

Syntax und Morphologie

Einführungskurs
8. Vorlesung



Strukturanalyse

- Aufgabe der **syntaktisch-funktionalen** Analyse ist es, alle Informationen bereitzustellen, die es der **semantischen** Analyse ermöglichen, ihre zentrale Aufgabe wahrzunehmen.
 - ▶ zwei Aspekte sind maßgebend: **Konstituenz** und **Dependenz**
 - ▶ beide sind einander **ergänzend** zu sehen
- (nicht-)konfigurationale Sprachen
 - ▶ besonders in sog. **nicht-konfiguralen** Sprachen, die eine relativ freie Wortstellung zulassen, kommt der Gewinnung von Dependenzinformation durch die Strukturanalyse besondere Bedeutung zu,
 - ▶ während in **konfiguralen** Sprachen wie dem Englischen sich diese Information in der Regel direkt aus der Wortstellung gewinnen läßt.



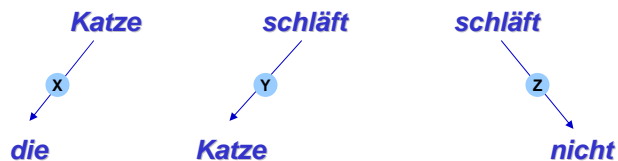
Konstituenz und Dependenz

- Dabei handelt es sich um zwei wesentliche **Grundprinzipien der Satzorganisation**, die sich gegenseitig bedingen bzw. als sich gegenseitig ergänzend zu betrachten sind.
- Das Prinzip der **Konstituenz** beruht auf der Teil-Ganzes-Beziehung zwischen Satzbestandteilen (z.B. "ein Satz besteht aus Subjekt und Prädikat").
- Das Prinzip der **Dependenz** oder Abhängigkeit beruht auf der Tatsache, dass zwischen Ausdrücken im Satz mehr oder weniger enge Beziehungen bestehen, so z.B. zwischen Attribut und Bezugswort (ein *schlauder Fuchs*) oder zwischen dem Verb und seinen Ergänzungen (Er *begegnete einem Zombie*).

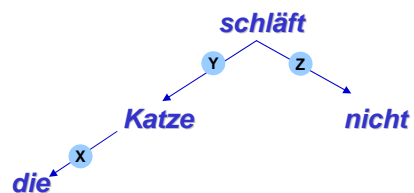


„Die Katze schläft nicht.“ (1)

- einzelne Dependenz (graphisch)

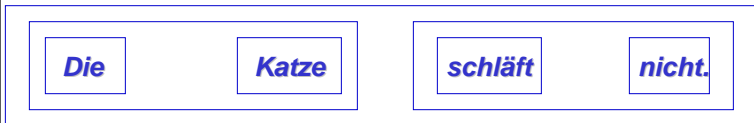


- Dependenz zusammengesetzt → **Dependenzbaum**



„Die Katze schläft nicht.“ (2)

■ Konstituenten



■ „Ausziehen“ → **Konstituentenbaum**



Dependenzsyntax vs. Konstituentensyntax

Dependenz-Relation

- geht über das Kriterium der bloßen Form hinaus
- Abhängigkeitsverhältnis sprachlicher Elemente
- aus dem Vorkommen eines Elements kann auf das Vorkommen anderer Elemente geschlossen werden

Dependenzgrammatik

- Satz: ein durch die Abhängigkeitsrelation festgelegtes, hierarchisch geordnetes Ganzes
- Ziel der Beschreibung: hinter der linearen Anordnung der Elemente eines Satzes eine hierarchische Struktur sichtbar zu machen
- dem Verb steht eine zentrale Position zu

Teil-Ganzes-Relation

- Aufeinanderfolge und grammatische Nachbarschaft
- Konstituenten: die einzelnen ausgezeichneten Teile einer derart zusammengesetzten Form

Phrasenstrukturgrammatik

- definiert ein System von komplexen Kategorien
- analysiert syntaktische Strukturen durch Segmentierung in kleinere Einheiten, die durch ihre Stellung im Ganzen zu erklären sind
- Rekursivität: das System der Kategorien beruht ausschließlich auf der Form immer wiederkehrender Muster
- Ökonomieprinzip: Minimalität und Redundanzfreiheit

Konstituentengrammatik ↔ Dependenzgrammatik

- Zwei konkurrierende Ansätze (?)
- Beide postulieren **Baumstrukturen** (um bestimmte Abhängigkeiten auszudrücken).
- Beide verwenden (heute) zur Ergänzung **Merkmalsstrukturen**.
- In vielerlei Hinsicht nähern sich beide Ansätze einander an.
- Bis zu einem gewissen Grad sind sie ineinander überführbar (~ schwach äquivalent).

Syntaktische Strukturierung durch Konstituenz

Ein Läufer erreichte das Ziel nach zwei Stunden

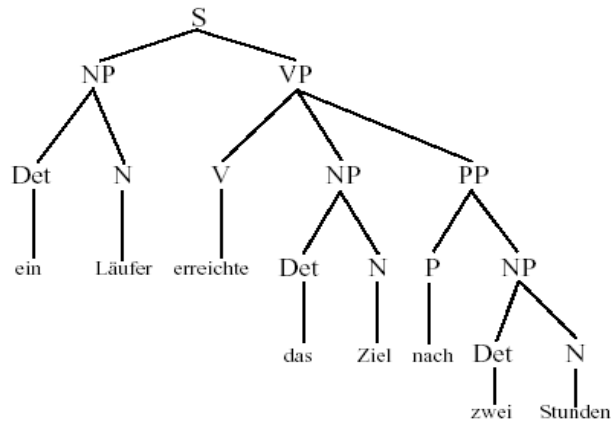
Ein Läufer erreichte das Ziel nach zwei Stunden

