

Einführung in die Computerlinguistik

Syntax I

Hinrich Schütze & Robert Zangeneid

Centrum für Informations- und Sprachverarbeitung, LMU München

14.12.2015

Take-away

- Begriff
- Themen
- Gegenstandsbereich
- Segmentieren von Sätzen – Konstituententests
- Konstituentenstruktur

Der Begriff “Syntax”

- griech.: *sýn* (zusammen) + *táxis* (Ordnung)
→ Anordnung, Regelung, Organisation, Verhältnis
- vgl.a. “Syntaktik”: in der Semiotik von Morris: Verknüpfungen und Beziehungen von Zeichen
- Syntaxtheorien: verschiedene Theorien, die immer nur Hilfskonstruktionen sein können

Themen zur Syntax

- Lehre von der Funktion der Wortklassen oder Wortarten bzw. Wortverbindungen/Syntagmen im Satz
- Teil der Grammatik, der die Zusammenordnung der Elemente regelt
- Lehre von den Regularitäten des Satzes
- Lehre von den Syntagmen, Sätzen und Texten
- Leistungen der einzelnen Wortformen und Wortarten beim Aufbau des Satzes, ihre Beziehungen zum Satzganzen, ihre wechselseitigen Beziehungen
- Wortstellung
- 3-Teilung: Syntax – Semantik – Pragmatik
- 4-Teilung: Phonologie – Morphologie – Syntax – Semantik

Gegenstandsbereich der Syntax (1)

- Regeln, nach denen Wörter zu kommunikativen Äußerungen (Benennung von Sachverhalten) verbunden werden
- begrenzte Zahl von Regeln → potentiell unbegrenzt viele Sätze
- Klassifizierung der beteiligten syntaktischen Einheiten
- Hierarchische Organisation der Konstruktionen (Syntagmen, Sätze); (vgl. z.B. Wortstellung):
 - (1) a. die Freunde ihrer Söhne
b. die Söhne ihrer Freunde
- vielfältige Bildungsweise eines Satzes (im Vgl. zur Wortbildung)
- unbegrenzte Möglichkeit zur Hinzufügung (im Dt. theoretisch auch bei Wortbildung möglich)

Gegenstandsbereich der Syntax (2)

- Grundeinheit der Syntax: Satz
- Prädikativität (entscheidendes Merkmal des Satzes): Tempus, Modus (Relation zur Wirklichkeit)
- Einfacher Satz – komplexer Satz (Prädikativität kommt mehrfach vor)
- Syntagma: auch nichtprädikative Konstruktionen
 - Wortfügungen (subordinierende Syntagmen), z.B. *kleines Kind, schnell wachsen*
 - Wortreihungen (koordinierende Syntagmen), z.B. *Mutter oder Sohn, kommen und sehen*
- Wortform: kleinster Bestandteil einer syntaktischen Konstruktion; wird verbunden zu syntaktischen Einheiten
- wichtig sind alle Erscheinungen, zwischen denen syntaktische Beziehungen bestehen
- auch synt. Beziehungen zwischen zwei Sätzen

Gegenstandsbereich der Syntax (3)

charakteristisch für Kommunikation: Satz steht nicht allein

- Satz bildet relativ geschlossene Äußerung, Sinneinheit: Mitteilungsabsicht des Sprechers; Hörer versteht etwas
- Ist *Ich* ein Satz? → Dialog!
- Monolog, vgl. *Gestern bin ich spät schlafen gegangen. Deswegen bin ich heute müde. Trotzdem muss ich . . .*
- Dialog, vgl. *Ich würde heute gerne ins Kino gehen. – Das ist eine gute Idee / Ich auch.*
→ 2. Version dieser Antwort: Satz unvollständig, aber durch Kontext verständlich

→ Satz ist diejenige sprachliche Einheit, durch die Verständigung gewährleistet wird

Kurzer Ausblick

- wichtig: Wortfolge, Intonation (auch schriftlich durch Interpunktion!), vgl. z.B. gespr.: *Du musst arbeiten ! / ?*
- dt. Satzbau: Zweigliedrigkeit (Satzsubjekt, Prädikat) wesentlich, auch ohne sem. Subjekt wichtig, z.B. *Es donnert* (vgl. russ. *Gremit*)
- zwei konkurrierende Grundprinzipien: Konstituenz und Dependenz

