

Universität Heidelberg

Peter Hellwig

Einführung in die Computerlinguistik*

* wie ich sie mir vorstelle

basierend auf der letzten Version

meiner Vorlesung

im WS 2004/05

Siehe auch Aufgaben und Lösungen

free courseware

Kontakt: hellwig@cl.uni-heidelberg.de

Abstract

The aim of this course (written in German) is a very detailed description of the components of the language system, how the structure of each component can be detected and how the subsystems work together. Phonology, intonation, graphemes, morphology, syntax and paradigmatic relations, formal grammar, lexical and referential semantics, the logical form of sentences, dialogs, cohesion and coherence of texts and speech acts are conflated into a consistent theory. On this basis each aspect is looked at for its potential of being processed by computer.

Von Babelfish.de übersetzt (2021):

Ziel dieses Kurses (in deutscher Sprache) ist eine sehr detaillierte Beschreibung der Komponenten des Sprachsystems, wie die Struktur jeder Komponente erkannt werden kann und wie die Subsysteme zusammenarbeiten. Phonologie, Intonation, Grapheme, Morphologie, Syntax und paradigmatische Beziehungen, formale Grammatik, lexikalische und referentielle Semantik, logische Satzform, Dialoge, Zusammenhang und Kohärenz von Texten und Sprechakten werden zu einer konsistenten Theorie zusammengeführt. Auf dieser Grundlage wird jeder Aspekt auf sein Verarbeitungspotential durch Computer untersucht.

Inhaltsverzeichnis (zum Direktzugang auf Zeilen klicken)

I 1 Grundlagen der Hardware	1
1. Rechenmaschinen	1
2. Programmierbare Rechenmaschinen.....	3
3. Computer und Geist? Kann man mit Sprache rechnen?	8
I 2 Grundlagen der Software und der Programmierung	9
1. Die Software-Schichten.....	9
2. Phasen bei der Erstellung eines Programms:	9
S 1.1 Phonologie und artikulatorische Phonetik.....	17
1. Aufgabe einer (strukturalistischen) Beschreibung der Laute	17
2. Die psychische und die physische Ebene der Sprache	18
3. Die strukturalistische Methode	18
S 1.2 Akustische Phonetik und maschinelles Erkennen und Erzeugen gesprochener Sprache	24
1. Auditive und akustische Phonetik.....	24
2. Maschinelle Erkennung gesprochener Sprache	28
S 2.1 Grundlagen der Graphemik und maschinelle Schrifterkennung	32
1. Grundlagen der Graphemik.....	32
2. Maschinelles Lesen (Optical Character Recognition, OCR)	38
S 2.2 Orthographie und maschinelle Rechtschreibhilfen.....	39
S 3.1 Silbenbau und Prosodie	47
1. Silbenbau.....	47
2. Prosodie	53
3. Funktion der Tonbrüche	54
S 3.2 Maschinelles Zeilenumbruch mit Silbentrennung	55
S 4.1 Grundlagen der Morphologie , Morphologische Phänomene	57
S 4.2 Maschinelle morpho-syntaktische Analyse und Klassifizierung	64
1. Komplexe Kategorien zur Klassifizierung von Morphen.....	64
2. Modelle für die maschinelle morphologische Analyse	66
S 5.1 Heuristische Grundlagen der Syntaxanalyse	72
S 5.2 Formale Syntaxbeschreibung	82
S 5.3 Maschinelle Syntaxanalyse (Parsing).....	89
S 6.1 Aussagenlogik.....	96
1. Gegenstand der Logik.....	96
2. Aussagenlogischer Kalkül	98
3. Logik aus pragmatischer Sicht (Sprachhandeln, Sprachspiel).....	100
4. Die möglichen Welten	102
S 6.2 Prädikatenlogik.....	105
1. Gegenstand der Prädikatenlogik.....	105
2. Prädikatenlogischer Kalkül – formale Definition	109

S 7.1 Grundlagen der lexikalische Semantik	112
1. Lexikalische Semantik.....	112
2. Beschreibungsansätze.....	116
S 7.2 Referentielle Semantik	123
1. Was ist referentielle Semantik?.....	123
2. Der Mechanismus des Referierens	128
S 7.3 Satzsemantik	133
1. Abgrenzung	133
2. Ausgewählte semantische Beziehungen aufgrund von Satzstrukturen.....	135
S 8.1: Grundlagen der Pragmatik und Sprechakttheorie	137
1. Sprechen als Handeln.....	137
2. Sprechaktanalyse	145
3. Pragmatische Bedingungen	150
S 8.2: Grundlagen der Textkonstitution und des Textverstehens	152
1. Textkonstitution.....	152
2. Sonderfall monologische Texte	154
4. Textverarbeitung	158
S 8.3 Wissensrepräsentationen	159
1. Grundlagen	159
2. Repräsentationsformate	161
A 1.1 Grundlagen der Sprachübersetzung und die Arbeitswelt des Übersetzers und Dolmetschers	169
A.1.2 Maschinelle Übersetzungssysteme	177
A.2.1 Grundlagen des Dokumentationswesens und des Information Retrieval	180
1. Dokumentationswesen.....	180
2. Dokument Retrieval mit Computer	183
A 2.2 Automatische Texterschließung und Textzusammenfassung	186
A 3.1 Grundlagen sogenannter Expertensysteme	195
A 3.2 Natürlichsprachige Frage-Antwort-Systeme (FAS).....	201
A.4.1 Computer in der Lehre und Ausbildung, "e-Learning"	205
A 4.2 Sprachlernsoftware	208
K 1 Wissenschaftstheorie und Methodologie der CL	210
K 2 Computer und Denken, Künstliche Intelligenz	224
K 3 Technologiefolgenabschätzung.....	230
Checkliste - Das Wichtigste aus der Vorlesung	238

I 1 Grundlagen der Hardware

Lernziel: Eine Vorstellung davon bekommen, was ein Computer ist und was er macht, mehr vom Prinzip her gesehen als unter technischen Gesichtspunkten.

1. Rechenmaschinen

Ein Computer ist eine Rechenmaschine. **Was ist Rechnen?**

$$\begin{array}{r} 21 \times 52 \\ \hline 105 \\ 42 \\ \hline 1092 \end{array}$$

Zum Beispiel eine Multiplikation

Regeln, z.B.

$$\begin{array}{l} 1 \times 5 = 5 \\ 2 \times 5 = 10 \\ 1 \times 2 = 2 \\ 2 \times 2 = 4 \end{array} \qquad \begin{array}{l} 5 + 4 = 9 \\ \\ \\ \text{usw.} \end{array}$$

Vorschrift zum Ausführen der Rechnung

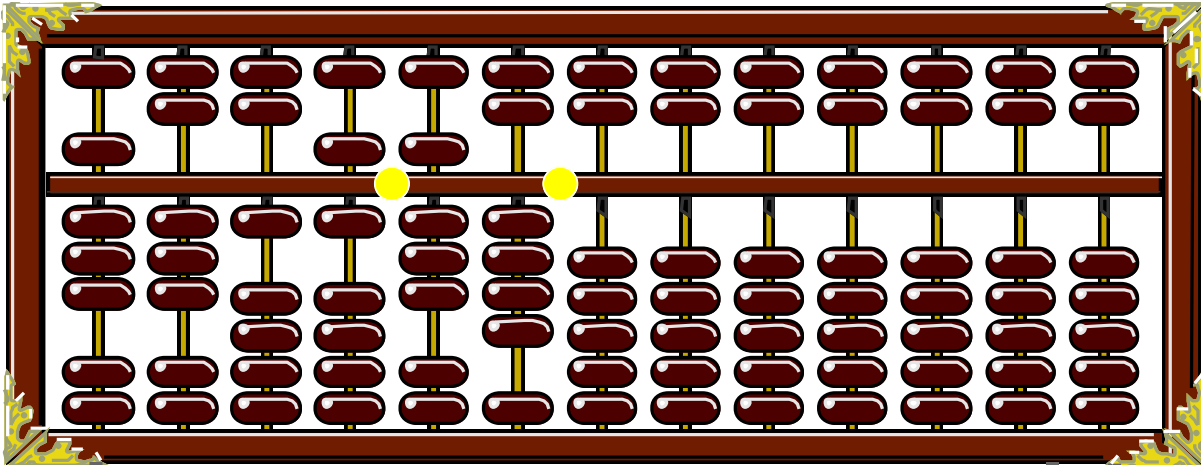
(vereinfacht): Schreibe die Aufgabe in bestimmter Symbolik aufs Papier. Stimmen die Ausgangszahlen in bestimmter Reihenfolge mit den Zahlen links in den Regeln überein, so schreibe die Zahl rechts in der Regel an bestimmter Stelle aufs Papier. Wende dabei erst die Regeln mit "x" an, später auf das Zwischenergebnis die Regeln mit "+".

➤ **Definition:** Rechnen ist Symbolmanipulation nach einer Vorschrift. Die Vorschrift heißt Algorithmus.

Symbole sind physikalische Gegenstände (Schallwellen, Zeichen auf Papier, elektrische Spannung ...)

Da Symbole physikalisch vorliegen, ist Rechnen mechanisierbar.

Beispiele aus der Geschichte:



Abakus - antike Rechenhilfe, Darstellung der Rechenaufgabe 8316×84

1623 Wilhelm **Schickard** stellt in Tübingen ein Modell auf der Basis von **Zählrädern** mit 10 Raststellungen zur Addition und Subtraktion im Dezimalsystem vor.

1673 entwickelte Gottfried Wilhelm **Leibniz** eine Rechenmaschine mit einer **Staffelwalze**.



Eine mechanische Registrierkasse

2. Programmierbare Rechenmaschinen

Eine Rechenvorschrift kann durch eine Vorschrift auf einer abstrakteren Ebene festgelegt werden. Mit den Befehlen

SPEICHERN (x an s)

ENTNEHMEN (x von s)

VERGLEICHEN (Stelle s mit Stelle t)

WENN (Bedingung) DANN (Befehl)

könnte man die Vorschrift zur Ausführung einer Multiplikation selbst als Symbolmanipulation nach einer (Meta)vorschrift realisieren. Man kann also gewissermaßen das Rechnen berechnen.

➤ **Definition: Die Vorschrift zur Ausführung einer Vorschrift ist ein Programm.**

Eine Rechenmaschine, die Vorschriften über Vorschriften zum Rechnen ausführt, ist programmierbar. Sie ist nicht auf einen Algorithmus festgelegt, sondern die Rechenvorschrift selbst kann wieder berechnet werden. Das Programm kann nun von außen eingegeben werden.

Geschichte:

1833 legte der Engländer Charles **Babbage** ein Konzept eines **Analytischen Rechenautomaten (Analytical Engine)** vor mit

- Speichereinheit (für 1000 Zahlen zu 50 Stellen)
- Rechenwerk mit dezimalen Zählern und Schaltgetrieben
- Steuereinheit zur Steuerung des Weiterrechnens in Abhängigkeit vom jeweiligen Rechenergebnis
- Ein- und Ausgabereinheit (unter Verwendung von Lochkarten)



Die erste Hollerith-Maschine 1890. Die Anlage besteht aus einem Pantographlocher, aus dem Zählwerk und der Sortiermaschine.



Lochkartenarchiv. Volkswagenwerk 1954

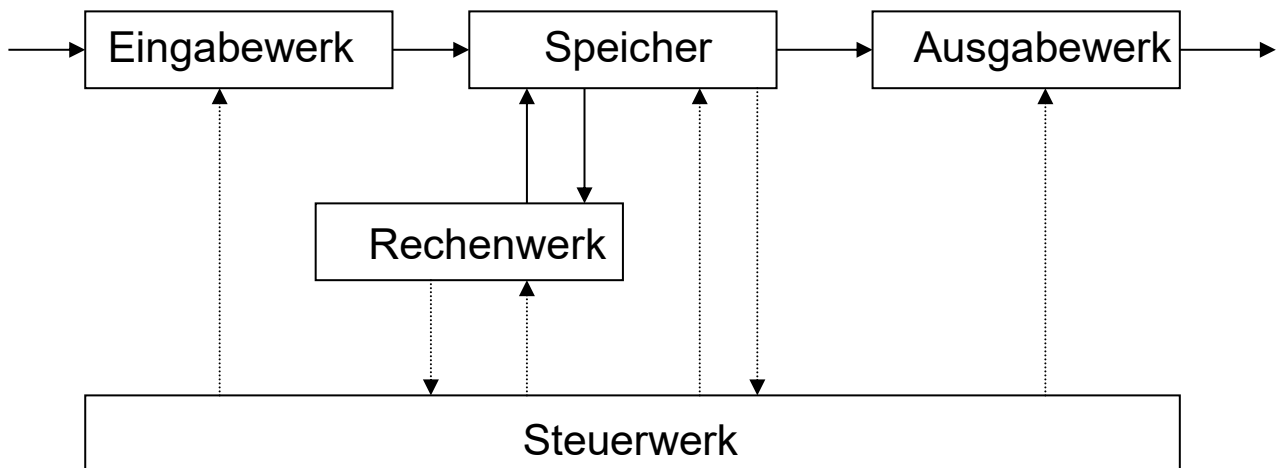
Konrad Zuse (1911-1995) übernimmt das Konzept eines "Aussagenkalkül" für die Arbeitsweise von dualen Rechenautomaten mit den Grundoperationen der Logik "Und", "Oder" und "Negation". Damit sind alle arithmetischen Operationen ausführbar.

1941 stellte Zuse die erste voll arbeitsfähige programmgesteuerte (elektromechanische) Rechanlage der Welt vor.

Wesentliche Merkmale waren ein Rechenwerk bestehend aus 600 Relais, Relaisspeicher für 64 Zahlen zu 22 Dualstellen und eine Eingabevorrichtung für Kinofilmlochstreifen (8 Spuren).



Typische Komponenten einer programmierbaren Rechenmaschine (nach von Neumann)

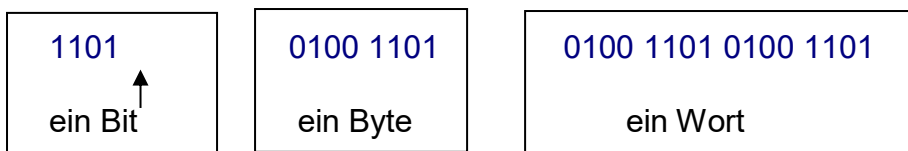


Von Neumann-Architektur eines Universalrechners

- **Eingabe, Ausgabe** auf irgendeinem Medium (ursprünglich Lochkarten, Lochstreifen)
- **Speicher** mit Adressen, darin Daten und Programme
- **Rechenwerk** mit logischen Schaltungen für Basisoperationen
- **Steuerwerk** um Inhalte von Adressen (Befehle für Rechenoperationen, Daten) ins Rechenwerk holen, um Befehle zu interpretieren und Funktionseinheiten steuern

Informationsdarstellung

binär / digital: nur zwei Symbole: 0, 1



1 KB = 2^{10} Bytes = 1.024

1 MB = 2^{20} Bytes = 1.048.576

1 GB = 2^{30} Bytes = 1.073.741.824

Symbole können gegeneinander ausgetauscht werden.

Darstellbar daher mit Kombinationen von 0, 1: Zahlen, beliebige Zeichen, Bilder ...

Zeichenvorrat

ASCII - 7 Bit = 255 Zeichen

CODE TRANSLATION TABLE (Contd)

Dec.	Hex	Instruction (RX)	Graphics and Controls			7-Track Tape BCDIC(2)	Card Code	Binary
			BCDIC	EBCDIC(1)	ASCII			
64	40	STH		Sp Sp	@	(3)	no punches	0100 0000
65	41	LA			A		12-0-1-9	0100 0001
66	42	STC			B		12-0-2-9	0100 0010
67	43	IC			C		12-0-3-9	0100 0011
68	44	EX			D		12-0-4-9	0100 0100
69	45	BAL			E		12-0-5-9	0100 0101
70	46	BCT			F		12-0-6-9	0100 0110
71	47	BC			G		12-0-7-9	0100 0111
72	48	LH			H		12-0-8-9	0100 1000
73	49	CH			I		12-1-8	0100 1001
74	4A	AH		⚡ ⚡	J		12-2-8	0100 1010
75	4B	SH	.	.	K	B A 8 2 1	12-3-8	0100 1011
76	4C	MH	□)	< <	L	B A 8 4	12-4-8	0100 1100
77	4D		[((M	B A 8 4 1	12-5-8	0100 1101
78	4E	CVD	<	+ +	N	B A 8 4 2	12-6-8	0100 1110
79	4F	CVB	#		O	B A 8 4 2 1	12-7-8	0100 1111
80	50	ST	& +	& &	P	B A	12	0101 0000
81	51				Q		12-11-1-9	0101 0001
82	52				R		12-11-2-9	0101 0010
83	53				S		12-11-3-9	0101 0011
84	54	N			T		12-11-4-9	0101 0100
85	55	CL			U		12-11-5-9	0101 0101
86	56	O			V		12-11-6-9	0101 0110
87	57	X			W		12-11-7-9	0101 0111
88	58	L			X		12-11-8-9	0101 1000
89	59	C			Y		11-1-8	0101 1001
90	5A	A		! !	Z		11-2-8	0101 1010

Verschiedene Zeichensätze übersetzt in Binärkode

Semantik

Symbole haben für den Benutzer eine Semantik. Auch das Verhältnis von Symbolen zueinander, die Umwandlung von Symbolen in andere hat für den Benutzer eine Semantik.

Ein Computer ist ein Symbolwandler. Er wandelt Symbole in der Eingabe in Symbole in der Ausgabe.

- Indem der Benutzer die Symbole ebenso wie die Umwandlung interpretiert, bekommt das Rechnen als rein mechanische Symbolmanipulation Bedeutung.

Rechenoperationen

0,1 technisch: Hohe Spannung, niedrige Spannung

Operationen im Rechenwerk : Eingabe, Vergleichen, Ausgabe

Takt des Computers: mit jedem Takt ändert sich der Ausgangszustand in Abhängigkeit vom Eingabezustand.

Schaltfunktionen, sog. Gatter = alle Möglichkeiten von 0 und 1 bei ein oder zwei Eingängen



Eingänge	Ausgang
0 0	0
0 1	0
1 0	0
1 1	1

a) UND-Schaltung

Eingänge	Ausgang
0 0	0
0 1	1
1 0	1
1 1	1

b) ODER-Schaltung

Eingang	Ausgang
0	1
1	0

c) NICHT-Schaltung

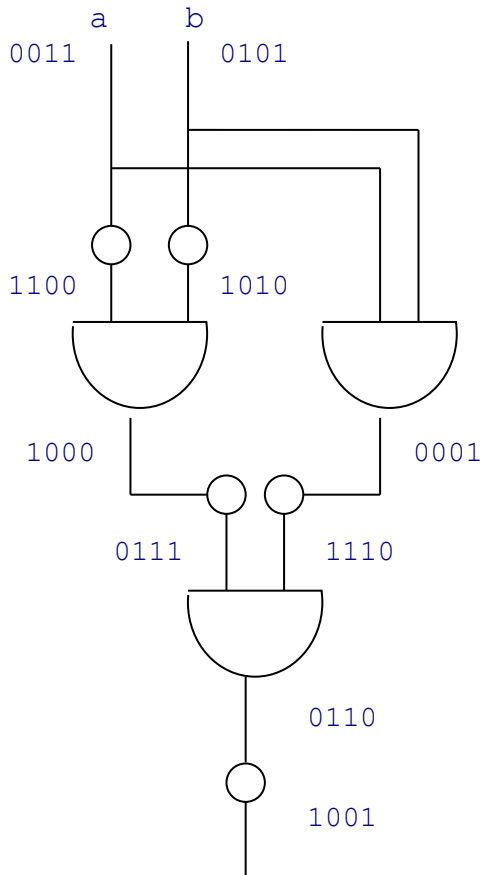
einfache Gatter

Alle anderen Schaltungen können durch Kombination aus einfachen hergestellt werden

Die GENAU-WENN-DANN- Schaltung ist definiert als:

Eingänge		Ausgang
a	b	
0	0	1
0	1	0
1	0	0
1	1	1

Schaltung nur aus UND und NICHT-Gattern:



3. Computer und Geist? Kann man mit Sprache rechnen?

- Alles und nur das lässt sich mit Computern simulieren, was auf Symbolmanipulation reduziert werden kann.

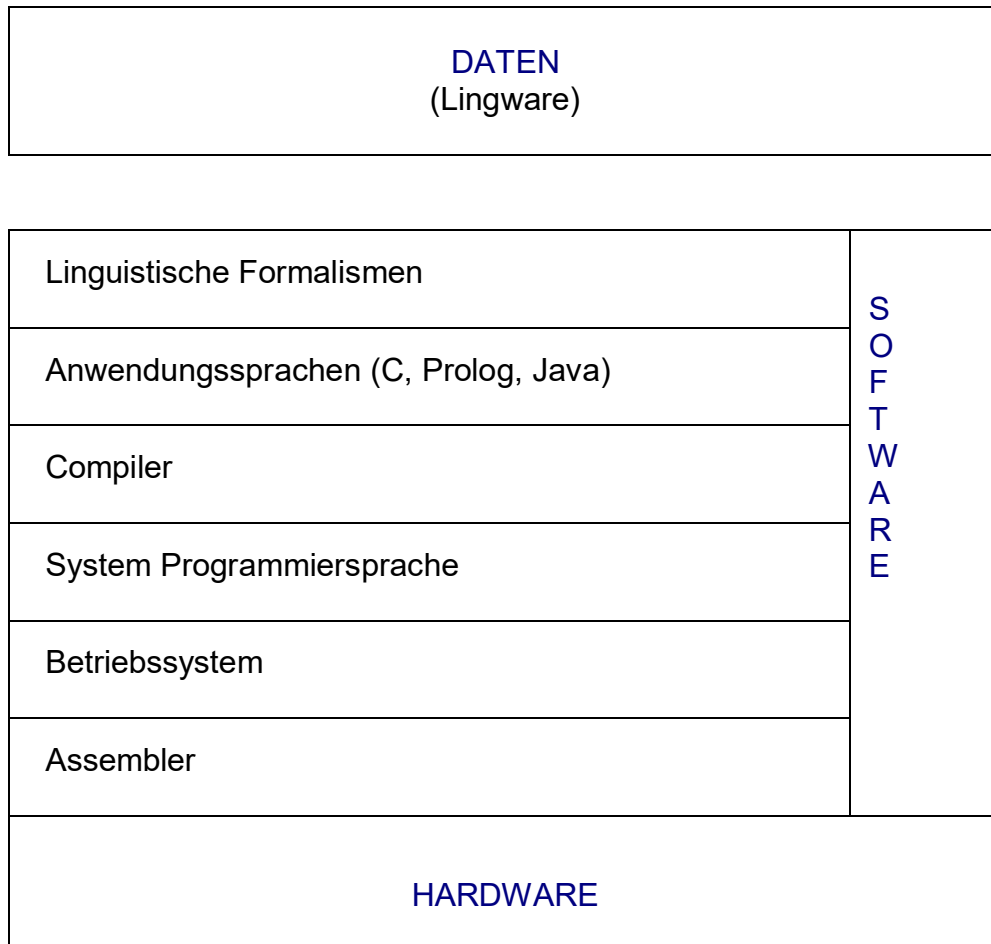
"Wenn man Charaktere oder Zeichen finden könnte, die geeignet wären, alle unsere Gedanken ebenso rein und streng auszudrücken, wie die Arithmetik die Zahlen oder die analytische Geometrie die Linien ausdrückt, könnte man offenbar bei allen Gegenständen, soweit sie dem vernünftigen Denken unterworfen sind, das tun, was man in der Arithmetik und der Geometrie tut. Und wenn jemand an dem, was ich vorgebracht haben würde, zweifelte, würde ich zu ihm sagen: "Rechnen wir, mein Herr!"

G.W. Leibniz

I 2 Grundlagen der Software und der Programmierung

Lernziel: Einführung in das algorithmische Denken und den Entwurf von Programmen

1. Die Software-Schichten



2. Phasen bei der Erstellung eines Programms:

ENTWURF

- Problemanalyse, Aufteilung in Teilprobleme
- Entwicklung einer Lösung in Form eines Algorithmus
- Spezifikation der Programme und Unterprogramme und ihrer Schnittstellen - Darstellung mittels Flussdiagramm oder mittels Nassi-Shneiderman Struktogramm

IMPLEMENTIERUNG

- Kodierung in einer Programmiersprache, mit erklärenden Kommentaren (!)
- Erprobung anhand von Daten, Debugging
- Erstellen einer Dokumentation mit Benutzeranleitung
- Evaluierung durch Benutzer
- ggf. Verbesserungen vornehmen

Beispiel: Entwickeln Sie ein Programm, das römische Zahlen in arabische Zahlen umformt.

Römische Zahlen

I = 1	IX = 9	XVII = 17	XXIX = 29	XC = 90	M = 1000
II = 2	X = 10	XVIII = 18	XXX = 30	C = 100	MDLVIII=1558
III = 3	XI = 11	XIX = 19	XL = 40	CI = 101	MDCXXV=1625
IV = 4	XII = 12	XX = 20	L = 50	CL = 150	MDCCCIV=1804
V = 5	XIII = 13	XXIII = 23	LIX = 59	CC = 200	MCM=1900
VI = 6	XIV = 14	XXIV = 24	LX = 60	CD = 400	MCMLXII=1962
VII = 7	XV = 15	XXVII = 27	LXX = 70	D = 500	MIM=1999
VIII = 8	XVI = 16	XXVIII = 28	LXXX = 80	DC = 600	MM=2000

Problemanalyse ?

Ziffern römisch: I, V, X, L, C, M

Ziffern arabisch: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

Bildung von Zahlen römisch: normalerweise Addition

$$XXVIII = 10+10+5+1+1+1=28$$

aber Subtraktion einer kleineren Zahl vor einer größeren

$$XIV = 10+ 5-1=14$$

Bildung von Zahlen arabisch: je nach Stelle Multiplikation mit

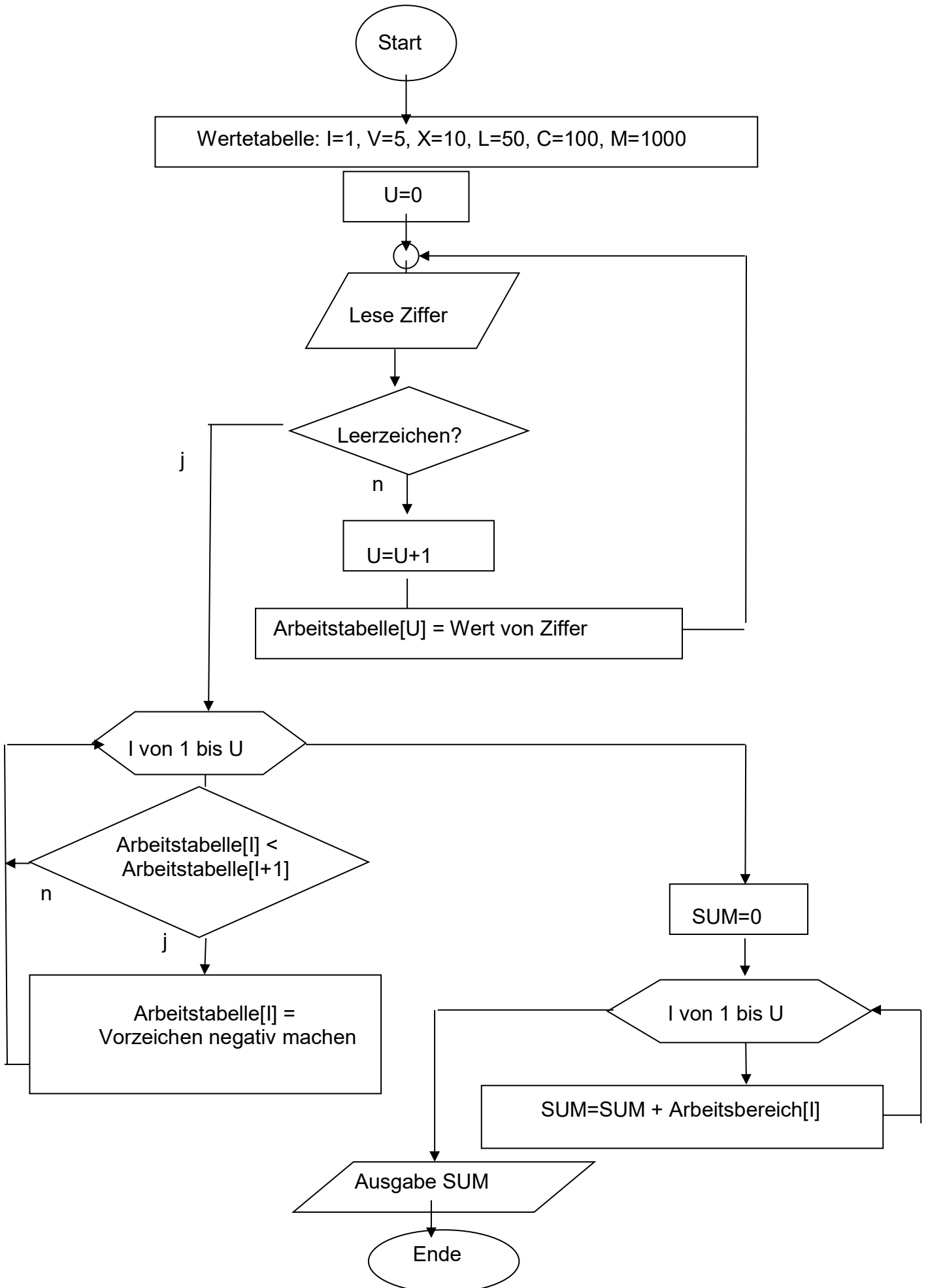
Potenz von 10^0 , 10^1 , 10^2 ... , 10^n

Eine mögliche Lösung:

1. Die einfachen römischen Ziffern bekommen den entsprechenden arabischen Zahlenwert zugeordnet: I=1, V=5, X=10, L=50, C=100, M=1000, D=500
2. Die römische Zahl wird gelesen, für jede Ziffer wird der entsprechende Wert gespeichert. Falls der Wert der nächsten römischen Ziffer größer ist, bekommt der Wert der kleineren ein negatives Vorzeichen.
3. Am Schluß werden alle gespeicherten Werte addiert und ergeben die arabische Zahl.

