

Syntaktische Funktionen

Dr. Benjamin Roth

CIS LMU München

Gliederung

1 Syntaktische Funktionen

2 Dependenzgrammatik

Der Begriff “Syntax”

- griech.: *sýn* (zusammen) + *táxis* (Ordnung)
- → Anordnung, Regelung, Organisation, Verhältnis
- Einteilung der linguistischen Beschreibung/Theorie in Ebenen/Levels:
 - ▶ 5-Teilung: Phonetik/Phonologie – Morphologie – Syntax – Semantik – Pragmatik
 - ▶ Interaktionen zwischen Phonetik/Phonologie und Syntax: “Du musst gehen?” vs “Du musst gehen!”
 - ▶ Syntax erfordert bestimmte Intonation: “Musst Du gehen?”
 - ▶ Interaktionen zwischen Semantik und Syntax: Ein Satz hat oft viele mögliche syntaktische Analysen, aber nur wenige sind semantisch sinnvoll.
 - ▶ “time flies like an arrow”

Gegenstandsbereich der Syntax

- Wortstellung, “Regeln” der Wortstellung
- Zusammensetzungen von Wörtern: Wortverbindungen, Syntagmen, Phrasen
- Rekursiv: Zusammensetzungen von Phrasen zu größeren Phrasen
- Der Satz als Zusammensetzung von Phrasen
- **Funktion der Wortarten und Phrasenarten beim Aufbau von größeren Phrasen und beim Aufbau des Satzes**
- **Wechselseitigen Beziehungen der Glieder des Satzes und ihre Beziehung zum Satzganzen**
- Zwei Sichtweisen:
 - ① Die Syntax beschreibt welche Sätze möglich (wohlgeformt) sind.
 - ② Die syntaktische Beschreibung eines Satzes stellt das Gerüst für seine semantische Interpretation.

Syntax in der Computerlinguistik

- Theoretische Linguistik:

- ▶ “Welche sprachlichen Konstrukte sind wohlgeformt?”
- ▶ “Welche Analysen sind kognitiv motiviert?”
- ▶ “Welche Beschreibungsmechanismen sind auf alle Sprachen anwendbar?”
- ▶ ⇒ Grenz- und Sonderfälle von besonderem Interesse.

- Sprachverarbeitung:

- ▶ “Welche linguistisch motivierte Repräsentation beschreibt die wesentlichen Zusammenhänge einer sprachlichen Äußerung (wohlgeformt oder nicht) zur weiteren Verarbeitung?”
- ▶ Tweets, Tippfehler, nicht-Muttersprachler, Korrekturen in gesprochener Sprache, ...
- ▶ ⇒ Von besonderem Interesse: Robustheit, Abdeckung der häufigsten Fälle, Konsistenz.

Outline

1 Syntaktische Funktionen

2 Abhängigkeitsgrammatik

Syntaktische Funktionen

Grammatische Relation zwischen zwei Ausdrücken

- bestimmt durch die morphologische Markierung
- und / oder die strukturelle Relation der Ausdrücke zueinander

Funktionen sind durch satzsyntaktische syntagmatische Relationen zwischen Teilen und Ganzem definiert.

- Subjekt_von_X
- Objekt_von_X
- Prädikativ_zu_X
- Attribut_von_X
- Adverbiale_von_X

Das Verb wird manchmal auch als Prädikat bezeichnet.

Syntaktische Funktion: Subjekt

• Wer? Was?

- ▶ Kasus: Nominativ
- ▶ Kongruenz mit dem finiten Verb
- ▶ Realisierung durch verschiedene Kategorien:

Der Kater *lässt das Mäusen nicht.*

Er *wittert Gefahr.*

Dass das nicht funktioniert, *wundert mich nicht.*

Es/das *überrascht niemanden, dass er schweigt.*

Syntaktische Funktion: Subjekt

- In einer Nominalisierung entspricht dem Subjekt normalerweise das Possesivum:
Paul rezensiert ein Buch
Pauls Rezension
- Im Deutschen muss immer ein syntaktisches Subjekt realisiert werden, auch wenn es semantisch leer ist.
- Dies geschieht z.B. durch das **Wetter-es**, ein sog. **Expletivum**.
es regnet
es graut mir vor der Prüfung

Syntaktische Funktion: Objekt

- Verbergänzungen im Akkusativ, Dativ oder Genitiv:
*Peter isst **einen Apfel*** (Akkusativobjekt)
*Peter hilft **seinem Freund*** (Dativobjekt)
*Peter gedachte **seiner Mutter*** (Genitivobjekt)
- Objektsatz: vom Verb geforderte Satzwertige Ergänzung:
*Ich habe versprochen, **dass ich mich beile**.*

