

# Computerlinguistik

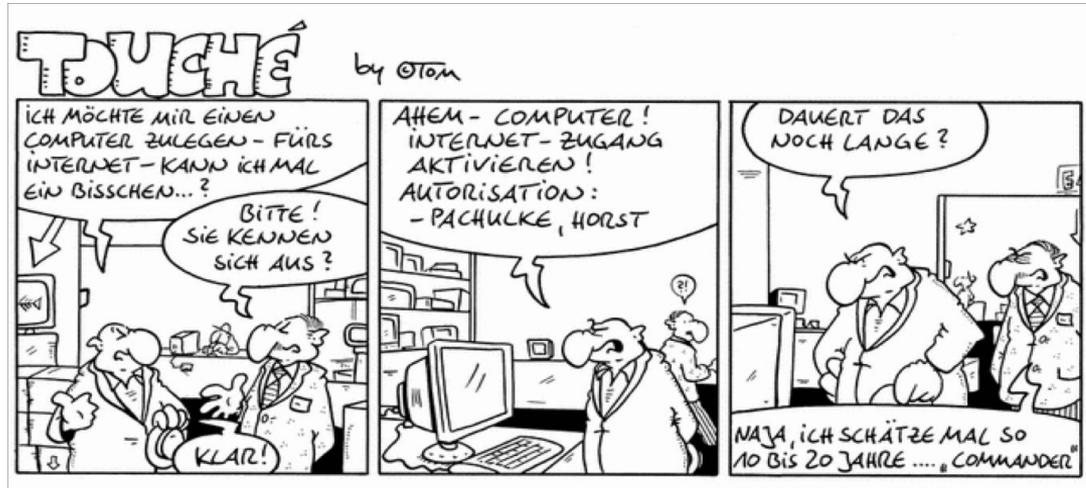
E01: Einführung in die CL

# Was ist Computerlinguistik?

- Definition
- Anwendungen
- Fragestellung
- Nachbardisziplinen
- Wissensbereiche
- Semesterprogramm

# Was ist Computerlinguistik?

- Computerlinguistik aus *wissenschaftlicher* Sicht:  
Schnittstelle zwischen Linguistik und Informatik
- Computerlinguistik aus *anwendungsbezogener* Sicht:  
Schnittstelle zwischen Sprache und Rechnersystemen



# Computerlinguistik in der Anwendung

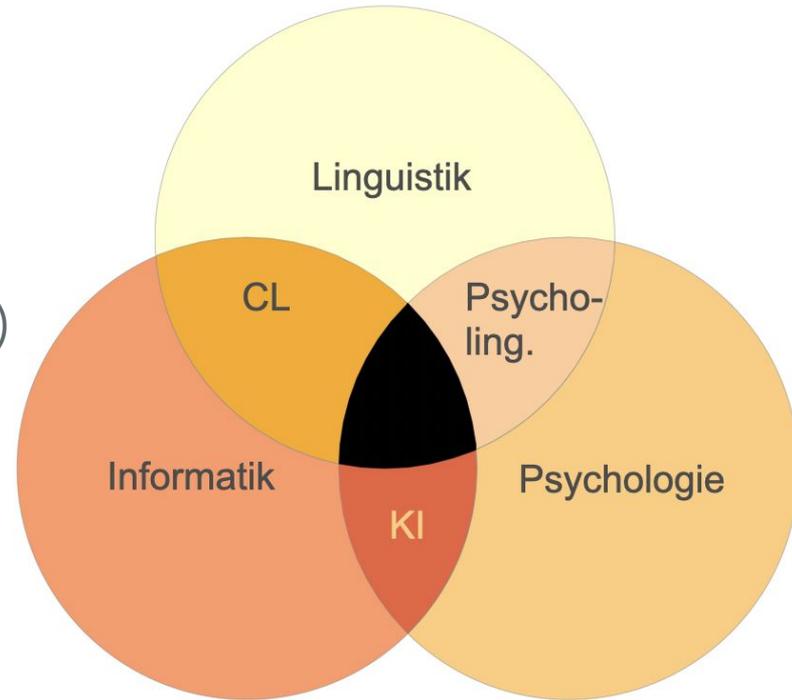
- Mensch-Maschine-Kommunikation
  - Spracherkennung, Sprachsynthese, Dialogsysteme
- Maschinelle bzw. computergestützte Übersetzung
  - Vollautomatische MÜ, Übersetzungsspeicher, Terminologie-Datenbanken
- Information Retrieval & Text Mining
  - Text Classification, Text Summarization, Information Extraction
- Aufbau und Nutzung von Ressourcen
  - Korpora, Lexika, Wortnetze, Baumbanken

# Fragestellungen

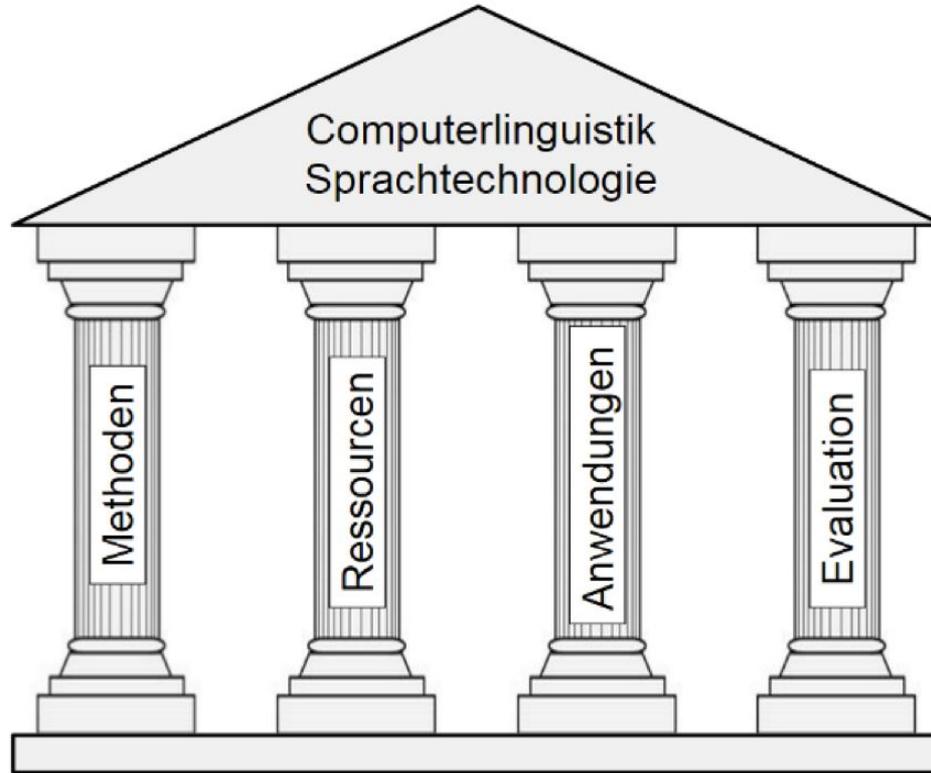
- Computerlinguistik als...
  - Teildisziplin der Linguistik
  - Linguistische Datenverarbeitung
  - Maschinelle Sprachverarbeitung
  - Sprachtechnologie