M	1000
C	-100
M	1000
I	-1
V	5
MCMIV	1904

Flußdiagramm (Programmablaufplan)

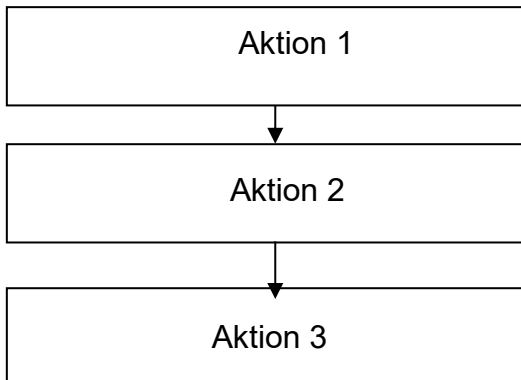


Die Darstellungsmittel in Flussdiagrammen:

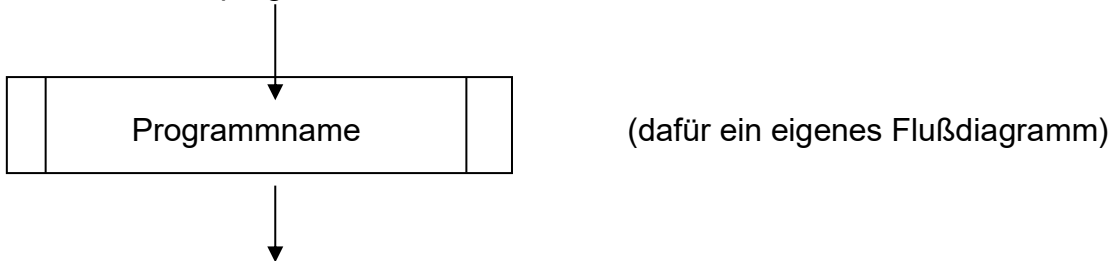
1. Anfang und Ende des Programms



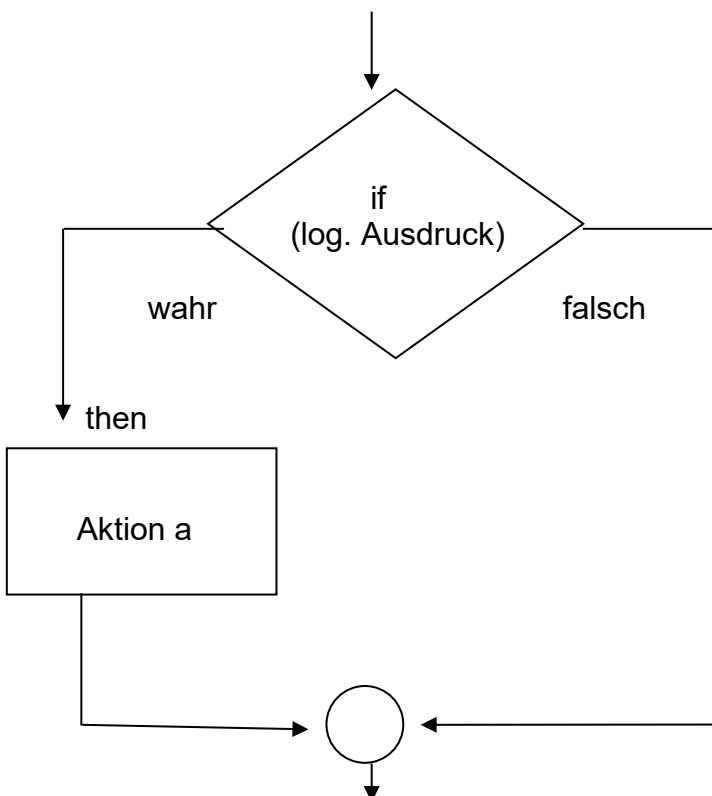
2. Sequenz unbedingte Aktionen



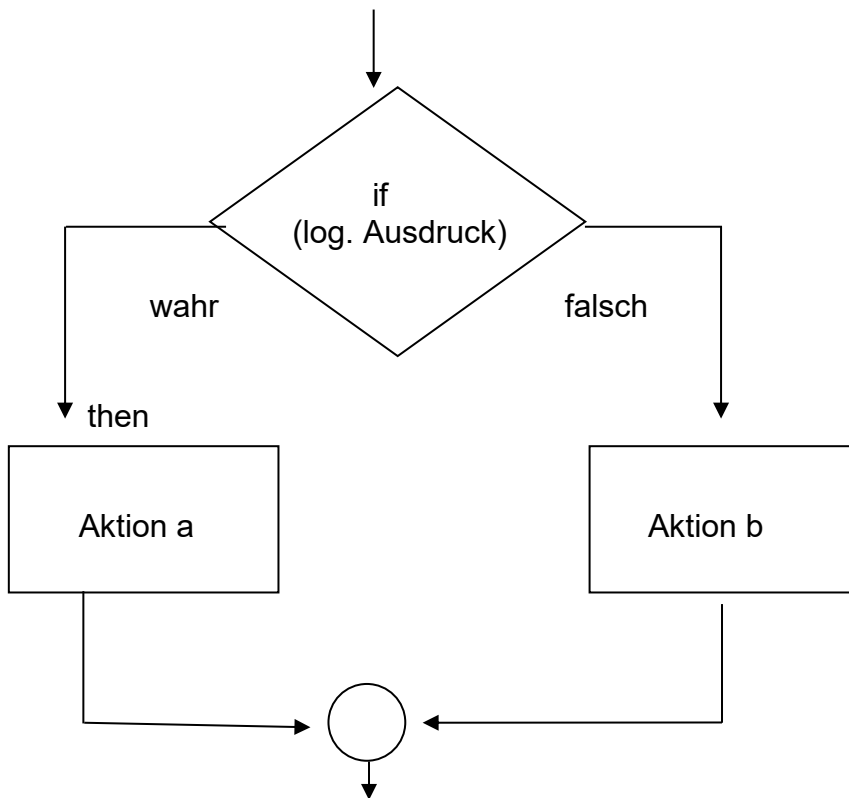
3. Aufruf eines Unterprogramms



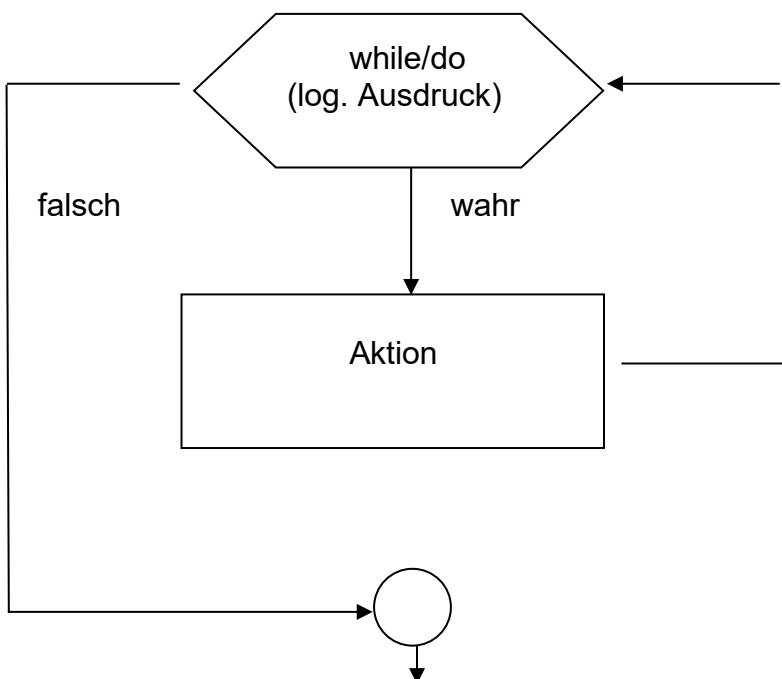
4. If- Abfrage, Verzweigung einseitig



4. If- Abfrage, Verzweigung zweiseitig



5. while- oder do- Schleife



Abstraktion, Finden einer generellen Lösung

Beispiel Kongruenz im Deutschen (Sternchen heißt "falsch")

der Mann *die Mann
 die Männer
 er kommt *er kommst
 der Mann kommt
 du kommst
 wir kommen
 die Männer kommen *der Mann kommen

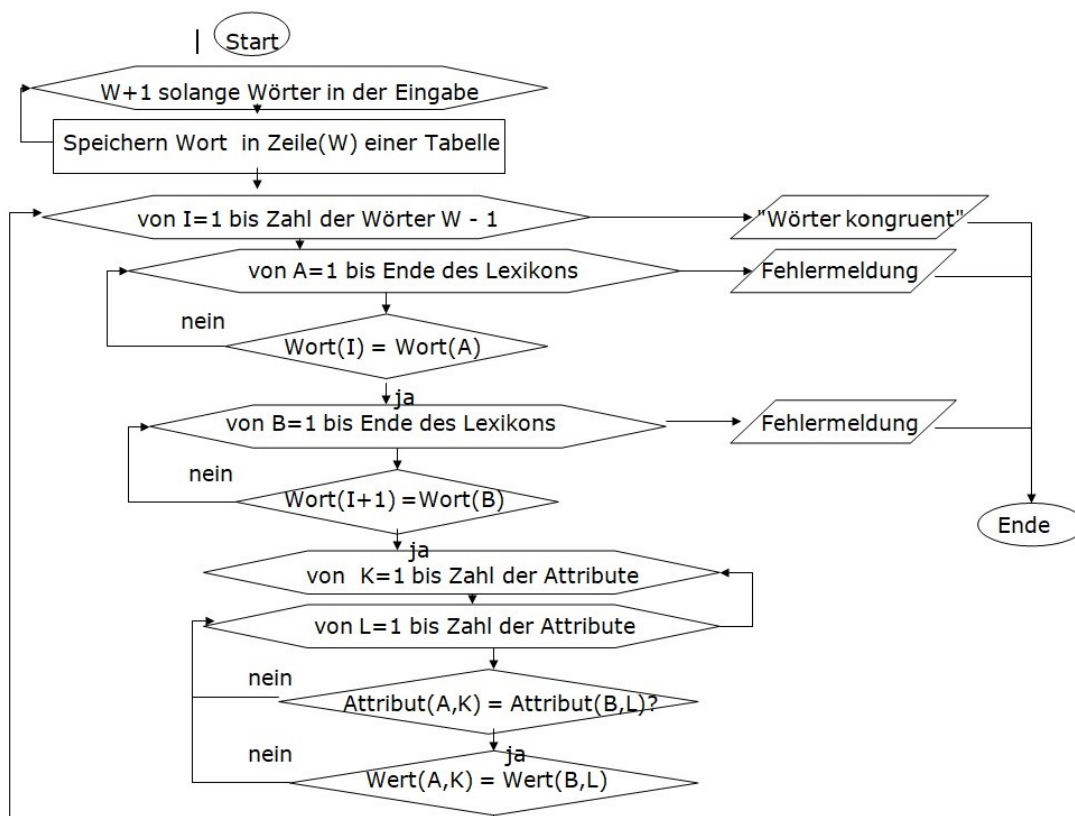
Kongruenz :

singular - singular
 plural - plural
 erste Person - erste Person
 zweite Person - zweite Person
 dritte Person - dritte Person

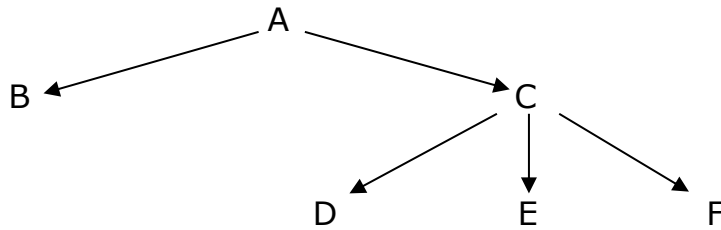
Verallgemeinerung: Wörter, Attribute und Werte, gespeichert in einem "Lexikon"

der	numerus[singular]	
Mann	numerus[singular]	person[dritte]
die	numerus[plural]	
Männer	numerus[plural]	person[dritte]
er	numerus[singular]	person[dritte]
kommt	numerus[singular]	person[dritte]
du	numerus[singular]	person[zweite]
kommst	numerus[singular]	person[zweite]
wir	numerus[plural]	person[erste]
kommen	numerus[plural]	person[erste]
kommen	numerus[plural]	person[dritte]

Kongruenzprüfprogramm (wie könnte es aussehen?)



Rekursion - Ein Programm ruft sich selbst auf

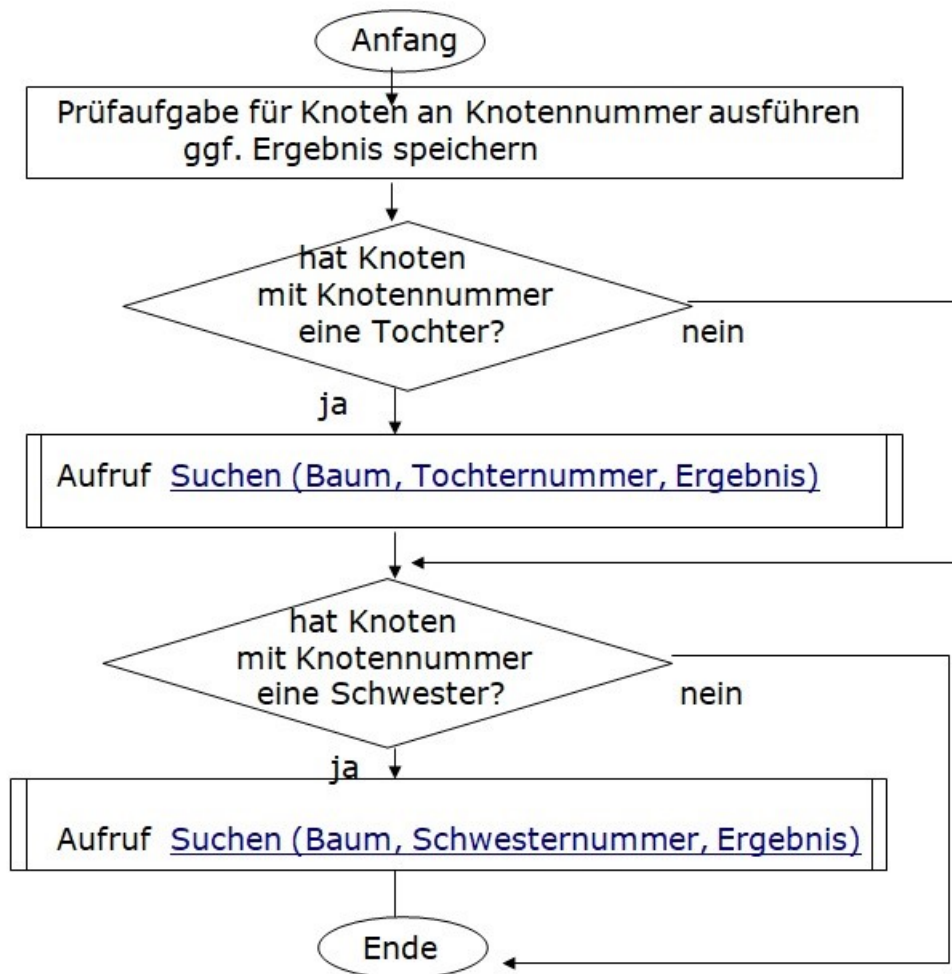


Gegeben ein Baumgraph. Aufgabe: den Baum durchsuchen

Lösung: Das Programm 'Suchen' mit drei Parametern 'Baum', 'Knotennummer', 'Ergebnis'

Knoten sind nummeriert. Am Anfang ist Knotennummer die Nr. des obersten Knotens und Ergebnis = leer

Suchen (Baum, Knotennummer, Ergebnis)



S 1.1 Phonologie und artikulatorische Phonetik

Lernziel: Einige Fakten über die Sprachlaute, vor allem aber Einsicht in die strukturalistische Methodik

1. Aufgabe einer (strukturalistischen) Beschreibung der Laute

Lautinventar aufstellen - mögliche Lautkombinationen angeben

- Physisch sind viel Laute möglich. Kinder in der Lallperiode probieren viele aus. Viel weniger Laute werden dann in einer einzelnen Sprache benutzt, und zwar in verschiedenen Sprachen der Welt, verschiedene Teilmengen der möglichen Laute.
- Physikalisch, meßbar nach Frequenz und Amplitude, sind die Laute auch bei Sprechern derselben Sprache sehr verschieden (Frauenstimme, Männerstimme, schreien, flüstern, lachen, schluchzen, normal oder heiser ...). Die meisten Unterschiede sind aber offenbar für das Verständnis irrelevant.
- Sprechen ist ein Lautkontinuum. Es gibt physikalisch oft keine Grenzen zwischen Lauten. Man muß wissen, wie das Kontinuum in Einzellaute zu zerlegen ist, um etwas korrekt zu hören.

Also: erstes Ziel der Erforschung der Laute einer Sprache:

Was macht überhaupt einen Unterschied?

Was sind demnach die Einheiten?

Was sind die "unterscheidenden Merkmale"?

2. Die psychische und die physische Ebene der Sprache

Sprache kennen	Sprache aktuell gebrauchen
psychologisch	physikalisch, physiologisch
Regeln	regelhafte Ausführung
virtuell	aktuell

langue Sprache als System	parole aktuell Sprechen	(De Saussure)
relevante Einheiten	+ irrelevantes Beiwerk	

PHONEM	PHON
Laut als geistige Einheit	Laut als physikalische Erscheinung

Phonologie/Phonemik	Phonetik
<i>Bestimmung der Lauteinheiten der langue</i>	<i>Messung und Realisierung von Lauteinheiten in der parole</i>

3. Die strukturalistische Methode (Trubetzkoi)

Experiment, Austausch, Umstellung -> distinktive Opposition

-> Distribution

-> System

Definitionen

PHONEM = kleinste bedeutungsunterscheidende Lauteinheit im Sprachbewußtsein

PHON = Schalleinheit, durch die ein Phonem realisiert wird

ALLOPHON = akustisch unterschiedliche Phone, die dasselbe Phonem realisieren, Phonemvarianten

Allophonie, bedingt durch Kontext (Ko-Artikulation) und durch viele äußere Faktoren

Minimalpaaranalyse

singen	sinken	/g/ - /k/
singen	ringen	/z/ - /r/
ringen	wringen	/0/ - /v/
ringen	rangen	/i/ - /a/
lassen	lasen	[as] - [a:z]
kann	Kahn	/a/ - /a:/
Muße	Muse	/s/ - /z/
Wahl	Wal	kein Bedeutungsunterschied,
Ehre	Ähre	also kein Lautunterschied

Jedes Sprachsystem ist gekennzeichnet durch eine eigene Menge von Oppositionen

Jeweils andere distinktive Oppositionen in jeder Sprache, jedem Dialekt

thing	sing	/t/ - /s/
think	sink	/t/ - /s/
the	[t] [ð]	/t/ - /d/ th - d wirklich eine distinktive Opposition?

Lautwandel - Verschiebungen im System

	/p/ -> /f/	/t/ -> /s/	/k/ -> /x/
Niederdeutsch	open	water	ik
Hochdeutsch	offen	wasser	ich

Distinktive Merkmale

z.B. Vokale im Deutschen:

7 Monophthonge lang:

liegen - legen - lagen - logen - lügen - lügen - lögen

7 Monophthonge kurz:

Wille - Welle - walle - Wolle, Stollen - Stullen, Hülle - Hölle

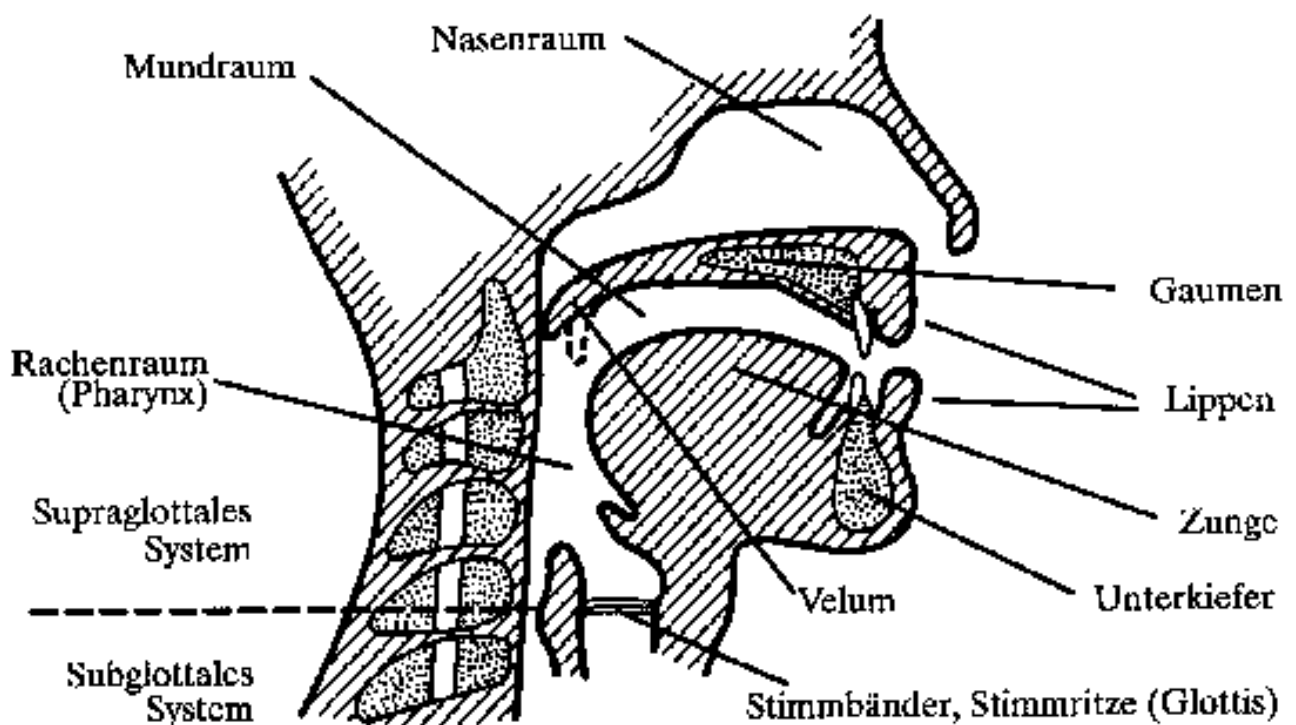
lang versus kurz:

Stiel	stehle	Stahl	Fohlen	Buhle	fühle	Höhle
still	Stelle	Stall	vollen	Bulle	Fülle	Hölle

Worauf kommt es in der jeweiligen Opposition an?

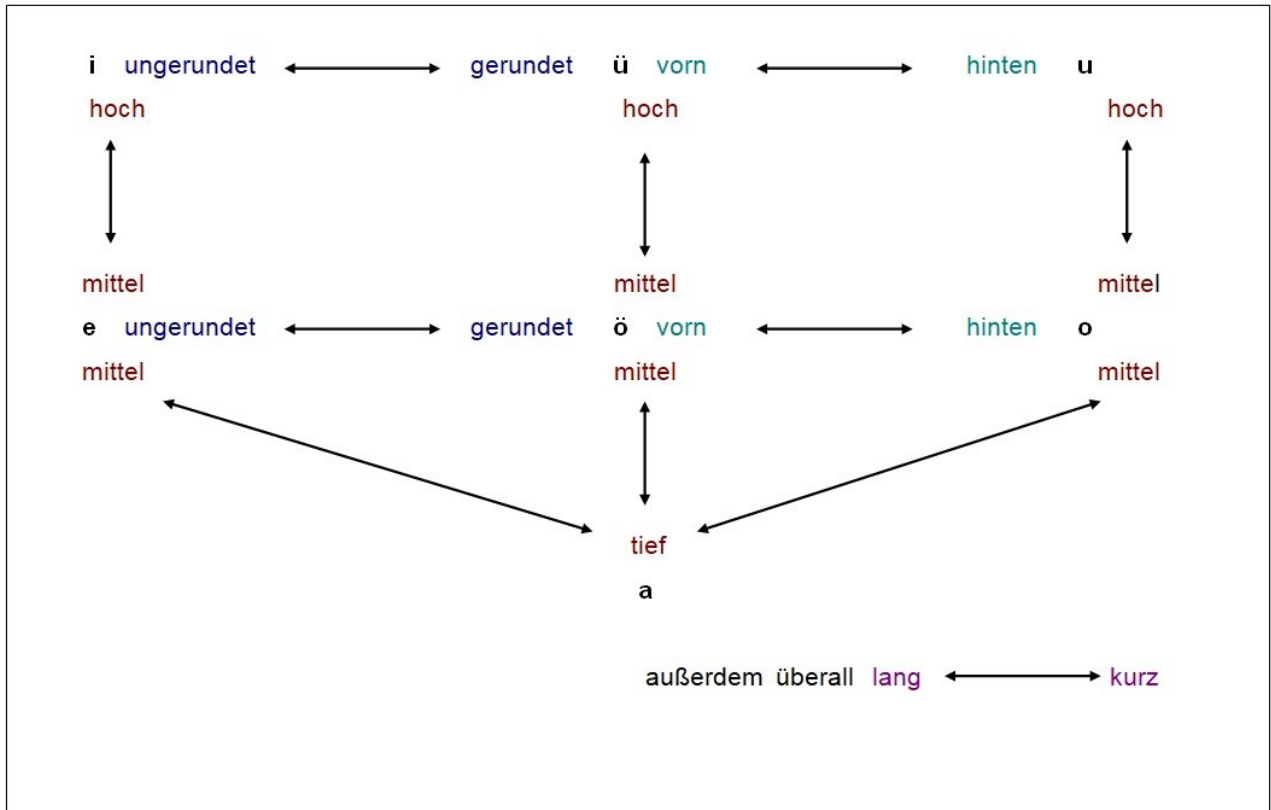
- akustische (-> akustische Phonetik S12)
- artikulatorische Merkmale

Die menschlichen Sprechwerkzeuge:



deutsche Vokale	Unterkiefer/Zunge :	tief - mittel - hoch
	Zunge:	vorn - hinten
	Lippen:	gerundet - ungerundet

Die deutschen Vokal-Phoneme und ihre artikulatorisch-distinktiven Merkmale

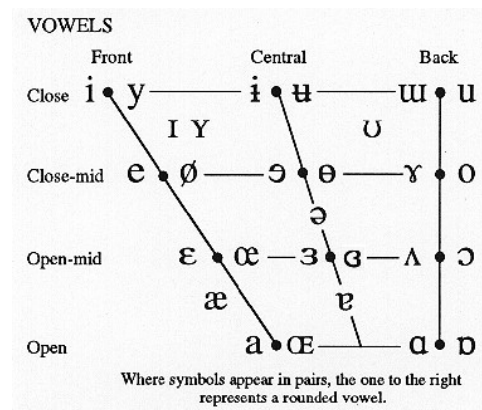


Notationen:

- Phonologisch - was zählt semantisch? /ɪɛ/ /ɑɛ/
- Phonetisch - wie klingt es natürlich? [ɪχ] [ɑɛ]
- Graphematisch - wie wird es geschrieben? <ich> <ach>

Das Internationale Phonetische Alphabet (IPA)

i	as in tree	/tri/
ɪ	as in insect	/ɪnsekt/
ɛ	as in bet	/bet/
æ	as in cat	/kæet/
ɑ	as in car	/kɑ/ British, /kɑr/ American
ɒ	as in bought	/bɒt/
ɔ	as in saw	/sɔ/
ʊ	as in as in foot	/fʊt/
u	as in boot	/but/
ʌ	as in up	/ʌp/
ɜ	as in fur	/fɜ/ British, /fɜr/ American
ə	as in ago	/əlgəʊ/
ɛr	as in play	/pleɪ/
əʊ	as in know	/nəʊ/
ɑɪ	as in sky	/skaɪ/
aʊ	as in how	/haʊ/
ɔɪ	as in boy	/bɔɪ/



Unterschied zwischen Vokalen und Konsonanten

Konsonanten	Vokale
Behinderung des Luftstroms oberhalb der Glottis	Keine Behinderung des Luftstroms

Distinktive artikulatorische Merkmale der Konsonanten

- **Artikulationsstelle:**
Bilabial | Labiodental | Dental-Alveolar | Palatal | Velar | Uvular | Glottal
- **Artikulationsart:**
Verschluß | Frikative | Nasale | Liquide: Laterale - Vibranten - Schlag
- **Stimme**
stimmhaft | stimmlos

Verteilung der (artikulatorischen) distinktiven Merkmale der Konsonanten

stimmhaft ● stimmlos ○

stelle Artikulations- modus	Bilabial	Labio- dental	Dental- Alveolar	Alveolar	Palatal	Velar	Uvular	Glottal
	● ○	● ○	● ○	● ○	● ○	● ○	● ○	● ○
Verschluß	b p		d t			g k		ʔ
Frikativ		v f	z s	ʒ ʃ	j ç	ɣ x		h
Nasal	m			n		ŋ		
Liquide	Lateral			l				
	Vibrans			r			R	
	Schlag			ɾ			R	

Auf den Rabenklippen bleichen Knabenrippen,
und der Mond verkriecht sich düster im Gewölk.
Ringsum kringeln, schnattern Ringelnattern,
und der Uhu naht sich mit Gebölk.

Mit den Tatzen kratzen schwarze Katzenfratzen
an dem Leichenstein der Mördergruft...

furchtbar, schrecklich, häßlich, gräuslich, gräßlich
hallt ihr Wehgewimmer durch die Luft.

Dumpf im Moore brodel't's, rings im Chore jodelt's,
in die kohlpechrabenschwarze Nacht hinaus.

Keine Brandungslücke, keine Landungsbrücke
gibt's in diesem Meer von Schreck und Graus.

Selbst ein dummer Stänker wird zum stummen Denker,
wenn er solches hört und schaut.

Trinkt noch schnell 'nen Bitter'n, sinkt zur Stell' mit Zittern,
mit 'ner Kreidehaut ins Heidekraut.

(Heinrich Seidel)

S 1.2 Akustische Phonetik und maschinelles Erkennen und Erzeugen gesprochener Sprache

Lernziel: Allgemeine Vorstellung davon, wie gesprochene Sprache mit Computern erkannt wird

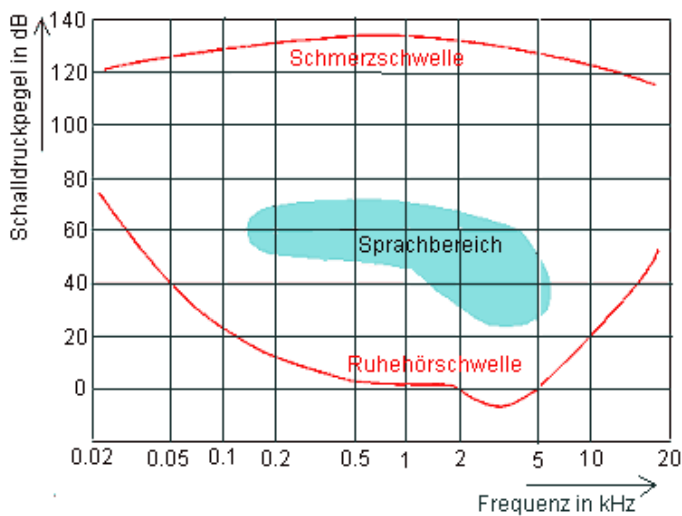
Phonetik = die Lehre von der physikalischen Seite der Laute

- artikulatorische Phonetik
- auditive Phonetik
- akustische Phonetik

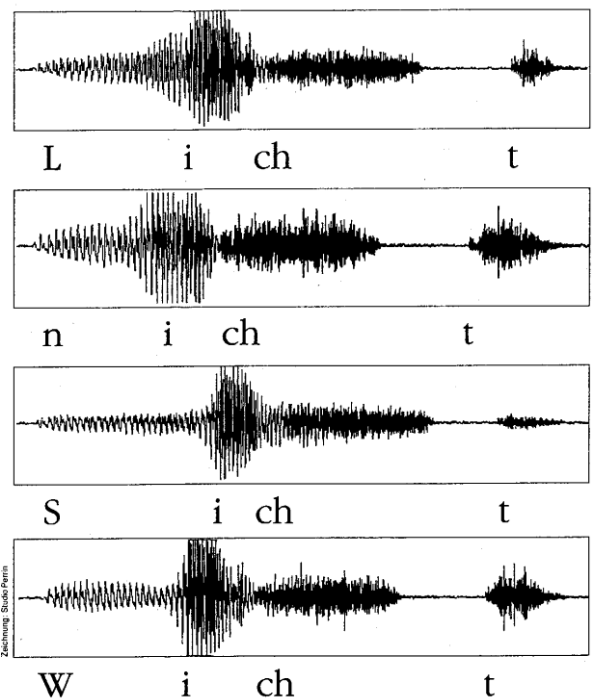
Mit der artikulatorischen Phonetik haben wir uns schon in S 1.1 befasst.

1. Auditive und akustische Phonetik

Laute - unter auditivem Aspekt



Laute - akustisch aufgezeichnet

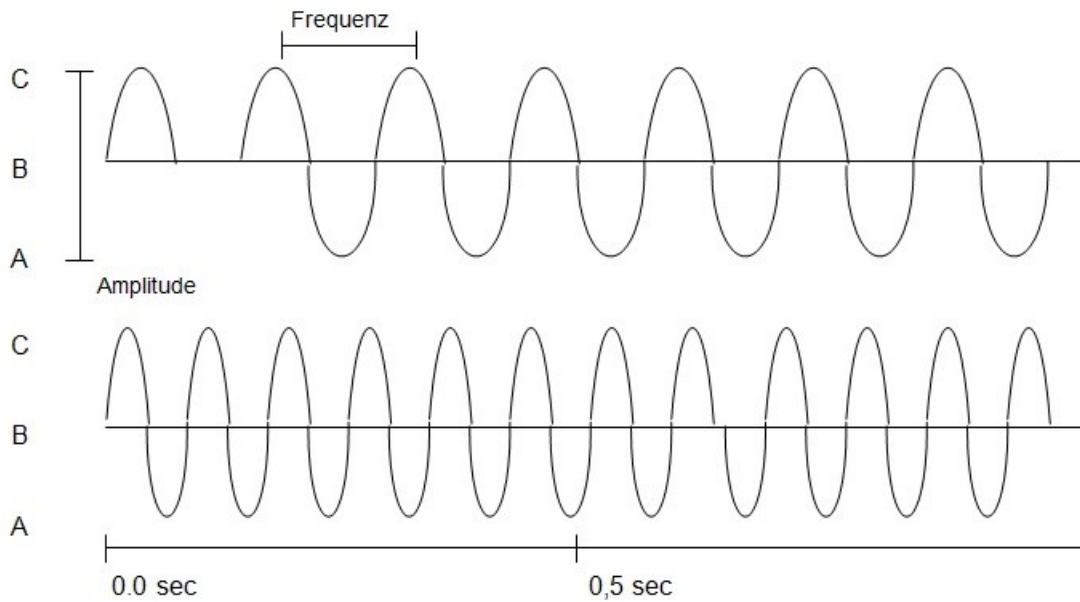


Akustische Phonetik: Schallmessung und Schallanalyse

Was ist ein Laut physikalisch?

Ein Laut ist wie jeder Schall eine Luftdruckschwankung, die sich wellenartig ausbreitet.

Frequenz = Druckänderung, **Amplitude** = Druckstärke - hörbar 20 Hz bis 20KHz



periodische Schallwellen, Sinusschwingungen = reine Töne

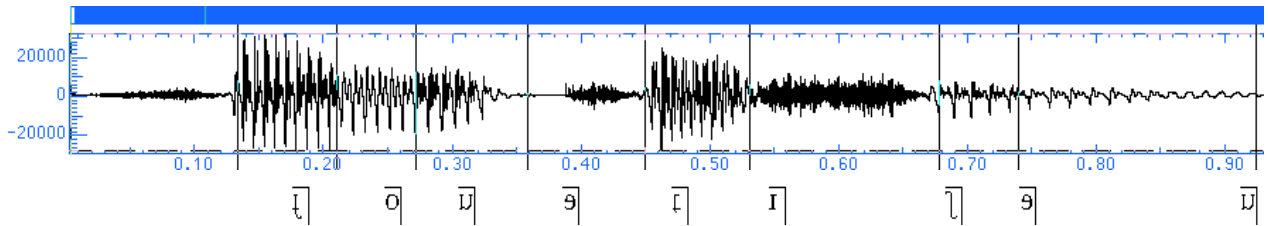
Das **Mikrophon** übersetzt Luft-Schankungen analog in **Spannungsschwankungen**

Anschließend wird das analoge Signal **digitalisiert** - ADC (analog to digital conversion)
Frequenz und Amplitude werden für einzelne Zeitpunkte gemessen,

wobei die **Sampling Frequenz** die Menge der Proben im Zeitintervall bedeutet,
z.B. 16000 x pro Sec = 16 kHz = 16 dots per MiliSec. (Telefon 8kHz, CD 44kHz).

Je nach Bitzahl pro Punkt - Genauigkeit der digitalen Darstellung der analogen Werte.