Ein Läufer erreichte das Ziel nach zwei Stunden

Ein Läufer erreichte das Ziel nach zwei Stunden

Ein Läufer erreichte das Ziel nach zwei Stunden

Syntaktische Strukturierung durch Konstituenz



PSG: Phrasenstrukturgrammatik

- **Phrasenstrukturregeln** sind Regeln, die unmittelbare Konstituenten von Ketten und deren Zugehörigkeit zu grammatischen Kategorien festlegen.
- Ihre allgemeine Form ist $A \rightarrow B$, wobei **A** (der **Regelkopf**) eine grammatische Kategorie ist und **B** (der **Regelkörper**) eine Kette aus einer oder mehreren grammatischen Kategorien darstellt.
- Ist **B** ein **lexikalisches** Element, drückt $A \rightarrow B$ eine *ist-ein*-Relation aus (**B ist ein A**)

PSG: Phrasenstrukturgrammatik

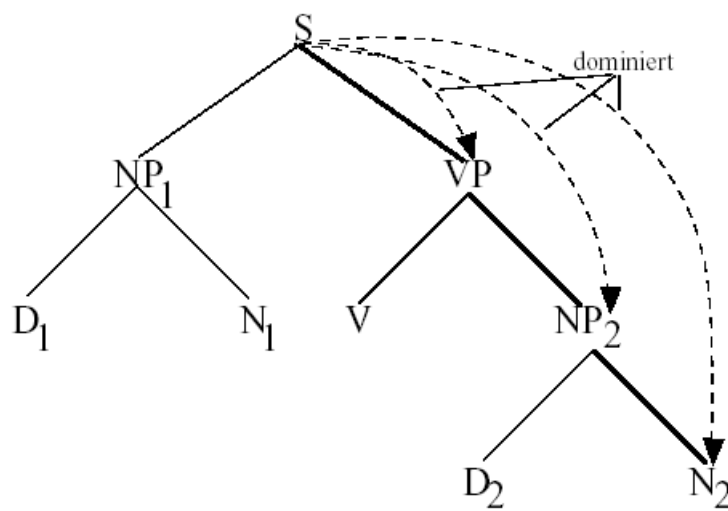
Die Menge der Phrasenstruktur-Regeln einer Sprache ist eine **PHRASENSTRUKTUR-GRAMMATIK** dieser Sprache.

Die **STRUKTURBESCHREIBUNG** eines Satzes ist eine Spezifikation seiner Konstituentenstruktur unter Angabe der Zugehörigkeit zu grammatischen Kategorien.

Die Menge der Regeln, die explizit festlegen, welche Wortketten Sätze einer Sprache sind und gleichzeitig jedem Satz eine Strukturbeschreibung zuweisen, ist eine **GENERATIVE GRAMMATIK** dieser Sprache.



PSG: Dominanzrelation



PSG: Dominanz und Präzedenz

Gegeben sei eine Regel $X \rightarrow Y_1 \dots Y_n$, wobei X und Y_i Kategorien sind. Dann gilt: X dominiert unmittelbar jedes Y_i . Die Regel $S \rightarrow NP VP$ spezifiziert, daß S in dem entsprechenden Phrasenmarker unmittelbar **NP** und **VP** dominiert.

DOMINIEREN ist eine binäre Relation zwischen zwei Kategorien X und Y mit folgenden Eigenschaften:

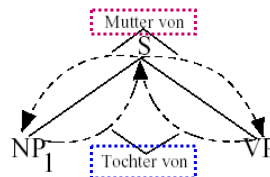
- (1) X dominiert Y , wenn X das Y unmittelbar dominiert.
- (2) X dominiert Y , wenn es ein Z gibt derart, daß X unmittelbar Z dominiert, und Z Y dominiert.

Eine PS-Regel der Form $A \rightarrow B C$ sagt auch aus, daß in der Kette A die Teilkette B der Teilkette C vorausgeht.

PSG: lokale Strukturen

Ein Knoten ist die **MUTTER** aller Knoten, die er direkt dominiert.

Knoten, die direkt von einem anderen Knoten dominiert werden, sind **TÖCHTER** dieses Knotens. Die Tochterrelation ist die Umkehrung der Mutterrelation, d.h. es gilt für zwei beliebige Knoten X und Y : wenn $Mutter(X, Y)$ dann $Tochter(Y, X)$

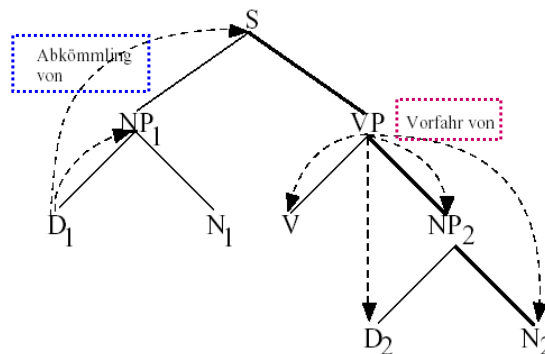


Knoten, die Töchter des gleichen Knotens (der gleichen Mutter) sind, sind **SCHWESTERN** von einander (sind GESCHWISTER).

PSG: lokale & nicht-lokale Strukturen

Ein Knoten ist **VORFAHR** aller Knoten, die er dominiert.

Knoten, die von einem anderen Knoten dominiert werden, sind **ABKÖMMLINGE** dieses Knotens. Die Abkömmling-Beziehung ist die Umkehrung der Vorfahr-Relation, d.h. es gilt für zwei beliebige Knoten X und Y: wenn *Vorfahr*(X, Y) dann *Abkömmling*(Y, X).



PSG: Subkategorisierung / Valenz

Subkategorisierung nennt man die **Untergliederung der lexikalischen Kategorien** (Nomen, Verb, etc.) in syntaktisch-semantisch motivierte Subkategorien um bestimmten Abhängigkeitsbeziehungen im Satz Rechnung zu tragen.