# Nachbardisziplinen

- Linguistik (Gegenstand, Terminologie)
- Informatik (Algorithmen, Datenstrukturen)
- Philosophie (Verhältnis Sprache-Denken)
- Künstliche Intelligenz (Wissensverarbeitung)
- Kognitionswissenschaft (Sprache-Denken)
- Mathematik (Logik, Graphen, Statistik)



# Grundpfeiler der Computerlinguistik



# Methoden

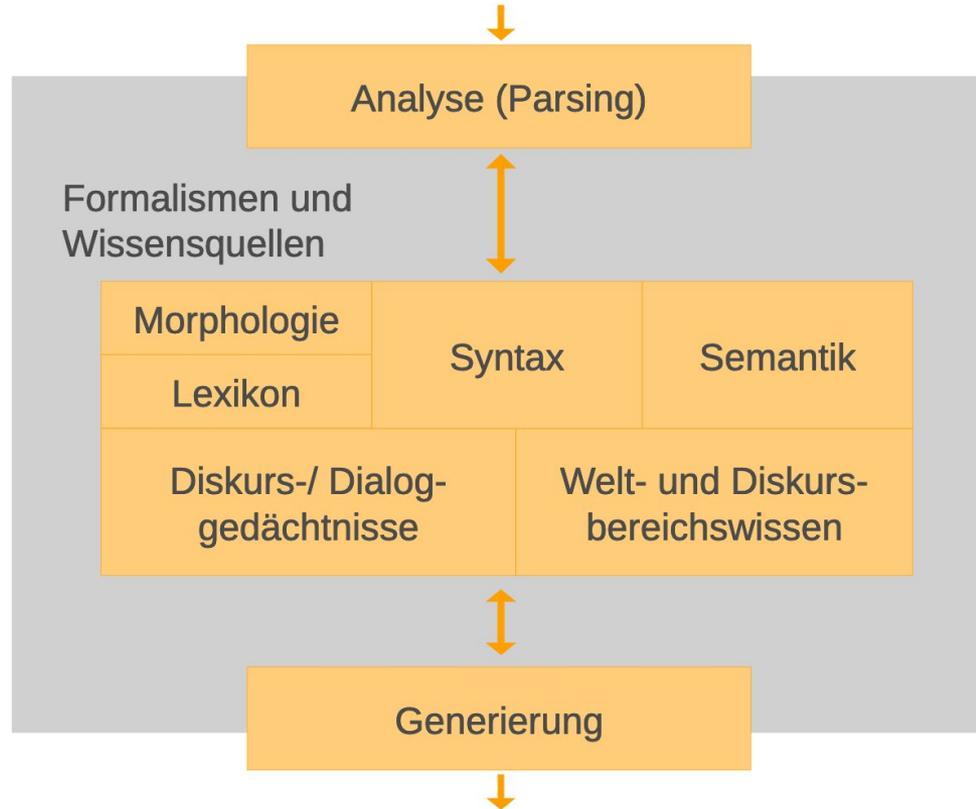
## ➤ Methoden der Linguistik

- Identifikation sprachlicher Einheiten auf unterschiedlichen Ebenen (Phonologie - Morphologie - Syntax)
- Untersuchung des Zusammenspiels der Ebenen mit der Ebene der Bedeutung (Semantik)

## ➤ Methoden der Informatik

- Nutzung etablierter Methoden zur Datenspeicherung und -verarbeitung
- Darunter Automaten- und Graphentheorie, formale Sprachen, Netz- und Vektorrepräsentationen, maschinelle Lernverfahren, Deep-Learning-Architekturen etc.

# Architektur natürlichsprachliches System



# Forschung und Anwendung

## ➤ Theoretische Computerlinguistik

- Suche nach komplexen Beschreibungsformalismen, die gleichzeitig handhabbar und effizient sind.
- Abwägung über das Verhältnis von linguistischer Beschreibung und Berechnungsoperationen.

## ➤ Praktische Computerlinguistik

- Entwicklung und Erforschung realitätsnaher Anwendungen.
- Natürlichsprachliche Kommunikation mit oder auf der Maschine

# Semesterprogramm

- Allgemeine Einführung in die Grundbegriffe von Computerlinguistik und Sprachwissenschaft
- Methoden der (Computer)Linguistik:  
Phonologie, Morphologie, Syntax, Semantik
- **Ressourcen für computerlinguistisches Arbeiten  
und ihre Anwendung**
- **Algorithmen und Formalismen für die Verarbeitung  
natürlichsprachlicher Daten**
- **(Begleitend) Computerlinguistische Anwendungen**

# Literatur / Hausaufgabe

➤ **Zur Nachbereitung:**

- Lesen Sie: Carstensen et al. (2009): Kapitel 1 (S. 1-23)

➤ **Zur Vorbereitung:**

- Lesen Sie: Vater (2002): Kapitel 1 (S. 2-24)

➤ Die Texte finden Sie im Ilias-Seminarordner.

# Computerlinguistik

E02: Einführung in die Linguistik

# Einführung in die Sprachwissenschaft

- **Strukturalismus**  
Sprachwissenschaft von Anfang bis Mitte des 20. Jahrhunderts
- **Kognitivismus**  
Sprachwissenschaft von Mitte bis Ende des 20. Jahrhunderts
- **Gebrauchsbasierte Linguistik**  
Sprachwissenschaft im frühen 21. Jahrhundert

Max Planck: “Eine neue wissenschaftliche Wahrheit pfl egt sich nicht in einer Weise durchzusetzen, dass ihre Gegner überzeugt werden und sich als belehrt erklären, sondern vielmehr dadurch, dass ihre Gegner aussterben.”

# Strukturalismus

- **Sprache als präzise erfassbares, formal exakt darstellbares relationales System von formalen Elementen.** (Zitiert nach Bußmann)
- Begründendes Werk der modernen Sprachwissenschaft deskriptiver Tradition: Ferdinand de Saussure: *Cours de linguistique général* (1916)

# Strukturalismus - Distinktionen I

Sprache kann unter drei verschiedenen Aspekten betrachtet werden:

- **Langue:** In allen Sprechern einer Sprache gespeichertes System (Sprachsystem)
- **Parole:** Aktuelle Sprachtätigkeit in bestimmten Situationen (Sprachgebrauch)
- **Faculté de langage:** Generelle Fähigkeit zum Erwerb und Gebrauch von Sprache

# Strukturalismus - Distinktionen II

Sprache im Sinne von Langue ist ein System von Zeichen (und damit Teildisziplin der *Semiotik*).