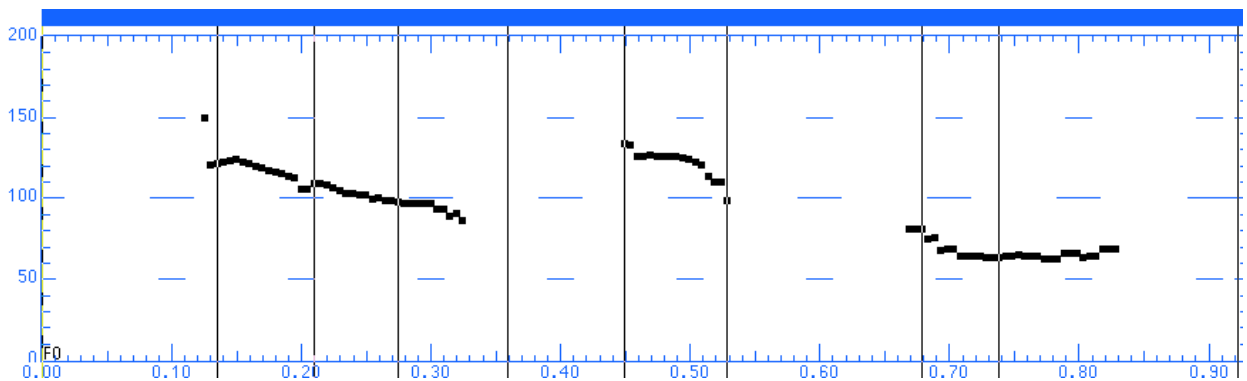
Bei gesprochener Sprache ändern sich Wellenlänge und Amplitude ständig. Außerdem besteht Sprache nicht aus einfachen Tönen, sondern aus einer Grundfrequenz und vielen Filtern. Deshalb sieht die Schwingung im zeitlichen Verlauf unregelmäßig aus.



Ocillogramm

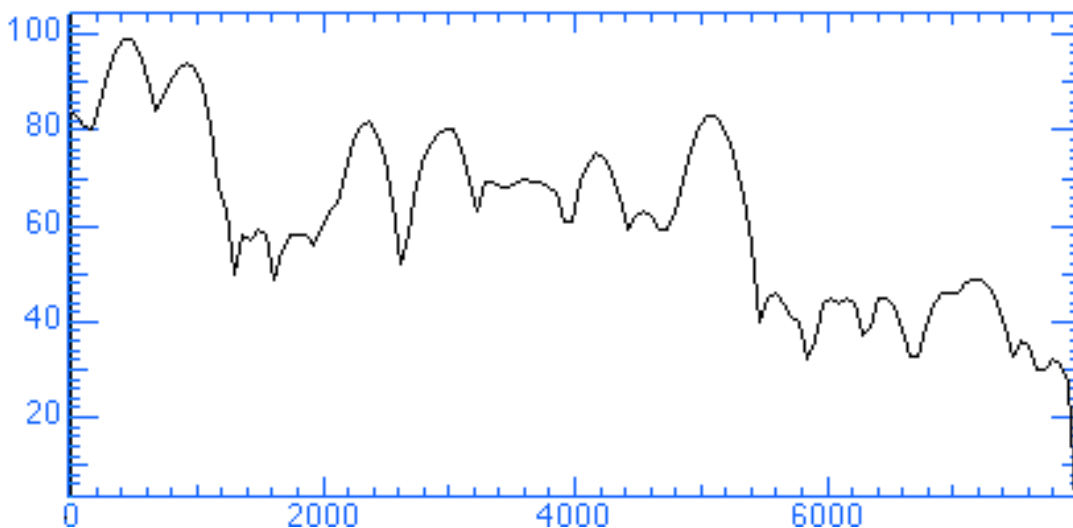
(diese und andere Graphiken in diesem Abschnitt aus *Speech Visualisation Tutorial, Leeds University*)

Aus dem Oscillogramm müssen die distinktiven Lautbestandteile herausgezogen werden. Dazu zum Teil schwierige Berechnungen.



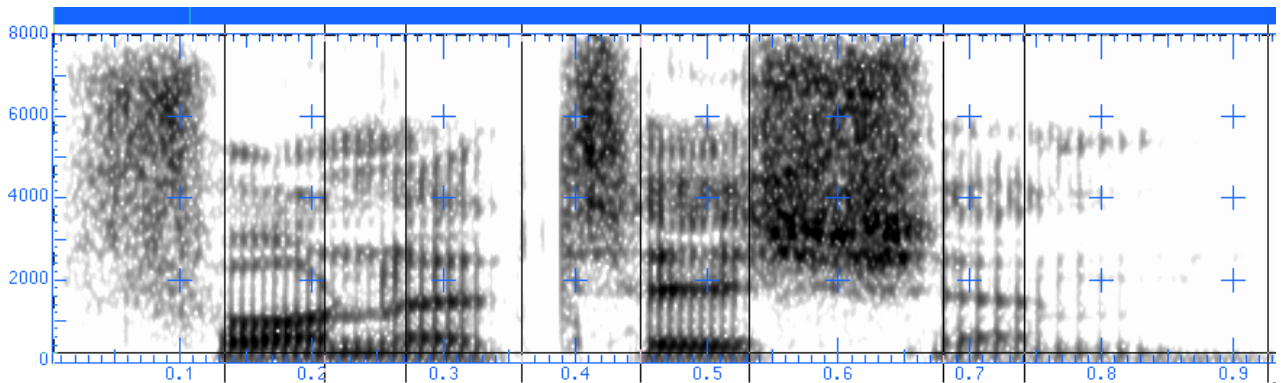
Grundfrequenz, Pitch (von den Stimmbändern erzeugt, bei Frauen doppelt so hoch wie bei Männern) ergibt die Intonation

Fourieranalyse: Das Oscillogramm wird für feste Zeiteinheiten (z.B. 0,15 sec) in Bezug auf Frequenzen und Amplituden untersucht, die Ausgabe ist ein **Spektrum**.

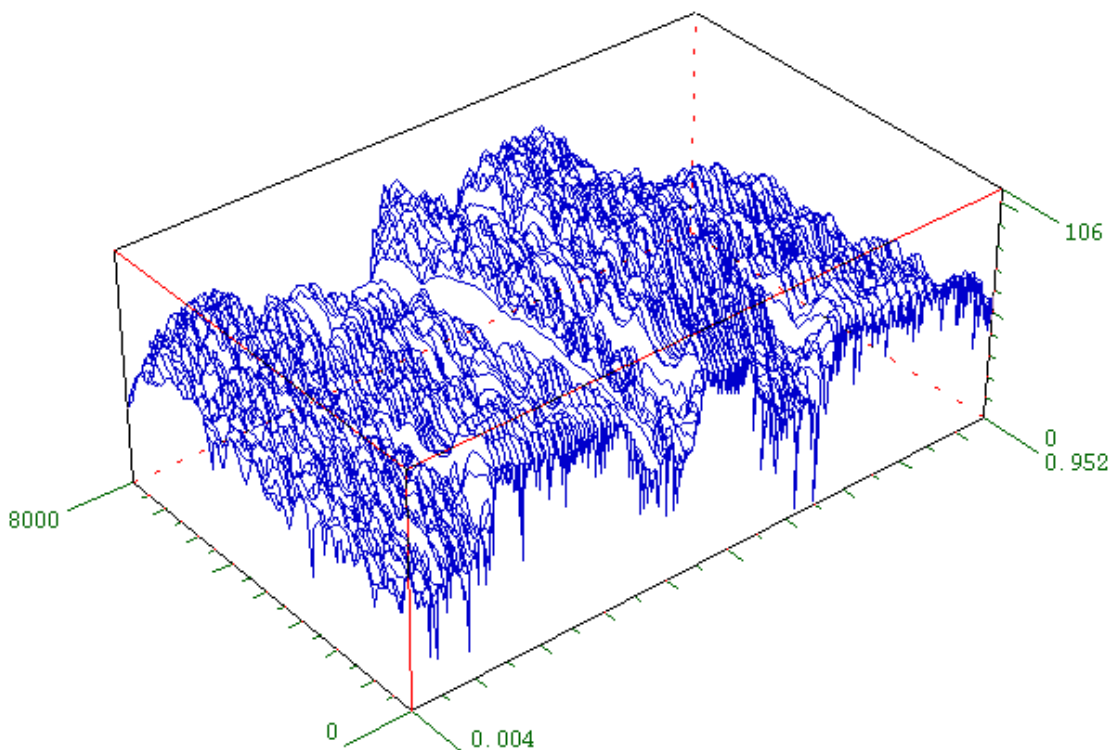


Spektrum - eine Matrix aus Frequenzen und Amplituden innerhalb einer festen Zeiteinheit.
Nimmt man die Zeitachse hinzu, ergibt sich ein dreidimensionales Gebilde: **das Spektrogramm**

(Zeit=horizontal, Frequenz vertikal, Amplitude = Schwärzung oder Farbe für abnehmend, gleichbleibend, zunehmend)



3-D Spektrogramm (Wasserfall, Frequenz und Zeit horizontal, Amplitude vertikal)



Hier sind Phoneme schon eher zu identifizieren - nämlich als typische Spektrogramm-Muster dabei Schallformen beachten

- stumme Schallform (stimmlose Verschußlaute)
- Explosionsschall (stimmhafte Verschußlaute)
- Friktionsrauschen, aperiodisch
- quasiperiodischer Schall, Klang (Vokale)

2. Maschinelle Erkennung gesprochener Sprache

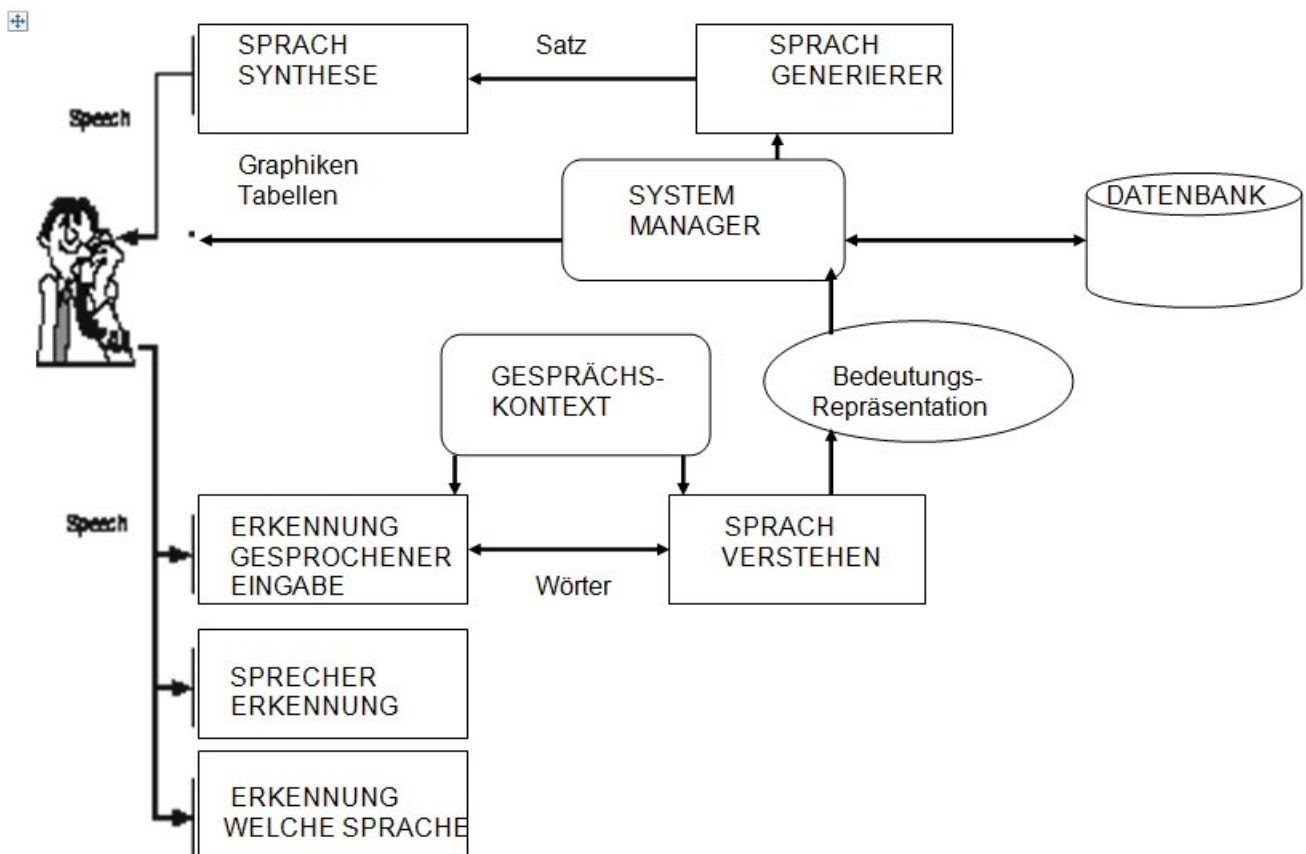
(speech recognition)

Aufgabe: Vom akustischen Signal zu sprachlichen (psychischen) Einheiten kommen, zum Phonem, zum Graphem, evtl. zur Silbe, zum orthographischen Wort

Typen von Spracherkennern

- Isoliertwörterkennner - Wort gleich kleinste Einheit, deutlich Pausen
- Verbundwörterkennner, kontinuierlich
- Ganzwörterkennner
- Erkennner von Wortuntereinheiten, z.B. Phonem/Phonemen
- sprecherabhängig, intensiv zu trainieren
- sprecherunabhängig, wenig Training

Wie funktioniert ein Spracherkennner? Welche Phasen, welche Module?



aus Survey of Human Language Technology

Phasen

AKUSTIK

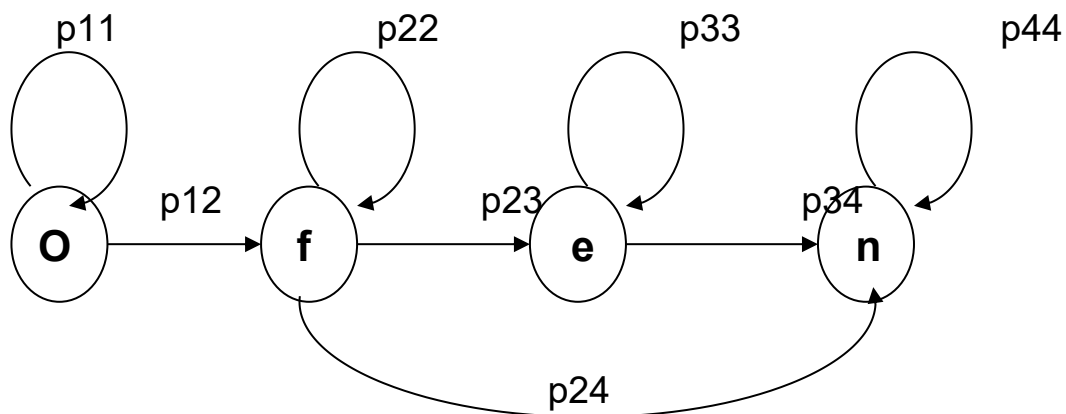
- Spracheingabe
- Aufnahme per Mikrofon, Digitalisierung
- Zerlegen in kleine Zeitsegmente (Sampling)
- Fourieranalyse für jedes Segment ergibt Spektrum (Amplitude-Frequenz pro Aufnahme-Takt),
- Zeitachse hinzugenommen ergibt Spektrogramm

PHONEME

- Gewinnung von Merkmalen für Wörter oder Phoneme, möglichst typische und robuste, gehörorientiert, Gruppen von Frequenzen, ergibt **Merkmalsvektoren**, z.B. alle 10 Millisekunden (reduzierte Information)
- vorher auf gleiche Weise erzeugt - sog. **Referenzmuster**, Training, sprecherabhängig
- dabei evtl. Ausgleich von Unterschieden auf der Zeitachse (dynamische Zeitnormierung) sowie weiterer Tricks

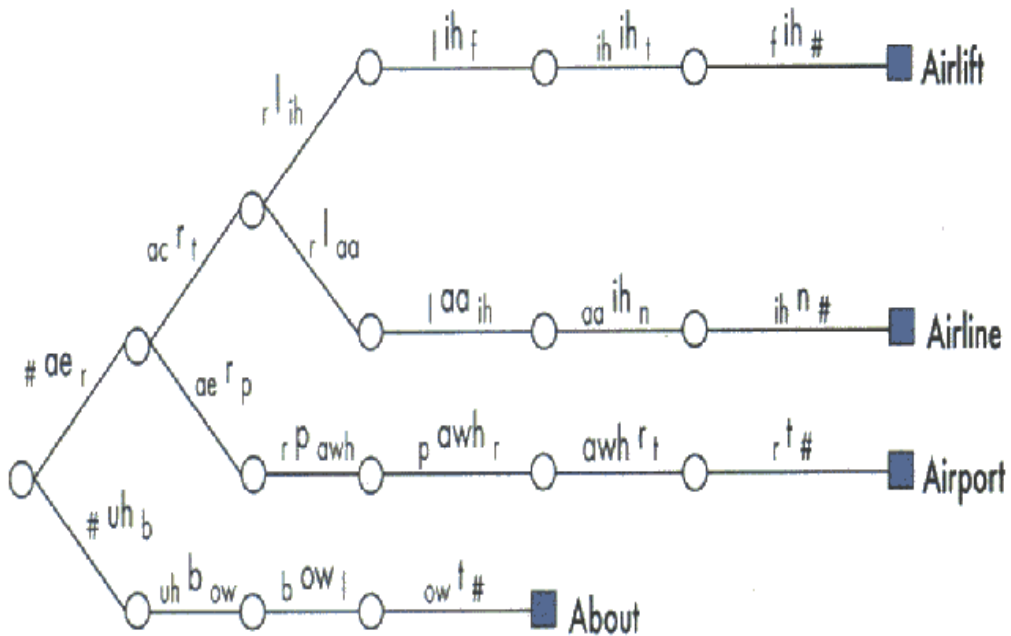
... ergibt Hypothesen, aber keine Eindeutigkeit

Hilfe vom Kontext



Hidden Markov-Kette - Wahrscheinlichkeiten von Phonemübergängen

Aussprachelexikon als Baum von Lautübergängen

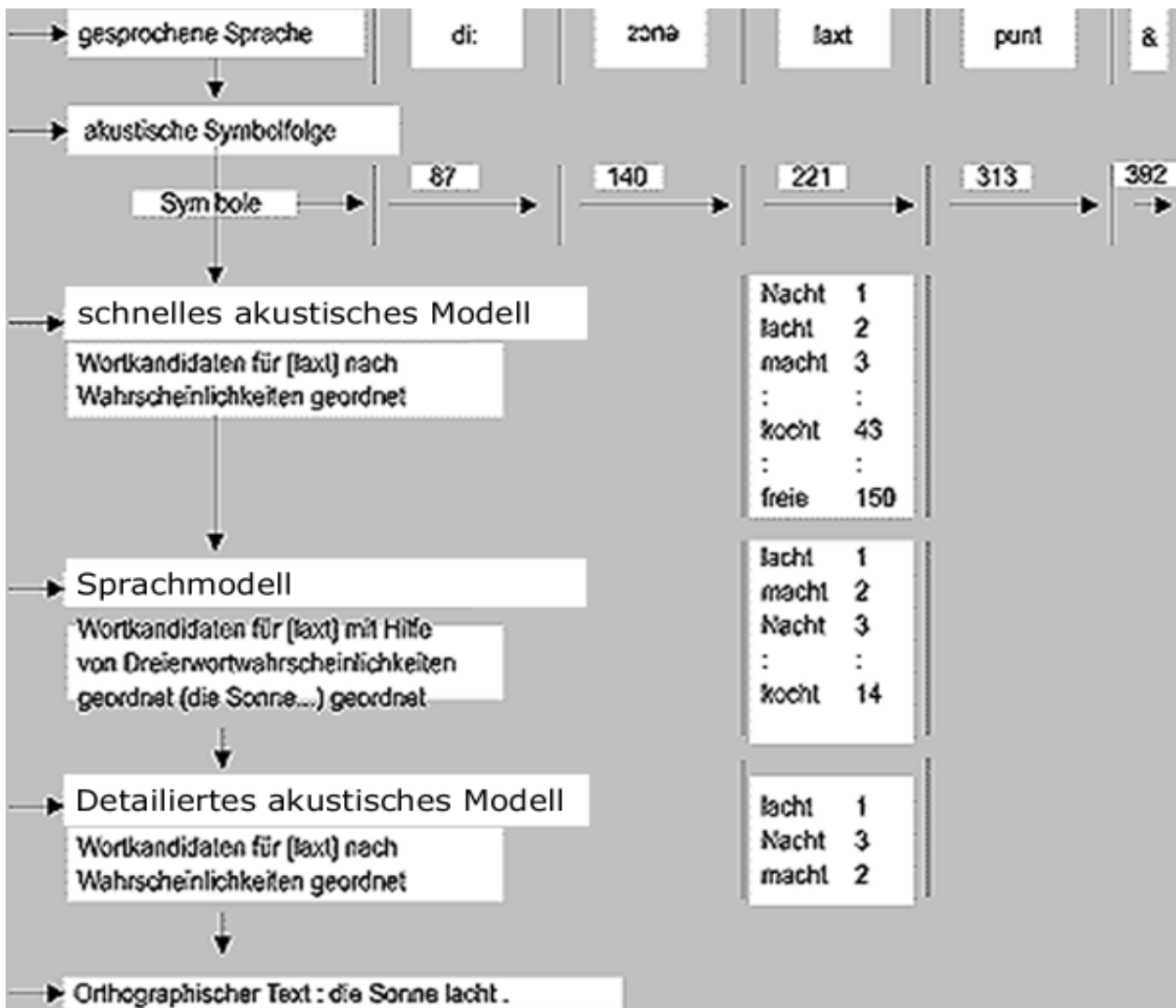


Wortebene - Trigramme mit Häufigkeit

Ich rufe an	80
von neun bis	77
Anfang nächster Woche	69
Ihnen das recht	67
lassen Sie uns	61
am Freitag den	59
am Dienstag den	59
wenn Ihnen das	58
rufe an wegen	58
ich weiß nicht	54
am Mittwoch den	54
am Donnerstag den	53
halten wir das	52
wäre Ihnen das	52
oder Anfang nächster	50
ich freue mich	49
es geht um	49
wenn es Ihnen	49
von mir aus	44
paßt es Ihnen	44

Alle Wahrscheinlichkeiten summiert ergibt großen Rechenaufwand.

Hybride Verfahren



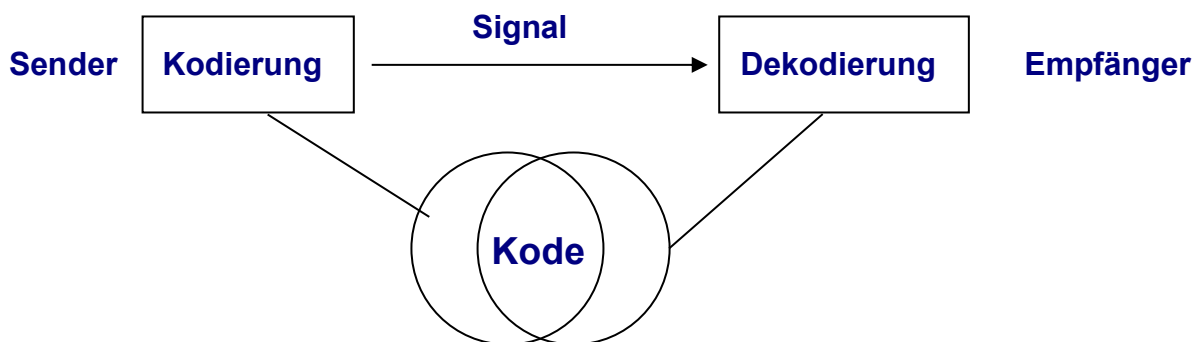
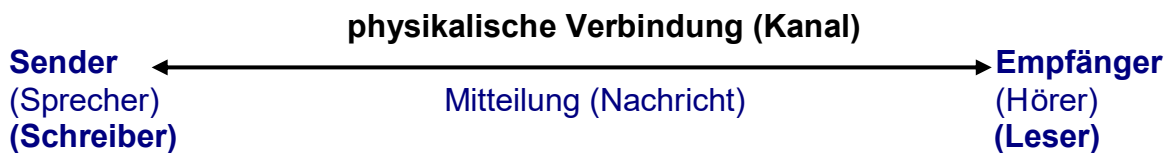
Architektur des Spracherkenners ViaVoice

S 2.1 Grundlagen der Graphemik und maschinelle Schrifterkennung

Lernziel: Schrift als eigenes Zeichensystem sehen, das mit strukturalistischen Methoden bearbeitet werden kann sowie eine allgemeine Vorstellung davon, was Schrifterkennung und Rechtschreibprüfung leisten muss.

1. Grundlagen der Graphemik

Verallgemeinerung: Semiotik - Lehre von den Zeichen - Kommunikationsmodell



Schrift (Graphemik) ist nur ein Sonderfall der Kodierung.

Verschiedene Schriftsysteme:

- Pictographie – Bilder
- Ideographie - Begriffszeichen
- Logographie - Symbole für Wörter, Sätze
- Phonographie - Zeichen für Sprachlaute

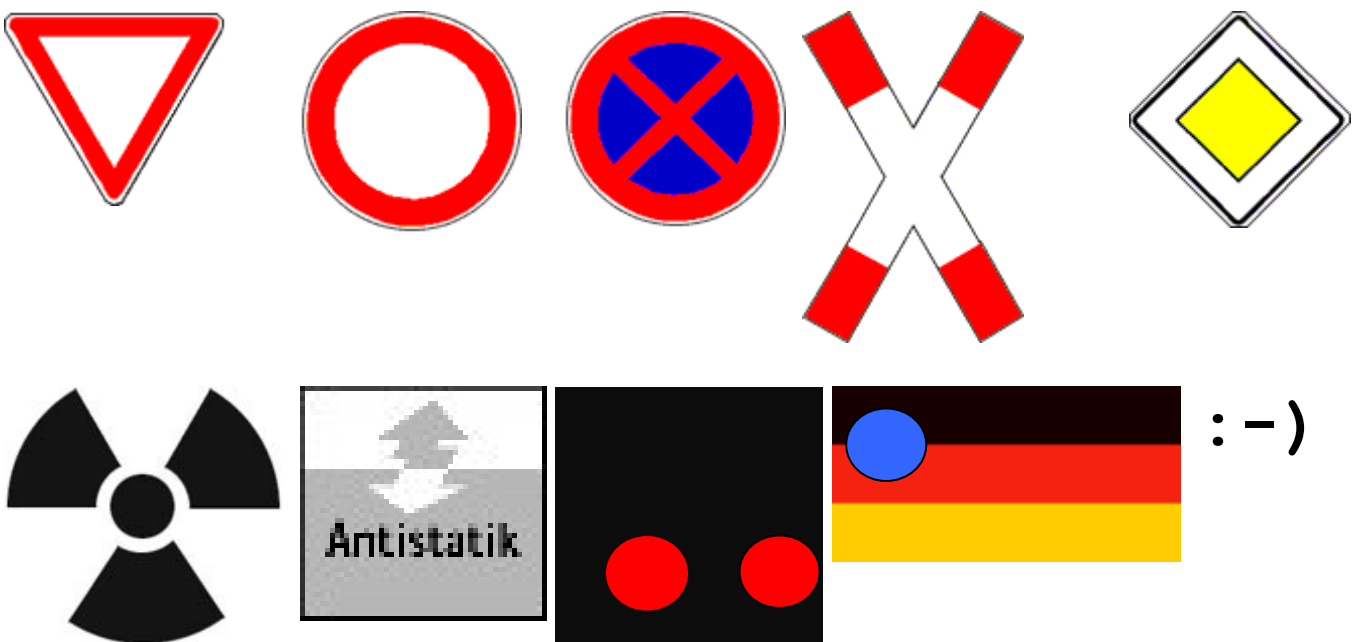
Piktogramme

Bildzeichen, Bedeutung unmittelbar einsichtig, daher auch sprachunabhängig und international verwendbar. Keilschriften, Hieroglyphen und die chinesische Schrift basieren in ihren Frühformen auf Piktogrammen.



Ideogramme

Ideogramme sind Graphiken, deren Bedeutung durch Konvention geregelt ist. Man muss lernen, was sie bedeuten.



Logographie

Ein Zeichen entspricht einem Wort oder einem Satz, z.B.

1, 2, 3, 4,

§ % €



Hieroglyphen, stehen für Wörter oder für Silben

风 日 月 星 云 雨

Wind

Sonne

Mond

Sterne

Wolken

Regen

Chinesische Wortzeichenzeichen

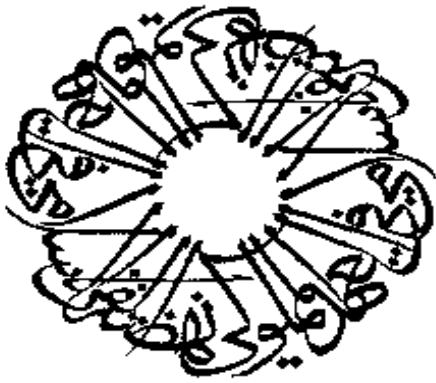
Phonographie

Die Schrift beruht auf der Aussprache, Erfindung des Alphabets, die kleinsten Zeichen sind nicht mehr bedeutungstragend.

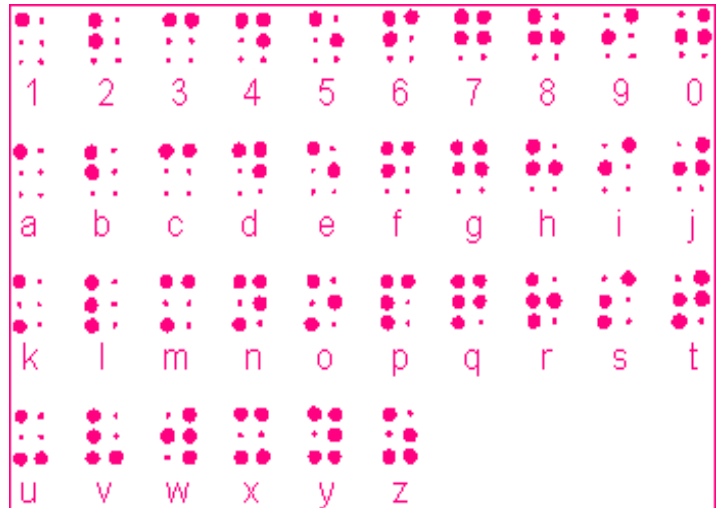
Übersicht : einige fremde Schriften © Peter Dörfling

russisch	ОТКЛОНИЛО ЛИНИЮ МОИХ ПОИСКОВ В ак говорил Мэнни. Его расчеты ило ли это, что ему вообще некого
griechisch	<i>Ἐρθα δὴ καὶ Ἰφικράτης εἰς Φλειοῦντα γοῖς δὲ λεηλατῶν, βοηθησάντων τῶν ἐκ τοσοῦτους ὥστε καὶ τοὺς Λακεδαιμονίους</i>
armenisch	Հայաստան աշխարհի հնամենի Նախիջևանի իրով ստեղծված-արարված բազմահազար կոթ վերաբանդակներից մեր օրերն հասած առավ
georgisch	რომი როდესაც მოკვდებო ვინმე უკვთუ ოქოდგან ვინმე წარვოდეს
hebräisch - Israel	יְהוָה אֱלֹהֵינוּ יְהוָה אֱלֹהֵינוּ יְהוָה אֱלֹהֵינוּ יְהוָה אֱלֹהֵינוּ
amharisch - Äthiopien	አሁን ፣ ከመዘ ፣ አፋቶ ፣ አገዛ ፣ ለዓለም ፣ አሰን ፣ ወልደ ፣ ዋሕደ ፣ ወከመ ፣ ከሉ ፣ ዘየአም አላ ፣ ደረቡ ፣ ሕይወት ፣ ዘለዓለም ጦጦ ፣ አይ
arabische Nesghi-Schrift	سفرات السندباد منقولة من كتاب الف ليلة وليلة حكاية السندباد قالت شهرزاد
hindi : Indien	ग्यारह बजे वहां पहुंचा था और पौने सन्देश लाया हूँ कि कमला अपनी भाः
gudjerati : Indien	ખસીને કવિ તરફ ગયેલું લાગે - ખાસ કરીને સાધના, કવિની શ્રદ્ધા, કવિનો સર્જનવ્યા
tamil : Indien, Sri Lanka	சுதந்திர புருஷர்களாய் இந்த மண்ணில் களின் திணைவு தோன்றி அவர்களைப்போ
singhalesisch : Sri Lanka	ගැනීමයි. එවිටයි සාහිත්‍ය කලාවෙන් කළ වෘත්ති විස්ලේෂ සර්වක වින්තෝ. සමී

<p>Burmesisch</p>	<p>ဒီဟာမြန်မာ့ပုံနှိပ်စာလုံးများပဲ။ လှတယ်မဟုတ်လား။</p>
<p>Laotisch</p>	<p>ເພາະວ່າພະເຈົ້າໄດ້ຮັບພະບຸລະໄວກພຽງປີ້ສົມ ດ້ວຍທາງແ ພິໄດ້ປະທານໃຫ້ປະຊາຊົນເບິ່ງເບິ່ງເບິ່ງວ່າ ສົມ ບໍ່ເປັນປິດ .</p>
<p>Thai</p>	<p>ได้ประดิษฐ์ขึ้นโดยดัดแปลงให้ทันสมัย แยกตัวพิมพ์แต่ละตัว จากแถวยาวของ</p>
<p>Kambodjanisch</p>	<p>លំនៅឡំនៅគ្រប់ដុះវៃល្អដំបូល ។ ដុះវៃល្អដំ សលមានលំនៅឡើង ។ លំនៅឡំមានបន្ទប់</p>
<p>Chinesisch</p>	<p>體育之宜提倡人盡皆知以其 能增進健康也德國各學校各</p>
<p>Japanisch</p>	<p>次第行ケバ深山モアサモヨ イ木曾路ノ旋ニ出デウヨ詞</p>
<p>Japanisch</p>	<p>のツドイ以んしライのあすきそ にゆ ハ人イサでだてンでるるてはなほえ</p>
<p>Koreanisch</p>	<p>눈때 있어서 결장적인 역할을 합은 누구나다 방치되어 있다. 큰 비풍을 두고 있다. 그러나,</p>
<p>Mongolisch</p>	
	<p>Der Stein von Rosetta, ca. 1960 vor Christus.</p> <p>Text in 3 verschiedenen Schriften: in Hieroglyphen, in Demotisch (eine einfachere Schrift, die sich aus den Hieroglyphen entwickelte) und in Griechisch.</p>



Arabisch, Kaligraphie



Braille Blindenschrift

Definitonen analog zur Phonemik:

GRAPHEM = kleinste bedeutungsunterscheidende Schrifteinheit im Sprachbewußtsein

GRAPH = aktuelles Schriftzeichen, durch das ein Graph realisiert wird

ALLOGRAPH = unterschiedliche Graphen, die dasselbe Graph realisieren, Graphemvarianten

Allographie, bedingt durch verschiedene Fonts und durch den Kontext (Ligaturen, Schreibschrift)

Wiederum strukturalistische Analyse

Experimente, Minimalpaare, Austausch, Umstellung

ergibt

- distinktive Opposition
- Distribution
- System

2. Maschinelles Lesen (Optical Character Recognition, OCR)

- **on-line Schreibschrifterkennung**

Schreiben mit einem speziellen Stift auf einer elektronischen Oberfläche.