Valenz ist ein aus der Chemie entlehnter Begriff und bezeichnet die Fähigkeit eines Lexems (z.B. eines Verbs, Adjektivs, Nomens), seine syntaktischen Umgebungen im Satz vorzustrukturieren, indem es anderen Konstituenten Bedingungen bezüglich ihrer **Anzahl** und ihrer **grammatischen Eigenschaften** (syntaktische Kategorie, sekundäre grammatische Kategorien wie Genus, Kasus etc.) auferlegt.

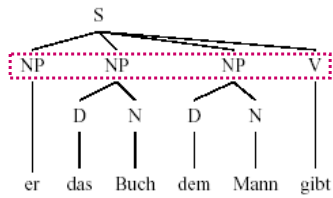
STRIKTE SUBKATEGORISIERUNG bedeutet **strikt lokale**, d.h. auf die Konstituenten der fraglichen Kategorie **innerhalb einer Konstituente** beschränkte, Subkategorisierung.

PSG: flache vs. binärverzweigte Darstellung

Struktur

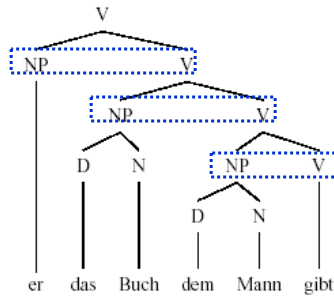
“flat”

“binary branching”



NP → D, N

S → NP, NP, NP, V



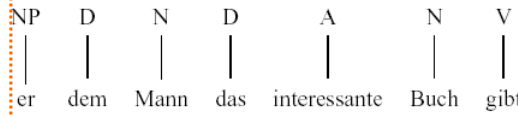
NP → D, N

V → NP, V

PSG: Beispielableitung bei Annahme flacher Strukturen

- S → V NP NP NP
- NP → D N
- N → A N
- N → N PP

- D → das
- D → dem
- A → interessante
- A → kluges
- NP → er
- N → Mann
- N → Buch
- V → gibt
- V → schläft
- V → hat



PSG: Beispielableitung bei Annahme flacher Strukturen

S → V NP NP NP

NP → D N

N → A N

N → N PP

D → das

D → dem

A → interessante

A → kluges

NP → er

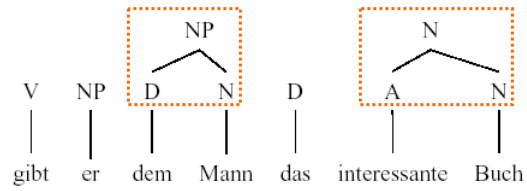
N → Mann

N → Buch

V → gibt

V → schläft

V → hat



PSG: Beispielableitung bei Annahme flacher Strukturen

S → V NP NP NP

NP → D N

N → A N

N → N PP

D → das

D → dem

A → interessante

A → kluges

NP → er

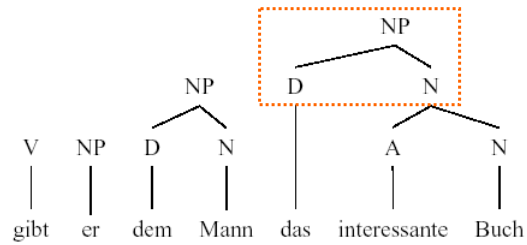
N → Mann

N → Buch

V → gibt

V → schläft

V → hat



PSG: Beispielableitung bei Annahme flacher Strukturen

S → V NP NP NP

NP → D N

N → A N

N → N PP

D → das

D → dem

A → interessante

A → kluges

NP → er

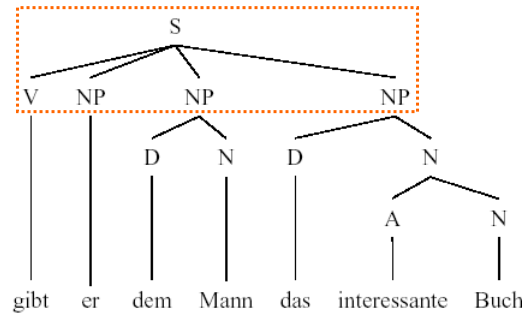
N → Mann

N → Buch

V → gibt

V → schläft

V → hat



PSG: Erweiterungen

Köpfe

Kopf bestimmt die wichtigsten Eigenschaften einer Projektion

- Karl **schläft**.
- Karl **liebt** Maria.
- in** diesem Haus
- ein **Mann**

ein Satz ist die Maximalprojektion eines finiten Verbs

Hauptkategorien sind:

Kategorie	projizierte Merkmale	Beispiel
Verb	Kategorie, Finitheitsmerkmale	hat, schläft, lieben
Nomen	Kategorie, Kasus, Numerus, Genus	Mann, Blume, Kind
Präposition	Kategorie, Form der Präposition	an, auf, in, neben
Adjektiv	Kategorie, Kasus, Numerus, Genus	klug, schön, faul

PSG: Erweiterungen

Abstraktion über Regeln: \bar{X} -Schemata

\bar{X} -Theorie (Jackendoff, 1977):

\bar{X} -Regel

mit Kategorien

$\bar{X} \rightarrow \overline{\text{Spezifikator}} \bar{X}$

$\bar{N} \rightarrow \overline{\text{DET}} \bar{N}$

$\bar{X} \rightarrow \bar{X} \overline{\text{Adjunkt}}$

$\bar{N} \rightarrow \bar{N} \overline{\text{REL_SATZ}}$

$\bar{X} \rightarrow \overline{\text{Adjunkt}} \bar{X}$

$\bar{N} \rightarrow \overline{\text{ADJ}} \bar{N}$

$\bar{X} \rightarrow X \overline{\text{Komplement}^*}$

$\bar{N} \rightarrow N \bar{P}$

X steht für beliebige Kategorie, '*' für beliebig viele Wiederholungen

PSG: Erweiterungen

\bar{X} -Schemata

$X(2) \rightarrow \text{Spezifikator}(2) X(1)$

$X(1) \rightarrow X(1) \text{Adjunkt}(2)$

$X(1) \rightarrow \text{Adjunkt}(2) X(1)$

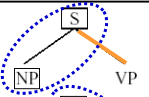
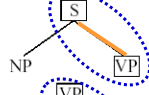
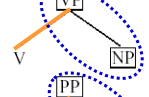
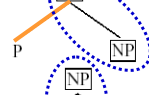
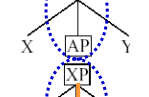
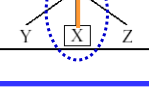
$X(1) \rightarrow X(0) \text{Komplement}(2) *$