Jedes Zeichen besteht aus zwei - sich gegenseitig bedingenden - Aspekten:

- **Signifiant:** Konkreter, materieller Zeichenkörper (Ausdrucksseite - Bezeichnendes)
- **Signifié:** Begriffliches Konzept (Inhaltsseite - Bezeichnetes)

# Strukturalismus - Distinktionen III

- Konzept des distinktiven Prinzips
- synchrone vs. diachrone Sprachwissenschaft
- syntagmatische vs. paradigmatische Ebene
- Segmentierung vs. Klassifizierung

# Kognitivismus (auch: Generative Grammatik)

- Sprache (besser: Sprachfähigkeit) als genetisch determiniertes (menschliches) Organ, das sich in Konfrontation mit Sprachdaten ausbildet.
- Abgrenzung vom Strukturalismus: Nicht nur Beobachtung und Beschreibung der menschlichen Sprache ist Gegenstand der Sprachforschung, sondern v.a. auch die mentalen Grundlagen des Spracherwerbs.
- Begründung der generativen Grammatik: Noam Chomsky - Syntactic Structures (1957)

# Kognitivismus - Grundkonzepte I

- **Universalgrammatik (UG):** Biologisch determinierter kognitiver Ausgangszustand des Sprachorgans
- **Kompetenz:** Fähigkeit eines idealisierten Sprecher/ Hörers, Sprache anzuwenden, entwickelt sich aus UG
- **Performanz:** Tatsächliche Sprachverwendung, teils fehlerhaft.

# Kognitivismus - Grundkonzepte II

- Organisation der Universalgrammatik in
  - **Prinzipien** (sprachübergreifende Wohlgeformtheitsbedingungen)
  - **Parameter** (sprachspezifische Parametrisierungen)
- Organisation der Kompetenz in Modulen
  - phonologisches Modul
  - morphologisches Modul
  - syntaktisches Modul
  - semantisches Modul
  - ...

# Gebrauchsbasierte Linguistik - Grundkonzepte

- Grammatische Strukturen sind nicht angeboren, sondern das Ergebnis von Geschichte und Psychologie.
  - Geschichte: Sprachen werden von einer Generation zur nächsten tradiert.
  - Psychologie: Jede Generation besitzt soziale und kognitive Fähigkeiten für den Spracherwerb, die aber nicht speziell für diesen Zweck entstanden sind.
- Sprachvermögen basiert auf gebrauchbasierten Mehrzweckmodulen: Bildung von Kategorien, Deutung von Absichten, Erkennen von Analogien (Eingrenzungsmechanismen)
- Mit empirischen Methoden (Auswertung von Sprachdatensammlungen) untersuchbar.

# Literatur / Hausaufgabe

## ➤ **Zur Nachbereitung:**

- Lesen Sie: Vater (2002): Kapitel 1 (S. 2-24)
- Ibbotson/Tomasello (2020, S. 14-19)

## ➤ **Zur Vorbereitung:**

- Lesen Sie: Vater (2002): Kapitel 2 (S. 25-60)

## ➤ Die Texte finden Sie im Ilias-Seminarordner.

# Computerlinguistik

E04: Phonetik und Phonologie

# Ebenen der sprachlichen Beschreibung

	Zeichenebene	Wortebene	Satzebene
Ausdrucksseite (signifiant)	<b>Phonetik / Phonologie</b>	<b>Morphologie</b>	<b>Syntax</b>
Inhaltsseite (signifié)		<b>Wortsemantik</b>	<b>Satzsemantik</b>

# Phonetik & Phonologie

## ➤ Abgrenzung

## ➤ Phonetik

- Unterdisziplinen
- Artikulationsorte & Artikulationsarten
- Konsonanten & Vokale

## ➤ Phonologie

- Phoneme & Allophone
- Phonologische Merkmale
- Phonologische Prozesse

# Abgrenzung von Phonetik und Phonologie

- **Phonetik:** Untersuchung der physiologischen und physikalischen Eigenschaften von Lauten.
- **Phonologie:** Untersuchung der Verhältnisse der Laute zueinander innerhalb eines Sprachsystems.

# Phonetik: Unterdisziplinen

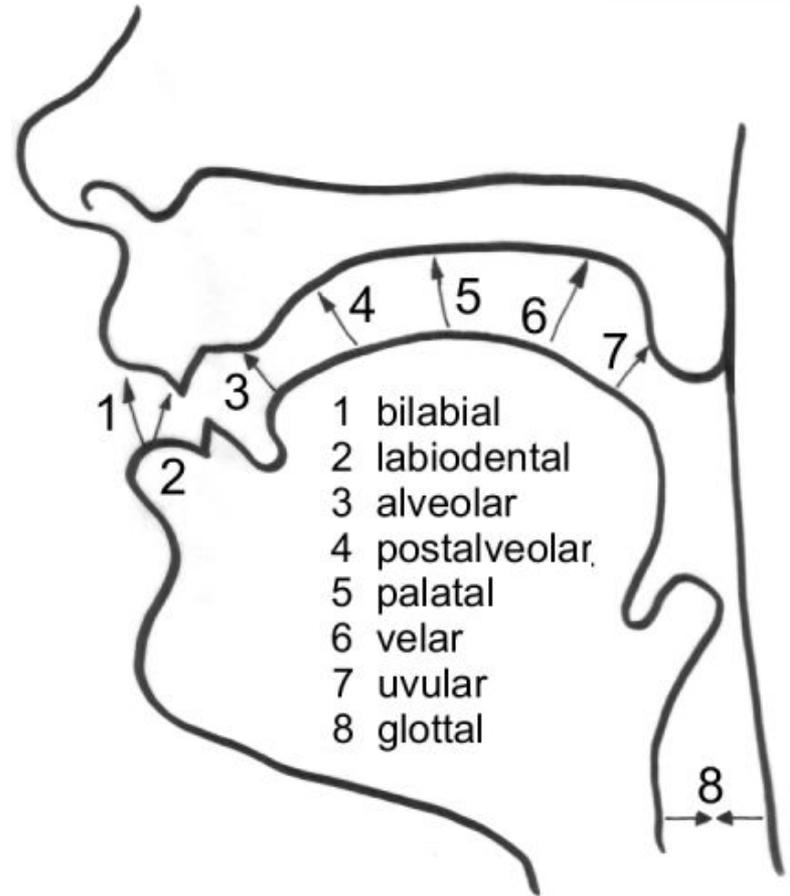
- **Artikulatorische Phonetik:** Untersuchung der Produktion von Lauten mithilfe von Sprechwerkzeugen.
- **Akustische Phonetik:** Untersuchung der physikalischen Eigenschaften von Lauten innerhalb eines Mediums.
- **Auditive oder perzeptive Phonetik:** Untersuchungen zur Wahrnehmung von Lauten.

