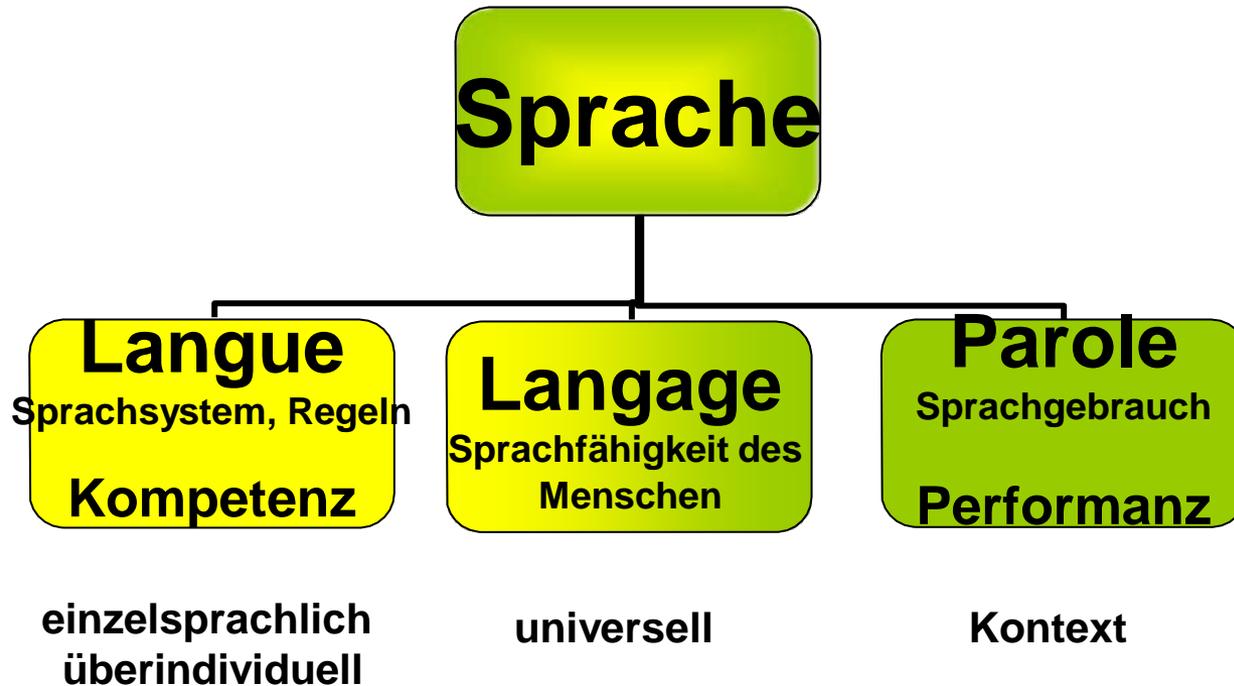


Interdisziplinäre
fachdidaktische Übung:
Sprachmodelle –
Formale Sprachen

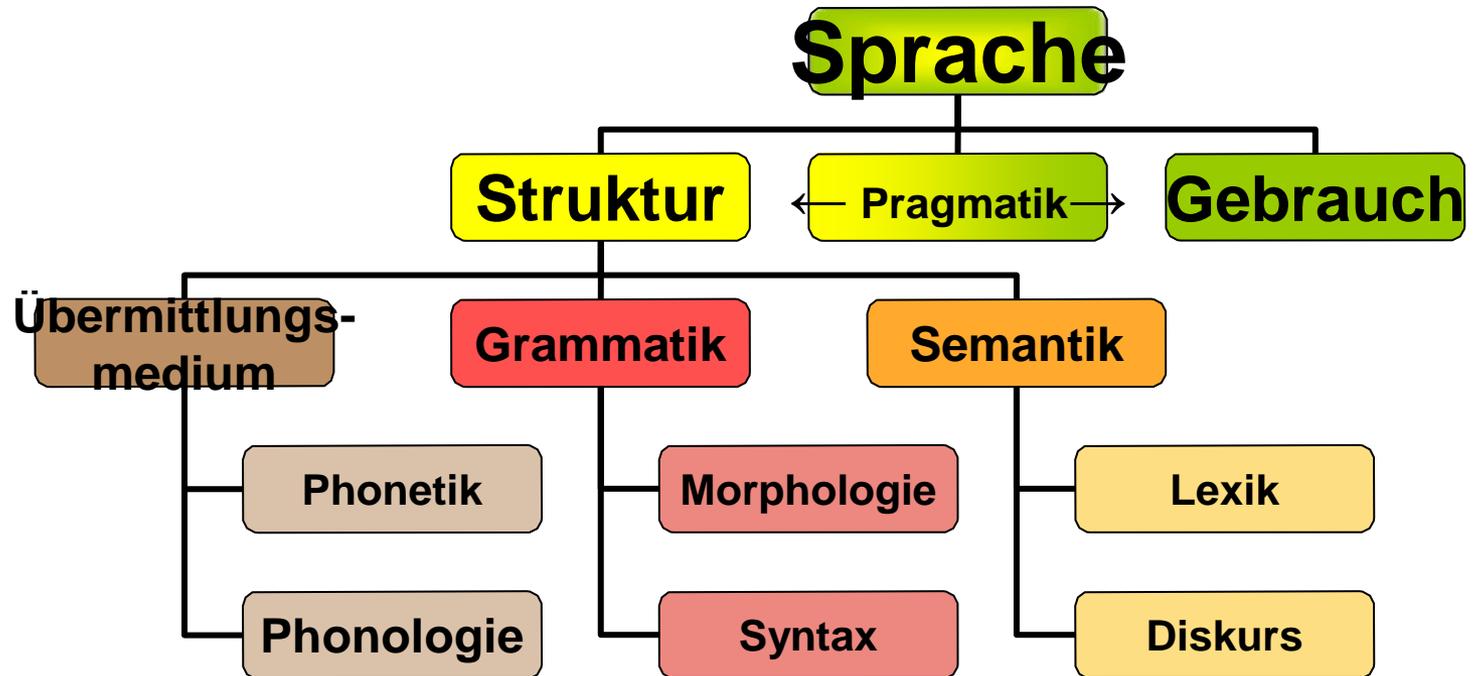
SS 2014: Grossmann, Jenko

Modell Natürliche Sprachen



Maschine hat keine Sprachfähigkeit !

Modell Natürliche Sprachen



- Welche Konzepte lassen sich wie auf formale Sprachen anwenden?

Struktur - Formale Sprache

Strukturbeschreibung formaler Sprachen wird wesentlich durch die Aufgabe „Kommunikation mit einer Maschine“ bestimmt

Da Maschine keine Sprachfähigkeit hat muss Sprache von einem abstrakten Modell her entwickelt werden

Struktur - Formale Sprache

Beispiel:

Frage: „Scheint am 20.4.2012 um 9:00 in
Wien die Sonne?“

Antwort: „ja“

Wie kann eine Maschine diesen Dialog
abbilden?

Struktur - Formale Sprache

Maschine erhält Instruktionen = Programm:

Eingabe: *Datum:20.4.2012, 9:00,Wien,
wolkenlos, Sonnenaufgang 6:00*

Instruktion: *Wenn Wetter wolkenlos und*

*Tageszeit nach Sonnenaufgang dann
antworte mit „ja“ ansonsten mit „nein“*

Struktur - Formale Sprache

Die Maschine entscheidet nur ob sie diese Kombination von Eingabe und Instruktion als eine für sie korrekte Zeichenkette erkennt und nachdem sie das erkannt hat reagiert sie mechanisch, im Prinzip nur mit „korrekt“ oder „nicht korrekt“

Struktur - Übermittlungsmedium

Das Übermittlungsmedium ist heute für formale Sprachen sehr vielfältig (Eingabemedien)

Im Prinzip wird aber jede Eingabe in eine Zeichenkette umgewandelt

Struktur - Grammatik

Grammatik konzentriert sich im wesentlichen auf die Syntax und ist präskriptiv (Regeln) und deskriptiv (Beschreibung aller Konstruktionen)

Wir müssen ein sehr einfaches und allgemeines Sprachmodell definieren

Struktur - Grammatik

Elemente der Sprache sind:

Eine endliche Menge von Symbolen, die man Alphabet bezeichnet. Sie bilden die Terminalen Symbole

Eine endliche Menge von Nicht-terminalen Symbolen

Ein Startsymbol

Produktionsregeln: Ausdruck \rightarrow Ausdruck

Struktur - Grammatik

Analogie zu natürlichen Sprachen ?