- Bar-Level des X auf der rechten Seite ist gleich oder kleiner als das der linken
- bei Regeln mit Merkmalen (num, per, o. ä.):
Werte beim X links und rechts gleich
- X ist der Kopf, die Merkmale sind Kopfmerkmale
- Adjunktion verändert das Level nicht → Iteration möglich

PSG: Kategorien

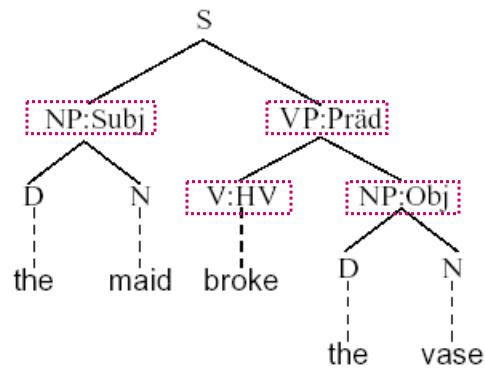
	GRAMMATISCHE KATEGORIEN	KATEGORIALSYMBOL
SYNTAKTISCHE KATEGORIEN	SATZ	S
	NOMINALPHRASE	NP
	VERBALPHRASE	VP
	ADJEKTIVPHRASE	AP
	PRÄPOSITIONALPHRASE	PP
	⋮	
LEXIKALISCHE KATEGORIEN	DETERMINATOR	D
	NOMEN	N
	VERB	V
	ADJEKTIV	A
	PRÄPOSITION	P
	KONJUNKTION	K
	⋮	

PSG: Funktionen

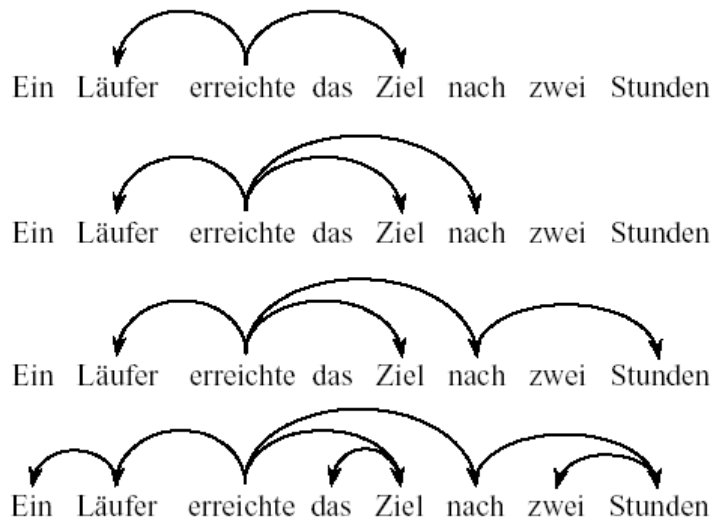
Funktion	Funktionsname	Struktur
(NP, S)	Subjekt (von)	
(VP, S)	Prädikat (von)	
(NP, VP)	Objekt (von)	
(NP, PP)	Objekt (von)	
(AP, NP)	Attribut (von)	
(X, XP)	Kopf (von)	

PSG: Funktionen

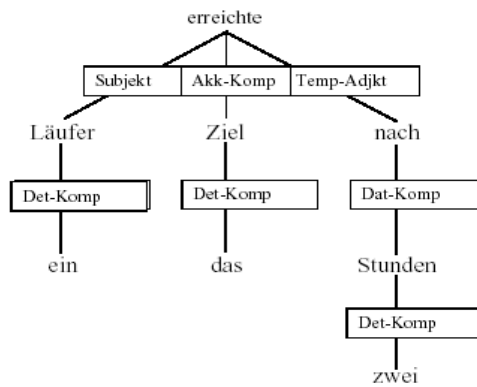
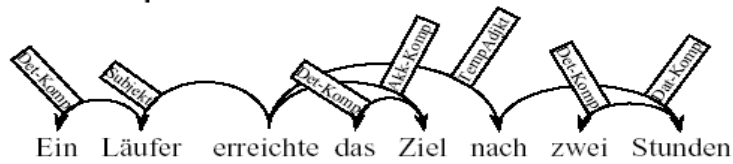
In den Phrasenmarkern, welche die Konstituentenstruktur von Sätzen beschreiben, könnten wir die relevanten **grammatischen Funktionen** durch zusätzliche Annotationen in der Form **KATEGORIE:FUNKTION** anzeigen:



Syntaktische Strukturierung durch Dependenz



Syntaktische Strukturierung durch Dependenz



Wie bestimmt man Dependenz?