# Klassifikation der Artikulatorischen Phonetik

Artikulatorisch gesehen können Laute basierend auf vier Eigenschaften klassifiziert werden:

- Zustand der Stimmbänder
- Weg des Luftstroms
- Artikulationsort
- Artikulationsart

# Artikulationsorte



# Artikulationsarten

- Verschlusslaute (Plosive)
- Frikative (Reibelaute)
- Nasalkonsonanten
- Laterale
- Vibranten

# Konsonanten

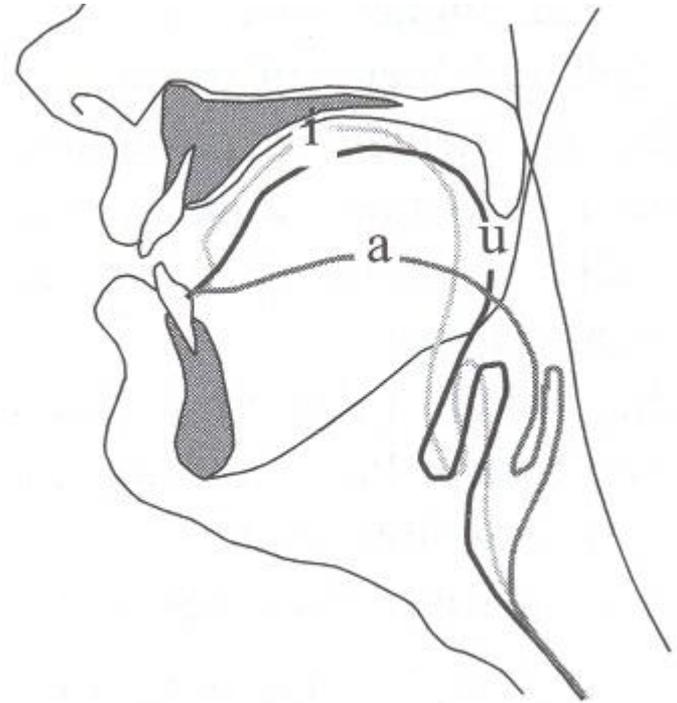
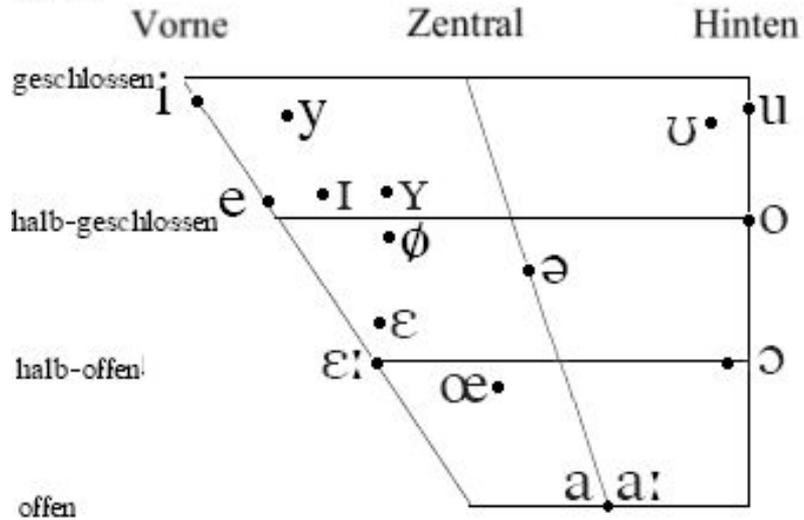
## IPA

	Labial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b		t d					k g			ʔ
Nasale	m		n					ŋ			
Trill			r						R		
Tap oder Flap											
Frikative		f v		s z	ʃ ʒ		ç	x	ʁ		h
Lateral-Frikative											
Approximanten							j				
Lateral-Approximanten			l								

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

# Vokale

## IPA



Bildquelle: [http://www.coli.uni-saarland.de/elaut/Languages\\_Sites/sampaDeutsch.htm](http://www.coli.uni-saarland.de/elaut/Languages_Sites/sampaDeutsch.htm)

# Phonologie - Phoneme & Allophone

- **Psychologischer Ansatz** - Der Ideallaut
- **Physikalischer Ansatz** - Die Lautfamilie
- **Abstrakter Ansatz** - Die Distribution
- **Funktionaler Ansatz** - Die Bedeutungs differenzierung  
→ Phoneme sind die kleinsten bedeutungsunterscheidenden Einheiten innerhalb eines Sprachsystems.
- Allophone sind unterschiedliche Ausprägungen eines Phonems (freie, regionale und komplementäre Allophone)

# Phonologische Merkmale

- Phoneme sind zerlegbar in Merkmale
- Merkmale sind binär
- Merkmale basieren auf artikulatorischen Unterschieden
- Jedes Phonem jeder Sprache lässt sich durch eine spezifische Merkmalsmatrix von anderen unterscheiden
- Wichtigste phonologische Merkmale: konsonantisch, sonorant, koronal, dauernd, stimmhaft, gerundet, gespannt, nasal, lateral, hinten, hoch, tief

# Merkmale deutscher Konsonanten

	p	b	f	v	m	t	d	s	z	n	l	r	ʃ	ç	j	k	g	x	ŋ	R
kons	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
son	-	-	-	-	+	-	-	-	-	+	+	+	-	-	-	-	-	-	+	+
dnd	-	-	+	+	-	-	-	+	+	-	+	+	+	+	+	-	-	+	-	+
kor	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
hoch	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-
hint	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
nas	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-
sth	-	+	-	+	+	-	+	-	+	+	+	+	-	-	+	-	+	-	+	+

# Phonologische Prozesse

- Realisierung von Phonemen in Sequenzen triggern phonologische Prozesse
- Diachronische vs. synchronische Prozesse
- Klassifikation phonologischer Prozesse
  - Assimilation vs. Dissimilation
  - Tilgung (Ellipse) vs. Insertion (Epithese)
  - Metathese
- Allgemeines Regelschema:  $A \rightarrow B / X\_Y$  (gesprochen: A wird zu B nach X vor Y)

# Literatur / Hausaufgabe

- **Zur Nachbereitung:**
  - Lesen Sie: Vater (2002): Kapitel 2 (S. 25-60)
- **Zur Vorbereitung:**
  - Lesen Sie: Vater (2002): Kapitel 3 (S. 61-74)
- Die Texte finden Sie im Ilias-Seminarordner.