Syntaktische Funktion: Objekt

- Präpositionalobjekt:

- ▶ Vom Verb geforderte obligatorische Ergänzung, die mit einer semantisch leeren Präposition realisiert wird:
*Peter denkt nur **an sich**.*
- ▶ Wird die Semantik der Präposition beibehalten, spricht man dagegen von einer adverbialen Ergänzung.
*Peter arbeitet/schläft/tanzt **im Garten**.*
- ▶ Obligatorische adverbiale Ergänzungen werden manchmal auch zu den Präpositionalobjekten gezählt:
*Peter wohnt **in Hamburg***

Syntaktische Funktion: Prädikativ

- Prädikative ordnen Satzgliedern Eigenschaften zu.
- Subjektsprädikativ bei Kopulaverben:
 - ▶ Kerstin ist **Informatikerin**.
 - ▶ Anna wird **reich und glücklich**.
- Objektsprädikative bei Verben wie *finden, nennen, heißen, schimpfen*
 - ▶ Sie fand das Buch **recht teuer**.
 - ▶ Sie hieß **ihn einen Versager**.
 - ▶ Er nannte sie **eine Lügnerin**.
- *Prädikativ* nicht mit *Prädikat* (Verb oder Verbphrase) verwechseln.

Syntaktische Funktion: Adverbiale

- Bei Adverbialen wird die Hauptunterscheidung semantisch getroffen, in einem Satz können verschiedene Kategorien vertreten sein:

*Sie liegt vor lauter Langeweile (**kausal**) den ganzen Tag (**temporal**) dösend (**Art und Weise**) im Bett (**lokal**).*

Syntaktische Funktion: Adverbiale

- Je nach Valenz des Verbs unterscheidet man:

- ① **Obligatorische** adverbale Ergänzungen sind in der Valenz des Verbs fest angelegt.

*Ich fühle mich **gut**.*

*Hans wohnt **in Berlin**.*

- ② **Fakultative** adverbale Ergänzungen sind nur bei bestimmten Verben möglich, wenn auch nicht notwendig.

*Anna schwimmt **schnell***

- ③ **Freie Angaben** können zu allen Verben ohne Beschränkungen hinzutreten.

Er arbeitet (am Wochenende)(gern)(in aller Ruhe)(im Garten).

Syntaktische Funktion: Adverbiale

- Adverbiale Ergänzungen sind typischerweise Adjektive (Adverbien) oder Präpositionalphrasen.
- Die adverbiale Funktion kann auch durch Sätze wahrgenommen werden:

Lena spielt, während Mama arbeitet. (Temporalsatz)

Unglückliche Menschen, wohin man schaut. (Lokalsatz)

Syntaktische Funktion: Attribut

- Beifügungen zur besonderen Bestimmung eines Substantivs (bzw. Nominalphrase).
- Können nur in Abhängigkeit dieses Substantivs im Satz auftreten.
- Syntaktisch können sie entweder als Teil der näher bestimmten Nominalphrase auftreten, oder als eigenes Satzglied.
- Als **Satzgliedteil** nur zusammen mit dem Bezugselement verschiebbar:

Er beantwortet [den Brief [des Freundes]] heute.

** [Des Freundes] beantwortet er [den Brief] heute.*

- Attribut als **umstellbares Satzglied**:

*Sie trinkt den Tee **mit Milch**.*

*Den Tee trinkt sie **mit Milch**.*

***Mit Milch** trinkt sie den Tee.*

Syntaktische Funktion: Attribut

- Adjektivattribut:
*ein **neues** Buch*
- Partizipialattribut:
***schlafende** Hunde*
- Präpositionalattribut:
*Der Mann **vom Mond***
- Genitivattribut:
*die Zerstörung **der Welt***
- Adverbattribut:
*Der Unterricht **gestern** war langweilig*
- Appositionen:
*Heiner, **der Chef der Firma***
- Attributsätze:
*die Dänen, **die Biertrinken**, ...*
*die Hoffnung, **dass alles gut wird***

Beispiel: Analyse eines Satzes Anhand von Wortarten und Syntaktischen Funktionen

	Opa	erzählte	gestern	eine	lange	Geschichte
Wortart:	Nomen	Verb	Adverb	Artikel	Adjektiv	Nomen
Funktion:	Subjekt	(Prädikat)	Temporal- adverbiale	< - - - - Objekt - - - - >		
					Adj.attribut	

Outline

1 Syntaktische Funktionen

2 Abhängigkeitsgrammatik

Vom syntaktischer Funktion zu Satzstruktur

- Wir haben gesehen, welche syntaktischen Funktionen zwischen Satzgliedern existieren können.
- Wie kann eine syntaktische Beschreibung eines Satzes aussehen, die
 - ▶ linguistisch motiviert ist, und gleichzeitig
 - ▶ formal genug ist, um von einem Computer verarbeitet werden zu können?