Wenn wir das analytische Verfahren der Tilgung (Weglaßprobe) auf bestimmte Beispiele anwenden, stellen wir fest, daß gewisse Wörter nur gleichzeitig mit anderen weggelassen werden können. Zum Beispiel können wir in dem Satz

The students solved extremely difficult problems

das Adjektiv *difficult* nur weglassen, wenn wir auch das Adverb *extremely* tilgen:

- * The students solved *extremely* ~~difficult~~ problems
- The students solved ~~extremely~~ *difficult* problems
- The students solved ~~extremely~~ *difficult* problems
- * The students solved (*extremely*) *difficult* ~~problems~~

Andersherum setzt das Adverb *extremely* die Anwesenheit des Adjektivs *difficult* voraus. Wir werden sagen, daß *extremely* vom Adjektiv *difficult* ABHÄNGIG ist. Das Adjektiv ist seinerseits vom Nomen *problems* abhängig. Diese Beziehung der Abhängigkeit wird DEPENDENZ

Wie bestimmt man Abhängigkeiten?

DEPENDENZ ist eine zweistellige Relation zwischen zwei Wörtern w_1 und w_2 in einer Kette, wobei das Vorkommen oder die Form oder allgemein das grammatische Verhalten von w_1 durch w_2 kontrolliert wird.

Ist w_1 von w_2 abhängig, dann nennt man w_1 das **DEPENDENS**.

REKTION ist die Umkehrung der Abhängigkeitsrelation. Gegeben seien zwei Elemente w_1 und w_2 : wenn w_2 von w_1 abhängig ist, dann **REGIERT** w_1 das w_2 .

Das kontrollierende Element in einer Abhängigkeitsrelation soll **REGENS** genannt werden.



Formalisierung des Abhängigkeitsbegriffs

- Abhängigkeitsrelationen können so formalisiert werden,

- ▶ dass ein Regens mehrere Dependents regiert

$$R \xrightarrow{dep} \{D_1, \dots, D_n\}$$

- ▶ oder dass eine iterative Anordnung von Abhängigkeiten ausgedrückt wird.

$$R \xrightarrow{dep} D_1 \xrightarrow{dep} D_2$$

- Diese Relationen müssen bestimmte Eigenschaften erfüllen.

- ▶ **Antisymmetrie**: ein von einem Regens R abhängiger Dependent D kann nicht selbst Regent von R sein:

$$R \xrightarrow{dep} D \Rightarrow \neg (D \xrightarrow{dep} R)$$

- ▶ **Antireflexivität**: ist ein Dependent D von einem Regens R abhängig, so können beide nicht gleich sein:

$$R \xrightarrow{dep} D \Rightarrow \neg (R = D)$$

- ▶ **Antitransitivität**: ist ein Dependent D_1 abhängig von einem Regens R und regiert selbst einen Dependenten D_2 , dann regiert der Regens R nicht den Dependenten D_2

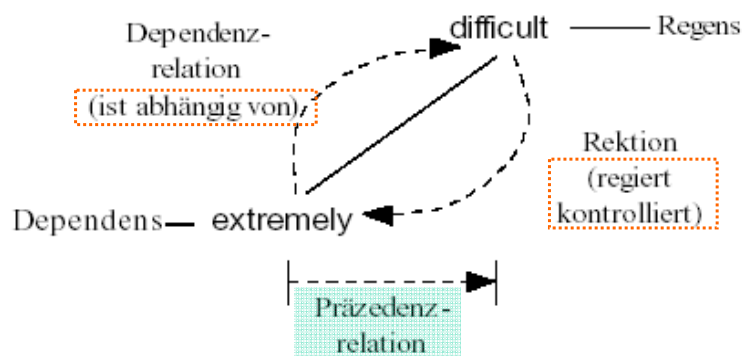
$$(R \xrightarrow{dep} D_1) \wedge (D_1 \xrightarrow{dep} D_2) \Rightarrow \neg (R \xrightarrow{dep} D_2)$$



DG: Dependenzgrammatik

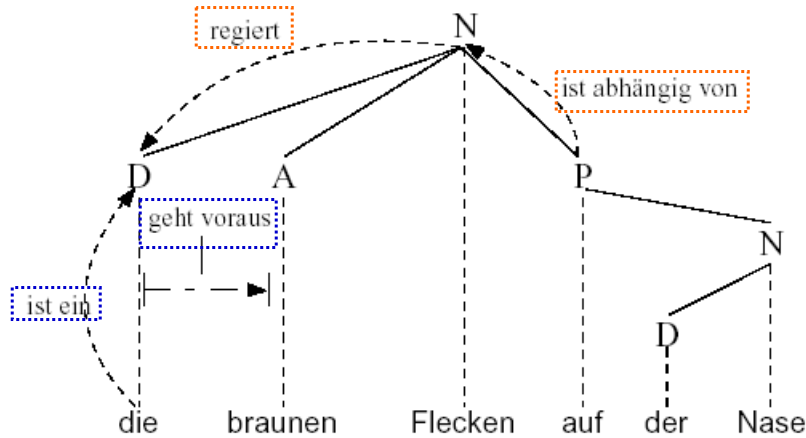
- In einer **Dependenzbeziehung** kann ein **Regens**, ein oder mehrere **Dependents** regieren, wohingegen ein **Dependens** immer nur von genau einem **Regens** regiert wird.
 - ▶ Die Dependents in einer Dependenzrelation können wiederum Regenten anderer Dependenzrelationen sein.
 - ▶ Sie können so **direkte** Dependents ihres Regens und **indirekte** Dependents weiterer "übergeordneter" Regenten sein.
- Ein **Satellit** eines Wortes wird von einem seiner direkten Dependents und allen von diesem regierten Dependents gebildet.

DG: Dependenz und Präzedenz



Verallgemeinerte Dependenzstruktur

Ein Regens und alle von ihm kontrollierten abhängigen Elemente bilden zusammen eine Konstituente.



Verallgemeinerte Dependenzstruktur

Eine regierende lexikalische Kategorie X zusammen mit all ihren Abhängigen konstituiert eine X-PHRASE; abgekürzt XP.

Die regierende lexikalische Kategorie X einer X-Phrase ist der (lexikalische) Kopf dieser Phrase.