# Computerlinguistik

E04: Morphologie

# Morphologie

- **Begriffe: Wort - Morphem - Allomorph**
- **Flexion und Wortbildung**
- **Sprachtypologie anhand der Wortstruktur**
- **Morphologische Analyse**
- **Generative Morphologie**
- **Morphologie mit endlichen Automaten**
- **Morphologie mit Default-Vererbungsnetzen**

# Begriffe: Wort - Morphem - Allomorph

- lexikalisches Wort (Lexem) vs. flexivisches Wort (Wortform)
- phonologisches Wort vs. syntaktisches Wort
- **Morphem**, funktionale Definition: Kleinste bedeutungstragende Einheit einer Sprache
  - freie vs. gebundene Morpheme
  - Wurzelmorpheme vs. Affixe
  - Derivationsaffixe vs. Flexive
- **Allomorphe**: Bedeutungs- und funktionsgleiche Varianten eines Morphems

# Flexion und Wortbildung

- **Flexion:** Bildung flexivischer Wörter
  - Konjugation (Verben) - Deklination (Nomen und Adjektive) - Komparation (Adjektive)
- **Wortbildung:** Bildung lexikalischer Wörter
  - Wichtigste Formen: Komposition - Derivation
  - Weitere Formen: Konversion - Abkürzung - Amalgamierung ( - Entlehnung)

# Sprachtypologie anhand der Wortstruktur

- **Isolierende Sprachen:** 1 Wort = 1 Bedeutung (keine Flexion)
- **Flektierende Sprachen:** Stämme + Flexive mit mehreren Bedeutungen
- **Agglutinierende Sprachen:** 1 Morphem = 1 Bedeutung (feste Flexive)
- **Polysynthetische Sprachen:** Verschmelzung von grammatischen Einheiten zu einer Worteinheit (inkorporierend).

# Analytische vs. generative Morphologie

- **Analytische Morphologie:** Identifikation von Wörtern / Morphemen in Äußerungen (Ausdrucksseite → Inhaltsseite)
- **Generative Morphologie:** Ableitung von Wörtern aus dem gegebenen Input (Inhaltsseite → Ausdrucksseite)

# Morphologische Analyse

- **Segmentierung:** Zerlegung des Inputs in Morpheme
- **Klassifizierung:** Zuordnung der identifizierten Morpheme zu Klassen
- **Strukturierung:** Konstruktion des hierarchischen Aufbaus der Morpheme bzw. Morphemkomplexe

# Grundmodelle der generativen Morphologie

- **Morphembasierter Ansatz:** Kombination einzelner Morpheme zu vollständigen flexivischen Wörtern
- **Wortbasierter Ansatz:** Bildung von flexivischen Wörtern aus Stammformen (z.B. dem Infinitiv)
- **Realisierungsbasierter Ansatz:** Realsierung flexivischer Wörter aus einer vorgegebener Bedeutung

# Morphologie mit endlichen Automaten

- Beispiel für den morphembasierten Ansatz
- Automatentheorie: Typ-3-Sprachen (beschränkte Syntax, einfache Prozessierung)
- Gebrauch von Finite-State-Transducern: Repräsentation des lexikalischen Wissens in Kanten, die Pfade übersetzen.
- Für Analyse und Generierung einsetzbar.
- Limitationen bei flektierenden Sprachen: Nicht-konkatenative Phänomene und nicht-lokale Abhängigkeiten

# Flexionsparadigmen

## WARTEN:

Ich wart-e wart-et-e

Du wart-est wart-et-est

Es wart-et wart-et-**e**

Wir wart-en wart-et-en

Ihr wart-et wart-et-et

Sie wart-en wart-et-en

Partizip: ge-wart-et

## HALTEN:

Ich halt-e hielt-**ø**

Du hält-st hielt-est

Es hält-**ø** hielt-**ø**

Wir halt-en hielt-en

Ihr halt-et hielt-et

Sie halt-en hielt-en

Partizip: ge-halt-**en**

# Morphologie mit Default-Vererbungsnetzen

- Beispiel für den realisierungsbasierten Ansatz
- Automatentheorie: Typ-0-Sprache (reiche Syntax, eingeschränkte Prozessierung)
- Repräsentationen des lexikalischen Wissens in Knoten, die Pfade Werte zuweisen.
- Gebrauch von Defaultmechanismen & Neutralisationen
- Limitationen: Nur für die Generierung, nicht für die Analyse einsetzbar.

# Wissensrepräsentation in DATR

**WARTEN:**

**<form präs sg eins> == warte**

**<form präs sg zwei> == wartest**

**<form präs sg drei> == wartet**

**(...)**

**.**

# DATR: Resolutionspfade

## WARTEN:

**<form> == <wurzel> <endung>**

**<wurzel> == wart**

**<endung präs sg eins> == e**

**<endung präs sg zwei> == est**

**<endung präs sg drei> == et**

**<endung präs pl eins> == en**

**<endung präs pl zwei> == et**

**<endung präs pl drei> == en**

.

# DATR: Redundanzen

## WARTEN:

<form> == <wurzel> <endung>

<wurzel> == wart

<endung **präs** sg eins> == e

<endung **präs** sg zwei> == est

<endung **präs** sg drei> == et

<endung **präs** pl eins> == en

<endung **präs** pl zwei> == et

<endung **präs** pl drei> == en

.

# DATR: Defaults und Neutralisierung

## WARTEN:

<b>&lt;form&gt;</b>	<b>==</b>	<b>&lt;wurzel&gt; &lt;endung&gt;</b>
<b>&lt;wurzel&gt;</b>	<b>==</b>	<b>wart</b>
<b>&lt;endung sg&gt;</b>	<b>==</b>	<b>e</b>
<b>&lt;endung sg zwei&gt;</b>	<b>==</b>	<b>est</b>
<b>&lt;endung präs sg drei&gt;</b>	<b>==</b>	<b>et</b>
<b>&lt;endung pl&gt;</b>	<b>==</b>	<b>en</b>
<b>&lt;endung pl zwei&gt;</b>	<b>==</b>	<b>et</b>
<b>&lt;endung präs&gt;</b>	<b>==</b>	<b>&lt;endung&gt;</b>
<b>&lt;endung prät&gt;</b>	<b>==</b>	<b>et &lt;endung&gt;</b>

.

# DATR: Klassenbildung

## VERB\_SCHWACH\_T:

<form> == <wurzel> <endung>

<endung sg> == e

<endung sg zwei> == est

<endung präs sg drei> == et

<endung pl> == en (...)

.

## WARTEN:

<> == VERB\_SCHWACH\_T

<wurzel> == wart

.

# DATR: Nichtlokale Abhängigkeiten

## VERB\_SCHWACH\_T:

<form> == <wurzel> <endung>

<form partizip> == ge <wurzel> et.

<endung sg> == e

<endung sg zwei> == est

<endung präs sg drei> == et

<endung pl> == en (...)

.