Die Koordinaten der fortlaufend geschriebenen Punkte werden in chronologischer Reihenfolge gespeichert.

Dies ergibt eindimensionale Daten des Schreibvorgangs,

- **off-line Erkennung von Druck- und Schreibschrift**

Einscannen einer beschriebenen oder bedruckten Seite als Bild.

Die zeitliche Komponente der Schriftentstehung fehlt hier.

Es muss ein zweidimensionales Bild analysiert werden

➤ Technisch gehört OCR zum Gebiet der Mustererkennung (ähnlich wie die Erkennung gesprochener Sprache)

- **Sequentiell ablaufende Erkennungsalgorithmen -**

arbeiten eine Liste von Beobachtungen ab, die sie mit einem bestimmten Wortmuster vergleichen, von links nach rechts .

➤ Am Schreibvorgang orientiert.

- **Wahrnehmungsorientierte Algorithmen -**

versuchen dagegen beliebige Buchstaben eines Wortbildes zu identifizieren.

Als nächstes wird ein Entscheidungsprozess durchlaufen, der die am sichersten erkannten Buchstaben mit möglichen Buchstaben für die Lücken zwischen ihnen auszufüllen versucht. Dieses Verfahren funktioniert analog zum menschlichen Lesen, bei dem auch anhand der hervorstechendsten sicher erkannten Buchstaben auf das ganze Wort geschlossen wird.

➤ Am Modell des menschlichen Lesevorgangs orientiert.

Interessant im Zusammenhang mit einer gleichzeitigen Rechtschreibkorrektur.

Vgl. nächstes Kapitel.

S 2.2 Orthographie und maschinelle Rechtschreibhilfen

Lernziel: Das Verhältnis von Laut zu Schrift ist in verschiedenen Sprachen verschieden. Welche Probleme ergeben sich daraus?

Zur Einstimmung

Betört in Nju Dschörsie

Ich fuhr mit Jacqueline nach **New Jersey**
 Für mich war es klar: Ich **betersey**.
 Doch schon beim Versuch,
 da sagte sie: "Huch -
 doch nicht im Jacquett, ich **beschwersey**."

Faxen in Taxen

In **Tucson**, da traf ich zwei **Sucson**,
 die hatten den Kopf voller **Fucson**.
 Die ließen bewußt
 sich auch auf der Brust
 zwei rassige Vollbärte **wucson**.

Gekurt in Scherbuhr

René **fourg zourg kourg** nach **Cherbourg** -
meat Vollbart und **Beatlefreasourg**.
 Dort gab man **eam nourg**
 eine **Chourg (meat Rasourg)**
und cheackte ean kourgzweg retourg.

von Alfred Schratzer

Das Verhältnis von Graphemen zu Phonemen ist nur eindeutig in phonetischer Lautschrift

p	as in play /pleɪ/	i	as in tree /tri/
b	as in boy /bɔɪ/	ɪ	as in insect /ɪnsekt/
t	as in tree /tri/	e	as in bet /bet/
d	as in dog /dɒg/	æ	as in cat /kæet/
k	as in cat /kæet/	ɑ	as in car /kɑ/ British, /kɑr/ American
g	as in got /gɒt/	ɒ	as in bought /bɒt/
tʃ	as in chin /tʃɪn/	ɔ	as in saw /sɔ/
dʒ	as in judge /dʒʊdʒ/	ʊ	as in as in foot /fʊt/
f	as in fine /faɪn/	u	as in boot /but/
v	as in visit /vɪzɪt/	ʌ	as in up /ʌp/
θ	as in thin /θɪn/	ɜ	as in fur /fɜ/ British, /fɜr/ American
ð	as in then /ðen/	ə	as in ago /ə'gəʊ/
s	as in sick /sɪk/	eɪ	as in play /pleɪ/
z	as in zoo /zu/	əʊ	as in know /nəʊ/
ʃ	as in shine /ʃaɪn/	aɪ	as in sky /skaɪ/
ʒ	as in vision /vɪʒn/	aʊ	as in how /haʊ/

In den natürlichen Sprachen dagegen entsprechen sich Phoneme und Grapheme meist nicht

- ein Laut - mehrere Grapheme
 - ein Laut - verschiedene Grapheme
 - ein Graphem - verschiedene Laute
- Die Schrift ist ein eigenes System mit eigenen Prinzipien. Lesen und Hören sind teilweise unabhängig voneinander.

Zuordnungen Phoneme <-> Grapheme

Phoneme	Grapheme bzw. Graphemvarianten	Phoneme	Grapheme bzw. Graphemvarianten
/a/	<a>	/b/	<b, bb>
/e/	<e, ä>	/d/	<d, dd>
/i/	<i>	/g/	<g, gg>
/o/	<o>	/f/	<f, ff, v>
/u/	<u>	/s/	<s, ss, ß>
/ö/	<ö>	/ʃ/	<sch, s>
/ü/	<ü>	/h/	<h>
/a:/	<a, aa, ah>	/x/	<ch, g>
/e:/	<e, ee, eh>	/v/	<w, v, u>
/i:/	<i, ie, ih, ieh>	/z/	<s>
/o:/	<o, oo, oh>	/j/	<j>
/u:/	<u, uh>	/m/	<m, mm>
/ä:/	<ä, äh>	/n/	<n, nn>
/ö:/	<ö, öh>	/ŋ/	<ng, n>
/ü:/	<ü, üh>	/l/	<l, ll>
/a/ + /e/	<ei, ai>	/r/	<r, rr>
/a/ + /o/	<au>	/k/ + /s/	<x>
/o/ + /ö/	<eu, äu>	/p/ + /f/	<pf>
/p/	<p, pp, b>	/t/ + /s/	<z, ts, tz>
/t/	<t, tt, dt, d>	/t/ + /ʃ/	<tsch>
/k/	<k, ck, q, ch, g>	/z/	<s>

Prinzipien der Rechtschreibung im Deutschen

- **Phonematisches Prinzip**

Ein Graph repräsentiert ein Phonem.

finden, aber, Institut

- **Morphematisches (etymologisches) Prinzip**

Morpheme bleiben in verschiedener Umgebung erhalten.

Haus - Häuser (Heuser), Fall - fällen (*fellen), Tage - Tag (*Tak)*

- **Grammatisches Prinzip**

Großschreibung von Eigennamen, Satzanfang, Substantiven

Satzzeichen z.B. zwischen Haupt- und Nebensatz

- **Semantisches Prinzip**

Unterscheidung von Homonymen

leeren - lehren, Saite - Seite, wieder - wider, ißt - ist

- **Historisches Prinzip**

überkommene Schreibweisen z.B. Dehnungs-e, ursprünglich 2 Laute

lieben, viel, Soest

besonders viele Beispiele historischer Schreibweise im Englischen!

Die Prinzipien für die Rechtschreibung sind z.T. unvereinbar und daher wird keins durchgehalten. Daraus resultieren immer wieder Reformversuche.

Hier einige Resultate der jüngsten Rechtschreib-Reform:

vor der Reform	nach der Reform
achtgeben preisgeben anheimgeben	Acht geben preisgeben anheim geben
allzumal allzubald	allzumal allzu bald
aufeinandertreffen zusammentreffen	aufeinander treffen zusammentreffen
bewußtmachen klarmachen haltmachen	bewusst machen klarmachen Halt machen
heute früh heute abend Dienstag abend	heute früh heute Abend Dienstagabend
im allgemeinen seit langem	im Allgemeinen seit langem

Maschinelle Rechtschreibprüfung und Korrektur

Bedeutung der Rechtschreibkorrektur

- andere Systeme können bei falscher Eingabe nicht funktionieren
- unsichere Eingaben (Spracherkennung, OCR) können mit denselben Mitteln verbessert werden

Typen von Fehlern

- Kompetenzfehler (Rechtschreibung wird nicht beherrscht)
- Performanzfehler (Vertipper: Auslassung, Einfügen, Vertauschen)
- Komplexe Fehler

Typische Kompetenzprobleme und die jüngste Rechtschreibreform

- Laut-Buchstaben-Zuordnung
ss/ß - hassen /Hass, müssen/muss, lassen/ließ
e/ä - Hand /behände
 drei Buchstaben *Schiffahrt*
- Getrennt- und Zusammenschreibung
 früher *Auto fahren - radfahren, sitzen bleiben - stehenbleiben*
- Schreibung mit Bindestrich
- Groß- und Kleinschreibung
 jetzt mehr Großschreibung, auch wenn es keine richtigen Substantive mehr sind
an Stelle -> anstelle von, mit Hilfe -> mithilfe von
- Zeichensetzung

Methoden der Fehlererkennung

- Auf Silbenebene. Zulässige Buchstabensequenzen prüfen *Hand *Hadn *Hnad*
- Wörter im Lexikon aufsuchen, wenn nicht vorhanden, Fehler annehmen
- Morphologiekomponente einsetzen (Flexion, Derivation, Komposition berücksichtigen – sonst würden viele korrekte Wörter als Fehler zählen)
- Kontext berücksichtigen (partieller Parser)

Korrektur

- von Nicht-Wörtern (d.h. nicht im Lexikon gefunden)
- von zwar existierenden aber im Kontext unpassenden Wörtern (mithilfe der Grammatik und eines Parsers)

*mir der Hand - * mit der Hund*

Verschiedene Korrektur-Algorithmen

- Distanzfunktionen, um (im Wörterbuch oder durch Veränderungen) die ähnlichsten Wörter zu finden

		Mengendistanz	nicht-übereinst. Positionen	minimale Zahl von Operationen
A	a b e n d s			
B1	a b h n d s	1	1	1 x tauschen
B2	a b n e d s	0	2	1 x vertauschen
B3	a b n d s	1	4	1 x einfügen
B4	a b w e n d s	1	5	1 x tilgen

außerdem

- Speichern häufig falsch geschriebener Wörter
- Speicherung von häufigen Verwechslungen
- n-Gramme

Vorschläge zur konsequenten Lösung der deutschen Rechtschreibproblematik

Erster Schritt:

Wegfall der Großschreibung: einer sofortigen Einführung steht nichts im Weg, zumal schon viele Grafiker und Werbeleute zur Kleinschreibung übergegangen sind.

Zweiter Schritt:

Wegfall der Dehnungen und Schärfungen. Diese Maßnahme eliminiert schon die größte Fehlerursache in der Grundschule; den Sinn oder Unsinn unserer Konsonantenverdopplung hat ohnehin niemand kapirt.

Dritter Schritt:

v und ph ersetzt durch f, z ersetzt durch s, sch ersetzt durch s; das Alphabet wird um zwei Buchstaben reduziert, Schreibmaschinen und Schreibmaschinen vereinfachen sich, wertvolle Arbeitskräfte können der Wirtschaft zugeführt werden.

Vierter Schritt:

q, c und ch ersetzt durch k, j und y ersetzt durch i, pf ersetzt durch f - iest sind schon sechs Buchstaben ausgespart. So kann die Suleit sofort von neun auf zwei Iare verkürzt werden; anstatt aksig präsent Rechtsreibunterricht können nützlichere Fächer wie Physik, Chemie, Rechnen mehr gepflegt werden.

Fünfter Schritt:

Wegfall von ä, ö und ü: alles überflüssige ist iest ausgespart. Die Orthografie wird slikt und einfach. Natürlich benötigt es einige Zeit, bis diese Vereinfachung überall richtig verstanden ist, fileikt sungsweise ein bis zwei Iare.

Sechster Schritt:

Anschließend nun dürfte als nächstes die Vereinfachung der noch schwierigeren und unsinnigeren Grammatik anvisiert werden.

ges.: die alternative Dudenredaktion

S 3.1 Silbenbau und Prosodie

Lernziel: Silbenbau und Prosodie als eigene Ebene des Sprachsystems verstehen

1. Silbenbau

Silbe regelhafte Folge von Phonemen, die bei der Aussprache – insbesondere bezüglich der Tonhöhe - eine Einheit bilden.

Silbenbau eine Art "Phonemsyntax" – nach sprachspezifischen Regeln

Vgl.

<i>un-ser</i>	<i>*rse</i>
<i>res-te</i>	<i>*sre</i>
<i>brau-chen</i>	<i>*nche</i>

Wieviel Phoneme bilden im Deutschen eine Silbe?

- 1 *A-pril, I-dee, e-ben, Ö-fen, U-fer*
- 2 *un-ser, ma-chen*
- 3 *un-ser, ma-chen*
- 4 *Volk, Land, mar-schert, wäh-rend*
- 5 *künst-lich, vor-wärts, Start*
- 6 *selbst*
- 7 *greifst*
- 8 *streichst*

Die Silbengipfel (Vokale) in Orthographie

V	Länge			Diphthonge					
	e	h	Doppel	a	e	i	o	u	y
a		ah	aa		ae	ai		au	ay
ä		äh					äo	äu	
e		eh	ee	ea		ei	eo	eu	ey
i		ih		ia			io		
	ie	ieh							
o		oh	oo	oa		oi		ou	oy
ö		öh					öö		
u		uh		ua		ui			
ou				oua		oui			
ü		üh							
y									

Die Silbenauslaute (Konsonanten) in Orthographie

E1= Auslaut in Wortmitte, E2=Auslaut am Wortende, E1/E2 gehört zu beiden Mengen

E1/E2	E2					
	Zweiter Konsonant	Flexionsendungen				
	Doppel	b ch d f g k l m n p pf sch v z	s	st	t	nd
b			bs	bst	bt	
		bsch				
ch			chs chts	chst	cht	
		chz			chzt	
d			ds	dst	dt	
f			fs	fst	ft	
	ff		ffs	ffst	fft	
			fts			
					fzt	

(Fortsetzung)

E1/E2	E2	
	Zweiter Konsonant	Flexionsendungen

	Doppel	b ch d f g k l m n p pf sch v z	s	st	t	nd
--	---------------	--	----------	-----------	----------	-----------

g			gs	gst	gt	
		gd				
	gg			ggst	ggd	
b			bs	bst	bt	
		bsch				
ch			chs	chst	cht	
			chts			
		chz			chzt	
d			ds	dst	dt	
f			fs	fst	ft	
	ff		ffs	ffst	fft	
			fts			
					fzt	
g			gs	gst	gt	
		gd				
	gg			ggst	ggd	
k			ks	kst	kt	
			kts			
c-k	ck		cks	ckst	ckt	
l			ls	lst	lt	lnd
		lb	lbs	lbst	lbt	
		ld	lds			
		lch				
		lf	lfs	lfst	lft	
	ll		lls	llst	llt	
		lg	lgs	lgst	lgt	
		lk	lks	lkst	lkt	
		lm	lms			
		ln				
		lp				
		lsch			lscht	
			lts	ltst		
		lz			lzt	

(Fortsetzung)

E1/E2	E2					
	Zweiter Konsonant	Flexionsendungen				
	Doppel	b ch d f g k l m n p pf sch v z	s	st	t	nd

m			ms	mst	mt	
			mts			
		mb		mbst	mbt	
		md				
		mf				
		mp	mps	mpst		
		mpf		mpfst	mpft	
	mm		mms	mmst	mmt	
		msch		mschst	mscht	
n			ns	nst	nt	
			nts			
		nch				
		nd	nds		ndt	
		nf			nft	
			nfts			
		ng	ngs	ngst	ngt	
		nk	nks	nkst	nkt	
			nkts			
	nn		nns	nnst	nnt	
		nsch				
		nz			nzt	
p			ps		pt	
			pts			
		pf	pfs	pfst	pft	
	pp		pps	ppst	ppt	
ph						rnd

(Fortsetzung)

E1/E2	E2	
	Zweiter Konsonant	Flexionsendungen

	Doppel	b ch d f g k l m n p pf sch v z	s	st	t	nd
--	--------	--	---	----	---	----

r			rs	rst	rt	
			rts	rts		
		rb	rbs	rbst	rbt	
		rch		rchst	rcht	
		rd	rds			
		rf	rfs	rfst	rft	
		rg	rgs	rgst	rgt	
		rk	rks	rkst	rkt	
		rl	rls	rlst	rlt	
		rm		rmst	rmt	
		rn	rns	rnst	rnt	
		rp	rps		rpt	
	rr			rrst	rrt	
		rsch		rrschst	rrscht	
		rv		rvst	rvt	
		rz			rzt	
s						
sch				schst	scht	
		sk				
st						
ß						
t			ts			
		tsch				
	tt		tts	ttst	ttst	
th						
		tz			tzt	
v			vs			
x					xt	
z					zt	
	zz					

2. Prosodie

Die Prosodie (auch. "Intonation") besteht aus (im Deutschen) aus

- **Tonhöhen** (Stimmbänder, Grundfrequenz)
- **Akzent** (Energie bei der Aussprache)
- **Pausen**

Die Prosodie dient u.a. zur Bedeutungsunterscheidung

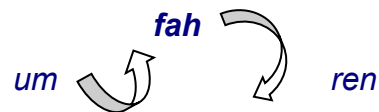
Beispiel mehrdeutig

" *Sollen wir den Laternenpfahl umfahren?*"

Mit Tonhöhenverlauf eindeutig::



dagegen fahren, so dass
er umfällt



drum herum fahren

Intonation nach Alexander V. Isacenko, Hans-Joachim Schädlich: *Untersuchungen über die deutsche Satzintonation*. Studia Grammatica VII, Berlin 1996

- Das wesentliche Merkmal der Prosodie / Intonation ist die **Tonhöhenbewegung**

grob 3 Tonhöhenbewegungen:



Iktus betonbare Silbe im Wort *um-fahren* versus *um-fah-ren*

Tonbruch ab diesem Punkt steigt oder fällt die vorher fallende oder steigende Stimme

- Jede abgeschlossene Äußerung hat mindestens einen Tonbruch.

3. Funktion der Tonbrüche

➤ Tonbrüche zeigen vor allem die Rolle eines Sprechaktes oder eines Satzes im Text an.

Test. Versuchen Sie die Beispiele möglichst monoton zu sprechen wobei Sie zwischen zwei Tonhöhen wechseln.

____ _

Der Iktus (vulgo Akzent) liegt bei *Berliner* auf der zweiten Silbe

- **prä-iktisch steigend**

Kündigt fallenden Tonbruch an. Bleibt dieser aus: **Satzfragment** oder **Weiterführung**

Ich bin ein Ber-  *lin-er* *aber kein So-*  *zial-* *hilfeempfänger*

- **post-iktisch steigend**

Kündigt fallenden Tonbruch an. Bleibt dieser aus: Normalintonation für **Frage**

Ich bin ein Ber lin-  *er*

- **prä-iktisch fallend**

Wenn es der letzter Tonbruch im Satz ist: **Aussage**

und *ich bin ein Ber-*  *lin-er*

Wenn es nicht der letzte Tonbruch im Satz ist: **Phrasierung**

- **post-iktisch fallend**

Wenn es der letzter Tonbruch im Satz ist: Normalintonation für **Aussage**

Ich bin ein Ber- *lin-*  *- er*

Wenn es nicht der letzte Tonbruch im Satz ist: **Kontrastierung des Wortes**

der *Ken-*  *nedy* *ist ein Berliner* *nicht der* *Bush*

S 3.2 Maschinelles Zeilenumbruch mit Silbentrennung

Lernziel: Nutzenanwendung der Silbenanalyse von S 3.1

PSG-Grammatik der Silben des Deutschen

WORT \rightarrow EINSILB

WORT \rightarrow ZWEISILB

WORT \rightarrow MEHRSILB

EINSILB \rightarrow A1 + V + E2

EINSILB \rightarrow A1 + V

EINSILB \rightarrow V + E2

ZWEISILB \rightarrow ASILB + ESILB

ASILB \rightarrow A1 + V + E1

ASILB \rightarrow V + E1

ASILB \rightarrow A1 + V

ASILB \rightarrow V

ESILB \rightarrow A2 + V + E2

ESILB \rightarrow V + E2

ESILB \rightarrow A2 + V

MEHRSILB \rightarrow ASILB + MSILB + ESILB

MSILB \rightarrow A2 + V + E1

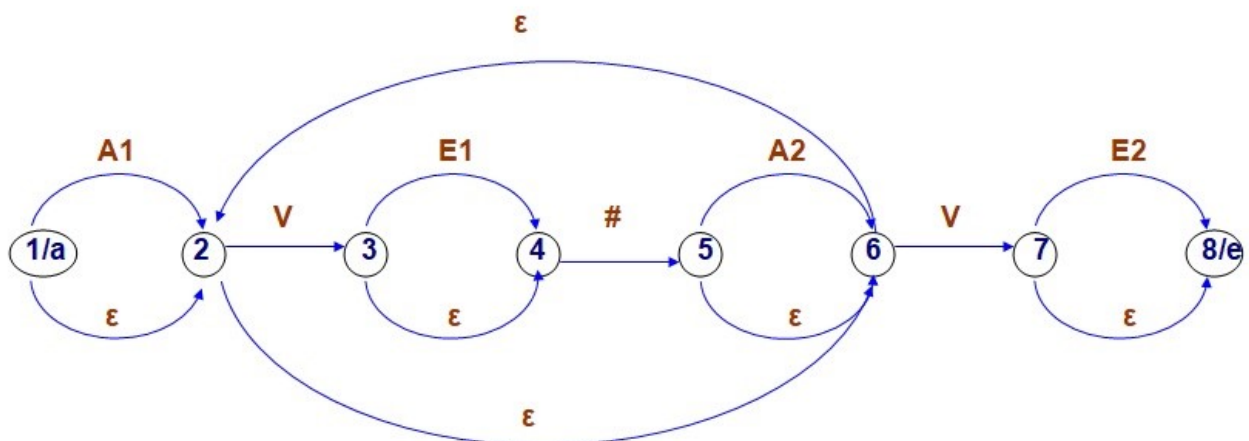
MSILB \rightarrow V + E1

MSILB \rightarrow A2 + V

MSILB \rightarrow MSILB + MSILB

Die Ersetzung der Konstituenten A1, A2, E1, E2, V erfolgt nach den Tabellen in S 3.1.

Darstellung als Übergangsnetzwerk



/a Eingang, /e Ausgang, # bedeutet Silbengrenze

Darstellung als sogenannter "regulärer Ausdruck":

$A1? (V E1? \# A2?)* V E2?$

Beispiel:

ver - zehrt

Mar - me - la - de

$A1 V E1 \# A2 V E2$

$A1 V E1 \# A2 V \epsilon \# A2 V \epsilon \# A2 V \epsilon$

Problem bei der Wortbildung

Leider überlappen sich bei der Wortbildung Auslaut **E2** und Anlaut **A1**. Deshalb gibt es zwei Möglichkeiten der Zerlegung:

WACHS#TUBE - WACH#STUBE
STAUB#BECKEN - STAUB#ECKEN

Bei den folgenden Beispiele wäre die alternative Zerlegung FALSCH.

INHALTS#ANALYSE	ARBEITS#PAPIER	BAUCH#REDNER
ABEND#ROETE	ARMIERUNGS#TRUPPE	BERG#LAND
ABGANGS#PRUEFUNG	ARSCH#LEDER	BLATT#ROLLER
AB/GAS#TURBINE	AUF#REIZUNG	BRAT#ROST
AB#LAGE	AUSDRUCKS#TANZ	DIENST/ENT#HEBUNG
AB#RUESTUNG	BARSCHE#LACHS	ENT#RICHTUNG
ADELS#TITEL	BART#HAAR	GAS/AUS#TAUSCH

Es muss also der Silbentrennung erst eine Wortzerlegungskomponente vorgeschaltet werden.

Beachte:

➤ **Silbengrenze** ist nicht immer gleich **Morphemgrenze**

Morpheme *Welt, Welt + en*

Silben *Wel # ten*

Kompositum *Welt +# en # de*

Auf jeden Fall zwei Schritte:

Zuerst Morphologische Zerlegung von Komposita und von Präfixen und Stämmen

Textverarbeitungssysteme

/ | \

Text verarbeitungs systeme

/ / | \

Text ver arbeitsungs systeme

Dann Silbentrennung der Fragmente des 1. Schritts nach phonologischen Regeln

S 4.1 Grundlagen der Morphologie , Morphologische Phänomene

Lernziel: Vor allem wieder die strukturalistische Methode (Segmentieren, Klassifizieren, Distribution) sowie die wichtigsten Grundbegriffe der Morphologie

Morphologie: Die Beschreibung der kleinsten bedeutungstragenden Einheiten der Sprache und ihren Kombination im Wort, Lehre von der Form der Wörter einer Sprache. Eine Ebene zwischen Phonologie und (Satz)syntax.

Syntax	<i>Lisa</i>	+	<i>studiert</i>	+	<i>Computerlinguistik</i>
Morphologie	<i>Lisa</i>		<i>studier - t</i>		<i>Computer-linguistik</i>
Phonologie	/li:s/a/		/□/t/u:d/i:r/t/		/k/]m/p/u:t/β/li/□g/v/i/s/t/i/k/

Wort: Einheit an der Schnittstelle zwischen Syntax und Morphologie, kleinste Einheit der Syntax – größte Einheit der Morphologie

Strukturalistische Methode:

segmentieren, kombinieren, vergleichen, austauschen

jetzt nicht Laute, sondern kleinste bedeutungstragende Teilchen: die Morphe

Beispiel Suaheli (Ost Afrika):

atanipenda	er wird mich mögen
atakupenda	er wird dich mögen
atampenda	er wird ihn mögen
atutupenda	er wird uns mögen
atawapenda	er wird sie mögen
nitakupenda	ich werde dich mögen
nitawapenda	ich werde sie mögen
utampenda	du wirst ihn mögen
tutampenda	wir werden ihn mögen
watampenda	sie werden ihn mögen
atakusumbua	er wird dich ärgern
unamsumbua	du ärgerst ihn
atanipiga	er wird mich schlagen
anakupiga	er schlägt dich
alinipiga	er schlug mich

Aus welchen Morphemen bestehen die Ausdrücke in Suaheli?

a-ta-ni-penda	er wird mich mögen
a-ta-ku-penda	er wird dich mögen
a-ta-m-penda	er wird ihn mögen
a-ta-tu-penda	er wird uns mögen
a-ta-wa-penda	er wird sie mögen
ni-ta-ku-penda	ich werde dich mögen
ni-ta-wa-penda	ich werde sie mögen
u-ta-m-penda	du wirst ihn mögen
tu-ta-m-penda	wir werden ihn mögen
wa-ta-m-penda	sie werden ihn mögen
a-ta-ku-sumbua	er wird dich ärgern
u-na-m-sumbua	du ärgerst ihn
a-ta-ni-piga	er wird mich schlagen
a-na-ku-piga	er schlägt dich
a-li-ni-piga	er schlug mich

Morpheme mit deutscher Bedeutung:

a	er (subjekt)
m	ihn (objekt)
ni	ich (subjekt), mich (objekt)
u	du (subjekt)
ku	dich (objekt)
tu	wir (subjekt), uns (objekt)
wa	sie (subjekt), sie (objekt)
ta	futur (tempus)
na	präsens (tempus)
li	präteritum (tempus)
penda	mögen (stamm)
sumbua	ärgern (stamm)
piga	schlagen (stamm)

Bildungsprinzip:

Subjekt - Tempus - Objekt - Stamm

wa li ku penda

sie <vergangenheit> dich mögen - sie mochten dich

Kriterien für Morphem

- Element hat Bedeutung, und zwar eine aus der sich die Bedeutung des Größeren gleichförmig zusammensetzt (Frege-Prinzip)
- Element kann nicht in kleinere Einheiten zerlegt werden, für die (a) gilt

mehrere Morpheme:

an-geben (reichen), ab-geben
Ein.fahrt, Aus-fahrt
Jung-tier, Alt-kanzler

nur ein Morphem:

angeben (wichtig tun)
Einsatz, Aussatz
Jungfrau, Jungfer

Sprachsystem und Realisierung genauso wie Phonem - Phon - Allophon

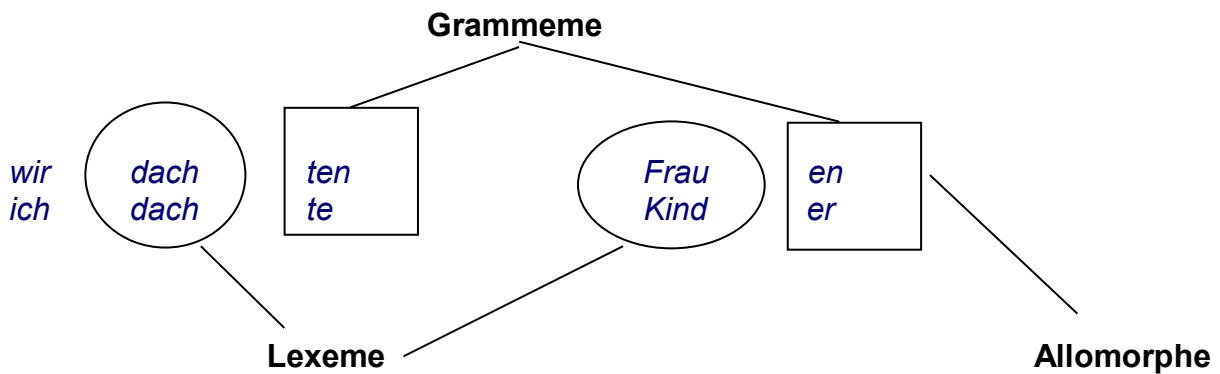
Morphem: die kleinste **bedeutungs-tragende** Einheit des Sprachsystems
(vgl. Phonem: die kleinste **bedeutungs-unterscheidende** Einheit des Sprachsystems)

Morph: die Realisierung eines Morphems

Allomorph: eine von alternativen Realisierungen eines Morphems

Morpheme nach ihrer Funktion:

lexikalische versus grammatische Morpheme - Lexeme und Grammmeme



Ein Gedicht mit korrekten englischen Grammmemen aber fremden Lexemem

Song of the Jabberwocky' in *Alice in Wonderland* by Lewis Carroll

*Tw'as brillig and the slithy toves
Did gyre and gimble in the wabe
All mimsy were the borogoves
And the mome raths outgrabe.*

Verdaustig wars und glasse Wieben
Rotterten gorkicht im Gemank;
Gar elump war der Pluckerwank,
Und die gabben Schwisel frieben.

Morphologische Operationen:

- **Flexion**

Markierung der Wörter mit grammatischen Merkmalen (Morphosyntax). Bei der Flexion entstehen keine neuen Lexeme, sondern nur neue Grammemen derselben Wortart (Wortart = Grammemklasse – nach syntagmatischen Kriterien)

machen, machst, machten, machtet, gemacht

- **Derivation**

Ableitung mithilfe von Wortbildungsmorphemen. Bei der Derivation entstehen neue Wortarten (d.h. neue Grammemklassen) und damit auch neue Lexeme

machen, Macher, gemachtes - Sand, sandig - gesund, Gesundheit – entsorgen, Entsorgung

- **Komposition**

Zusammensetzung aus mehreren Lexemen gleicher oder verschiedener Wortart.

Informations-technologie, Arbeits-beschaffung, Un-recht, Donaudampfschiffahrtsgesellschaftskapitänsmütze, dreihundrtvierundsiebzigtausendachthundertundneunzig

Komposition sollte man besser schon zur Syntax rechnen. Das zeigt der Vergleich Deutsch – Französisch:

<i>Leberwurst</i>	saucisse de foie
<i>Kalbsleberwurst</i>	saucisse de foie de veau
<i>Kräuterkalbsleberwurst</i>	saucisse de foie de veau aux herbes
<i>Katen-(oder Land)kräuterkalbsleberwurst</i>	saucisse "fermière" de foie de veau aux herbes
<i>Billigkatenkräuterkalbsleberwurst</i>	saucisse fermière de foie de veau aux herbes en promotion

Wichtiger Begriff

Paradigma: Gruppe von Ausdrücken mit einem bestimmten gemeinsamen Merkmal, auf Grund dessen sie in einem Austauschverhältnis zueinander stehen.

Einige Verbparadigma des Deutschen (Kombination mehrerer grammatischer Merkmale):

mach-e	Verb	1.PersSg	Präsens	Indikativ	finit
mach-st		2.PersSg			
mach-t		3.PersSg			
mach-en		1.PersPl			
mach-t		2.PersPl			
mach-en		3.PersPl			

mach-t-e	Verb	1.PersSg	Präteritum	Indikativ	finit
mach-t-est		2.PersSg			
mach-t-e		3.PersSg			
mach-t-en		1.PersPl			
mach-t-e		2.PersPl			
mach-t-en		3.PersPl			

mach-e	Verb	1.PersSg	Präsens	Konjunktiv	finit
mach-est		2.PersSg			
mach-e		3.PersSg			
mach-en		1.PersPl			
mach-et		2.PersPl			
mach-en		3.PersPl			

mach-en	Verb		Singular	Infinitiv		
ge-mach-t					Plural	Partizip 1
mach-end						Partizip 2
mach						Imperativ
mach-t						Imperativ

Deutsch – Substantivparadigmen

Staat	Substantiv	Nominativ	Singular	Maskulinum
Staat-es		Genitiv		
Staat		Dativ		
Staat		Akkusativ	Plural	
Staat-en		Nominativ		
Staat-en		Genitiv		
Staat-en		Dativ		
Staat-en		Akkusativ		

Frau	Substantiv	Nominativ	Singular	Femininum
Frau		Genitiv		
Frau		Dativ		
Frau		Akkusativ	Plural	
Frau-en		Nominativ		
Frau-en		Genitiv		
Frau-en		Dativ		
Frau-en		Akkusativ		

- Jedes Merkmal bedeutet Gleichheit in einem bestimmten Verhältnis und Unterschied in einem anderen.