Beschreibung der Satzstruktur durch Dependenzrelationen

- Beschreibt die Abhängigkeiten im Satz anhand **binärer Relationen zwischen Wörtern**.
- Grundprinzip bei der Verbindung von Wortpaaren:
Ein Wort ist der **Kopf** (engl. *head*), das andere Wort ist das **Dependent**.

Beispiele von Kopf-Dependent Relationen

- Kopf verlangt Dependent als obligatorische Ergänzung (Argument):
"Ich trinke Wasser"
trinke[Kopf]-ich[Dep.], trinke[Kopf]-Wasser[Dep.]
- Dependent modifiziert Kopf als optional Ergänzung:
"Ich trinke gerne viel Wasser"
trinke[Kopf]-gerne[Dep.], Wasser[Kopf]-viel[Dep.]
- *"Primacy of content words"*
 - ▶ In syntaktischen Konstruktionen aus Hilfswörtern (Partikel, Hilfsverben, ...) und Inhaltswörtern (Vollverben, Nomen, ...) ist das Hilfswort Dependent zum Inhaltswort.
"Ich habe getrunken"
getrunken[Kopf]-habe[Dep.]
 - ▶ Determinierer ist ein Dependent zum Kopf:
"Das Llama trinkt Wasser"
Llama[Kopf]-das[Dep.]

Bestimmung der Wortpaare

- Kopf und Dependent bestehen jeweils immer nur aus einem Wort.
- Wenn ein komplettes Satzglied (Phrase aus mehreren Wörtern) Argument oder Modifikator ist:
 - ▶ Bestimme zunächst Kopf innerhalb des Satzgliedes.
 - ▶ Der phraseninterne Kopf ist dann Dependent zum einbettenden (modifizierten) Wort.
- Das Verb nimmt die zentrale Position im Satz ein.

Dependenzanalyse (Beispiel)

“Die Katze schläft nicht.”

Katze
X
die

schläft
Y
Katze

schläft
Z
nicht

Hinweis: Verschiedene Dependenzformalismen

- Verschiedene Dependenzformalismen kommen teils zu sehr unterschiedlichen Analysen.
- **Rein syntaktisch** motivierte Formalismen.
Entscheidungskriterien sind typerweise:
 - ▶ Welches Wort bestimmt die möglichen Kontexte in denen die Verbindung aus Kopf und Dependent auftreten Kann?
 - ▶ Zwischen welchen Wörtern besteht Kongruenz in bestimmten syntaktischen Merkmalen?
 - ▶ Welches Wort fordert bestimmte syntaktischen Eigenschaften eines anderen Wortes (z.B. Kasus)?
- **Teils semantisch** motivierte Formalismen.
Entscheidungskriterien sind typerweise:
 - ▶ Welches sind die Inhaltswörter und ihre Argumente?
 - ▶ Welche Elemente haben nur syntaktische Funktionen?

Hinweis: Verschiedene Dependenzformalismen

- Unterschiede zwischen syntaktisch und semantisch motivierten Formalismen typischerweise bei der Analyse von:
 - ▶ Konstruktionen mit Hilfsverben,
 - ▶ Subjunktionen (eingebettete Nebensätze),
 - ▶ Präpositionen
- Weitere Argumente, die für verschiedene Analysen vorgebracht werden:
 - ▶ Argumentstruktur / Semantische Rollen (z.B. bei Hilfsverben, Präpositionalobjekte)
 - ▶ Phrasenstatus (z.B. X-bar Theorie bei Subjunktionen)
 - ▶ Parallelität zu morphologischer Markierung in anderen Sprachen (Präpositionen)
- Wichtig für Computerlinguisten ist, dass die Analysen innerhalb des Formalismus konsistent sind.
- Nützlich ist auch eine Darstellung, die eine semantische Verarbeitung erleichtert.
- In dieser VL: **Stanford Dependency Formalismus**, der als Köpfe vorrangig die **Inhaltswörter** wählt.

Primacy of Content Words

(from Universal Dependencies guidelines)

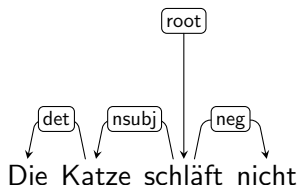
- “**Dependency relations hold primarily between content words**, rather than being indirect relations mediated by function words.
- Preferring content words as heads **maximizes parallelism** between languages because content words vary less than function words between languages.
- In particular, one commonly finds the same grammatical relation being **expressed by morphology** in some languages or constructions and **by function words** in other languages or constructions, while some languages may not mark the information at all (such as not marking tense or definiteness).
- We prefer to view the relations between content words and function words [...] as operations that modify the grammatical category of the content word so that it can participate in different dependency relations with other content words.”