# DATR: Nichtkonkative Phänomene

## VERB\_STARK:

<form> == <wurzel> <endung>  
<wurzel> == <w\_anf> <w\_vok> <w\_end>  
<endung sing> == Ø  
<endung präs sing eins> == e  
<endung sing zwei> == st  
<endung plur> == en  
<endung plur zwei> == et

.

## HALTEN:

<> == VERB\_STARK  
<w\_anf> == h  
<w\_end> == lt  
<w\_vok> == a  
<w\_vok prät> == ie <prät>  
<v\_vok präs sg zwei> == ä <präs sg zwei>  
<v\_vok präs sg drei> == ä <präs sg drei>

.

# Literatur / Hausaufgabe

## ➤ **Zur Nachbereitung:**

- Lesen Sie: Vater (2002): Kapitel 3 (S. 61-74)
- Lesen Sie: Carstensen et al. (2010): Kapitel 3.2 (S. 190-217)
- Bearbeiten Sie die schriftlichen Aufgaben in ILIAS

## ➤ **Zur Vorbereitung:**

- Lesen Sie: Ramers (2000): Kapitel 1 (S. 11-34)
- Die Texte finden Sie im Ilias-Seminarordner.

# Computerlinguistik

E05: Syntax

# Syntax

- **Begriffe: Grammatik – Satz – Phrase**
- **Konstituenten – Ermittlung und Klassifikation**
- **Phrasenstrukturen**
- **Dependenzstrukturen**
- **Referat: Grammatikprüfung**
- **Referat: Textgenerierung**



# Konstituenten: Ermittlung

- Ersetzungsprobe
- Pronominalisierungstest
- Fragetest
- Permutationstest (Verschiebeprobe)
- Tilgungstest

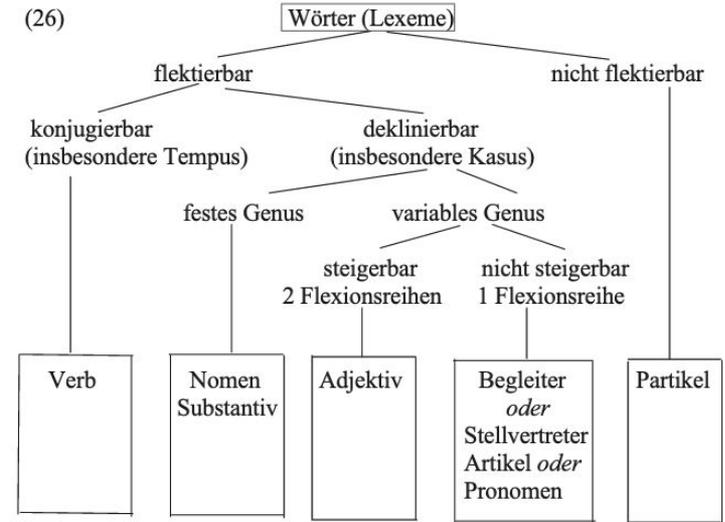
# Konstituenten: Klassifikation

## ➤ Wortartenklassifikation

- morphosyntaktische Eigenschaften (Merkmale)
- syntaktische Eigenschaften (Distribution)

## ➤ Klassifikation von Phrasen

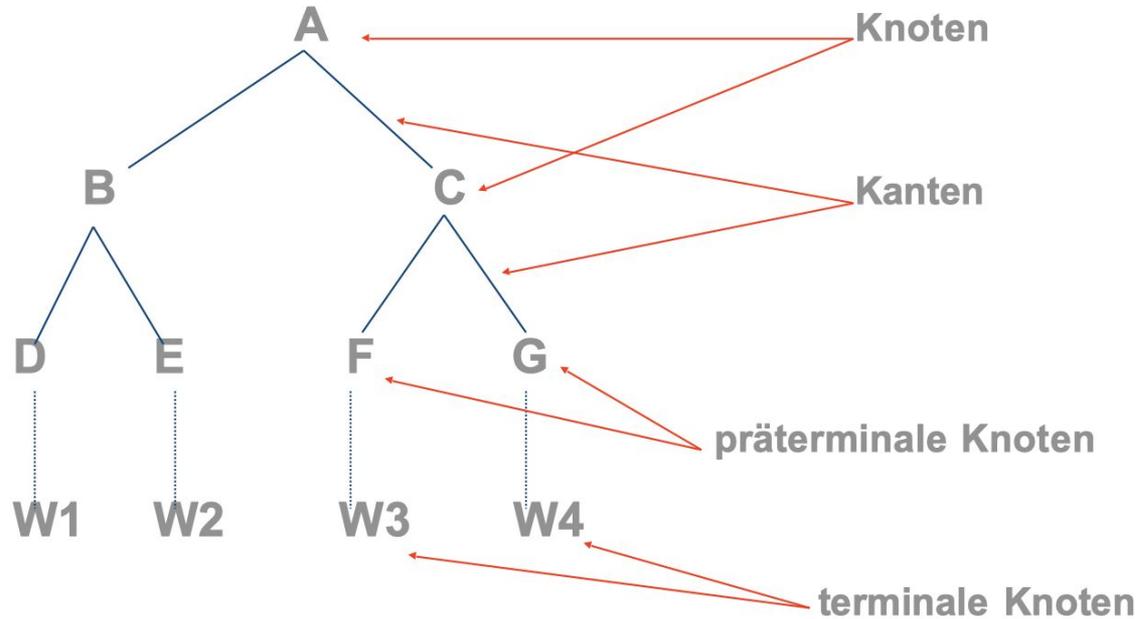
- Nominalphrasen
- Verbalphrasen
- Adjektivphrasen
- Präpositionalphrasen



aus Ramers (2007:24)

# Phrasenstrukturbäume

Baumdarstellung:



Klammerstrukturdarstellung: [A[B[D[W1] E[W2]] C[F[W3] G[W4]]]]

# Phrasenstrukturregeln - Beispiele

## ➤ Syntaktische Regeln:

- $S \rightarrow NP VP$
- $NP \rightarrow D N$
- $VP \rightarrow V NP$

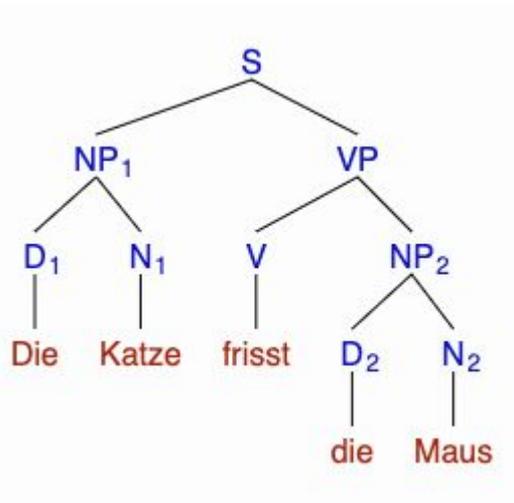
## ➤ Lexikalische Regeln:

- $D \rightarrow \text{die}$
- $N \rightarrow \text{katze}$
- $N \rightarrow \text{maus}$
- $V \rightarrow \text{frisst}$

# Phrasenstrukturregeln - Beispiele

## ➤ Syntaktische Regeln:

- $S \rightarrow NP VP$
- $NP \rightarrow D N$
- $VP \rightarrow V NP$



## ➤ Lexikalische Regeln:

- $D \rightarrow \text{die}$
- $N \rightarrow \text{katze}$
- $N \rightarrow \text{maus}$
- $V \rightarrow \text{frisst}$

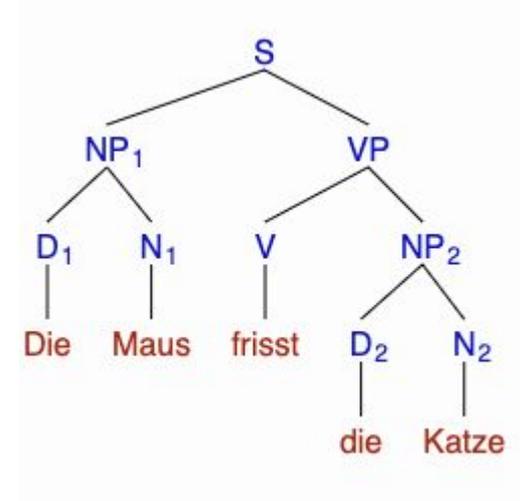
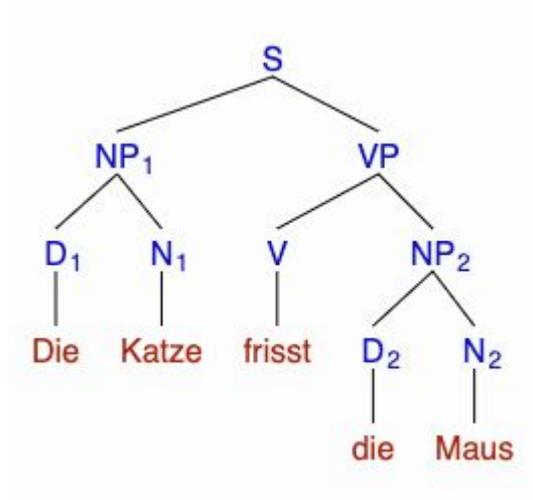
# Phrasenstrukturregeln - Beispiele

## ➤ Syntaktische Regeln:

- $S \rightarrow NP VP$
- $NP \rightarrow D N$
- $VP \rightarrow V NP$

## ➤ Lexikalische Regeln:

- $D \rightarrow \text{die}$
- $N \rightarrow \text{katze}$
- $N \rightarrow \text{maus}$
- $V \rightarrow \text{frisst}$



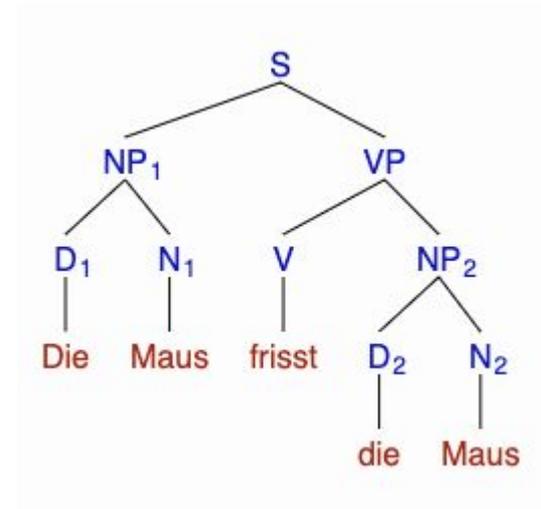
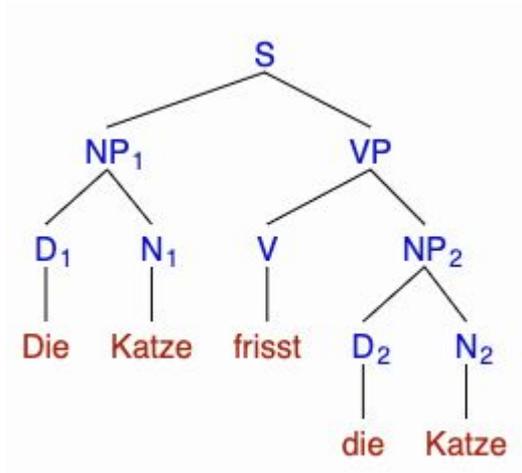
# Phrasenstrukturregeln - Beispiele

## ➤ Syntaktische Regeln:

- $S \rightarrow NP VP$
- $NP \rightarrow D N$
- $VP \rightarrow V NP$

## ➤ Lexikalische Regeln:

- $D \rightarrow \text{die}$
- $N \rightarrow \text{Katze}$
- $N \rightarrow \text{Maus}$
- $V \rightarrow \text{frisst}$



# Phrasenstrukturregeln - Schema

## ➤ Kontextfreie Regeln

- Links muss genau ein Kategorialesymbol stehen
- Rechte muss mindestens ein Kategorialesymbol- oder ein lexikalisches Symbol stehen

## ➤ Rekursive Strukturen sind zulässig (z.B. NP → D N NP)

## ➤ Subkategorisierungsrahmen für Verben muss eingehalten werden

- wohnen[PP] → “Das Kind wohnt in Düsseldorf.” vs. “\*Das Kind wohnt.”
- schenken[NP, NP] → “Die Mutter schenkt dem Kind den Ball.” vs. “\*Die Mutter schenkt dem Kind.”
- rufen[NP] → “Der Mann ruft die Frau.” vs. “\*Der Mann ruft die Frau den Hund.”