Einige Beispielsätze zur Einführung (1)

(u.a. aus H. Ramers: Einführung in die Syntax. Paderborn: Fink 2007)

- (1)
- a. Hans liest.
 - b. Es wird dunkel.
 - c. In Brasilien wird gut Fussball gespielt.
 - d. Vorsicht!
- (2)
- a. Hans *hat* uns gesehen.
 - b. Uns *hat* Hans gesehen.
 - c. Gesehen *hat* uns Hans.
 - d. *Hans uns gesehen *hat*.
 - e. *Hans uns *hat* gesehen.

Einige Beispielsätze zur Einführung (2)

- (3) a. Ein Mann *hat* uns gesehen.
b. *Ein *hat* Mann uns gesehen.
c. Der große Kerl mit dem roten Schal *hat* uns gesehen.
d. Peter, den Maria gestern um 4 Uhr zum Tee treffen wollte, *hat* uns gesehen.
e. *Peter *hat*, den Maria gestern treffen wollte, uns gesehen.
- (4) a. Sie lesen dieses Buch.
b. Lesen Sie dieses Buch?
c. Lesen Sie dieses Buch!
d. Dass Sie (mir ja) dieses Buch lesen!

Was ist ein Satz?

- “Aussage”? → Aussagen (im aristotelischen Sinn) sind wahr oder falsch (vgl. dagegen z.B. Fragesätze)
- “zweiteilig”? → vgl.a. *Wasser!* oder *Feuer!*
→ Syntax muss alle (syntaktisch) wohlgeformten (akzeptablen?) Sätze einer Sprache erzeugen/beschreiben (→ Algorithmus)
- vgl. *Farblose grüne Ideen schlafen aufgebracht.*
→ Trennen von Semantik!
- Syntax kann umgekehrt auch aus der Menge aller Strings diejenigen auswählen, die wohlgeformte Sätze sind

Zur inneren Struktur des Satzes

- Wörter nicht nur aneinandergereiht, sondern gruppiert
→ hierarchische Strukturen
- Tests zum Feststellen syntaktischer Einheiten
- intuitiv klare Gruppen bei z.B. *Der kleine Junge fährt mit dem neuen Fahrrad, das ihm der Opa geschenkt hat, zum Kindergarten.*
→ [*Der kleine Junge*] fährt [*mit dem neuen Fahrrad, das ihm der Opa geschenkt hat,*] [*zum Kindergarten*]. (nicht z.B. [*Junge fährt*])
- Regeln zum Gruppieren?
→ strukturelle Linguistik (Saussure → Bloomfield) → Konstituententests

Permutationstest

(Verschiebeprobe)

- Konstituente: Wörter, die zusammen umgestellt werden können, so dass der Satz grammatisch bleibt
- z.B. *An der Ecke steht ein Haus.* – *Ein Haus steht an der Ecke.*

Substitutionstest (1)

(Ersetzungsprobe)

- Konstituente: Wortgruppe, die durch eine andere ausgetauscht werden kann
- z.B. *Ein Haus steht an der Ecke.*
 - *Karl ...*
 - *Das Hotel Post ...*

Substitutionstest (2)

Varianten:

- Pronominalisierungstest
 - Pronomen (allgemeiner: Proform) dient als Ersatz
 - z.B. *Es/jemand steht an der Ecke.*
 - auch: Pronominaladverbien (*darüber, hierfür, ...*), Adverbien (*dort, ...*), Pro-Verben („Das macht/tut X auch gerade“)
- Fragetest
 - wie Pronominalisierungstest, nur mit Fragepronomen (z.B. *wer, wen, was?*)
- Problem z.B: *Ich möchte nach Paris fahren.* Der Teil *möchte nach Paris fahren* kann nicht pronominalisiert werden (ist aber syntaktische Einheit!)