Wort - Wortformen:

Ein Wort ist ein Paradigma mit identischem Lexem und Wortart-Grammem

Wortformen sind die einzelnen konkreten Wörter im Paradigma

Einige morphologische Bildungsprinzipien

I. Hinzufügung (Affix)

- **Suffigierung:** ein Suffix ist ein Affix, das nach dem Basismorphem steht.
Schön+heit, schäd+ig+en
- **Präfigierung:** ein Präfix ist ein Affix, das vor dem Basismorphem steht
un+schön
- **Infigierung:** ein Infix ist ein Affix, das innerhalb des Basismorphems steht.
Tagalog: sulat - schreiben, s-um-ulat - geschrieben haben
takbo - laufen, t-um-akbo - gelaufen haben
- **Zirkumfigierung:** bei Zirkumfixen steht ein Teil des Morphems vor, ein anderer nach dem Basismorphem.
ge-mach-t

II. Ersetzung (Substitution)

- **Ersetzung** eines oder mehrere Phoneme des Basismorphems durch andere Phoneme.
Mutter. Mütter (Umlaut), singen, sang (Ablaut).

III. Weglassung (Elision)

- Die **Weglassung** eines Phonems des Basismorphems.
flexibel, flexible

Der strukturalistischer Ansatz passt hauptsächlich zu agglutinierenden Sprachen (im Gegensatz zu fusionierenden Sprachen)

Einige Probleme und entsprechende theoretische Konstrukte:

Null-Allomorph (z.B. eine "Nicht-Endung" im Austausch mit Endungsmorphen)

Frau-en

Frau-

Leib-er

Leib

Garten

Gärten

Portmanteau-Morph (mehrere Morpheme verschmelzen in einem Morph)

geb-en, gab

im

zur

diskontinuierliches Morphem (mehrere Morphe zusammen repräsentieren ein Morphem)

ge-macht-t (ge + t = Partizip)

S 4.2 Maschinelle morpho-syntaktische Analyse und Klassifizierung

Lernziel: Eine prinzipielle Vorstellung davon, wie morphologische Phänomene mit dem Computer gemeistert werden. Ein weiteres Beispiel für „Rechnen mit Sprache“.

Aufgabe: Die Wörter in einzelne Morphe zerlegen und den Morphen Morpheme zuordnen.

- **Lemmatizierer** - erkennt das lexikalische Morphem (Lemma, Grundform)
- **Tagger** - versieht Wörter im laufenden Text mit grammatischen Morphemen (Kategorien)

1. Komplexe Kategorien zur Klassifizierung von Morphen

➤ **Einem Morph entspricht eher ein Bündel von Morphemen statt ein einzelnes Morphem.**

Terminologie: Statt von Morphemen sprechen wir hier von (lexikalischen und grammatischen) Merkmalen. Jedes Merkmal wird als **komplexe Kategorie** notiert, d.h. es besteht aus

- **Typ** engl. attribute z.B. Numerus z.B. Kasus
- **Wert(en)** engl. value z.B. Singular, Plural z.B. Nominativ, Genitiv, Dativ, Akkusativ

Der Merkmalstyp steht voran. Die Werte folgen in eckigen Klammern. Alternative Werte werden durch Komma getrennt:

ärgere wortart[verb] lexem[ärgern] person[1] numerus[singular]
ärgerst wortart[verb] lexem[ärgern] person[2] numerus[singular]
ärgert wortart[verb] lexem[ärgern] person[3] numerus[singular]
ärgern wortart[verb] lexem[ärgern] person[1, 3] numerus[plural]
ärgert wortart[verb] lexem[ärgern] person[2] numerus[plural]

Unifikation - Operationen über den Werten von Merkmalstypen

Merkmalstypen als Variable, Merkmalswerte als Belegungen.

Berechnung der Kompatibilität für komplizierte Strukturformeln.

Beispiel Kongruenz

wir ärgern uns

(akzeptieren)

sie ärgern sich

(akzeptieren)

** wir ärgern sich*

(nicht akzeptieren)

Formel:

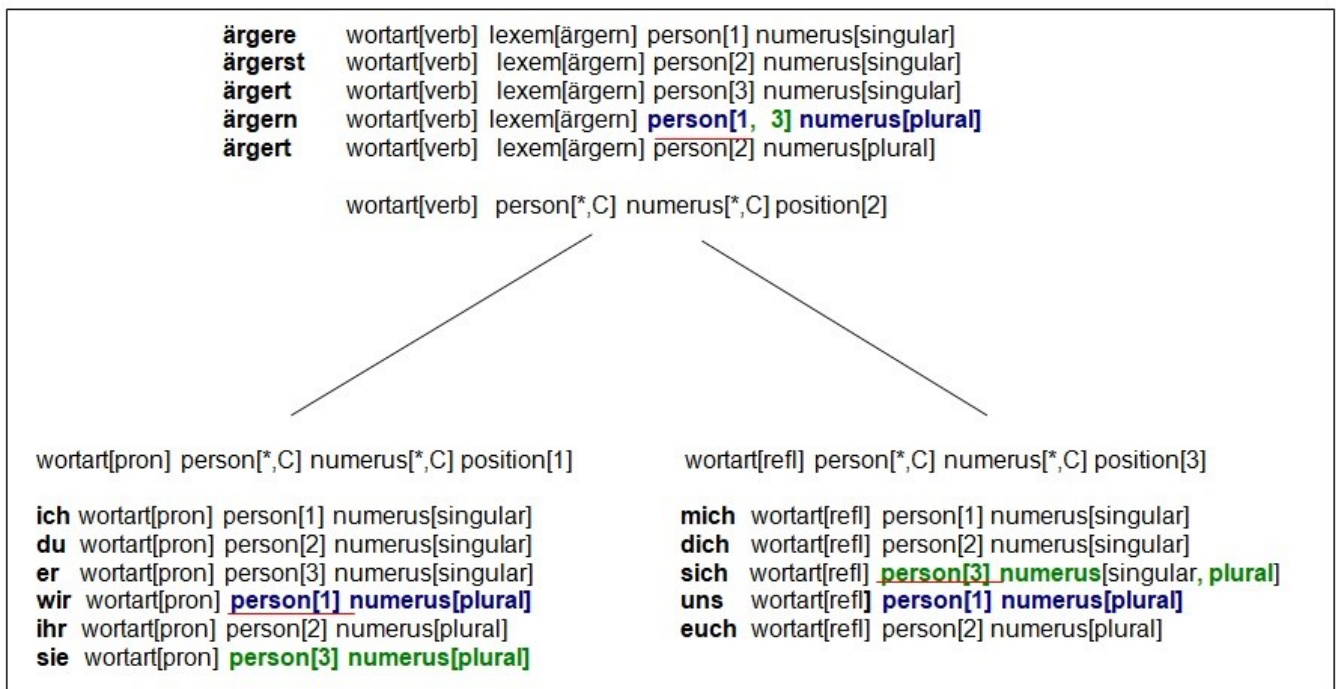
(wortart[verb] person[* ,C] numerus[* ,C] position[2])

(wortart[pron] person[* ,C] numerus[* ,C] position[1])

(wortart[refl] person[* ,C] numerus[* ,C] position[3])

* = beliebige Belegung C = es muss einen nicht-leeren Durchschnitt der Werte geben

Belegung der Elemente der Formeln mit den komplexen Merkmalen der Wörter:



wir ärgern uns

(akzeptiert)

sie ärgern sich

(akzeptiert)

** wir ärgern sich*

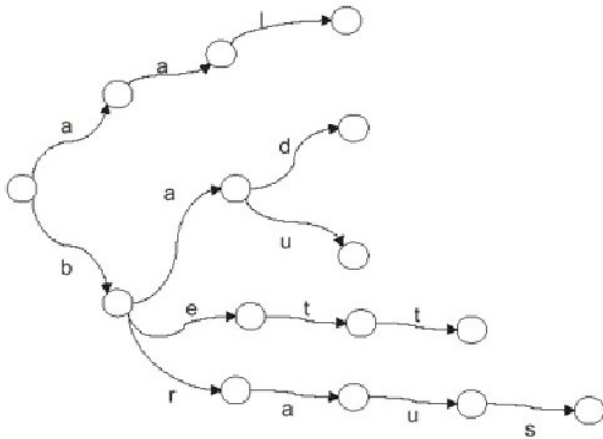
(nicht akzeptiert)

2. Modelle für die maschinelle morphologische Analyse

- **item and arrangement model** - Wortsegmente werden aneinandergereiht
ge-seh-en-es
- **item and process model** - auf Wörter werden bestimmte Operationen angewendet
gast, g(a>ä)st-e
lauf-en, l(au>ie)f
- **word and paradigm model** - Wörter werden als ganze zu einem Paradigma angeordnet
bin, bist, ist, sind, seid, sind

Item-and-arrangement model

wird implementiert als **finites Übergangsnetzwerk** ("finite state transition network", FTN)
 Dabei sind die Übergänge jeweils ein Buchstabe. Gleiche Übergänge werden nur einmal notiert.
 Das Lexikon der Wörter hat daher die Form eines Buchstabenbaums.

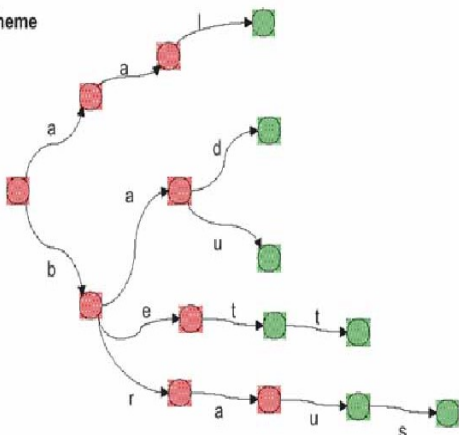


Dieses Netz enthält die Wörter: **aal, bad, bau, bett, braus**

Zusammensetzung von Netzwerken für Stämme und Endungen

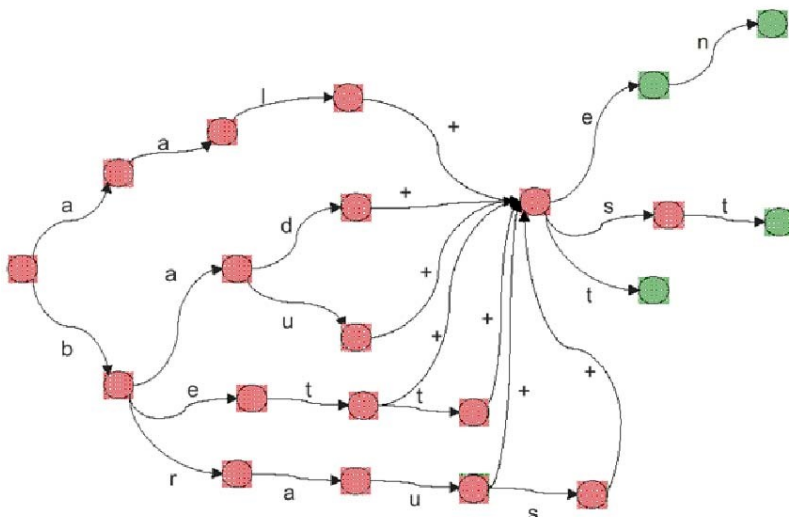
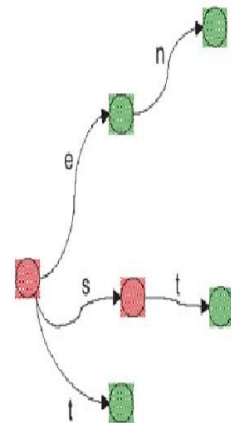
Verbstamm-Morpheme

aal
 bad
 bau
 bet
 bett
 brau
 braus



Verbsuffix-Morpheme

-e
 -en
 -st
 -t



Probleme	Lösung
<p>Sproßvokal</p> <p>reden --> redete (nicht: redte) atmen --> atmete (nicht: atmte)</p> <p>Tilgung des Vokals nach r,l</p> <p>handeln (nicht: handelen)</p> <p>Konsonantenschwung</p> <p>rasen --> rast (nicht: rasst)</p>	<p>verschiedene Endungsnetze:</p> <p>red - ete, etest, ete, eten atm -</p> <p>handel - e, st, t, n, t, n</p> <p>ras - e, t, t, en, t, en</p>
<p>Umlaut</p> <p>Mann, Männer Buch, Bücher Sohn, Söhne Baum, Bäume</p>	<p>mehrere Stämme pro Wort im Netz:</p> <p>Mann, Buch, Sohn, Baum - 0, es Männ, Büch - er, ern Söhn, Bäum - e, en</p>

Deutsche Verben mit Ablaut

Stämme		Endungen						
denk	prs-st	prs-st	e	st	t	en	t	en
dach	prt-te	prt-te	te	test	te	ten	tet	ten
ge-dach	ptz-t	ptz-t	t					
fass	prs-146	prs-146	e			en		en
faß	prs-235-t	prs-235-t		t	t		t	
faß	prt-te							
ge-faß	ptz-t							
sprech	prs-1456-t	prs-1456-t	e			en	t	en
sprich	prs-23-st	prs-23-st		st	t			
sprach	prt-st	prt-st	-	st	-	en	t	en
ge-sproch	ptz-en	ptz-en	en					
tret	prs-1456-et	prs-1456-et	e			en	et	en
tritt	prs-23-st0	prs-23-st0		st	-			
trat	prt-est	prt-est	-	est	-	en	et	en
ge-tret	ptz-en							

Item and Process Model

Es gibt Sprachen mit anderen Mechanismen als dem schlichten Aneinanderreihen von Morphen, z.B. verschiedene Arten von Reduplikation:

(1) Warlpiri

<i>kurdu</i>	child	<i>kurdu-kurdu</i>	children
<i>kamina</i>	girl	<i>kamina-kamina</i>	girls
<i>mardukuja</i>	woman	<i>mardukuja-mardukuja</i>	women

(2) Agta

<i>takki</i>	leg	<i>taktakki</i>	legs
<i>labáng</i>	patch	<i>lablabáng</i>	patches
<i>uffu</i>	thigh	<i>ufuffu</i>	thighs

(3) Samoan

<i>taa</i>	strike sg	<i>tataa</i>	strike pl
<i>nofo</i>	sit sg	<i>nonofo</i>	sit pl
<i>maliu</i>	die sg	<i>maliliu</i>	die pl

(4) Chamorro

<i>dankolo</i>	big	<i>dankololo</i>	very big
<i>bunita</i>	pretty	<i>bunitata</i>	very pretty
<i>metgot</i>	strong	<i>metgogot</i>	very strong

Hier eignet sich ein Item-and-Process Modell.

Im Prinzip Transformationen

- Beschreibung der Eingabe
- Beschreibung der Ausgabe

Item and Process z.B.. mit sog. **regulären Ausdrücken**

- beliebige Buchstaben: **a,b,c,d,e.... A,B,C,D**
- Klammern: **(,), {, }**
- "oder": **|**
- "keinmal oder beliebig oft": *****
- "einmal oder beliebig oft": **+**
- optional: **?**

Prozess am Beispiel Agta

<i>takki</i>	leg	<i>taktakki</i>	legs
<i>labáng</i>	patch	<i>lablabáng</i>	patches
<i>uffu</i>	thigh	<i>ufuffu</i>	thighs

1. Definition von Konsonanten, Vokalen und beliebigen Buchstaben:

 $\setminus C = \{t \mid k \mid l \mid b \mid n \mid g \mid f\}$
 $\setminus V = \{a \mid i \mid u\}$
 $\setminus X = \{C \mid V\}$

2. Beschreibung der Eingabe (d.i. die Grundform):

1. Klasse $(\setminus C) (\setminus V) (\setminus C) (\setminus X^+)$

2. Klasse $(\setminus V) (\setminus C) (\setminus X^+)$

3. Beschreibung der Veränderung (für jeden Klammerausdruck in der Eingabe steht eine Variable \$):

1. Klasse $\$1 \$2 \$3 \$1 \$2 \$3 \$4$

2. Klasse $\$1 \$2 \$1 \$2 \$3$

Item-and-Paradigm Model

Keine Zerlegung in Stamm und Endung, sondern komplette Wörter im Paradigma,
Dies ist geht immer und ist bei irregulären Formen nötig.

item-and-arangement		item-and-paradigm
mach-	-e -st -t -en -t -en	bin bist ist sind seit sind

Konzeptuell ist das Item-and-Paradigm Model grundlegend.

➤ Die Morphologie läuft immer auf die Formalisierung von Paradigmen hinaus.

S 5.1 Heuristische Grundlagen der Syntaxanalyse

Lernziel: Die heuristischen Tests für die Syntaxanalyse beherrschen, in jeder linearen Abfolge von Wörtern die impliziten syntaktischen Strukturen sehen und operational nachweisen können.

Syntax = die Lehre von der Syntagmatik d.h. den Verbindungen,

hier der Verbindung der Wörter im Satz.

Oder m.a.W. der Aufbau von zulässigen Ausdrücken einer Sprache aus Grundelementen, speziell der Aufbau der Sätze einer natürlichen Sprache (Satzsyntax)

Was ist ein Satz?

Nach Aristoteles: eine Verbindung mehrerer Wörter, die einen vollständigen Gedanken ausdrücken.

Unter kommunikativem Gesichtspunkt: eine potentiell autonome Äußerung (*ein Sprechakt*)

➤ Wesentlich ist eine semantische Zweigliederigkeit: in Referenz und Prädikation, selbst bei einem Einwortsatz

	Referenzobjekt	Prädikat
<i>Die Vorlesung fällt aus.</i>	die Vorlesung	ausfallen
<i>Wie viel Uhr ist es?</i>	Uhrzeit	jetzt-gegeben-sein
<i>Halt die Klappe!</i>	du	Klappe-halten
<i>Rauchen verboten</i>	Rauchen	verboten-sein
<i>Keine Werbung</i>	Werbung	nicht-einwerfen
<i>Hilfe</i>	ich, jemand	zur Hilfe kommen

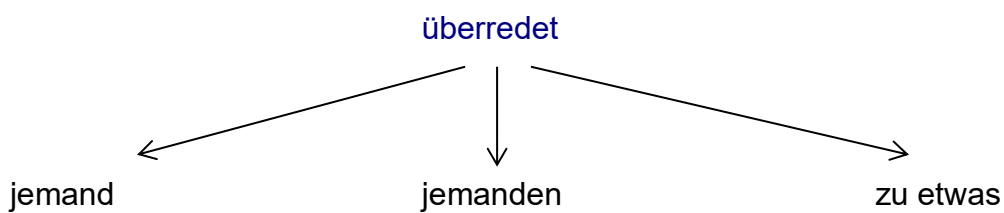
Prädikate allein sind „ungesättigt“ (Frege), sie müssen durch Bezeichnungen für Referenzobjekte ergänzt werden,

Man sagt: Prädikate haben **Ergänzungen** (in anderer Terminologie: Argumente)

Die Anzahl der Ergänzungen kann je nach Prädikat verschieden sein:

- 1 Er schläft.
- 2 Er verpasst die Vorlesung.
- 3 Anni hilft mir bei der Hausaufgabe.
- 4 Fritz legt seinem Freund die Hand auf die Schulter.

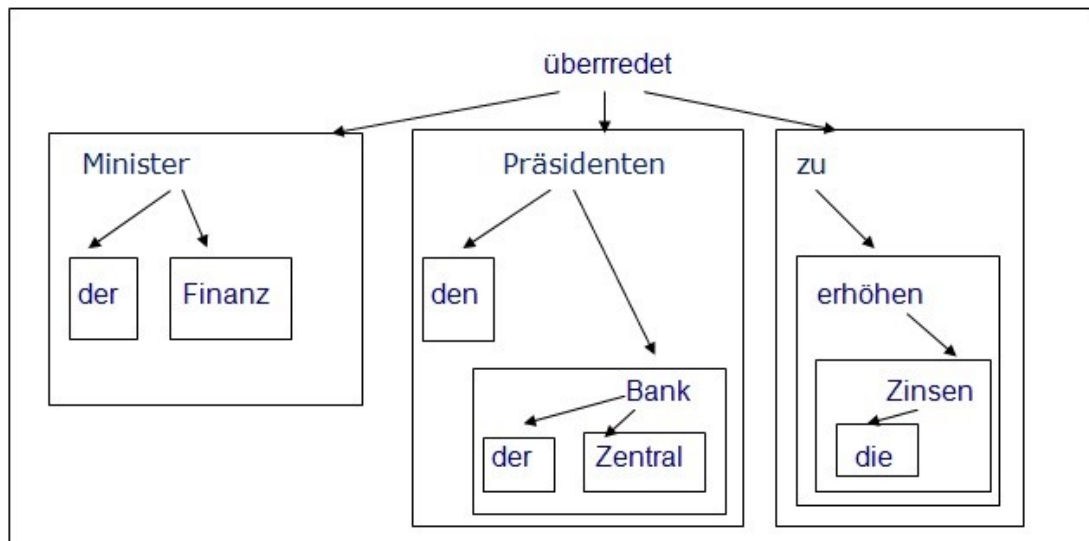
Syntaktischer Aufbau



<i>Lisa</i> <i>Wer zu feige ist</i> <i>Der Finanzminister</i>	<i>ihren Freund</i> <i>jemand anderen</i> <i>den Präsidenten der</i> <i>Zentralbank</i>	<i>zu einer Erdbeerbowle</i> <i>(dazu) die Suppe auszulöffeln</i> <i>(dazu) die Zinsen zu erhöhen</i>
---	--	---

Ergänzungen

E1 der „Überreder“	E2 der „Überredete“	E3 der Überredungsinhalt
--------------------	---------------------	--------------------------



Heuristische Tests

- Linguistik ist eine experimentelle Wissenschaft!

(1) Verschiebeprobe (Permutation)

Welche Einheiten können nur als ganze im Satz verschoben, z.B. vor das finite Verb im Hauptsatz gestellt werden?

Könnte ich das geblünte Kleid im Schaufenster ausprobieren

Das geblünte Kleid im Schaufenster könnte ich ausprobieren

Das geblünte Kleid könnte ich im Schaufenster ausprobieren

Zweck: Zur Bestimmung von Satzgliedern, wie Subjekt, Objekte, Adverbialbestimmungen im Unterschied zu Attributen und Partikeln.

(2) Austauschprobe (Kommuation)

Lassen sich bestimmte Elemente (Wörter) an ihrer syntaktischen Stelle systematisch gegen andere ihrer Wortklasse austauschen, oder können nur bestimmte Lexeme dort stehen?

*Das Gesetz tritt in Kraft (*in Macht)*

Er gab seine Bemühungen auf

**Er verstärkte seine Bemühungen auf*

Zweck: Zur Bestimmung von festen Syntagmen, z.B. des lexikalischen mehrteiligen Prädikats, im Unterschied zu offenen syntagmatischen Positionen.

Welche Wortgruppen lassen sich durch ein einziges Wort ersetzen (z.B. durch ein Personalpronomen oder ein Fragepronomen)?

Ich meine, daß er versprochen hat, das Handout zu kopieren

Ich meine das.

Könnte ich das geblünte Kleid im Schaufenster ausprobieren

Könnte ich das dort ausprobieren?

Was könnte ich ausprobieren?

Zweck: Zur Bestimmung von syntaktischen Einheiten: was sich durch ein Wort ersetzen lässt, muss auch eine Einheit bilden.

Im Rahmen welcher morpho-syntaktischen Merkmale (Wortart, Satzart, Merkmale wie Kasus, Genus usw.) lässt sich eine gegebene syntaktische Einheit kommutieren? Hängt dieser Rahmen vom Vorhandensein einer anderen Einheit ab?

Er/sie/der Freund/wer müde ist schläft

Er liebt seine Freundin/Geld/was aufregend ist

Er gibt ihr eine Mark

**Er schläft seine Freundin*

**Er liebt seine Freundin eine Mark*

Zweck: Zur Bestimmung des hierarchischen Aufbaus, Feststellen von Valenz und Rektion: die im bestimmten Rahmen kommutierende Einheit ist von der Einheit, welche die Merkmale bestimmt, abhängig; ist ihr untergeordnet.

Lassen sich bestimmte problematische Einheiten gegen ein Element mit einer festgelegten Semantik austauschen?

Ich wasche mich /dich / mein Auto

*Ich freue mich / *dich/ *mein Auto*

Wir waschen uns / einander

**Wir freuen einander*

Zweck: Zur semantischen Differenzierung: um z.B. festzustellen, wann ein reziprokes Verhältnis vorliegt, prüft man, ob sich das (mehrdeutige) Reflexivpronomen gegen das (eindeutige) Wort einander austauschen lässt.

(3) Weglassprobe, ggf. Implikationstest

Welche von zwei vermuteten syntaktischen Einheiten kann weggelassen werden? Welche kann allein stehen bleiben, welche kann nicht ohne die andere stehen bleiben? Welche Einheiten können nur entweder beide stehen, oder beide wegfallen?

Wir haben etwas gegen Frost

Wir haben etwas

**Wir haben gegen Frost*

Zweck: Zum Feststellen syntaktischer Beziehungen und zur Bestimmung des hierarchischen Aufbaus: Setzt eine Einheit die andere voraus, so besteht zwischen beiden eine syntaktische Relation; dabei ist die nicht-weglassbare der weglassbaren übergeordnet.

Impliziert ein Satz ohne Weglassung den Satz mit Weglassung der Konstituente (d.h. muss der zweite Satz immer auch als wahr gelten, wenn man den ersten Satz als wahr akzeptiert)?

Er kommt schnell -> er kommt

*Er kommt vielleicht -> *er kommt*

Zweck: Zur semantischen Differenzierung: hierin unterscheiden sich z.B. die Modalbestimmung von den Satzadverbialen.

(4) Exklusions- und Koordinationsprobe

Können zwei syntaktische Konstituenten ohne weiteres zusammen in einem Satz vorkommen oder nur wenn man sie mit und, oder oder einer anderen Konjunktion koordiniert?

**Er schlug die Schaufensterscheibe und den Weg zum Bahnhof ein*

Er spielte die Flöte und die Gitarre

**Er spielte die Flöte und den ganzen Abend*

Er spielte die Flöte den ganzen Abend

Zweck: Zur Bestimmung von syntagmatischen Positionen: Elemente, die ohne Koordination im Satz gleichzeitig vorkommen können, gehören naturgemäß verschiedenen Positionen an (für die man dann je eine andere Kategorie vorzusehen hat); Elemente, die nur koordiniert vorkommen können, zählen zur selben syntaktischen Position und Kategorie; Elemente, die nicht koordiniert werden können, gehören zu verschiedenen Kategorien.

(5) Paraphraseprobe (Transformation)

Wie kann man den Inhalt eines Satzes anders, u.U. expliziter ausdrücken? Man forme dabei der Satz syntaktisch um, behalte das lexikalische Material aber möglichst bei. Lässt sich nun ein anderer Satz, dessen Verhältnis zu dem ersten zur Debatte steht, völlig analog umformen, und erhält man auch hier eine Paraphrase?

Sie ist leicht zu lieben

Es ist leicht, sie zu lieben

Sie ist bereit zu lieben

**Es ist bereit, sie zu lieben*

schwer, schwerlich, unmöglich, innig?

Zweck: Zur Herausarbeitung syntaktischer Kategorien und Strukturen: ist eine völlig analoge Umformung möglich, ist bis auf weiteres (d.h. bis u.U. irgendwo eine divergierende Umformbarkeit in Beziehung auf eine andere Paraphrase entdeckt wird) eine identische Struktur beider Sätze anzunehmen; divergieren die Paraphrasemöglichkeiten, so haben die Sätze verschiedenen Aufbau, auch wenn sie oberflächlich gleich aussehen.

(6) Weitere semantische Tests (Erfragbarkeit, Negationstest, Imperativtest u.a. Spezialtests

Kann eine bestimmte syntaktische Einheit auf natürliche Weise erfragt werden? (d.h. ist ein Dialog akzeptabel, in dem eine Frage gestellt wird, in der an der Position der syntaktischen Einheit ein Fragepronomen steht. Und kann als Antwort die betreffende syntaktische Einheit selbst gegeben werden?) Mit welchem Fragewort kann die Einheit erfragt werden?

*Das Gesetz tritt in Kraft - *in was tritt das Gesetz*

Was tritt in Kraft?

Sie fahren mit Abstand am besten.

Wie fahren Sie am besten - mit Abstand

Wie fahren Sie? - mit Abstand am besten

Zweck: Zur semantischen Differenzierung der Einheiten, z.B. der Modalbestimmungen (erfragbar) und der Satzadverbien (nicht erfragbar).

Lässt sich die Negationspartikel nicht einer bestimmten Einheit voranstellen? Lässt sich ein Satz unter Beibehaltung bestimmter Einheiten in einen Befehl umformen? usw.

Er schläft tief - er schläft nicht tief

*Er schläft hoffentlich - *er schläft nicht hoffentlich*

**Er schläft tief nicht - er schläft hoffentlich nicht*

*Schlafe tief! *Schlafe hoffentlich!*

Zweck: Zur weiteren semantischen Differenzierung.

Einige wichtige syntaktische Phänomene

Dimensionen:

- (i) die lexikalische Bedeutung,
- (ii) die morpho-syntaktische Form,
- (iii) die syntagmatische Funktion.

Eine Dimension umfasst jeweils gleiche Bedeutung oder Form oder Funktion:

- (i) *geben, gibst, gabst*
- (ii) *gibt, liebt, singt*
- (iii) *Jens lügt*
der Mann lügt
wer das gesagt hat, lügt
wer lügt ?

Mehrdeutigkeiten

- (i) Ober: *Wie fanden Sie denn das Kotelett?*
Gast: *Indem ich die Kartoffel beiseite schob*
- (ii) Kunde: *Ich möchte gern RUMKUGELN*
Verkäufer: *Von mir aus. Aber nicht hier im Laden!*
- (iii) *Dem Junggesellen fehlt zum Glück die Frau*
- (iv) Kunde: *Ich brauche ein Paar Unterhosen*
Verkäufer: *Lange?*
Kunde: *Wieso? Ich will doch keine mieten sondern welche kaufen.*

Valenz

schlafen - jmd

lieben - jmd, jmden

geben - jmd, jmdem, etwas

wissen - jmd, dass etwas der Falls ist

ärgern - jmd, sich, über etwas

gespannt - sein, jmd, auf etwas

Liebe - zu ...

Brücke - über ...

Kongruenz

der alte Herr, der sich im Wartezimmer befindet

die alten Damen, die sich im Wartezimmer befinden

die Staubecken - das Staubecken

Idiomatische Wendungen

Das Gesetz tritt in Kraft.

Er schlug zunächst die Schaufensterscheibe und dann den Weg zum Bahnhof ein.

Wir haben etwas gegen Frost.

Hierarchische Struktur

Ich meine, dass er versprochen hat, das Handout zu kopieren.

Könnte ich das geblümete Kleid im Schaufenster anprobieren?

Sie fahren mit Abstand am besten

Sie ist leicht zu lieben

Sie ist bereit zu lieben.

Lineare Gliederung

Jens hat zweimal ein Examen gemacht

Jens hat ein Examen zweimal gemacht

Diskontinuierliche Konstituenten

Er gab ihr das Geld

Er gab seine Bemühungen auf

Er gab seine Hausaufgaben ab

Äpfel haben wir heute keine.

What do you think Gudrun said she feeds her cats?

Ellipsen

Die linke Seite hat nach oben die rechte nach unten eine Öffnung

Das Geld hat ihm mehr genützt als uns.

Das Geld hat ihm mehr genützt als die Empfehlung.

(Was sollen wir der Antje mitbringen?)

Jens meint Vanilleeis

Jens meint, daß wir der Antje Vanilleeis mitbringen sollen

S 5.2 Formale Syntaxbeschreibung

Lernziel: Verständnis für das Modell der generativen Grammatik und erste Erfahrung davon, wie man eine schreibt.

Sätze bestehen offenbar aus Teilen, die Teile bestehen wiederum aus Teilen, und so weiter bis zu den kleinsten Elementen (normalerweise Wörtern).

Satzanalyse genauso: Man zerlegt den Satz in größtmögliche Teile, diese wiederum in größtmögliche Teile, usw. bis man bei den Wörtern angelangt ist. Als Kriterium für die Zerlegungen wendet man die heuristischen Tests aus S 5.1 an.

Die	Kundin	will	das	geblünte	Kleid	im	Schaufenster	anprobieren
Die	Kundin	will	das	geblünte	Kleid	im	Schaufenster	anprobieren
		will	das	geblünte	Kleid	im	Schaufenster	anprobieren
			das	geblünte	Kleid	im	Schaufenster	
Die	Kundin	will	das	geblünte	Kleid	im	Schaufenster	anprobieren

Wenn man die Teile, die bei der Zerlegung entstehen, benennt, erhält man eine sog.