Eine regierende lexikalische Kategorie N zusammen mit all ihren Abhängigen konstituiert eine N-Phrase, abgekürzt NP, genannt Nominalphrase; Das N ist der Kopf dieser Phrase.

Eine regierende lexikalische Kategorie V zusammen mit all ihren Abhängigen konstituiert eine V-Phrase, abgekürzt VP, genannt Verbalphrase; Das V ist der Kopf dieser Phrase.

Verallgemeinerte Dependenzstruktur

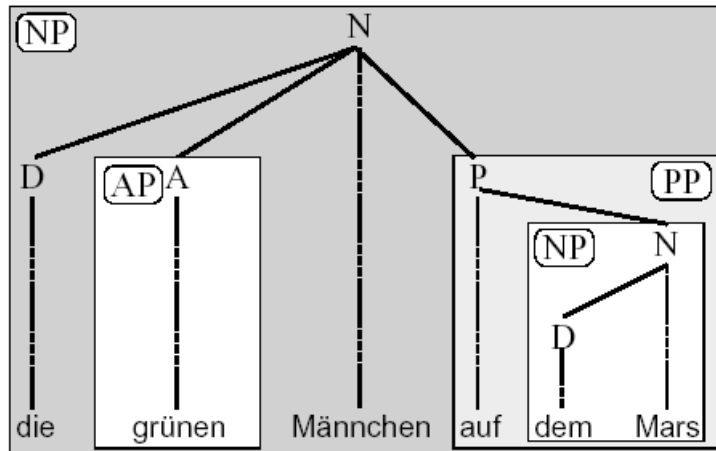
Eine regierende lexikalische Kategorie A zusammen mit all ihren Abhängigen konstituiert eine A-Phrase, abgekürzt AP, genannt: Adjektivphrase. Das A ist der Kopf dieser Phrase.

Eine regierende lexikalische Kategorie P zusammen mit all ihren Abhängigen konstituiert eine P-Phrase, abgekürzt PP, genannt: Präpositionalphrase. Das P ist der Kopf dieser Phrase.

Wichtig: Alle X-Phrasen können auch aus dem Kopf alleine bestehen. Ein einzelnes Wort bildet eine X-Phrase, wenn es die gleiche syntaktische Position wie die entsprechende X-Phrase einnimmt, d.h. mit einer X-Phrase kommutiert. Das Adjektiv in *a tall boy* kommutiert mit *very tall*: *a very tall boy*. Das Wort *tall* ist also einerseits ein Element der lexikalischen Kategorie Adjektiv: $[\text{tall}]_A$, es ist andererseits der Kopf einer Adjektivphrase, die keine weiteren Abhängigen hat: $[[\text{tall}]_A]_{AP}$.

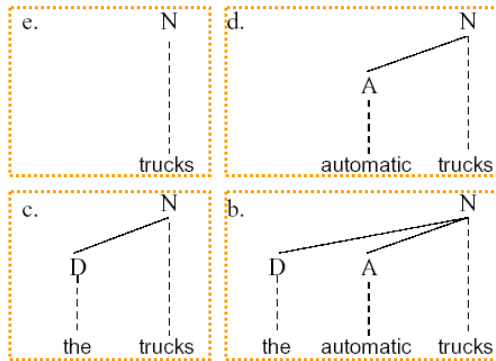


Dependenz und Konstituenz



Beispiel 1: Dependenz und Konstituenz (NP)

a. The automatic trucks from the factory	carry coal up the sharp incline
b. The automatic trucks	
c. The trucks	
d. Automatic trucks	
e. Trucks	



Beispiel 1: Dependenz und Konstituenz (NP)

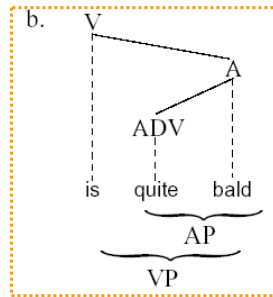
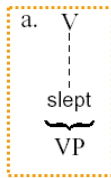
All diese Syntagmen können durch Anfügen der Kette *from the factory* (einer PP) erweitert werden. Die strukturellen Möglichkeiten einer Nominalphrase lassen sich wie folgt tabellarisch darstellen, wobei fakultative Elemente eingeklammert sind:

NP					
(D)	(AP)		N	(PP)	
	(Adv)	A		P	NP
					(D)

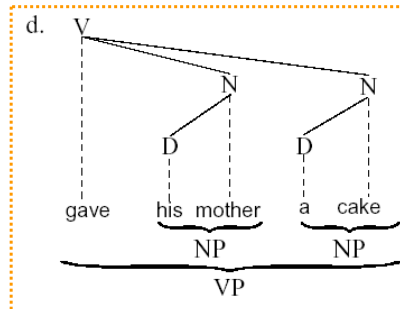
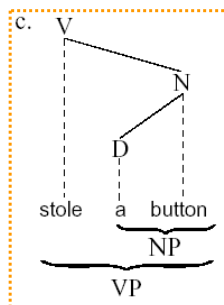


Beispiel 2: Dependenz und Konstituenz (VP)

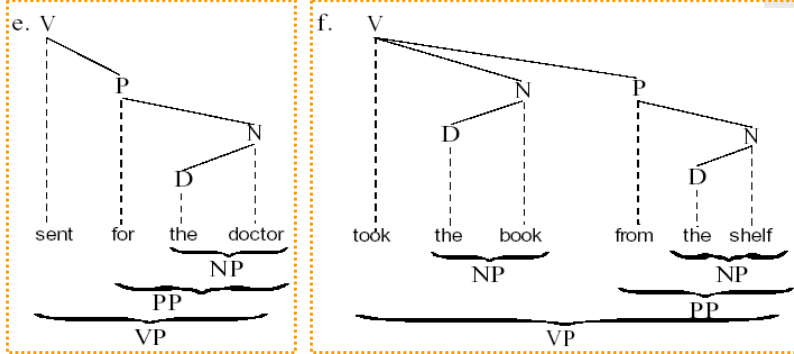
The policeman	a. slept
	b. is quite bald
	c. stole a button
	d. gave his mother a cake
	e. sent for the doctor
	f. took the book from the shelf



