~dm/papers/acl02.html>, 08.01.2004.
- Müller, Stefan. 1999. *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Linguistische Arbeiten, Nr. 394, Tübingen: Max Niemeyer Verlag. <http://hpsg.fu-berlin.de/~stefan/Pub/hpsg.html>, 15.04.2009.
- Müller, Stefan. 2002. *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German*. Studies in Constraint-Based Lexicalism, Nr. 13, Stanford: CSLI Publications. <http://hpsg.fu-berlin.de/~stefan/Pub/complex.html>, 15.04.2009.
- Müller, Stefan. 2003. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1), 29–62. <http://hpsg.fu-berlin.de/~stefan/Pub/mehr-vf-ds.html>, 15.04.2009.
- Müller, Stefan. 2004. Continuous or Discontinuous Constituents? A Comparison between Syntactic Analyses for Constituent Order and Their Processing Systems. *Research on Language and Computation, Special Issue on Linguistic Theory and Grammar Implementation 2(2)*. <http://hpsg.fu-berlin.de/~stefan/Pub/discont.html>, 15.04.2009.

Einführung in die Grammatikentwicklung

└ Literatur

- Müller, Stefan. 2005. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203, 297–330. <http://hpsg.fu-berlin.de/~stefan/Pub/mehr-vf-lb.html>, 15.04.2009.
- Müller, Stefan. 2008. *Head-Driven Phrase Structure Grammar: Eine Einführung*. Stauffenburg Einführungen, Nr. 17, Tübingen: Stauffenburg Verlag, zweite Auflage. <http://hpsg.fu-berlin.de/~stefan/Pub/hpsg-lehrbuch.html>, 15.04.2009.
- Pollard, Carl J. und Sag, Ivan A. 1987. *Information-Based Syntax and Semantics*. CSLI Lecture Notes, Nr. 13, Stanford: CSLI Publications.
- Pollard, Carl J. und Sag, Ivan A. 1994. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics, Chicago, London: University of Chicago Press.
- Sag, Ivan A., Wasow, Thomas und Bender, Emily M. 2003. *Syntactic Theory: A Formal Introduction*. CSLI Lecture Notes, Nr. 152, Stanford: CSLI Publications, zweite Auflage.
- Saussure, Ferdinand de. 1916. *Grundfragen der allgemeinen Sprachwissenschaft*. Berlin: Walter de Gruyter & Co, 2. Auflage 1967.
- Soehn, Jan-Philipp und Sailer, Manfred. 2003. At First Blush
- on Tenterhooks. About Selectional Restrictions Imposed by Nonheads. In Gerhard Jäger, Paola Monachesi, Gerald Penn und Shuly Wintner (Hrsg.), *Proceedings of Formal Grammar 2003, Vienna, Austria*, Seiten 149–161. <http://www.sfs.uni-tuebingen.de/hpsg/archive/bibliography/papers/soehnsailer03.ps.gz>, 30.05.2004.
- Steedman, Mark. 2000. *The Syntactic Process*. Language, Speech, and Communication, Cambridge: Massachusetts, London: England: The MIT Press.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Librairie C. Klincksieck.
- Tesnière, Lucien. 1980. *Grundzüge der strukturalen Syntax*. Stuttgart: Klett-Cotta.
- Uszkoreit, Hans. 1987. *Word Order and Constituent Structure in German*. CSLI Lecture Notes, Nr. 8, Stanford: CSLI Publications.
- von Stechow, Arnim und Sternefeld, Wolfgang. 1988. *Bausteine syntaktischen Wissens. Ein Lehrbuch der Generativen Grammatik*. Opladen/Wiesbaden: Westdeutscher Verlag.
- Weber, Heinz J. 1992. *Dependenzgrammatik. Ein Arbeitsbuch*. Narr Studienbücher, Tübingen: Gunter Narr Verlag.