Konstituentenstruktur:

S								
NP1		VP						
		Aux	NP2					
NP2				PP				
Det1	N1		Det2	Adj2	N2	Prep	N3	
Die	Kundin	kann	das	geblünte	Kleid	im	Schaufenster	anprobieren

Konstituente: jeder aus der sinnvollen Zerlegung einer Zeichenkette herrührende Teil

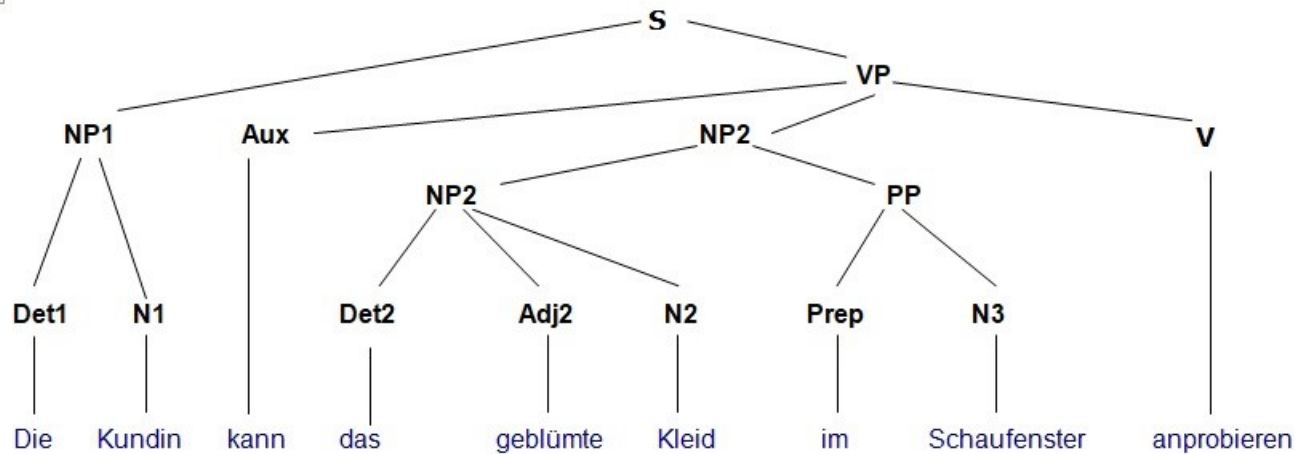
unmittelbare Konstituente (immediate constituents, IC): diejenige Konstituenten, die in einem einzigen Zerlegungsschritt aus einer größeren Konstituente gewonnen werden.

terminale Konstituente : ein Wort der analysierten Sprache

nicht-terminale Konstituente: (wie der Name sagt)

Konstituentenstrukturbaum:

(zwischen allen Konstituenten und ihren unmittelbaren Konstituenten werden Linien gezogen)



Konstituentenstrukturregeln:

Für jede Zerlegung eine Regel.

(links steht der Name einer Konstituente, dann folgt ein Pfeil, dann folgen die unmittelbaren Konstituenten)

S	→	NP1	VP	Det1	→	die
NP1	→	Det1	N1	N1	→	Kundin
VP	→	Aux	NP2	V	→	kann
NP2	→	NP2	PP	Det2	→	das
NP2	→	Det2	Adj2	N2	→	geblümete
PP	→	Prep	N3	N2	→	Kleid
				Prep	→	im
				N3	→	Schaufenster
				V	→	anprobieren

Die Idee der generativen Grammatik

- Eine generative Grammatik ist ein (gedachter) Automat, der bei S beginnt und solange Symbole ersetzt, bis er bei Wörtern der Sprache angekommen ist.
- Dabei ersetzt der immer Symbole auf der linken Seite der Regeln durch die Symbole auf der rechten Seite. Dies entspricht der Zerlegung einer syntaktischen Einheit in ihre Bestandteile.
- Wenn es alternative Regeln gibt, die der Automat noch nicht angewendet hat, fängt er erneut bei S an und benutzt eine neue Regel.
- Der Automat generiert so theoretisch alle möglichen Sätze der Sprache (wegen der Rekursivität natürlicher Sprachen eine unendliche Menge)

Formale Definition einer sog. kontextfreien Phrasenstrukturgrammatik G

$G = (T, NT, R, S)$

T - Menge von Terminalsymbolen (=die Elemente der Sprache)

NT - Menge von nicht-terminalen Symbolen (= Kategorien)

R - Menge von Regeln der Form $\alpha \rightarrow \beta$ (= Zerlegungen)

$$\alpha \in NT, \quad \beta \subseteq (NT \cup T)$$

S - Startsymbol (= beschriebene Einheit, Satz) $S \in NT$

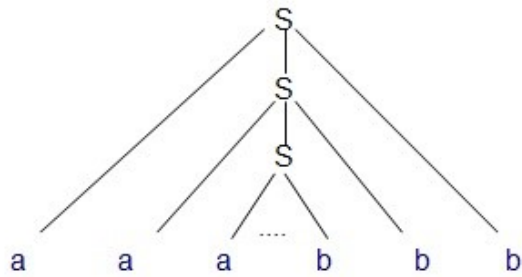
Formale Definition einer Sprache L

$L = G(L) = \{ S \mid S \text{ wird durch } G \text{ generiert} \}$

Beispiele (Kunstsprachen mit besonderen strukturellen Eigenschaften):

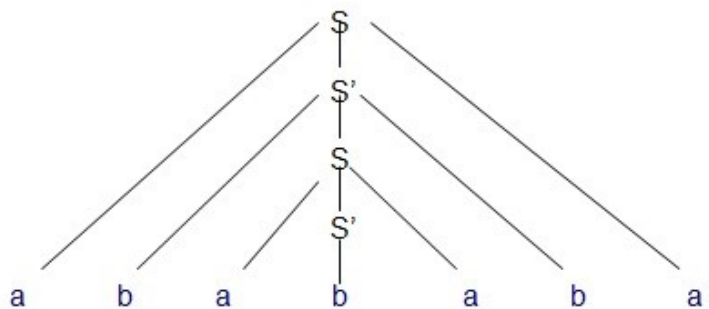
$S \rightarrow a b$
 $S \rightarrow a S b$

n-mal a und n-mal b



$S \rightarrow a S' a$
 $S \rightarrow a$
 $S' \rightarrow b S b$
 $S' \rightarrow b$

a und b abwechselnd



Schreiben einer Phrasenstrukturgrammatik des Deutschen

Welche der folgenden Sätze sind strukturell verwandt?

(Zur Beantwortung dieser Frage muss man im Kopf die heuristischen Tests von S 5.1 anwenden.)

- (1) Er vertraut jedem
- (2) Wer vertraut diesem Gauner
- (3) Er bringt mir die Milch
- (4) Kein Mensch glaubt ihm
- (5) Jeder Lehrer hilft seinem Schüler
- (6) Wer versteckte die Eier
- (7) Dem hilft kein Mensch
- (8) Die Schildkröte bewegt den Kopf

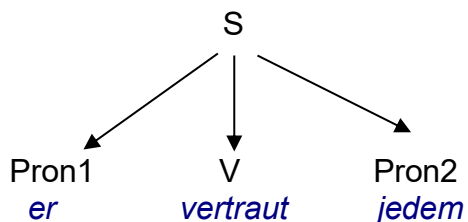
Lösung: (1), (2), (4), (5), (7)

Wir schreiben eine Phrasenstrukturgrammatik für diese Sätze

Typische Aufgabe dabei: Zerlegen, Klassifikation, Regeln schreiben

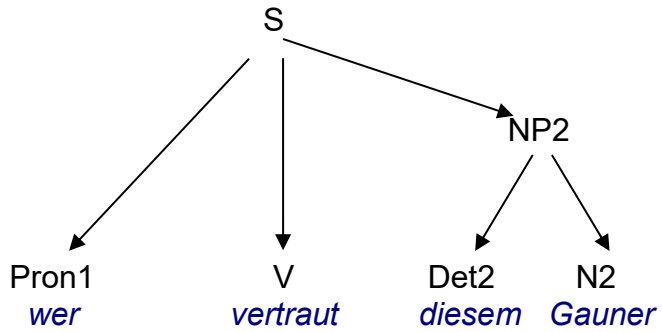
(1) *er vertraut jedem*

S	->	Pron1 + V + Pron2
Pron1	->	<i>er</i>
V	->	<i>vertraut</i>
Pron2	->	<i>jedem</i>

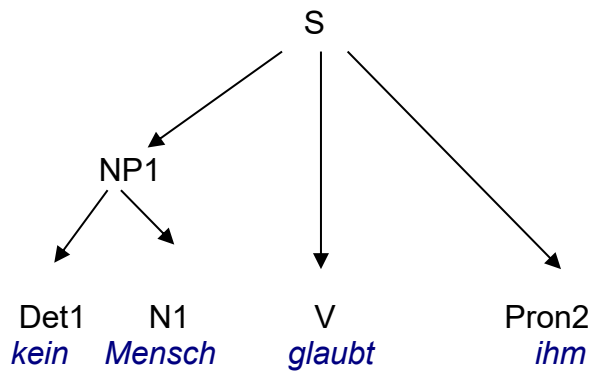


(2) *wer vertraut diesem Gauner*

Pron1	->	wer
S	->	Pron1 + V + NP2
NP2	->	Det2 + N2
Det2	->	diesem
N2	->	Gauner

(3) *kein Mensch glaubt ihm*

V	->	glaubt
Pron2	->	ihm
S	->	NP1 + V + Pron2
NP1	->	Det1 + N1
Det1	->	kein
N1	-->	Mensch



Verallgemeinerung der bisherigen Analyse, neue Grammatik:

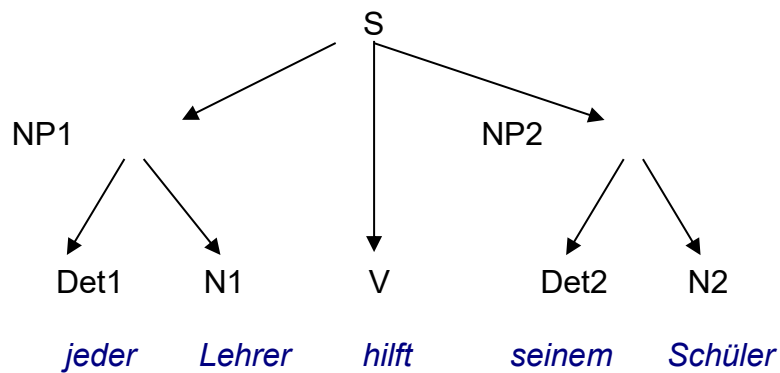
Regeln:

S	->	NP1 + V + NP2
NP1	->	Pron1
NP1	->	Det1 + N1
NP2	->	Pron2
NP2	->	Det2 + N2

Lexikon:

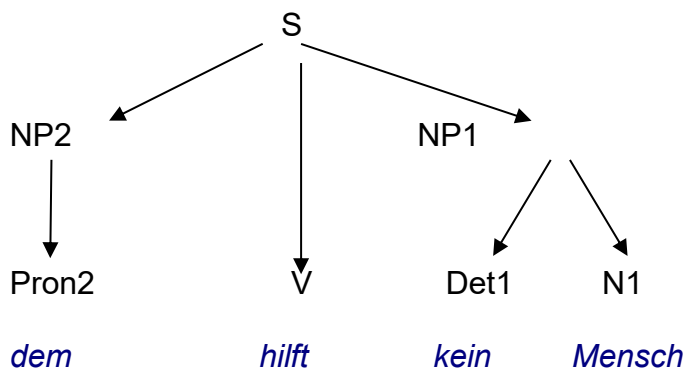
V	=	<i>vertraut, glaubt</i>
Pron1	=	<i>er, wer</i>
Pron2	=	<i>jedem, ihm</i>
Det1	=	<i>kein</i>
Det2	=	<i>diesem</i>
N1	=	<i>Mensch</i>
N2	=	<i>Gauner</i>

(4) *jeder Lehrer hilft seinem Schüler*



(5) *Dem hilft kein Mensch*

S -> NP2 + V + NP1



Bisherige Grammatik:

S -> NP1 + V + NP2
 S -> NP2 + V + NP1
 NP1 -> Pron1
 NP1 -> Det1 + N1
 NP2 -> Pron2
 NP2 -> Det2 + N2

Äquivalente Grammatik mit komplexen Kategorien (vgl. S 4.2)

kategorie[satz] -> kategorie[np] kasus[nominativ] person[C] numerus[C] +
 kategorie[verb] person[C] numerus[C] +
 kategorie[np] kasus[dativ]

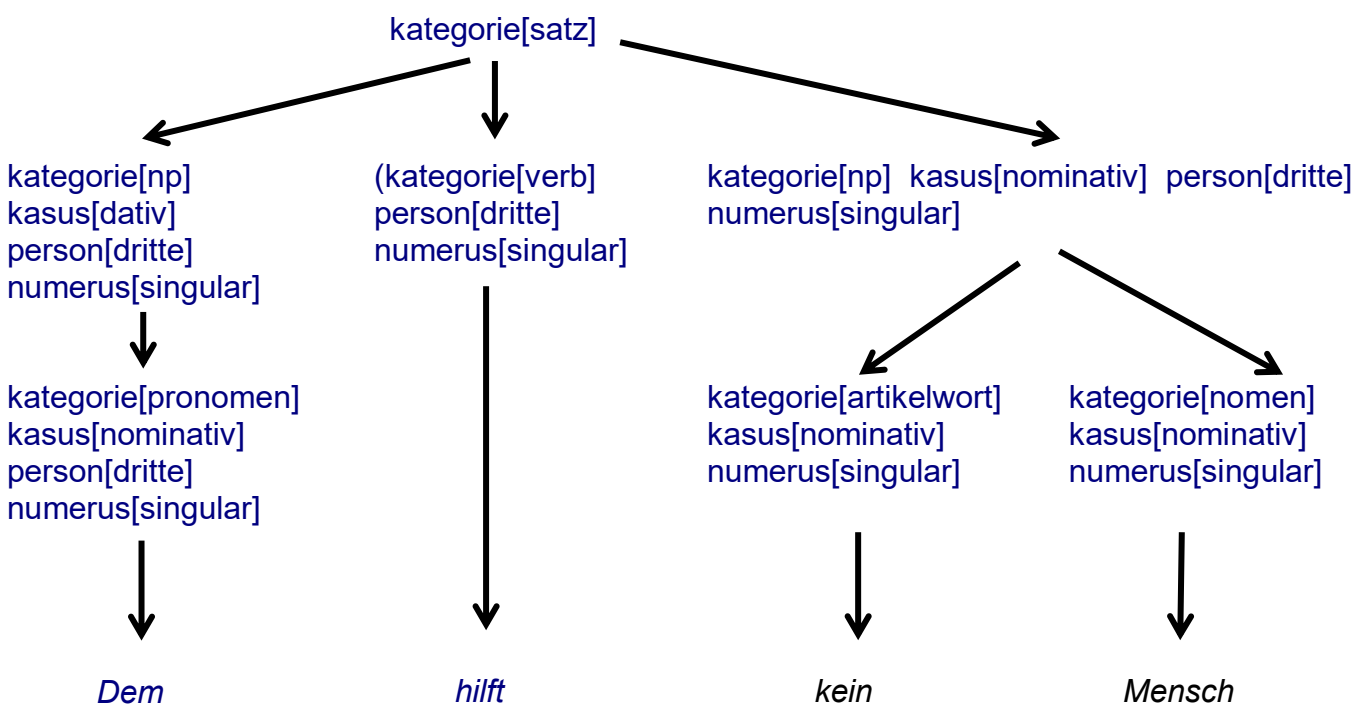
kategorie[satz] -> kategorie[np] kasus[dativ] +
 kategorie[verb] person[C] numerus[C] +
 kategorie[np] kasus[nominativ] person[C] numerus[C]

kategorie[np] kasus[C] person[C] numerus[C] ->
 kategorie[pronomen] kasus[C] person[C] numerus[C]

kategorie[np] kasus[C] person[dritte] numerus[C] ->
 kategorie[artikelwort] kasus[C] numerus[C] +
 kategorie[nomen] kasus[C] numerus[C]

Nur noch 4 Regeln, die obendrein weiterreichend und genauer sind!

(5) *Dem hilft kein Mensch*



S 5.3 Maschinelle Syntaxanalyse (Parsing)

Lernziel : Ein erster Eindruck davon, wie ein Parser funktionieren kann.

Parser Definition

- Ein Parser akzeptiert auf der Grundlage einer Grammatik G alle wohlgeformten Ausdrücke einer Sprache $L(G)$ und weist alle nicht-wohlgeformten Ausdrücke zurück
- Der Parser ordnet den wohlgeformten Ausdrücken von $L(G)$ eine Strukturbeschreibung entsprechend der Grammatik G zu.

Beispiel: Top-down Parser mit paralleler Expansion

- Ausgehend von S werden nicht-terminale Symbole solange nach den Regeln ersetzt, bis ganz links ein Symbol aus dem Lexikon (eine sog. präterminale Kategorie) steht.
- Wenn es mehrere Regeln für ein Symbol gibt, werden mehrere Ableitungen erzeugt.
- Sodann wird geprüft in welchen Ableitungen das linke Symbol mit der präterminalen Kategorie des aktuellen Eingabewortes übereinstimmt. Alle anderen Ableitungen werden entfernt.
- In den verbleibenden Ableitungen wird das präterminale Symbol entfernt. Dadurch gerät das nächste Symbol der Ableitung an den Anfang. In der Eingabe wird gleichzeitig zum nächsten Wort übergegangen. Der Prozess der Ersetzungen beginnt von neuem.
- Wenn das letzte Wort in der Eingabe erreicht ist, sollte gerade auch das letzte Symbol in wenigstens einer Ableitung verarbeitet sein.

Beispiel-Grammatik (aus S 5.2, erweitert um Relativsätze)

Regeln:		Lexikon:	
(1)	S -> NP1 + V + NP2	V	= <i>vertraut, glaubt, hilft</i>
(2)	S -> NP2 + V + NP1	Pron1	= <i>er, keiner</i>
(3)	NP1 -> Pron1	Relpron1	= <i>der</i>
(4)	NP1 -> Det1 + N1	Pron2	= <i>jedem, ihm, niemandem</i>
(5)	NP1 -> Pron1 + RelS	Relpron2	= <i>dem</i>
(6)	NP2 -> Pron2	Det1	= <i>kein, ein, unser</i>
(7)	NP2 -> Det2 + N2	Det2	= <i>diesem, dem, jedem</i>
(8)	NP2 -> Det2 + N2 + RelS	N1	= <i>Mensch, Student, Hirsch</i>
(9)	RelS -> Relpron1 + NP2 + V	N2	= <i>Gauner, Menschen, Förster</i>
(10)	RelS -> Relpron2 + NP1 + V		

Eingabesatz: *keiner der jedem Menschen vertraut glaubt diesem Gauner*

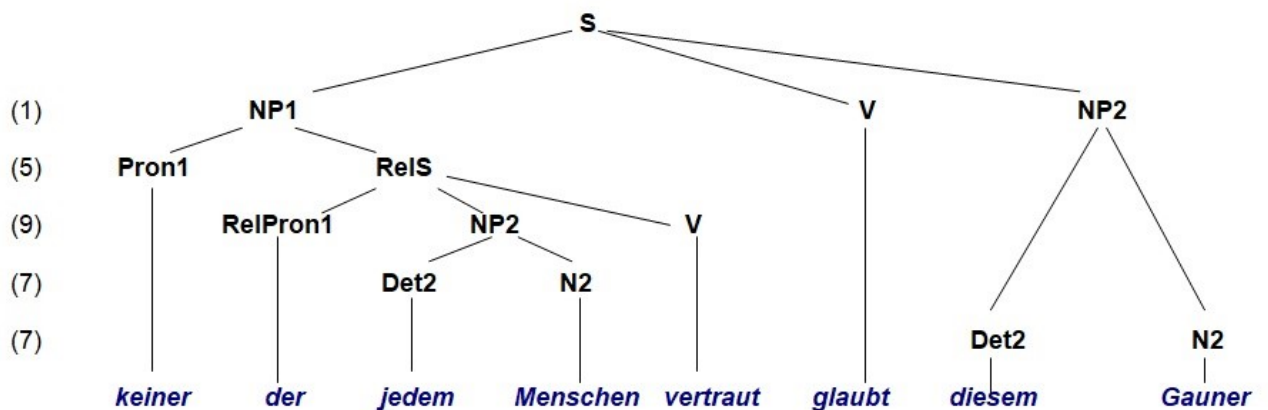
Ablauf des Parsers:

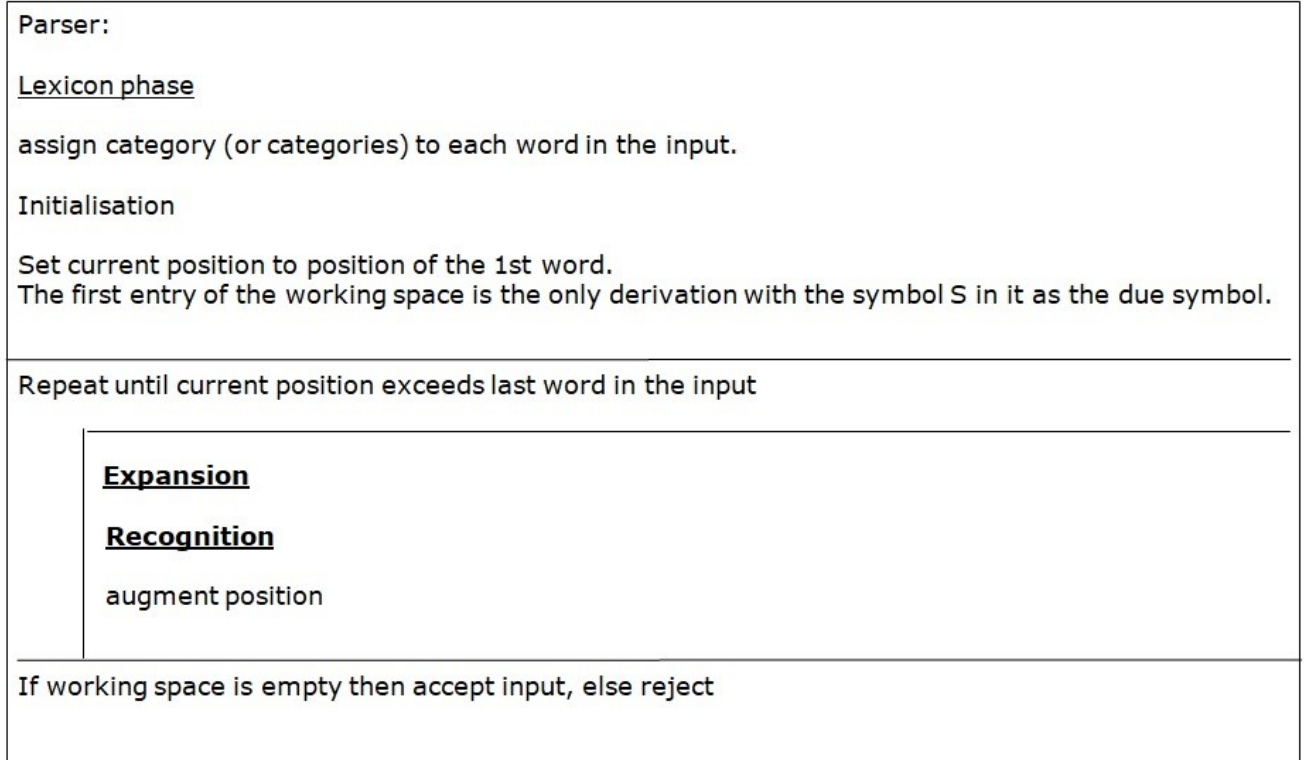
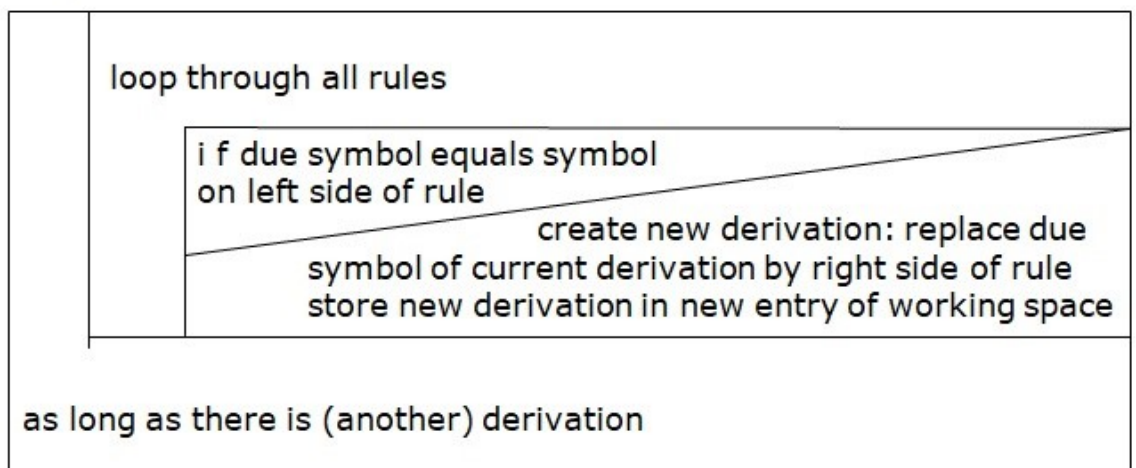
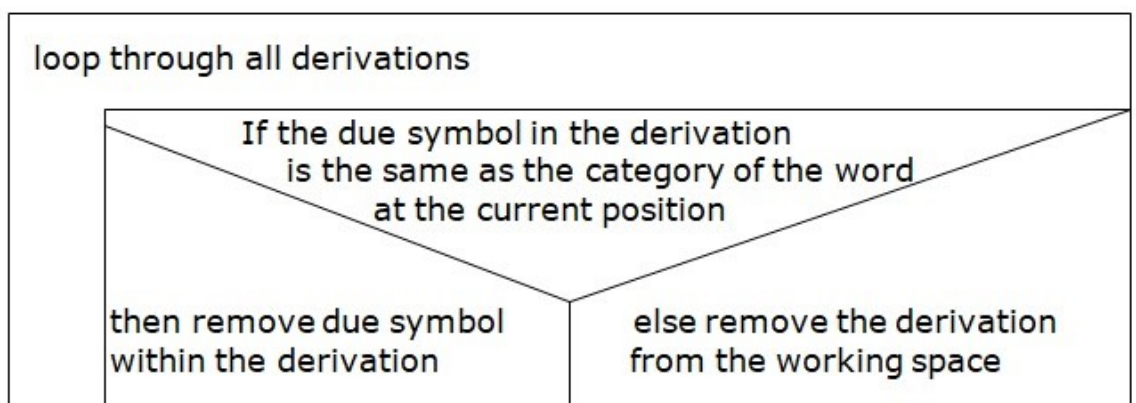
Position	Eingabe	Lexikon	Ableitung	angewandte Regeln
1	<i>keiner</i>	Pron1	S NP1 + V + NP2 NP2 + V + NP1 Pron1 + V + NP2 Det1 + N1 + V + NP2 Pron1 + RelS + V + NP2 Pron2 + V + NP1 Det2 + N2 + V + NP1 Det2 + N2 + RelS + V + NP1	start 1 2 1,3 1,4 1,5 2,6 2,7 2,8
			V + NP2 RelS + V + NP2	1,3 1,5
2	<i>der</i>	Relpron1	V + NP2 RelS + V + NP2 Relpron1 + NP2 + V + V + NP2 Relpron2 + NP1 + V + V + NP2	1,3 1,5 1,5,9 1,5,10
			NP2 + V + V + NP2	1,5,9

3	<i>jedem</i>	Pron2 Det2	Pron2 + V + V + NP2 Det2 + N2 + V + V + NP2 Det2 + N2 + RelS + V + V + NP2	1,5,9,6 1,5,9,7 1,5,9,8
			V + V + NP2 N2 + V + V + NP2 N2 + RelS + V + V + NP2	1,5,9,6 1,5,9,7 1,5,9,8
4	<i>Mensch</i>	N2	V + V + NP2 N2 + V + V + NP2 N2 + RelS + V + V + NP2	1,5,9,6 1,5,9,7 1,5,9,8
			V + V + NP2 RelS + V + V + NP2	1,5,9,7 1,5,9,8
5	<i>vertraut</i>	V	V + V + NP2 RelS + V + V + NP2 Relpron1 + NP2 + V + V + V + NP2 Relpron2 + NP1 + V + V + V + NP2	1,5,9,7 1,5,9,8 1,5,9,8,9 1,5,9,8,10
			V + NP2	1,5,9,7
6	<i>glaubt</i>	V	V + NP2	1,5,9,7
			NP2	1,5,9,7
7	<i>diesem</i>	Det2	NP2 Pron2 Det2 + N2 Det2 + N2 + RelS	1,5,9,7 1,5,9,7,6 1,5,9,7,7 1,5,9,7,8
			N2 N2 + RelS	1,5,9,7,7 1,5,9,7,8
8	<i>Gauner</i>	N2	N2 N2 + RelS	1,5,9,7,7 1,5,9,7,8
			- RelS	1,5,9,7,7 1,5,9,7,8

Erfolgreiche Analyse: Regeln **1,5,9,7,7**

Nach diesen Regeln aufgebauter Strukturbaum:



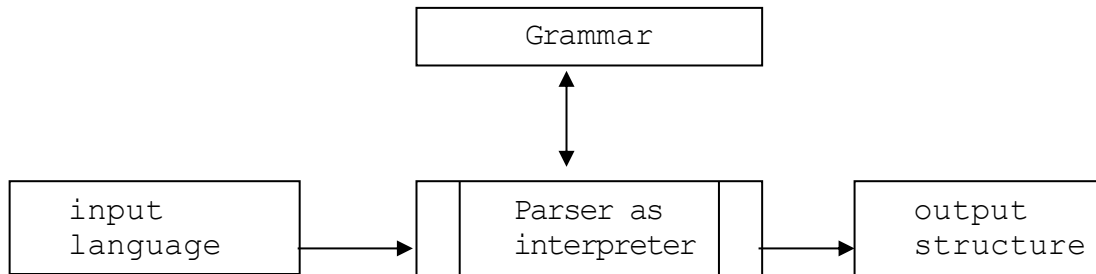
Struktogramm:**Expansion:****Recognition:**

Voraussetzungen für einen Parser

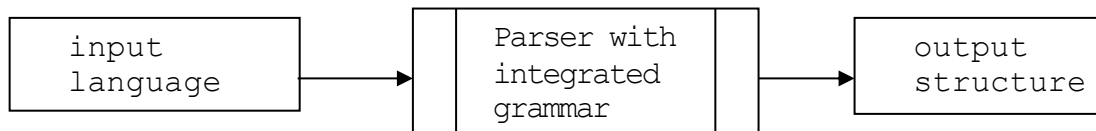
- ein Grammatikformalismus das ist. **Wissensrepräsentation**
- eine Grammatik das ist **Wissen**
- ein Algorithmus das ist **Wissensverarbeitung**

Verhältnis Parser und Grammatik

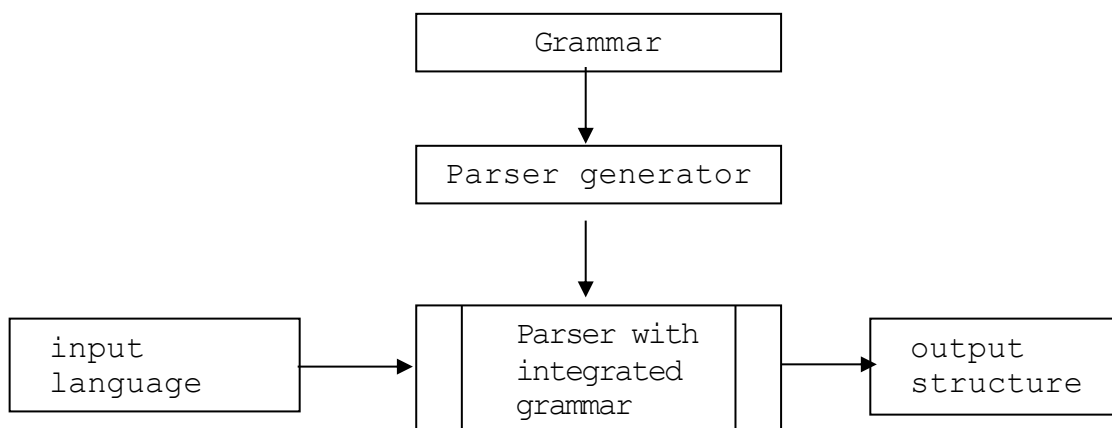
Grammatik und Parser getrennt (*Interpreting parser*)

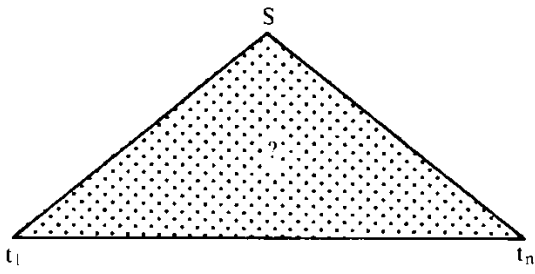


Grammatik in den Parser integriert (*Procedural parser*)



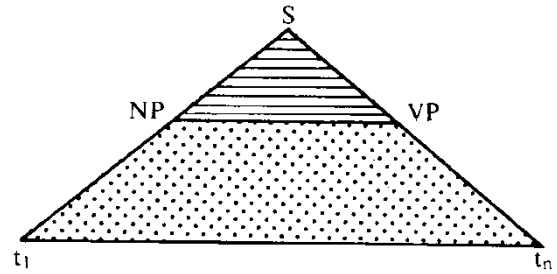
Grammatik zunächst von Parser getrennt, anschliessend automatisch integriert (*Compiled parser*)



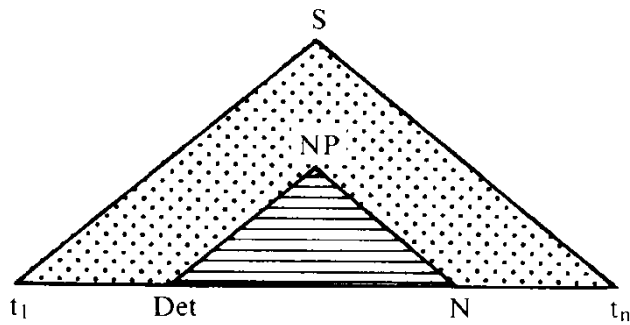
Erkennungsstrategie

<< der Suchraum

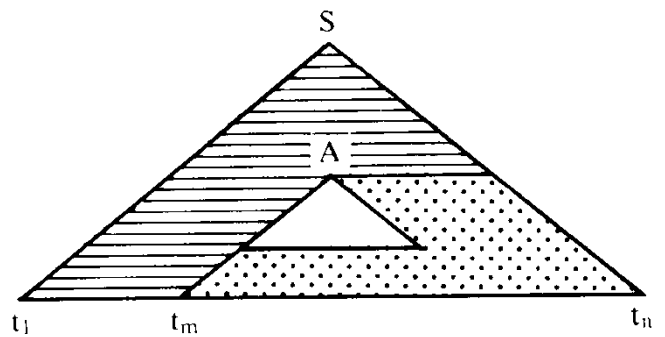
top-down Strategie >>



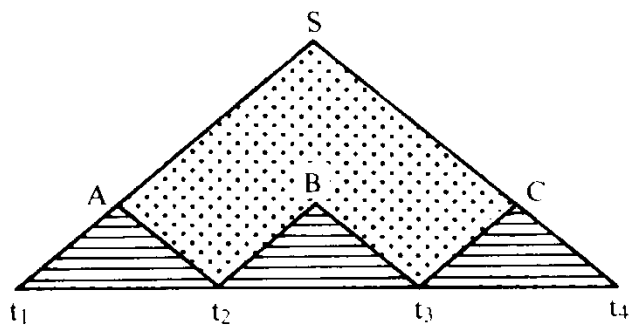
<< bottom-up Strategie



Tiefe zuerst >>



<< Breite zuerst



Verfahren bei alternativen Regeln und mehrdeutigen Wörtern

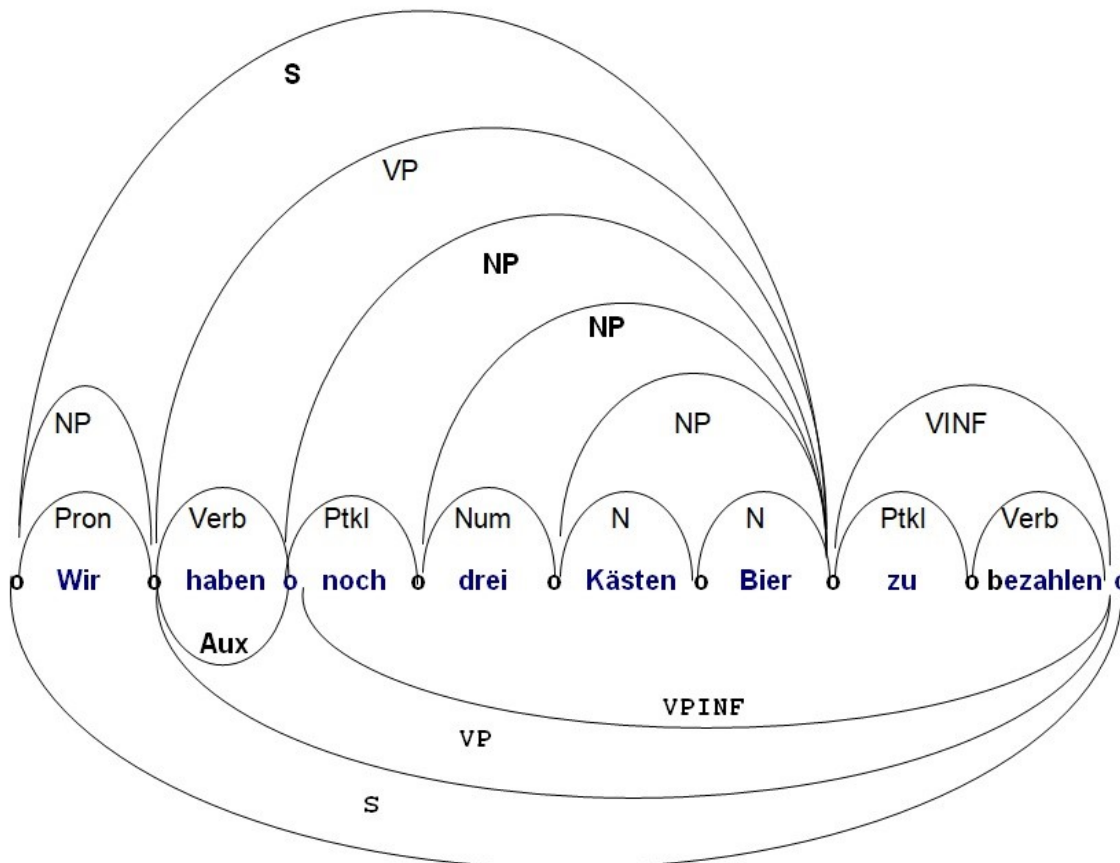
parallel: Verfolgen aller Alternativen

back-tracking: Zurückversetzen in den vorhergehenden Stand, Wahl der nächsten Alternative

Vorausschau: Regelanwendung vom Kontext abhängig machen

chart parsing: Speicherung und Wiederverwendung von Zwischenergebnissen

Bottom-up Chart Parser



Effizienz

Zeit und Speicherplatz

Behandlung von Alternativen

Übergenerierung

Reichweite

Welcher Grammatiktyp, welche Strukturen?