Koordinationstest

- Konstituente: Satzelement, das sich mit einem anderen koordinieren lässt
- z.B. Konjunktion *und*: [*Die Musik ertönt*] und [*der Vorhang hebt sich*].
- *Peter kauft neue Kartoffeln.* → *Peter kauft [neue Kartoffeln] und [saftige Äpfel].* (*kauft neue*: nicht koordinierbar (außer elliptisch, wie *P. kauft neue und verkauft alte K.*) → keine Konstituente!)
- Koordination ist grammatischer Prozess, der Konstituenten betrifft

Eliminierungstest

(Weglassprobe, Tilgungstest)

- welche Elemente sind syntaktisch notwendig?
- Konstituente: das, was zusammen weggelassen werden kann
- z.B. *Maria singt in der Badewanne.* – *Maria singt.*

Variante:

- Reduktionstest: *Den einen interessiert [die Kunst], den anderen langweilt die Kunst.* (Ellipse, die aus dem Kontext erschließbar ist)

Europäische Grammatik-Tradition

- lange Zeit verbreiteter Ansatz:
- zwei gleichberechtigte Zentren im Satz:
 - nominales Subjekt
 - verbales Prädikat
- bestimmen ihre Form gegenseitig
 - Verb verlangt Nominativ des Satzsubjekts
 - Nomen verlangt Kongruenz des Verbs in Person und Numerus (z.B. im Dt.)
- Lit.: K.H. Ramers: Einführung in die Syntax. München: Fink 2007

Konstituenz

- L. Bloomfield: Language (1933) (amerikanischer Strukturalismus), N. Chomsky
- Zerlegung und Klassifizierung führen zu immediate constituents
→ Beziehung zwischen Wortgruppen (Wortverbindungen)
- Fragen dazu:
 - welche Wörter oder Wortgruppen sind in einem Satz enger miteinander verbunden als mit anderen?
 - welche Wörter oder Wortgruppen in einem Satz gehören zur gleichen grammatischen Klasse?
- Hierzu verschiedene Tests (für potentielle Konstituenten; vgl. Abschnitt zum Segmentieren: Permutation, Substitution, Koordination, Eliminierung)

→ Wörter nicht als lineare Kette, sondern als Wortgruppen/Konstituenten

Klassifikation von Konstituenten/Phrasen (1)

→ Wortartenklassifikation

- jede Wortgruppe enthält als obligatorischen Kern ein Wort einer bestimmten Wortart
- im Wesentlichen:

(i) Nominalphrasen (NP)

- Kinder, die in Stuttgart leben

(ii) Adjektivphrasen (AP)

- ... auf seine Tochter stolz

Klassifikation von Konstituenten/Phrasen (2)

(iii) Verbalphrasen (VP)

- Eva *hustet*.
- Eva *besucht die Nachbarn*.
- Eva *schenkt ihrem Mann ein Buch*.
- Eva *glaubt, dass Peter krank ist*.

Klassifikation von Konstituenten/Phrasen (3)

(iv) Präpositionalphrasen (PP)

- Wir fahren *mit dem Fahrrad*.
- Sie stehen *vor der Tür*.
- Wir spazieren *die Straße entlang*.
- Sie kommen *von dort*.
- Wir brachten sie *bis vor die Tür*.

Klassifikation von Konstituenten/Phrasen (4)

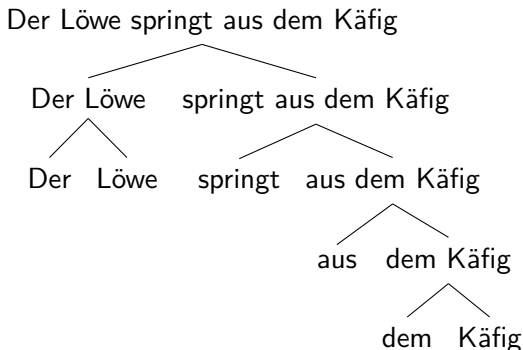
(v) Adverbialphrasen (AdvP)

- heute früh
- gestern Morgen
- sehr oft

Darstellung von Konstituentenstrukturen (1)