Alphabet, Sprachlaute

Worte

Grammatische Kategorien

Struktur - Grammatik

Im Prinzip gibt es nur drei Regeln um Ausdrücke zu bilden:

- Alle Elemente Alphabets sind Ausdrücke
- Verkettung: $(x, y) \rightarrow xy$
- Alternative: $x|y$

Wesentlich für die Charakterisierung von formalen Sprachen sind Produktionsregeln
Jede Produktion wird **Wort** genannt

Struktur – Grammatikhierarchien

Allgemeinste Form sind Typ 0 Grammatiken

Keine Einschränkung an Produktionsregeln

Sie heißen rekursiv aufzählbare Sprachen

Man kann ein Modell einer Maschine

entwickeln, die bei jedem in der Sprache

gebildeten Wort mit „richtig“ antwortet

Allerdings kann vorkommen, dass bei Worten

die nicht in der Sprache gebildet werden keine

Entscheidung getroffen wird

Struktur – Grammatikhierarchien

Typ 1 Grammatiken heißen kontextsensitive Grammatiken

Produktionsregeln ersetzen ein nicht terminales Symbol der linken Seite durch neue Symbol, eventuell in Abhängigkeit der Umgebung des nicht terminalen Symbols

Bei diesen Sprachen kann man entscheiden ob ein Wort zu der Sprache gehört oder nicht

Struktur – Grammatikhierarchien

Typ 2 Grammatiken heißen kontextfreie Grammatiken

Produktionsregeln enthalten auf der linken Seite nur nicht terminale Symbole, Ersetzung hängt also nicht vom Kontext ab

Grundlage der meisten Programmiersprachen

Sie können entscheiden ob ein Wort zur Sprache gehört und ob es nur endlich viele Worte in der Sprache gibt

Struktur – Grammatikhierarchien

Typ 3 Grammatiken heißen reguläre Grammatiken

Produktionsregeln enthalten auf der rechten Seite maximal ein terminales Symbol und ein nicht terminales Symbol, das immer auf der linken Seite steht

Sie können auch entscheiden ob zwei Sprachen äquivalent sind und werden durch endlichen Automaten realisiert

Struktur – Beschreibung von Grammatiken

Grammatiken werden meist durch

- Angabe der Produktionsregeln (Backus – Naur Form)
- Angabe der Struktur regulärer Ausdrücke

beschrieben

Struktur – Semantik

Semantik formaler Sprachen beschäftigt sich wie bei natürlichen Sprachen mit dem Aufbau und der Bedeutung der durch die Grammatik gebildeten Ausdrücke

Man unterscheidet statische und dynamische Semantik

Beschreibung der Semantik oft in natürlicher Sprache

Struktur – Semantik

Statische Semantik

Regeln, die in der Sprache erfüllt sein müssen und formal überprüft werden können

Beispiel: Addition nur für Zahlen sinnvoll

Mojca+Wilfried = ?

Struktur – Semantik

Statische Semantik

Oft ist Typisierung der Ausdrücke sinnvoll

Addition von ganzen Zahlen ist anders sehen als Addition von Dezimalzahlen (Rechengenauigkeit)

Nichtbeachtung kann zu schweren Fehlern führen

Struktur – Semantik

Dynamische Semantik

Semantik, die sich aus der Abarbeitung der Befehle ergibt

Meist in einem eingeschränkten Kontext behandelt: Ist Programm korrekt?

Korrekt bedeutet dabei nur es beendet seine Analyse in endlicher Zeit mit Antwort „richtig“ oder „falsch“

Struktur – Semantik – RDF

Zur Darstellung der Semantik wird im ***Semantic Web*** das Modell des ***Resource Description Framework*** verwendet (RDF)

Jede Semantik ist von der Form:
(Subjekt, Prädikat, Objekt)

Sprache – Gebrauch

Wir können im Gebrauch dieses einfachen Modells unterschiedliche Szenarien verwenden

Sprachparadigmen

Wichtige Paradigmen:

- Imperative Sprachen: Handlungsanweisung, Folge von Befehlen
- Deklarative Sprachen: Spezifikation was berechnet werden soll