Beispiel 2: Dependenz und Konstituenz (VP)



Beispiel 2: Dependenz und Konstituenz (VP)



Interdependenz

Eine besondere Situation liegt bei der **Beziehung zwischen Subjekt und Hauptverb** eines Satzes vor. Man kann zum Beispiel argumentieren, daß hinsichtlich bestimmter grammatischer Eigenschaften (z.B. NUMERUS, PERSON) das Subjekt eines Satzes die Form des finiten Verbs kontrolliert:

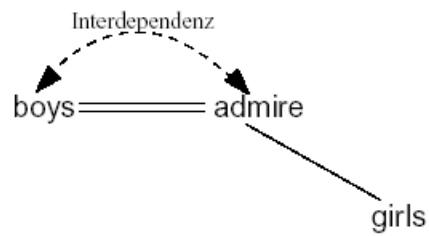
Ich singe	(1. Person Singular)
Du singst	(2. Person Singular)
Er singt	(3. Person Singular)
Sie singen	(3. Person Plural)

In dieser Hinsicht kann man sagen, daß das **Verb vom Subjekt abhängt**. Es gibt jedoch andere Merkmale, wo die Annahme plausibler ist, daß das **Verb das Subjekt kontrolliert**. So verlangen z.B. bestimmte Verben, daß das Nomen des Subjekts zu einer besonderen semantischen Klasse gehört, z.B. die Klasse der Nomina, die Lebewesen bezeichnen, die intentional handeln können.

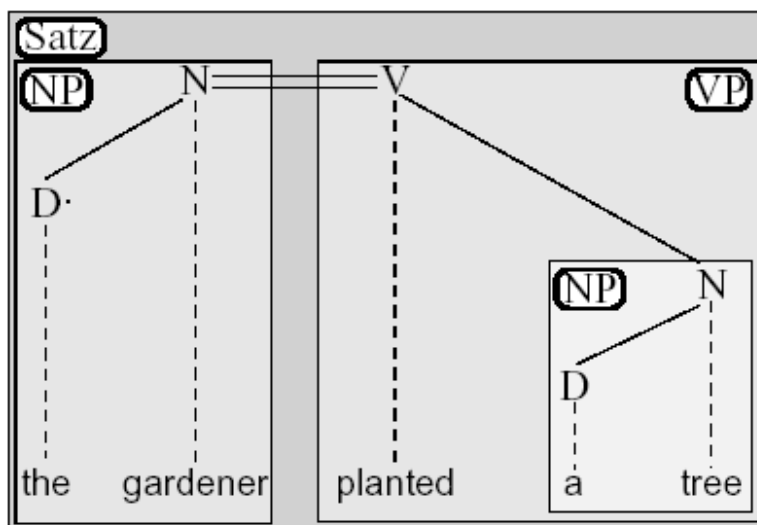
- (a) The boy admires sincerity
- (b) * Sincerity admires the boy
- (c) The mother admires the boy

Interdependenz

Wenn bei zwei Elementen w_1 und w_2 sowohl w_1 von w_2 als auch w_2 von w_1 abhängig ist, dann sind w_1 und w_2 **INTERDEPENDENT**.



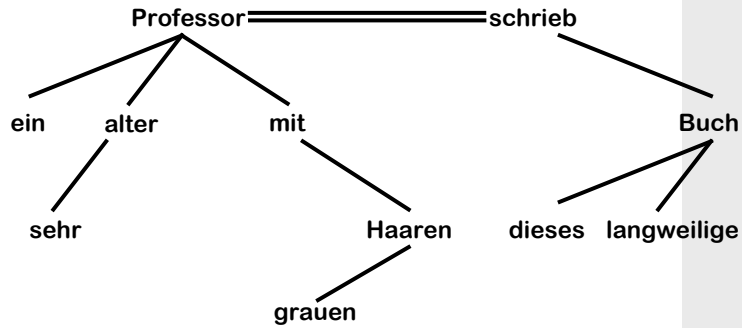
Dependenz und Konstituenz (Satz)



Dependenzanalyse

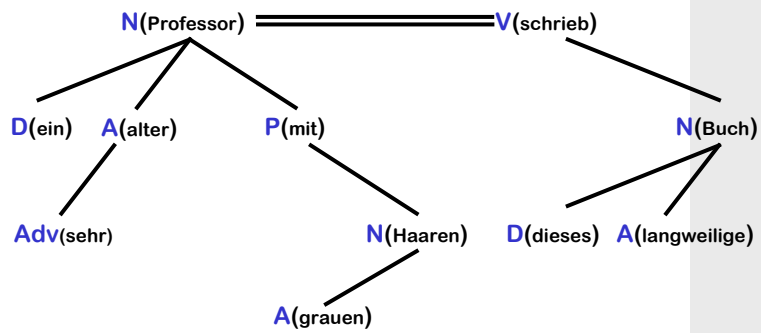
Ein sehr alter Professor mit grauen Haaren schrieb dieses langweilige Buch.