Lingwareentwicklung

Übersichtlichkeit

Seiten-Effekte

Verwendbarkeit von Ressourcen

Eignung für die jeweilige Anwendung

S 6.1 Aussagenlogik

Lernziel: Das Wesen der Aussagenlogik als Lehre vom Folgern begreifen; einfache Fälle formalisieren können

1. Gegenstand der Logik

Gegenstand der Logik ist die Folge d.h. die Abhängigkeit von Aussagen untereinander.

Zum Beispiel sind folgende Sätze nicht unabhängig voneinander:

wahr < --- > falsch?

- | | |
|--|--|
| • <i>Fritz ist ledig</i> | <i>Fritz hat mir gestern seine Gattin vorgestellt</i> |
| • <i>Der Schaden ist klein</i> | <i>Der Schaden ist groß</i> |
| • <i>Der Schaden ist nicht klein</i> | |
| • <i>Sein Vater ist nicht arm</i> | <i>Sein Vater ist reich</i> |
| • <i>Fritz kauft das Auto von Karl für 1000 DM</i> | <i>Karl verkauft dem Fritz das Auto für 1000 DM
Fritz bezahlt 1000 DM für das Auto</i> |
| • <i>Karl verlangt 1000 Mark für das Auto</i> | <i>Das Auto kostet 1000 Mark</i> |
| • <i>Die Lampe brennt</i> | <i>Es ist Strom in der Leitung</i> |
| • <i>Die Lampe ist aus</i> | <i>Es ist kein Strom in der Leitung</i> |
| • <i>Der Fernseher läuft.
Jemand hat den Fernseher angeschaltet.</i> | <i>Es war niemand im Zimmer</i> |
| • <i>Wir haben den Motor ausgebaut</i> | <i>Der Kunde fährt mit dem Wagen weg</i> |

Man kann argumentieren:

Wenn jemand ledig ist dann hat er keine Gattin!. Also
Wenn jemand keine Gattin hat, dann kann er sie dir nicht vorgestellt haben!

Entweder ein Schaden ist klein oder groß! Also
Ein Schaden kann nicht klein und groß (oder mittelgroß) zugleich sein!

Nur wenn Strom da ist, brennt die Lampe!. Also
Wenn kein Strom in der Leitung ist, ist die Lampe aus.

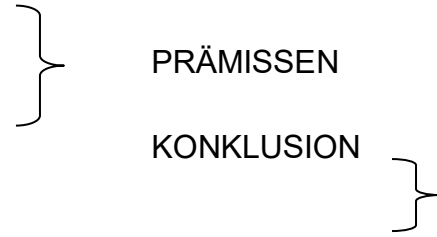
usw.

Wichtige Begriffe

Schluss

*Wenn jemand ledig ist, dann hat er keine Gattin.
Fritz ist ledig.*

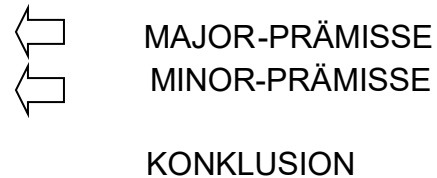
Also: Fritz hat keine Gattin

**Schlussfigur**

wenn a dann b

a

also b



Prämissen müssen beide wahr sein, dann ist auch die Konklusion wahr

Argumentation, d.h. Gründe vorbringen, die den anderen zur Annahme der Konklusion bewegen. Der macht Einwände, **indem er wenigstens eine Prämisse** widerlegt oder indem er weitere Schlüsse mit anderen Prämissen ins Spiel bringt.

Tip

Die Majorprämisse (die allgemeine Regel) wird in alltäglichen Argumentationen oft nicht verbalisiert. Meist wird darum gestritten, ob die Minor Prämisse ein Fakt ist. Die Majorprämisse wird still vorausgesetzt. Dabei ist die Majorprämisse oft schwach belegt. Eine gute Chance des Opponenten ist es, die implizite Majorprämisse eines Schlusses aufzudecken und anzugreifen.

- Die „formale Logik“ ist eine Sprache, in der die Folgebeziehung besonders klar hervortritt, indem bestimmte Symbole und Rechenvorschriften vorgegeben werden.

2. Aussagenlogischer Kalkül

Elemente:

Atomare Aussagen: $a, b, c \dots$

Aussagevariablen: $\alpha, \beta, \gamma \dots$

Junktoren: $\neg, \wedge, \vee, \rightarrow, \leftarrow, \leftrightarrow$

Syntax:

Der aussagenlogische Formalismus besteht aus einer Menge von Formeln.

1. Jede atomare Aussage ist eine Formel
2. Ist α eine Formel so ist auch $\neg \alpha$ eine Formel
3. Sind α und β Formeln so sind auch
 $(\alpha \wedge \beta), (\alpha \vee \beta), (\alpha \rightarrow \beta), (\alpha \leftarrow \beta)$ und $(\alpha \leftrightarrow \beta)$ Formeln

Semantik:

Wahrheitswertebelegung jeder Formel

Jeder atomaren Aussage in der Formel wird ein Wert aus $\{0, 1\}$ zugewiesen

α	$\neg \alpha$
0	1
1	0

nicht
Negation

α	β	$(\alpha \wedge \beta)$
1	1	1
1	0	0
0	1	0
0	0	0

und
Konjunktion

α	β	$(\alpha \vee \beta)$
1	1	1
1	0	1
0	1	1
0	0	0

oder
Disjunktion

α	β	$(\alpha \rightarrow \beta)$
1	1	1
1	0	0
0	1	1
0	0	1

wenn-dann
Implikation

α	β	$(\alpha \leftarrow \beta)$
1	1	1
1	0	1
0	1	0
0	0	1

nur wenn-dann
Replikation

α	β	$(\alpha \leftrightarrow \beta)$
1	1	1
1	0	0
0	1	0
0	0	1

genau dann-wenn
Äquivalenz

Übersetzung von natürlichsprachigen Sätzen in formale Logik

Wer nichts erwartet, wird nicht enttäuscht, und wer nicht enttäuscht wird ist glücklich, deshalb ist jeder glücklich, der nichts erwartet.

jemand erwartet nichts: $\neg e$

jemand wird nicht enttäuscht: $\neg t$

jemand ist glücklich: g

Formel: $(\neg e \rightarrow \neg t) \wedge (\neg t \rightarrow g) \Rightarrow (\neg e \rightarrow g)$

Wahrheitstafel dazu:

e	t	g	$\neg e$	$\neg t$	$\neg g$	$(\neg e \rightarrow \neg t)$	$(\neg t \rightarrow g)$	\wedge	$(\neg e \rightarrow g)$	\Rightarrow
1	1	1	0	0	0	1	1	1	1	1
1	1	0	0	0	1	1	1	1	1	1
1	0	1	0	1	0	1	1	1	1	1
1	0	0	0	1	1	1	0	0	1	1
0	1	1	1	0	0	0	1	0	1	1
0	1	0	1	0	1	0	1	0	0	1
0	0	1	1	1	0	1	1	1	1	1
0	0	0	1	1	1	1	0	0	0	1

Die Argumentation ist korrekt (sie ist eine Tautologie), da die Formel immer 1 (wahr) ergibt.

Vereiste Straßen sind gefährlich, daher ist diese Straße nicht gefährlich, da sie nicht vereist ist.

Die Straße ist vereist: e

Die Straße ist gefährlich: g

Formel: $((e \rightarrow g) \wedge (\neg e)) \Rightarrow (\neg g)$

e	g	$\neg e$	$\neg g$	$(e \rightarrow g)$	$(e \rightarrow g) \wedge (\neg e)$	$((e \rightarrow g) \wedge (\neg e)) \Rightarrow (\neg g)$
1	1	0	0	1	0	1
1	0	0	1	0	0	1
0	1	1	0	1	1	0
0	0	1	1	1	1	1

Wahrheitstafel ergibt keine Tautologie. Es gibt den Fall, in dem die Straße zwar eisfrei, aber aus anderen Gründen gefährlich ist. Die Argumentation ist also nicht gültig.

3. Logik aus pragmatischer Sicht (Sprachhandeln, Sprachspiel)

wahr / falsch – gottgegeben – naturgegeben – oder ein Sprachspiel?

- ❖ Eine Aussage ist wahr, wenn das der Fall ist, was sie aussagt. Eine Aussage ist falsch, wenn das nicht der Fall ist, was sie aussagt. Darüber, ob etwas der Fall ist oder nicht, haben Menschen unterschiedliche Meinungen. Dementsprechend werden Aussagen für wahr oder falsch gehalten (Wunderlich: *für wahr beansprucht*).
 - ❖ Wird eine Aussage geäußert, so behauptet der Sprecher (daß er meint, daß das, was er aussagt, der Fall ist) daß er die Aussage für wahr hält.
 - ❖ Wird eine Aussage akzeptiert, so behauptet der Hörer (daß er meint, daß das was ausgesagt worden ist, der Fall ist) daß er die Aussage ebenfalls für wahr hält.
 - ❖ Wird eine Aussage nicht akzeptiert, so behauptet der Hörer (daß er meint, daß das was ausgesagt worden ist, nicht der Fall ist) daß er die Aussage nicht für wahr hält.
 - ❖ Wenn jemand eine Aussage behauptet, so kann er das Negat der Aussage nicht akzeptieren.
 - ❖ Wenn jemand eine Aussage nicht akzeptiert, so muß er das Negat der Aussage behaupten.
 - ❖ Statt zu notieren, ob jemand eine Aussage für wahr hält oder nicht, braucht man nur festzuhalten, ob er die Aussage oder das Negat behauptet.
- **Logik ist eine Lehre, was man unter bestimmten Umständen sagen kann/muß d.h. eine Sprachwissenschaft!**

Das Argumentationsspiel

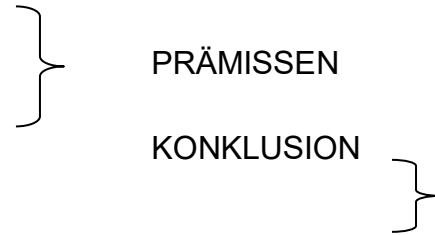
- Verabredung in welcher Sprache gespielt wird
- Verwendung der Wörter liegen damit fest
- Die Sätze sind die Züge
- Ziel: einen bestimmten Satz, oder mehrere, zu halten
- Strategien: Behauptung eines Satzes
- Gegenzug: Widersprechen
- Gegenzug: Satz beweisen durch andere Sätze
- Gegenzug: Nachweis daß der Beweis nicht gültig ist
- usw.

noch einmal

Schluss

*Wenn jemand ledig ist, dann hat er keine Gattin.
Fritz ist ledig.*

Also: Fritz hat keine Gattin

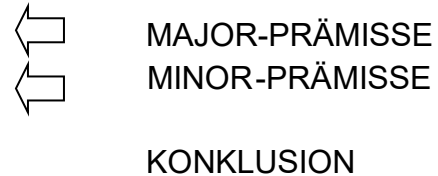


Schlussfigur

wenn a dann b

a

also b



Akzeptiert der Gegner beide (!) Prämissen, so muss er die Konklusion akzeptieren.

Beweisen: Um den Gegner dazu zu bringen, die Konklusion zu akzeptieren, muss ich ihn dazu bringen, beide Prämissen zu akzeptieren.

Will der Gegner die Konklusion nicht akzeptieren, so muss er versuchen, eine(!) Prämisse zu widerlegen.

Bei der Minorprämisse geht es dabei um einfache Fakten, die man vielleicht empirisch prüfen kann.

Der Schluss steht und fällt aber mit der Majorprämisse. Sie ist eine Generalisierung über viele Fälle.

Aber wie kommt man zu den Majorprämissen?

4. Die möglichen Welten

Wie kommt man zu den Aussageverknüpfungen?

Man teile gedanklich die Welt in die Zustände, in denen eine bestimmte Aussage (a) zutreffen würde und in die Zustände, in denen die verneinte Aussage (nicht a) zutreffen würde. Das ergibt zwei mögliche Welten.

Daselbe mache man für eine andere Aussage (b) und die entsprechende verneinte Aussage (nicht b). Das ergibt zwei weitere mögliche Welten.

Man ordne gedanklich die vier Welten in einer Matrix an, in der jeweils die beiden Aussagen und ihren Verneinungen gelten bzw. nicht gelten können.

Man erhält vier Felder, welche vier Weltzuständen entsprechen:

- a und b gilt,
- nicht a und b gilt,
- a und nicht b gilt,
- nicht a und nicht b gilt.

Welche dieser Zustände sind (nach meiner Vermutung oder, besser aufgrund empirischer Untersuchung) in unserer Realität möglich und welche sind unmöglich? Man markiere sie in der Matrix entsprechend mit M und U.

Es gibt 16 mögliche Zustandskonstellation, alle Belegungen von viermal M bis viermal U in der Matrix. Folgerungen (\vdash) von einer Aussage (a oder nicht a) auf die andere (b oder nicht b) und umgekehrt lassen sich ziehen, wenn es in einer Spalte oder in einer Zeile ein M und ein U gibt. In diesem Fall ist für die eine Aussage nicht sowohl die andere Aussage wie auch ihre Negation möglich, sondern nur eins von beiden. Man kann dann eine Majorprämisse mit den beiden Aussagen als Elementen formulieren, wobei der Junktor für die Matrix steht (also für die besondere Konstellation möglicher Welten in Bezug auf a, nicht a, b und nicht b).

<p style="text-align: center;">$\langle a \vdash b \rangle$ $\langle \text{nicht } b \vdash \text{nicht } a \rangle$</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="text-align: center;">nicht</td></tr> <tr><td style="text-align: center;">a</td><td style="text-align: center;">a</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">U</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">M</td></tr> </table> <p style="text-align: center;">IMPLIKATION</p> <p>$a \rightarrow b$</p> <p>wenn a, dann b</p> <table style="margin-left: auto; margin-right: auto;"> <tr><td>1</td><td>1</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>1</td><td>0</td><td style="border-left: 1px solid black;">0</td></tr> <tr><td>0</td><td>1</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>0</td><td>0</td><td style="border-left: 1px solid black;">1</td></tr> </table>	nicht		a	a	b	M	nicht b	U	b	M	nicht b	M	1	1	1	1	0	0	0	1	1	0	0	1	<p style="text-align: center;">$\langle \text{nicht } a \vdash \text{nicht } b \rangle$ $\langle b \vdash a \rangle$</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="text-align: center;">nicht</td></tr> <tr><td style="text-align: center;">a</td><td style="text-align: center;">a</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">U</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">M</td></tr> </table> <p style="text-align: center;">REPLIKATION</p> <p>$a \leftarrow b$</p> <p>nur wenn a, dann b</p> <table style="margin-left: auto; margin-right: auto;"> <tr><td>1</td><td>1</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>1</td><td>0</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>0</td><td>1</td><td style="border-left: 1px solid black;">0</td></tr> <tr><td>0</td><td>0</td><td style="border-left: 1px solid black;">1</td></tr> </table>	nicht		a	a	b	M	nicht b	U	b	M	nicht b	M	1	1	1	1	0	1	0	1	0	0	0	1
nicht																																																	
a	a																																																
b	M																																																
nicht b	U																																																
b	M																																																
nicht b	M																																																
1	1	1																																															
1	0	0																																															
0	1	1																																															
0	0	1																																															
nicht																																																	
a	a																																																
b	M																																																
nicht b	U																																																
b	M																																																
nicht b	M																																																
1	1	1																																															
1	0	1																																															
0	1	0																																															
0	0	1																																															

a: Sokrates ist ein Mensch

b: Sokrates wird sterben

a: Sauerstoff ist vorhanden

b: Die Kerze brennt

<p style="text-align: center;">$\langle a \vdash \text{nicht } b \rangle$ $\langle b \vdash \text{nicht } a \rangle$</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="text-align: center;">nicht</td></tr> <tr><td style="text-align: center;">a</td><td style="text-align: center;">a</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">U</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">M</td></tr> </table> <p style="text-align: center;">EXKLUSION</p> <p>$a \mid b$</p> <p>a oder b (aber nicht beides)</p> <table style="margin-left: auto; margin-right: auto;"> <tr><td>1</td><td>1</td><td style="border-left: 1px solid black;">0</td></tr> <tr><td>1</td><td>0</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>0</td><td>1</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>0</td><td>0</td><td style="border-left: 1px solid black;">1</td></tr> </table>	nicht		a	a	b	U	nicht b	M	b	M	nicht b	M	1	1	0	1	0	1	0	1	1	0	0	1	<p style="text-align: center;">$\langle \text{nicht } a \vdash b \rangle$ $\langle \text{nicht } b \vdash a \rangle$</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="text-align: center;">nicht</td></tr> <tr><td style="text-align: center;">a</td><td style="text-align: center;">a</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">M</td></tr> <tr><td style="text-align: center;">b</td><td style="text-align: center;">U</td></tr> <tr><td style="text-align: center;">nicht b</td><td style="text-align: center;">U</td></tr> </table> <p style="text-align: center;">DISJUNKTION</p> <p>$a \vee b$</p> <p>a oder b (oder beides)</p> <table style="margin-left: auto; margin-right: auto;"> <tr><td>1</td><td>1</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>1</td><td>0</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>0</td><td>1</td><td style="border-left: 1px solid black;">1</td></tr> <tr><td>0</td><td>0</td><td style="border-left: 1px solid black;">0</td></tr> </table>	nicht		a	a	b	M	nicht b	M	b	U	nicht b	U	1	1	1	1	0	1	0	1	1	0	0	0
nicht																																																	
a	a																																																
b	U																																																
nicht b	M																																																
b	M																																																
nicht b	M																																																
1	1	0																																															
1	0	1																																															
0	1	1																																															
0	0	1																																															
nicht																																																	
a	a																																																
b	M																																																
nicht b	M																																																
b	U																																																
nicht b	U																																																
1	1	1																																															
1	0	1																																															
0	1	1																																															
0	0	0																																															

a: Karl ist in Deutschland geboren

b: Karl ist in Frankreich geboren

a: In der Schulklasse sind Mädchen

b: In der Schulklasse sind Jungen

<p> $\langle a \mid b \rangle$ $\langle \text{nicht } b \mid \text{nicht } a \rangle$ $\langle \text{nicht } a \mid \text{nicht } b \rangle$ $\langle b \mid a \rangle$ </p> <p> <table border="1"> <tr> <td></td> <td>a</td> <td>nicht a</td> </tr> <tr> <td>b</td> <td>M</td> <td>U</td> </tr> <tr> <td>nicht b</td> <td>U</td> <td>M</td> </tr> </table> </p> <p>ÄQUIVALENZ</p> <p> $a \Leftrightarrow b$ <table border="1"> <tr> <td>1</td> <td>1</td> <td>1</td> </tr> <tr> <td>1</td> <td>0</td> <td>0</td> </tr> <tr> <td>0</td> <td>1</td> <td>0</td> </tr> <tr> <td>0</td> <td>0</td> <td>1</td> </tr> </table> </p> <p> a genau dann, wenn b a dann und nur dann, wenn b wenn a dann b und umgekehrt </p>		a	nicht a	b	M	U	nicht b	U	M	1	1	1	1	0	0	0	1	0	0	0	1	<p> $\langle a \mid \text{nicht } b \rangle$ $\langle b \mid \text{nicht } a \rangle$ $\langle \text{nicht } a \mid b \rangle$ $\langle \text{nicht } b \mid a \rangle$ </p> <p> <table border="1"> <tr> <td></td> <td>a</td> <td>nicht a</td> </tr> <tr> <td>b</td> <td>U</td> <td>M</td> </tr> <tr> <td>nicht b</td> <td>M</td> <td>U</td> </tr> </table> </p> <p>KONTRAVALENZ</p> <p> $a \vee \neg b$ <table border="1"> <tr> <td>1</td> <td>1</td> <td>0</td> </tr> <tr> <td>1</td> <td>0</td> <td>1</td> </tr> <tr> <td>0</td> <td>1</td> <td>1</td> </tr> <tr> <td>0</td> <td>0</td> <td>0</td> </tr> </table> </p> <p>entweder a oder b</p>		a	nicht a	b	U	M	nicht b	M	U	1	1	0	1	0	1	0	1	1	0	0	0
	a	nicht a																																									
b	M	U																																									
nicht b	U	M																																									
1	1	1																																									
1	0	0																																									
0	1	0																																									
0	0	1																																									
	a	nicht a																																									
b	U	M																																									
nicht b	M	U																																									
1	1	0																																									
1	0	1																																									
0	1	1																																									
0	0	0																																									

a: *Hi. Abend fällt heuer auf Donnerstag*
b: *Neujahr fällt heuer auf Freitag*

a: *Der Klassensprecher ist ein Junge*
b: *Der Klassensprecher ist ein Mädchen*

S 6.2 Prädikatenlogik

Lernziel: Einfache Sätze natürlicher Sprache in prädikatenlogische Notation übersetzen können.

1. Gegenstand der Prädikatenlogik

Gegenstand der Logik allgemein ist die Folge d.h. die Abhängigkeiten von Aussagen untereinander.

Die Abhängigkeit kann auf aussagenlogischen Verknüpfungen beruhen (siehe S 6.2)

Schlussfigur

wenn a dann b

a

also b



MAJOR-PRÄMISSE

MINOR-PRÄMISSE

KONKLUSION

Für a und b können beliebige atomare oder komplexe Aussagen eingesetzt werden.
Der Schluss bleibt immer gültig.

Modus Darii

Alle Schwaben verstehen Hochdeutsch

Einer der Reiseteilnehmer ist ein Schwabe

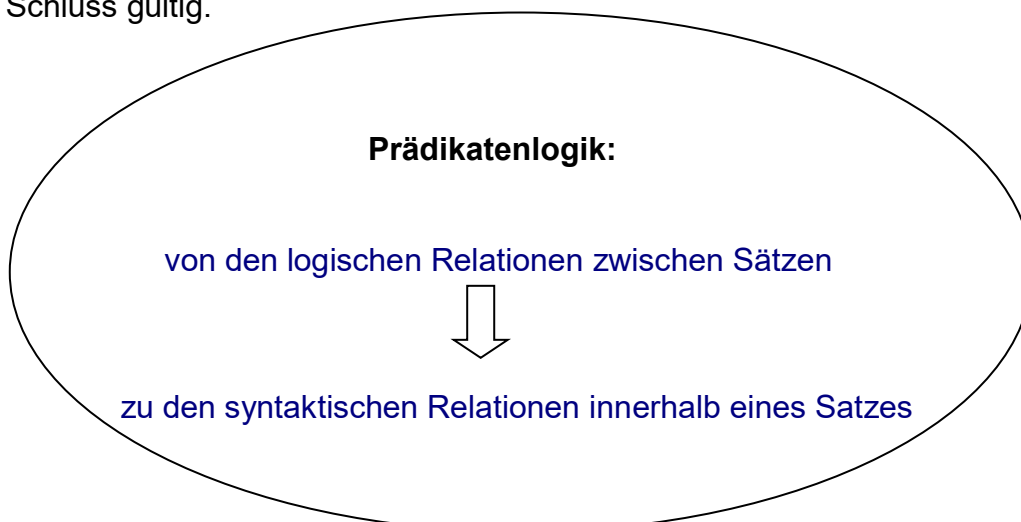
Einer der Reiseteilnehmer versteht Hochdeutsch

alle x P

ein y ist ein x

ein y P

Bei dieser Schlussfigur müssen TEILE von Aussagen ausgetauscht werden, d.h. der Innere Aufbau der Aussagen spielt eine Rolle. Wenn die richtigen Teile ausgetauscht werden, bleibt der Schluss gültig.



Erinnerung an S 5.1:

Eine Aussage macht man, indem man bestimmte **Objekte** sprachlich bezeichnet (*Referenz*) und ihnen ein **Prädikat** zuspricht (*Prädikation*). Prädikate sind in der natürlichen Sprache meist Verben oder auch Adjektive mit einem Hilfsverb, viele andere Wendungen oder auch manchmal gänzlich implizit:

Die Vorlesung fällt aus.

Die Vorlesung wird verschoben.

Die Vorlesung ist interessant.

Die Vorlesung ist eine Katastrophe.

Heute keine Vorlesung.

Wesentlich ist eine semantische Zweigliederigkeit: **Referenz** und **Prädikation**, selbst bei einem Einwortsatz

Prädikate allein sind „ungesättigt“ (Frege), sie müssen durch Bezeichnungen für Referenzobjekte ergänzt werden.

Man sagt: Prädikate haben **Ergänzungen** (in anderer Terminologie: **Argumente**)

Die Anzahl der Ergänzungen kann je nach Prädikat verschieden sein:

- 1 *Er schläft.*
- 2 *Er verpasst die Vorlesung.*
- 3 *Anni hilft mir bei der Hausaufgabe.*
- 4 *Fritz legt seinem Freund die Hand auf die Schulter.*

Ausdrucksmöglichkeiten in der Prädikatenlogik

Sie ist gegenüber der natürlicher Sprache (absichtlich) stark vereinfacht. Übersetzen aus natürlicher Sprache in prädikatenlogische Notation ist umständlich.

Es gibt nur **Prädikate**, **Eigennamen**, die **Quantoren** „für alle x gilt“ und „für wenigsten ein x gilt“, und die Verknüpfungen aus der Aussagenlogik

Die Argumente (in S 5.1 hießen sie *Ergänzungen*), denen ein Prädikat zugesprochen wird, stehen hinter dem Prädikat in Klammern, durch Komma voneinander getrennt). Im einfachsten Fall wird mit einem Eigennamen (einer sog. *Individuenkonstanten*) auf einen Gegenstand referiert. So erhält man Aussagen:

<i>Peter schläft</i>	\Rightarrow	schlafen (Peter)
<i>Maria liebt Peter</i>	\Rightarrow	lieben (Maria, Peter)

Gegenstände, für die es keinen Eigennamen gibt, müssen mithilfe eines Prädikats identifiziert werden, wobei für die Gegenstände selbst eine Variable eingesetzt werden muss, die durch einen Quantor gebunden ist, also z.B.

<i>ein Auto, das Auto</i>	$\exists x$ auto(x)	es gibt ein x, für das gilt: x ist ein Auto
<i>alle Autos</i>	$\forall x$ auto(x)	für alle x gilt: x ist ein Auto

Aussagen

<i>Es gibt Studenten in Heidelberg</i>	$\exists x$ studieren(x, heidelberg)	es gibt ein x für das gilt x studiert in Heidelberg
--	---	---

<i>Die Vorlesung fällt aus</i>	$\exists x$ vorlesung(x) \wedge ausfallen(x)	es gibt ein x für das gilt x ist eine Vorlesung und x fällt aus
--------------------------------	--	---

<i>Ein Maikäfer liebt eine Gazelle</i>	$\exists x \exists y$ maikäfer(x) \wedge gazelle(y) \wedge lieben(x,y)	es gibt ein x für das gilt x ist ein Maikäfer und es gibt ein y für das gilt y ist eine Gazelle und x liebt y
--	---	---

<i>Lydia liebt einen Maikäfer</i>	$\exists y$ maikäfer(y) \wedge lieben(Lydia ,y)	
-----------------------------------	---	--

Um die einzelnen Prädikationen zu verknüpfen, benutzt man also die Junktoren aus der Aussagenlogik.

Alle Elche sind klug $\forall x \text{ elch}(x) \rightarrow \text{klug}(x)$ für alle x gilt, wenn x ein Elch ist dann ist x klug

Hier muss als Verknüpfung die Implikation \rightarrow genommen werden. Die Konjunktion \wedge ergäbe einen falschen Sinn:

Alle sind Elche und klug $\forall x \text{ elch}(x) \wedge \text{klug}(x)$ für alle x gilt, x ist ein Elch und x ist klug

Aussagen mit mehreren Quantoren (Skopus beachten)

Jeder Elch liebt einen Maikäfer $\forall x \exists y \text{ elch}(x) \wedge \text{maikäfer}(y) \rightarrow \text{lieben}(x,y)$ für alle x gilt, es gibt ein y sodass wenn x ein Elch und y ein Maikäfer ist, dann liebt x y

Ein Maikäfer wird von allen Elchen geliebt $\exists y \forall x \text{ elch}(x) \wedge \text{maikäfer}(y) \rightarrow \text{lieben}(x,y)$ es gibt ein y sodass für alle x gilt, wenn x ein Elch und y ein Maikäfer ist, dann liebt x y

2. Prädikatenlogischer Kalkül – formale Definition

Elemente:

Individuenkonstanten (Eigennamen): $m, n, e12, karl, \dots$

Prädikate (Konstanten): $P, Q, R, schlafen, \dots$

Individuenvariablen: x, y, z, \dots

Quantoren: $\{ \forall, \exists \}$

Plus die Elemente der Aussagenlogik

Syntax:

Der prädikatenlogische Formalismus besteht aus einer Menge Formeln.

1. Jede Individuenkonstante und jede Individuenvariable ist ein Term.
2. Sind t_1, \dots, t_n Terme und ist P ein n -stelliges Prädikat, so ist $P(t_1, \dots, t_n)$ eine Formel.
3. Sind t_1 und t_2 Terme so ist $t_1 = t_2$ eine Formel
4. Ist φ eine Formel und x eine Individuenvariable, so sind auch $\forall x \varphi$ und $\exists x \varphi$ Formeln. φ ist der Skopus des Quantors. Kommt die Variable x innerhalb von φ vor, so ist sie gebunden. Sonst ist sie ungebunden.
5. Eine Formel φ ist ein Aussage (Satz), wenn in φ keine freie Variable vorkommt.
6. Plus die Syntax der Aussagenlogik für Formeln.