Benennung der Abhängigkeitsrelationen

- Abhängigkeitsrelationen werden entweder durch die **Wortart** des Dependents oder seine **syntaktische Funktion** gekennzeichnet.
- Siehe z.B. *Stanford Universal Dependencies*:
<http://universaldependencies.org/>
- **Wortart oder phrasale Kategorie** z.B. wenn Dependent: Nominal (n), Verb/Satzeinbettung (c – *clause*), Adjektiv (a), Adverb (adv), Artikel/Determinierer (det), Negation (neg), Hilfswort (aux)
- **Syntaktische Funktion** z.B. wenn Dependent: Subjekt (subj), Objekte (obj), Modifikator/Adverbiale (mod), Attribut (attr), Apposition (appos)
- Beide Kategorisierungen können auch kombiniert werden:
Schwimmen erfrischt
erfrischt[Kopf]–schwimmen[csubj] (clausal subject)

Dependenzbäume

- Die Kanten enthalten die Namen der Dependenzrelationen.
- Pfeile werden immer vom Kopf zum Dependenten gezeichnet.
- Das nicht-abhängige Verb des Satzes wird oft als Wurzel des Dependenzbaums (*root*) gekennzeichnet.



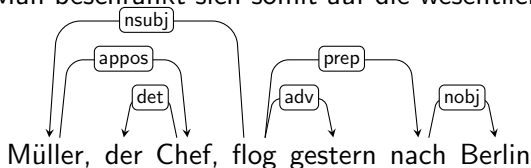
Präpositionen

Bei Präpositionalobjekten muss man sich z.B. für eine der folgenden Konventionen entscheiden:

- ① Die Präposition ist das Dependent zum Verb, und das Nomen ist Dependent zur Präposition.
 - ▶ Analyse gemäß der ursprünglichen Variante der Stanford Dependencies. Widerspricht eigentlich dem Prinzip der “Primacy of content words”.
 - ▶ **Beachte:** Wird von vielen automatischen Analyse-Tools so ausgegeben.
- ② Das Nomen der Präpositionalphrase ist das direkte Dependent zum Verb, und die Präposition fungiert als Kasus-Markierer (**case**).
 - ▶ Analyse gemäß der neueren der Stanford *Universal* Dependencies.
 - ▶ Argumente mit Präposition werden auch als oblique (**obl**) bezeichnet.
 - ▶ Es wird keine Unterscheidung zwischen Argumenten und Adjunkten, getroffen, da der Unterschied oft fließend ist.

Praktische Verwendung von Dependenzbäumen

- Dependenzstrukturen spielen eine wichtige Rolle bei der automatischen Informationsextraktion.
- Typischerweise wird
 - 1 Eine automatische Dependenz-Analyse durchgeführt (**dependency parsing**).
 - 2 Die kürzeste Verbindung (Pfad) zwischen zwei Entitäten ermittelt.
 - 3 Die Wörter und Dependenzrelationen auf dem Pfad mit einer Liste an bekannten Mustern verglichen.
- Man beschränkt sich somit auf die wesentliche Information:



⇒ `travelled_to(Müller, Berlin)`

Dependenz-Parsing

- Software zur automatischen Dependenz-Analyse: Dependenz-Parser
 - ▶ <http://demo.ark.cs.cmu.edu/parse>
 - ▶ <https://nlp.stanford.edu/software/lex-parser.shtml>
- Dependenz-Parsern liegt meist ein statistisches Modell zugrunde.
- Die Dependenzrelationen werden so gewählt, dass die Wahrscheinlichkeit der Analyse für den Satz maximiert wird. (Kombinatorisches Problem!)
- **Fehleranfällig:** Typischerweise 90% oder weniger Genauigkeit (pro Kante).
- Automatische Dependenz-Analysen sind oft der Ausgangspunkt für weitere regelbasierte oder statistische Verfahren.

Dependenz-Parsing: Demo

Zusammenfassung

- Die Dependenzanalyse zeigt die Beziehungen der Wörter in einem Satz auf
- Wenn man ein automatisches Dependenz-Analyse-Tool verwenden will, muss man
 - ▶ wissen, was die Grundzüge des unterliegenden Formalismus sind
 - ▶ beachten, dass automatische Analysen immer fehleranfällig sind
- Die von uns betrachtete Sichtweise ist an die Stanford Universal Dependencies angelehnt
 - ▶ Grundlage vieler computerlinguistischer Analyse-Tools
 - ▶ Inhaltswörter stehen im Zentrum: bessere Generalisierung zwischen verschiedenen Sprachen und Konstruktionen
 - ▶ weicht teils von anderen linguistischen Theorien ab