Professor schrieb Buch

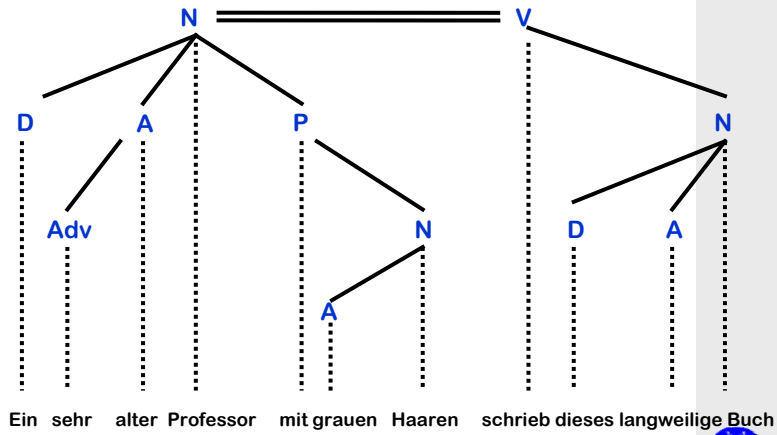


Dependenzanalyse

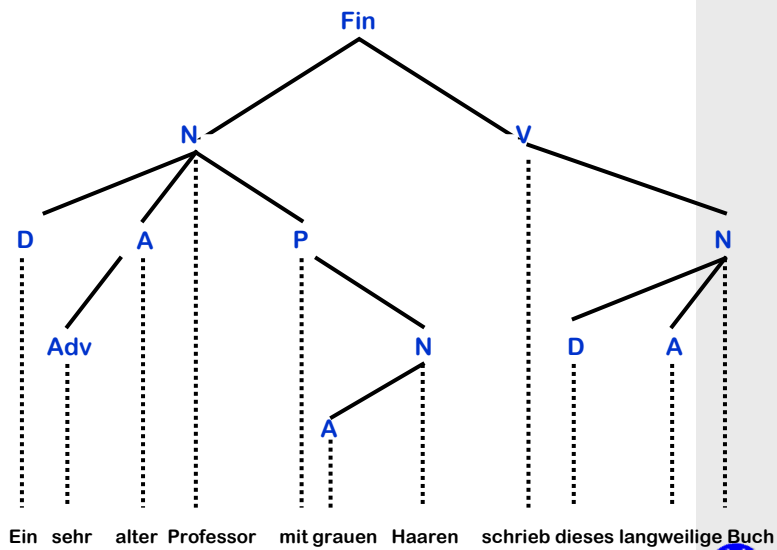
Ein sehr alter Professor mit grauen Haaren schrieb dieses langweilige Buch.



Dependenzanalyse



Dependenzanalyse

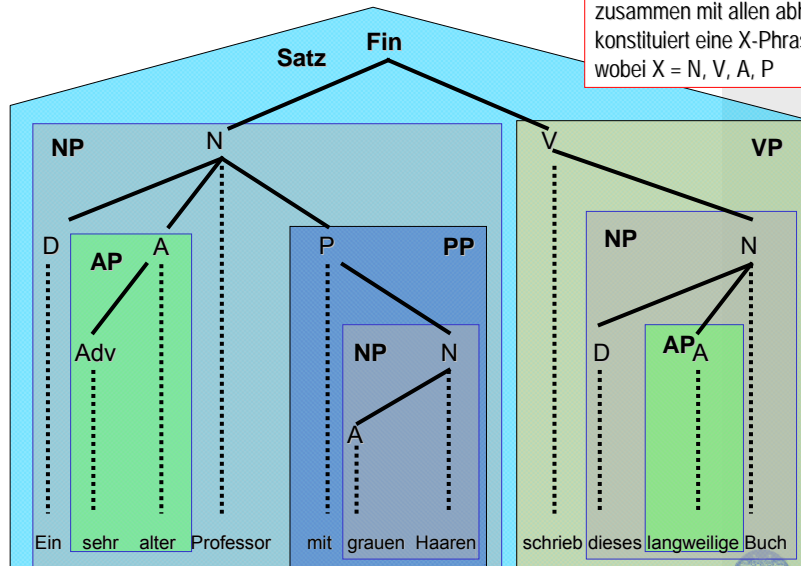


Dependenzregeln

1. Fin → N * V
2. N → D (A) * (P)
3. V → * (N) (N) (P)
4. P → * N
5. A → Adv *

- | | | |
|---------|---|-------------------------------|
| 6. D | → | {ein, dieses} |
| 7. N | → | {Professor, Haaren, Buch ...} |
| 8. V | → | {schrieb, ...} |
| 9. A | → | {alter, langweilige, ...} |
| 10. P | → | {mit, ...} |
| 11. Adv | → | {sehr, ...} |

Dependenzstruktur – Phrasenstruktur



Eine regierende Kategorie X zusammen mit allen abhängigen konstituiert eine X-Phrase (XP), wobei X = N, V, A, P

Konstituentenregeln (Phrasenstrukturregeln)

1. S → NP VP
2. NP → D (AP) N (PP)
3. VP → V (NP)(PP)
4. PP → P NP
5. AP → Adv A

- | | | |
|---------|---|-------------------------------|
| 6. D | → | {ein, dieses} |
| 7. N | → | {Professor, Haaren, Buch ...} |
| 8. V | → | {schrieb, ... } |
| 9. A | → | {alter, langweiliges, ...} |
| 10. P | → | {mit, ... } |
| 11. Adv | → | {sehr, ... } |

Syntaktische Strukturierung: Vergleich

- | | |
|---|---|
| <ul style="list-style-type: none">■ Konstituenz<ul style="list-style-type: none">▶ Prinzip: Zerteilen■ Wortfolge:<ul style="list-style-type: none">▶ starre Wortstellung■ aus Baum ableitbar:<ul style="list-style-type: none">▶ Menge von Regeln■ Erweiterungen:<ul style="list-style-type: none">▶ X-bar-Schema,▶ Transformationen etc. | <ul style="list-style-type: none">■ Dependenz<ul style="list-style-type: none">▶ Prinzip: Verbinden■ Wortfolge:<ul style="list-style-type: none">▶ keine Wortstellungsinformation■ aus Baum ableitbar:<ul style="list-style-type: none">▶ Wortvalenzen■ Erweiterungen:<ul style="list-style-type: none">▶ Markierung von Konstituenten,▶ komplexe Knoten,▶ Wortstellungsregeln |
|---|---|