Baumdiagramm

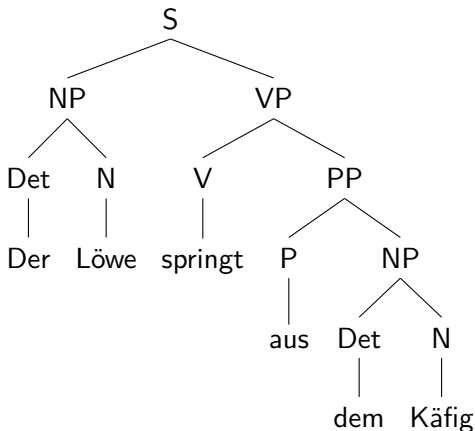
→ Teilkonstituenten



Darstellung von Konstituentenstrukturen (2)

- Ersetzen der konkreten Wortketten durch Symbole für syntaktische Kategorien
→ Zerlegung in unmittelbare Konstituenten (NP + VP) und weiter bis zur Ebene der Wortkategorien:

Darstellung von Konstituentenstrukturen (3)



- (schräge) Kanten: “Konstituente von ...”
- senkr. Linien: Zuordnung Wortkategorialsymbol – konkr. Wort

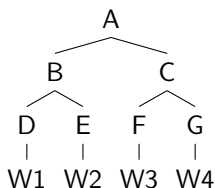
Darstellung von Konstituentenstrukturen (4)

alternative Darstellung:

[S [NP [Det Der] [N Löwe]] [VP [V springt] [PP [P aus] [NP [Det dem]
[N Käfig]]]]]]

- Konstituenten dürfen nicht linear auseinandergerissen sein

Allgemeine Darstellung: System der Konstituenten



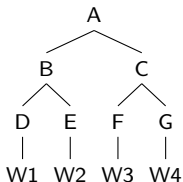
- Schrägstriche: Kanten, branches, Äste, Zweige
- Variablen A–G (Knoten, nodes): Kategoriensymbole
- speziell A–C: phrasale Kategorien
- speziell D–G: präterminale Kette; Wortarten
- W1–W4: terminale Kette; konkrete Wortformen
- einzelne Wortformen können nicht unmittelbare Konstituenten eines Satzes sein → präterminale Kette nötig!

zur Terminologie (1)

Strukturelle Beziehungen zwischen Konstituenten:

(unmittelbare) Konstituente:

- ein Knoten bildet eine Konstituente eines höheren Knotens, wenn er durch Kanten nach oben im Baum mit diesem verbunden ist
- unmittelbare Konstituente: Verbindung mit dem nächsthöheren Knoten

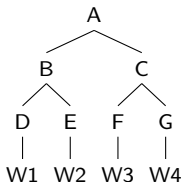


- z.B.: D ist eine Konstituente von B und A
- W1 ist ein Exemplar der Menge D

zur Terminologie (2)

(unmittelbare) Dominanz:

- Umkehrrelation zur Konstituenz
- ein Knoten dominiert alle tieferen Knoten, mit denen er durch Kanten verbunden ist
- ein Knoten dominiert unmittelbar nur die nächsttieferen Knoten
- allg.: ein Knoten X dominiert einen Knoten Y, wenn Y ein Teil von X ist



- z.B.: A dominiert alle anderen Knoten im Baum; unmittelbar dominiert A nur B und C

zur Terminologie (3)

Schwester von ... (Mutter von ..., Tochter von ...):

- Schwestern sind Konstituenten, die vom gleichen Knoten (ihrer Mutter) unmittelbar dominiert werden
- z.B. oben: D und E sind Schwestern, weil sie von der gleichen Mutter dominiert werden

c-command (constituent command):

- X kommandiert Y, wenn
 - (i) Y eine Schwester von X ist, oder
 - (ii) Y eine – auch mittelbare – Konstituente einer Schwester von X ist

→ Konstituente: Kette von Symbolen, die vom gleichen Knoten ausschließlich dominiert wird

Take-away

- Begriff
- Themen
- Gegenstandsbereich
- Segmentieren von Sätzen – Konstituententests
- Konstituentenstruktur