# Abhängigkeitsstrukturen - Schema

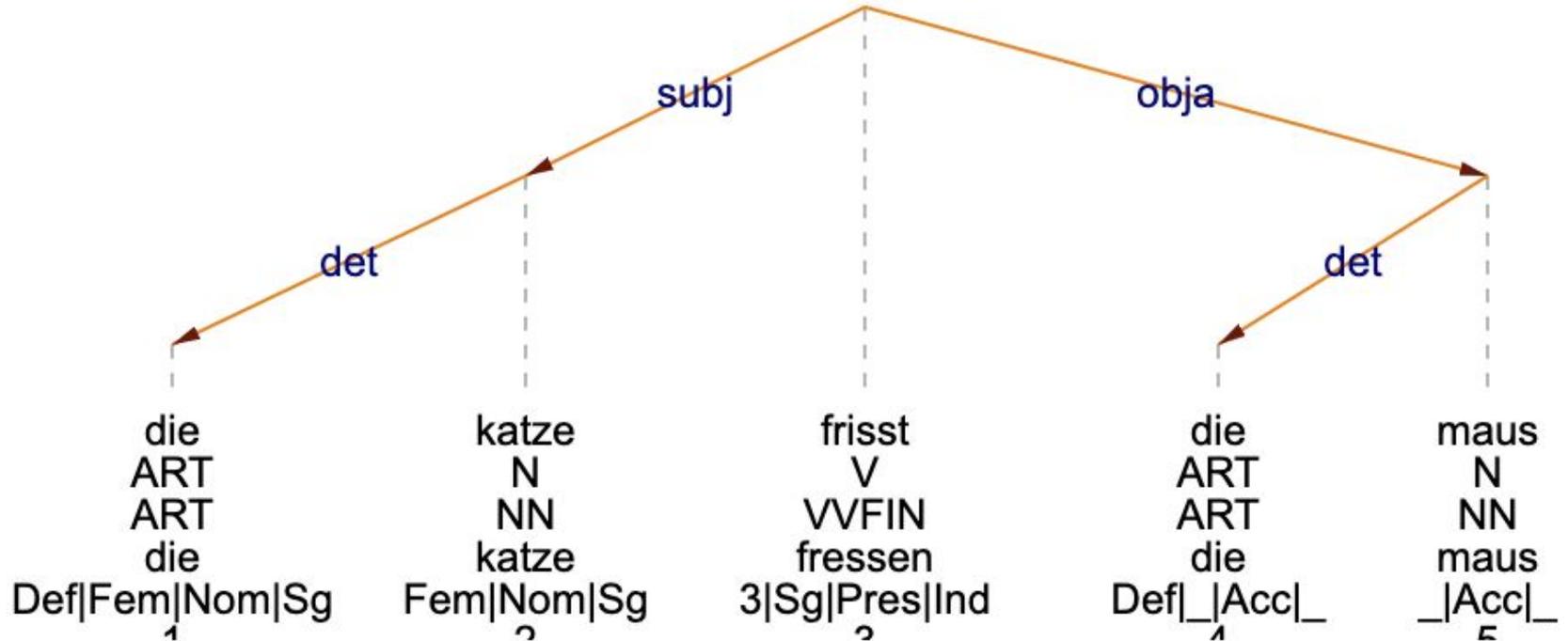
## ➤ Valenzgrammatik (Tesnière 1959)

- Abhängigkeitsrelationen (Stemmata): Regens - Dependens (1:n)
- Verb als Zentralnexus des Satzes mit
  - Aktanten (Akteure des Geschehens) und
  - Angaben (Umstände des Geschehens)

## ➤ Typisierte Abhängigkeiten

- Relationen zwischen Knoten werden durch Kanten ausgedrückt
- Typ der Relation ist Label der Kante

# Dependenzstrukturen - Beispiel



# Literatur / Hausaufgabe

- **Zur Nachbereitung:**
  - Lesen Sie: Ramers (2007): Kapitel 1, 2 und 4 (S. 11-55 und 77-88)
  - Bearbeiten Sie die schriftlichen Aufgaben in ILIAS
  - Testen Sie den ParZu-Parser <https://pub.cl.uzh.ch/demo/parzu/>
- Die Texte (bzw. Links) finden Sie im Ilias-Seminarordner.

# Computerlinguistik

E06: Semantik

# Semantik

- Einführung
- Wortsemantik vs. Satzsemantik
- Merkmaltheorie
- Prototypentheorie
- Semantische Relationen

	Zeichenebene	Wortebene	Satzebene
Ausdrucksseite (signifiant)	<b>Phonetik / Phonologie</b>	<b>Morphologie</b>	<b>Syntax</b>
Inhaltsseite (signifié)		<b>Wortsemantik</b>	<b>Satzsemantik</b>

# Semantik - Einführung

- Sprachliches Zeichen → Semiotisches Dreieck
- Mentales Lexikon: Teil des Langzeitgedächtnisses
- Konzepte: Bausteine unseres Wissens, basieren auf Erfahrungen
  - Kategorienkonzepte (Types): Repräsentanten von Informationen über Klassen -  
Relationsbeziehung: Äquivalenz
  - Partikularkonzepte (Token): Repräsentanten von Informationen über Einzelnes -  
Relationsbeziehung: Identität

# Wortsemantik vs. Satzsemantik

## ➤ Lexikalische Semantik (oder Wortsemantik):

- Kontextunabhängige Bedeutungen von Wörtern
- Relationen zwischen Bedeutungen (Sinnrelationen)

## ➤ Satzsemantik:

- Prinzip der Kompositionalität:

$$\begin{array}{r} \text{Wortbedeutung} \\ + \quad \text{Art der Zusammensetzung} \\ \hline = \quad \text{Satzbedeutung} \end{array}$$

# Merkmalsemantik

- **Merkmalthypothese:** Bedeutungen sind nicht ganzheitlich, sondern aus elementaren Inhaltselementen zusammengesetzt. Diese Inhaltselemente nennt man *semantische Merkmale* oder *Seme*.
- **Merkmalbündel:** Innere Struktur von Bedeutungen
- **Semantische Opposition:** Gegenüberstellung distinktiver Bedeutungsmerkmale
- Ermittlung wesentlicher Merkmale: Kriterium der Notwendigkeit
- Probleme der Merkmalthypothese: Relationale Merkmale - Verben - kognitive Plausibilität

# Prototypensemantik

- **Prototypenhypothese:** Prototypen sind mentale Repräsentationen von Kategorien, die nicht klar umgrenzt werden können. Kontinuum der Kategoriezugehörigkeit.
- Entstehung von Prototypen: Frequenz und Relevanz
- Problem: Mentale Bilder oder abstrakte Repräsentationseinheiten
- Alternative oder Ergänzung zur Merkmalsemantik?

# Semantische Relationen

- **Synonymie:** Bedeutungsgleichheit (gleiche denotative Merkmale, ggfs. unterschiedliche konnotative Merkmale)
- **Referenzidentität:** Gleicher Referent, ggfs. unterschiedliche Bedeutungen
- **Ambiguität:** Unterschiedliche Bedeutungen einer Wortform (Polysemie vs. Homonymie; Homophonie vs. Homographie)
- **Hyponymie** (Unterordnung) vs. **Hyperonymie** (Überordnung) vs. **Kohyponymie** (Nebenordnung)
- **Kontradiktion** (keine Zwischenstufen, nicht steigerbar) vs. **Antonymie** (Zwischenstufen, steigerbar)

# Literatur / Hausaufgabe

- **Zur Nachbereitung:**
  - Lesen Sie: Schwarz/Chur (1996) Kap. 1 - 2.2 (S. 13-60)
  - Lösen Sie die Hausaufgabe in ILIAS (testen Sie v.a. den Dokumentupload)
- Die Texte (bzw. Links) finden Sie im Ilias-Seminarordner.