Semantik:

ein sog. Modell $M=(D,F)$

D ist eine nicht-leere Menge von Objekten (die Domäne).

F ist eine Funktion, die

- jeder Individuenkonstante ein Element aus D zugeordnet,
- jedem n -stelligem Prädikatensymbol eine n -stellige Relation über D zuordnet

$D = \{\text{Peter, Simone, Hans, LCL, Heidelberg, Leimen}\}$

$F(\text{peter}) = \text{Peter}$

$F(\text{simone}) = \text{Simone}$

$F(\text{hans}) = \text{Hans}$

$F(\text{lieben}) = \{(\text{Peter, Simone}), (\text{Simone, Hans})\}$

$F(\text{schlafen}) = \{\text{Peter}\}$

$F(\text{wohnen}) = \{(\text{Peter, Heidelberg}), (\text{Simone, Heidelberg}), (\text{Hans, Leimen})\}$

$F(\text{studieren}) = \{(\text{Peter, LCL}), (\text{Simone, LCL})\}$

Tautologien, logische Gesetze

$$\forall x \varphi \Leftrightarrow \neg \exists x \neg \varphi$$

Alle Vögel fliegen \Leftrightarrow *Es gibt keinen Vogel der nicht fliegt*

für alle x gilt P \Leftrightarrow es gibt kein x für das gilt nicht P

$$\exists x \varphi \Leftrightarrow \neg \forall x \neg \varphi$$

Ein Student ist fleißig \Leftrightarrow *Nicht alle Studenten sind faul*

es gibt ein x für das gilt P \Leftrightarrow nicht für alle x gilt das nicht P

$$(\forall x \varphi \wedge \forall x \delta \Leftrightarrow \forall x (\varphi \wedge \delta))$$

Alle Studenten lösen Hausaufgaben und alle Studenten schreiben eine Klausur \Leftrightarrow

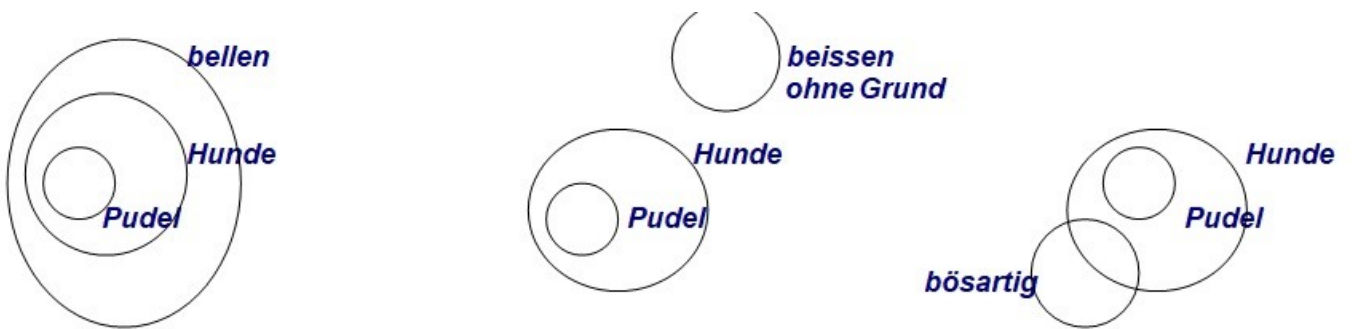
Alle Studenten lösen Hausaufgaben und schreiben eine Klausur

für alle x gilt P und für alle x gilt Q \Leftrightarrow Für alle x gilt P und Q

Prädikatenlogische Schlussfiguren – sog. Syllogismen

<i>Alle Pudel sind Hunde</i>	$\forall x (P(x) \rightarrow H(x))$	
<i>Alle Hunde bellen</i>	$\forall x (H(x) \rightarrow B(x))$	
<i>Alle Pudel bellen</i>	$\forall x (P(x) \rightarrow B(x))$	
<i>Alle Pudel sind Hunde</i>	$\forall x (P(x) \rightarrow H(x))$	$\forall x (P(x) \rightarrow H(x))$
<i>Kein Hund beißt ohne Grund</i>	$\forall x (H(x) \rightarrow \neg BG(x))$	$\neg \exists x (H(x) \wedge BG(x))$
<i>Kein Pudel beißt ohne Grund</i>	$\forall x (P(x) \rightarrow \neg BG(x))$	$\neg \exists x (H(x) \wedge BG(x))$
<i>Alle Pudel sind Hunde</i>	$\forall x (P(x) \rightarrow H(x))$	$\forall x (P(x) \rightarrow H(x))$
<i>Kein Pudel ist böartig</i>	$\forall x (P(x) \rightarrow \neg BÖ(x))$	$\neg \exists x (P(x) \wedge BÖ(x))$
<i>Nicht alle Hunde sind böartig</i>	$\exists x (H(x) \wedge \neg BÖ(x))$	$\neg \forall x (H(x) \rightarrow BÖ(x))$

mengentheoretisch:



<i>Alle Pudel sind Hunde</i>	$\forall x (P(x) \rightarrow H(x))$	
<i>Alte Damen lieben Pudel</i>	$\forall y \exists x (D(y) \wedge P(x) \rightarrow L(y,x))$	
<i>Alte Damen lieben manche Hunde</i>	$\forall y \exists x (D(y) \wedge H(x) \rightarrow L(y,x))$	
<i>Alle Pudel sind Hunde</i>	$\forall x (P(x) \rightarrow H(x))$	
<i>Alle Hunde sind Tiere</i>	$\forall x (H(x) \rightarrow T(x))$	
<i>Manche Tiere sind Pudel</i>	$\exists x (T(x) \wedge P(x))$	
<i>Ein Bündnisgrüner ist ein Kommunist</i>	$\exists x (BÜ(x) \wedge K(x))$	$\exists x (BÜ(x) \wedge K(x))$
<i>Kein Bündnisgrüner befürwortet Gewalt</i>	$\forall x (BÜ(x) \rightarrow \neg BG(x))$	$\neg \exists x (BÜ(x) \wedge BG(x))$
<i>Nicht alle Kommunisten befürworten Gewalt</i>	$\exists x (K(x) \rightarrow \neg BG(x))$	$\neg \forall x (K(x) \rightarrow BG(x))$
<i>Alle DKP-Mitglieder sind Kommunisten</i>	$\forall x (DKP(x) \rightarrow K(x))$	$\forall x (DKP(x) \rightarrow K(x))$
<i>Kein Unternehmer ist Kommunist</i>	$\forall x (U(x) \rightarrow \neg K(x))$	$\neg \exists x (U(x) \wedge K(x))$
<i>Kein DKP-Mitglied ist Unternehmer</i>	$\forall x (DKP(x) \rightarrow \neg U(x))$	$\neg \exists x (DKP(x) \wedge U(x))$
<i>Alle Vegetarier leben gesund</i>	$\forall x (V(x) \rightarrow LG(x))$	$\forall x (V(x) \rightarrow LG(x))$
<i>Alle Raucher leben ungesund</i>	$\forall x (R(x) \rightarrow \neg LG(x))$	$\neg \exists x (R(x) \wedge LG(x))$
<i>Kein Vegetarier ist Raucher</i>	$\forall x (V(x) \rightarrow \neg R(x))$	$\neg \exists x (V(x) \wedge R(x))$

S 7.1 Grundlagen der lexikalische Semantik

Lernziel: Einsicht, worum es in der lexikalischen Semantik geht, welche Wege der formalen Beschreibung offen stehen und schon einige Beziehungen im Wortschatz einer Sprache auseinanderhalten können.

1. Lexikalische Semantik

Welche Bedeutung haben die Wörter?

Welche Bedeutungsbeziehungen existieren zwischen den Wörtern?

Lexikologie = die Lehre von der lexikalischen Semantik

Lexikographie = die Kunst Wörterbücher zu schreiben

Beispiel für Lexikographie

ein|stellen <sw. V.; hat> **1. a)** *in etw. (als den dafür bestimmten Platz) stellen, einordnen:* die Bücher [in das Regal] e.; **b)** *an einem [dafür bestimmten] Platz vorübergehend, zeitweilig abstellen, unterstellen:* das Auto [in eine/einer Garage] e.; falsch (nicht an der richtigen Stelle) eingestellte Bücher. **2.** *in ein Arbeitsverhältnis nehmen. anstellen:* die Firma stellt vorläufig keine neuen Arbeitskräfte ein. **3. a)** *(ein technisches Gerät o. ä. in bestimmter Weise stellen, regulieren:* ein Fernglas scharf, eine Kamera auf die richtige Entfernung, den Zeiger auf eine Marke e.; das Radio, den Fernsehapparat leiser, schärfer, auf Zimmerlautstärke, auf einen bestimmten Sender e.; **b)** *(bei einem technischen Gerät) durch Betätigen der Armaturen o. ä. etw. Bestimmtes regulieren od. zum Arbeiten bringen:* die Lautstärke, die Entfernung, einen bestimmten Sender e.; **c)** *justieren:* die Scheinwerfer, die Zündung [neu] e. **4.** *[vorübergehend] nicht fortsetzen, mit einer Tätigkeit o. ä. aufhören:* die Produktion, Zahlungen, ein Gerichtsverfahren e.; der Feind stellte das Feuer ein. **5.** (e. + sich) **a)** *zu bestimmter Zeit an einen bestimmten Ort kommen:* ich stellte mich pünktlich bei ihm ein; **b)** *(als Folge von etw.) eintreten:* starke Schmerzen stellten sich ein; Zweifel stellten sich bei uns ein. **6.** (e. + sich) **a)** *sich innerlich od. durch entsprechendes Verhalten, durch bestimmte Maßnahmen auf etw. vorbereiten:* sich, auf jmds. Besuch, auf die neue Situation e.; **b)** *sich an jmdn. anpassen, sich in seinem Verhalten nach jmdm. richten:* sie hat sich ganz auf ihren Mann, sie haben sich gut aufeinander eingestellt. **7.** (Sport) *egalisieren* (1 a). **8.** (Sport) *eine Mannschaft, einen Sportler in bestimmter Weise auf den Gegner vorbereiten:* der Trainer hat die Mannschaft defensiv, auf Defensive eingestellt.

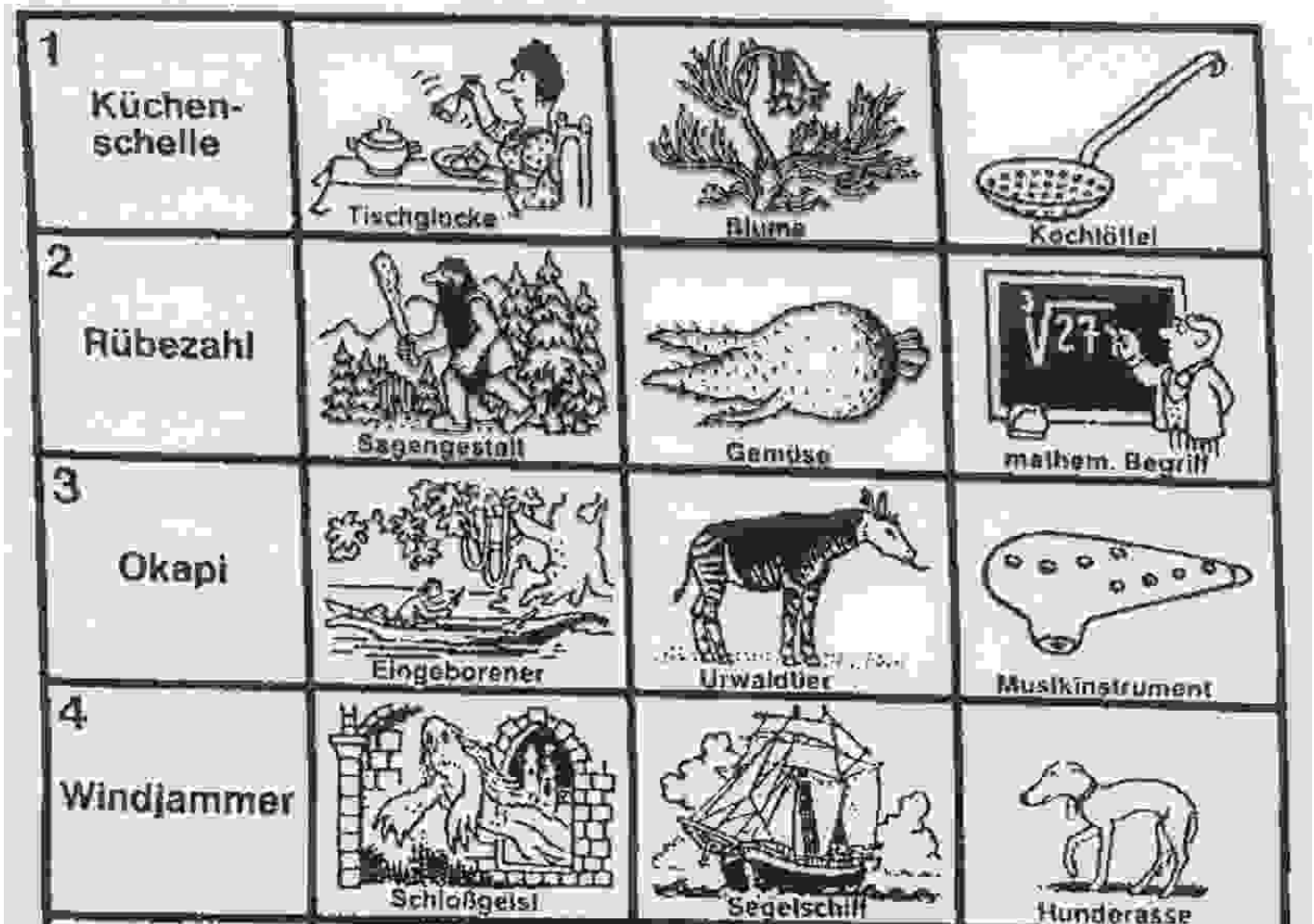
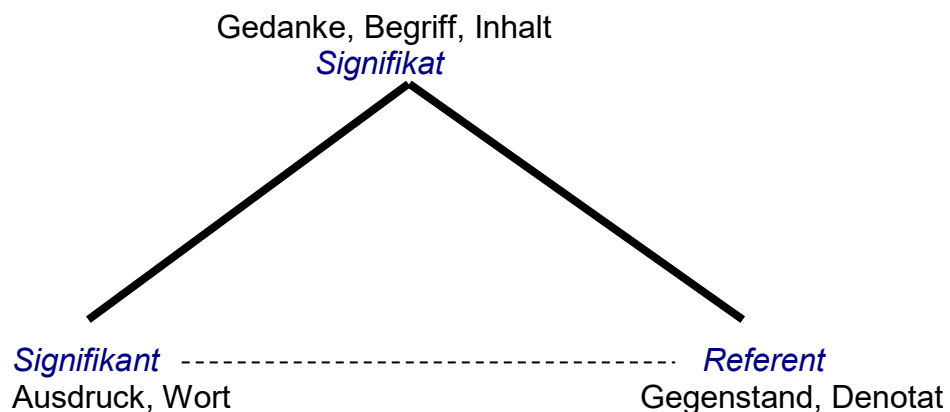
(Duden. Deutsches Universalwörterbuch)

Semasiologie - Onomasiologie

von den Wörtern zu den Sachen

–

oder von den Sachen zu den Wörtern

**Semiotisches Dreieck**

Lexikalische Semantik: Die Beziehung zwischen Wort und Begriff und der Begriffe untereinander.

Referentielle Semantik: Wie funktioniert die Bezugnahme mithilfe von Wörtern auf konkrete Gegenstände?

Polysemie

Das Wort hat verschiedene Bedeutungen, z.B.

- Flügel** 1. Körperteil von einem Vogel, Insekt e.t.c.
 2. Konzertflügel.

Siehe oben die 8 verschiedenen Bedeutungen von *einstellen*.

Homonymie

eher zufälliger Zusammenfall von Aussprache oder Orthographie von ansonsten aber unabhängigen Wörtern

- **Homophon: *Lehre, Leere***
 verschiedene Paradigmen: *Lehre, Lehren, lehren – Leere, Leeren, leer*
- **Homonym: *Ungebeiztes Holz wird modern***
 verschieden Paradigmen: Verb *modern, moderst, moderte*
 Adjektiv *modern, modernes, moderner*

Ambiguität = Mehrdeutigkeit in einem konkreten Fall



Zeichnung: Löffler

Weitere Beispiele für Polysemie (Homonymie, Metapher, Homophone, Homographe)

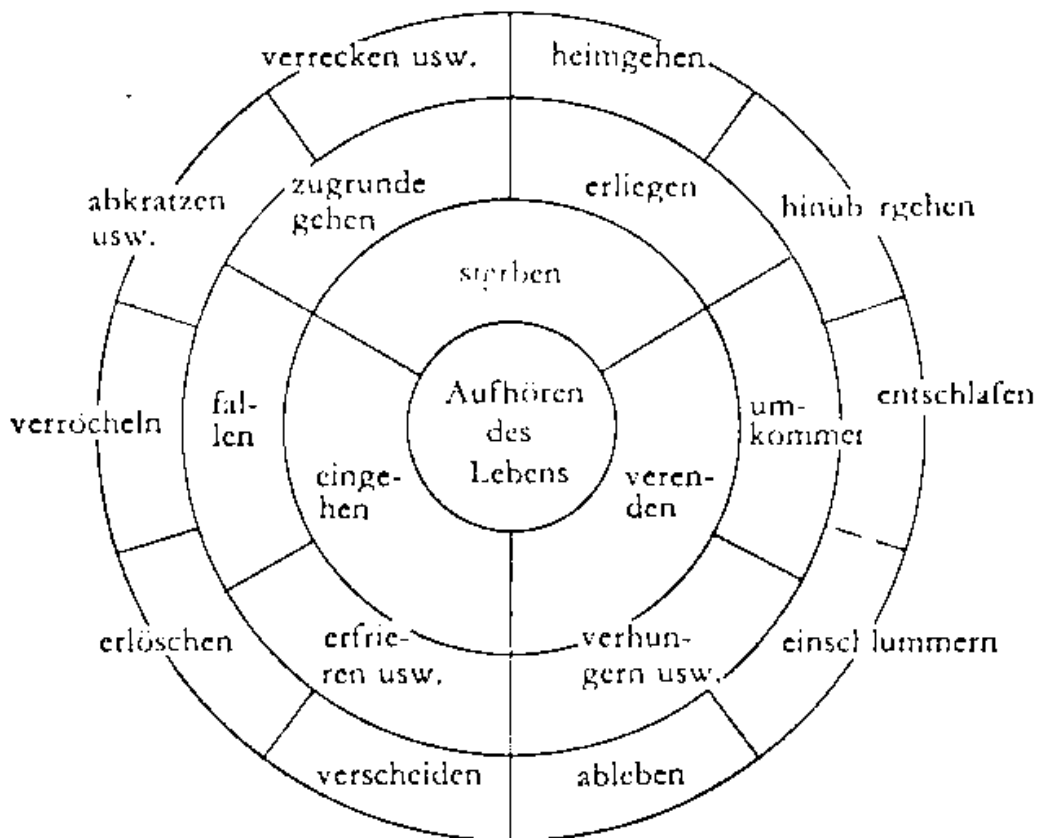
- *Warum haben Fische Schuppen?
Wo sollen sie denn sonst ihre Fahrräder unterstellen?*
- Inserat für ein Haarfärbemittel:
Mit unserer neuen Tönung fällt ihr Haar schon nach dem ersten Versuch gleichmäßig aus.
- Der Pfarrer in der Kirche:
Unser Organist kann heute nicht spielen. Ich stimme daher jetzt das Lied an, danach fällt die ganze Kirche ein.
- Glasermeister zum Kunden:
Nehmen Sie die Scheibe so mit oder soll ich sie Ihnen einschlagen
- *"Warum haben Sie ihrem Nachbarn auf einer Postkarte geschrieben, er sei ein Betrüger?"*
fragt der Richter. Der Angeschuldigte rechtfertigt sich: *"Andere schreiben ja auch Ansichtskarten."*
- Beim **Familienausflug** merkt die Mutter, dass das junge Ehepaar verschwunden ist. Sie fragt ihren Mann: *"Was werden die Kinder wohl machen?" "Nachkommen", brummt er.*
- *Lieber fernsehmüde als radioaktiv.*

2. Beschreibungsansätze

(1) Die Wortfeldidee:

Abhängigkeit der semantischen Reichweite der Wörter von der Existenz anderer (Jost Trier)
 z.B. *rot, blau, grün, gelb, weiß, ...* gehören zu dem Wortfeld der Farbbegriffe
Tisch, Stuhl, Sessel, Schrank, Spiegel, ... zum Wortfeld „Möbel“.

Aufhören des Lebens



(2) Der strukturalistische Ansatz - die Komponentenanalyse:

wie üblich Oppositionbildung, Minimalpaarverfahren, distinktive Merkmale

Minimalpaare auf phonetischer Ebene:

frau ↔ grau	/f/ ↔ /g/
frau ↔ flau	/r/ ↔ /l/
frau ↔ frei	/au/ ↔ /ei/

ebenso auf lexikalischer Ebene:

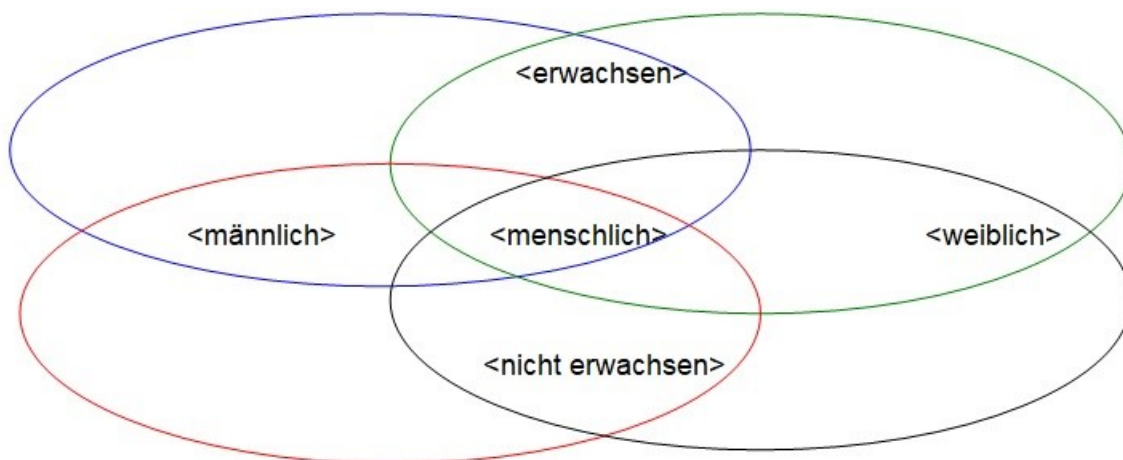
„Frau“ <-> „Mann“	[+,- WEIBLICH]
„Frau“ <-> „Mädchen“	[+,- ERWACHSEN]
„Frau“ <-> „Weibchen“	[+,- MENSCHLICH]

Wörter mit gleichen und unterschiedlichen Merkmalen :

Mensch	Mann	Frau	Kind	Eunuch
Rind	Bulle	Kuh	Kalb	Ochse
Pferd	Hengst	Stute	Fohlen	Wallach
Huhn	Hahn	Henne	Küken	Kapaun

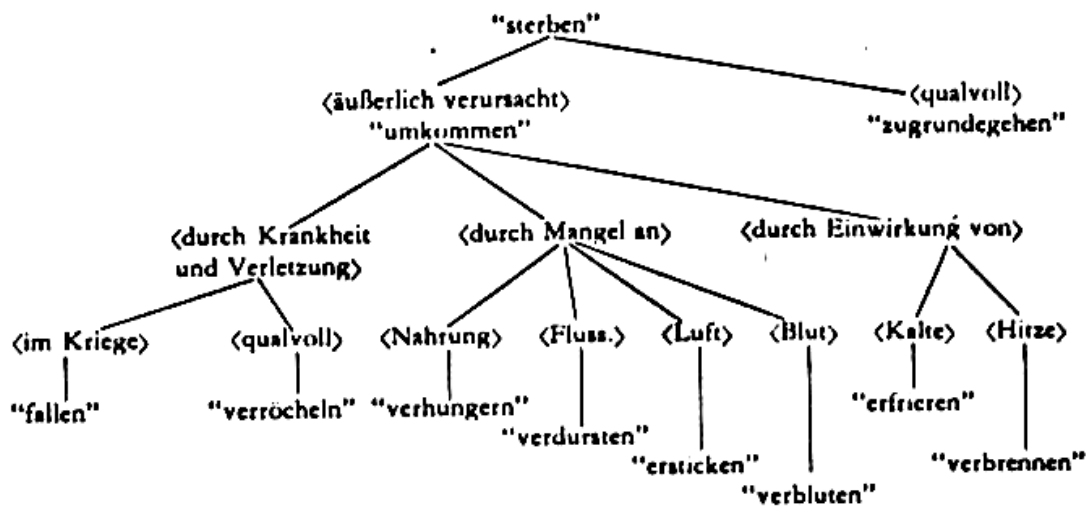
Die Merkmale können in einer **Merkmalsmatrix** dargestellt werden.

	<männlich>	<weiblich>	<erwachsen>	<kastriert>	<menschlich>
Mann	+	-	+	0	+
Frau	-	+	+	0	+
Kind	0	0	-	0	+
Eunuch	+	-	0	+	+
Junge	+	-	-	0	+
Mädchen	-	+	-	0	+
Mensch	0	0	0	0	+
Junges	0	0	-	0	-



- Die Bedeutung eines Wortes (**Semems**) kann mit der Menge seiner Oppositionsmerkmale (**Semen**) gleichgesetzt werden.

Semantische Merkmale als Entscheidungsbaum:



Aber ist diese wissenschaftliche Theorie nicht eine Illusion?

In der Praxis sind die Merkmale nicht rein formale Oppositionen (für die man Nummern vergeben könnte) sondern selbst wieder Wörter derselben Sprache, auch wenn man sie in spitze Klammern setzt. Daher greifen wir statt zur Komponentenanalyse lieber gleich zu Sinnrelationen!

(3) Sinnrelationen:

Bei der sinnrelationalen Semantik wird die Bedeutung des einzelnen Wortes, als Summe seiner Beziehungen zu anderen (echten) Wörtern betrachtet. Für die Beziehungen (Synonymie, Hyperonymie, Antonymie, Meronymie) werden hier logische Kriterien eingesetzt.

Synonym gleichbedeutend
x impliziert y und y impliziert x

1. Testsatz: "Wenn etwas/jemand (ein) X ist, dann ist es/er auch (ein) Y"
"Wenn jemand/etwas ein X hat, dann hat er/es auch ein Y"

2. Testsatz: "Wenn etwas/jemand ein Y ist, dann ist er/er auch ein X"
"Wenn jemand/etwas ein Y hat, dann hat er/es auch ein X"

3. Test: ".....X." "Dieses Y"
Mit dem Hyperonym kann man im Text sich auf dasselbe Objekt beziehen, ein Objekt wiederaufnehmen.

Beispiel *Abfall - Müll*

Hyperonym/Hyponym Ober- und Unterbegriff

x impliziert y aber nicht umgekehrt

1. Testsatz: "(ein) X ist ein Y (mit besonderen Eigenschaften)"
 "alle X sind Y"
 "X ist eine besondere Art von Y"

2. Testsatz: "nicht nur X ist ein Y"
 "nicht alle Y sind X"

3. Test: ".....X." "Dieses Y"
 Mit dem Hyperonym kann man im Text sich auf dasselbe Objekt beziehen, ein Objekt wiederaufnehmen.

Beispiele: 1. *"Liebe ist ein Gefühl"*
 2. *"Nicht nur Liebe ist ein Gefühl"*
 3. *"Sie spürte tiefe Liebe für ihn. Dieses Gefühl überwältigte Sie fast."*

 1. *"Alle Autos sind Fahrzeuge"*
 2. *"Nicht alle Fahrzeuge sind Autos"*
 3. *"Er stellt das Auto in die Garage. Am anderen Morgen war das Fahrzeug weg"*

Antonym Gegenteil

x impliziert nicht-y und y impliziert nicht-x

Testsatz: " X ist das Gegenteil von Y "
 "Beide sind/ beides ist ein (bestimmtes) Z"

Beispiele: *"Mann ist das Gegenteil von Frau"*
 "Beides ist eine bestimmte Geschlechtsrolle"
 "Zunahme ist das Gegenteil von Abnahme"
 "Beides ist Maß-Veränderung"

Meronym / Holonym Teil - Ganzes

Testsätze: ein X besteht (unter anderem) aus n Y
 ein X hat (normalerweise) n Y
 Y ist u.a. ein Bestandteil von einem X
 (dabei ist n die typische Anzahl von Teilen)

Beispiele: *"ein Baum besteht aus einem Stamm"*
 "ein Baum hat einen Stamm"
 "ein Stamm ist ein Bestandteil eines Baumes"

(4) Prototypensemantik:

Oppositionsbildung lenkt das Augenmerk zu sehr auf die Grenzen und damit auf die Zweifelsfälle. Die Prototypentheorie beschäftigt sich mit dem **Kern** der jeweiligen Bedeutung, indem sie einen typischen Vertreter beschreibt.

Außerdem kann man Realwissen und Wortwissen beim Sprachverstehen nicht trennen. Wenn jemandem eine Bedeutung erklärt wird, erfährt er etwas über die Welt und den Ausdruck zugleich.

Begriffe bilden sich aufgrund der Begegnung mit vielen Exemplaren eines Gegenstandes im täglichen Leben. Dabei kristallisiert sich das Typische, immer Wiederkehrende heraus, ein sog. **Prototyp**. Ein Prototyp ist ein idealisierter Gegenstand, dem die tatsächlichen Gegenstände aber doch mehr oder weniger nahe kommen. Bei der Herausbildung von Begriffen spielt übrigens die Kommunikation mit anderen eine wichtige Rolle. Wir verständigen uns untereinander ständig über die Welt und das darin Typische und indem wir darüber reden (oder auch streiten), bilden und ändern wir Begriffe und verknüpfen sie mit den Ausdrücken der Sprache.

Im einfachsten Fall kann Prototypensemantik aus einer Menge von (sog. generischen) Aussagen bestehen, mit denen man den typischen Vertreter des Begriffs beschreibt, so wie die Bedeutungsangaben in den meisten Wörterbüchern. Diese Aussagen wirken wie Prämissen, aufgrund derer man folgern kann.

Beispiel von Carnap, sog. meaning postulates:

Ein Junggeselle ist ein unverheirateter Mann

$\forall x \text{ Junggeselle}(x) \rightarrow \text{unverheiratet}(x) \wedge \text{Mann}(x)$

So wie Aussagen allgemein, beziehen sich auch Prototypbeschreibungen auf einen Referenzbereich, und der ist nichts anderes als unsere tatsächliche Welt.

- Die Aufstellung einer Prototypensemantik mündet in die Bemühungen um eine Kodifizierung des Weltwissens ein.

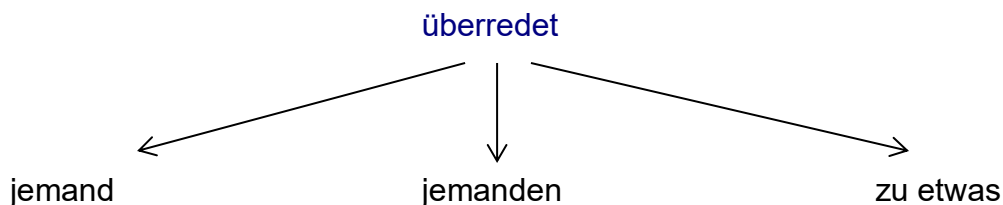
Es lassen sich viele Formate denken: Wörterbücher, Ontologien, Thesauri , Wikipedia...

Wir nehmen den Faden noch einmal in der Stunde **S 8.3 Wissensrepräsentationen** auf.

Noch ein Hinweis zur Systematik

Zur Bedeutung von Wörtern, welche als Prädikate gebraucht werden, gehören ihre Ergänzungen untrennbar hinzu. Vgl. das Beispiel aus S 5.1

Syntaktischer Aufbau



<i>Lisa</i> <i>Wer zu feige ist</i> <i>Der Finanzminister</i>	<i>ihren Freund</i> <i>jemand anderen</i> <i>den Präsidenten der</i> <i>Zentralbank</i>	<i>zu einer Erdbeerbowle</i> <i>(dazu) die Suppe auszulöffeln</i> <i>(dazu) die Zinsen zu erhöhen</i>
---	--	---

Ergänzungen

E1 der „Überreder“	E2 der „Überredete“	E3 der Überredungsinhalt
--------------------	---------------------	--------------------------

Der lexikalistischer Ansatz in der Syntax:

- Die syntagmatische Dependenzstruktur ist Teil der lexikalischen Semantik!

Zum Schluss nochmals zum Wortfeld „Aufhören des Lebens“

Wie die Leute aus dem Leben scheiden

DER ADVOKAT - tritt vor einen höheren Richter
DER GELEHRTE - gibt den Geist auf
DER FÄRBER - ist verblichen
DER KONDUKTEUR - hat seine letzte Reise angetreten
DER MAURER - kratzt ab
DER SCHRIFTSTELLER - endet
DER MATROSE - läuft in den letzten Hafen ein
DER PFARRER - segnet das Zeitliche
DER STRAßENKEHRER - kehrt zum Staub zurück
DER SCHAUSPIELER - tritt von der Bühne ab
DER SCHWERATHLET - hat ausgerungen
DER VEGETARIER - beißt ins Gras
DER CHAUFFEUR - fährt ab
DER MUSIKER - geht flöten
DER POSTBOTE - wird ins Jenseits befördert
DER SCHAFFNER - liegt in den letzten Zügen
DER STRAßENKEHRER - kehrt nie wieder
DER LANGSCHLÄFER - ist entschlafen
DER KOCH - gibt den Löffel ab
DER WEITSICHTIGE - hat die Augen geschlossen
DER FÄHRMANN - ist über den Jordan gegangen

S 7.2 Referentielle Semantik

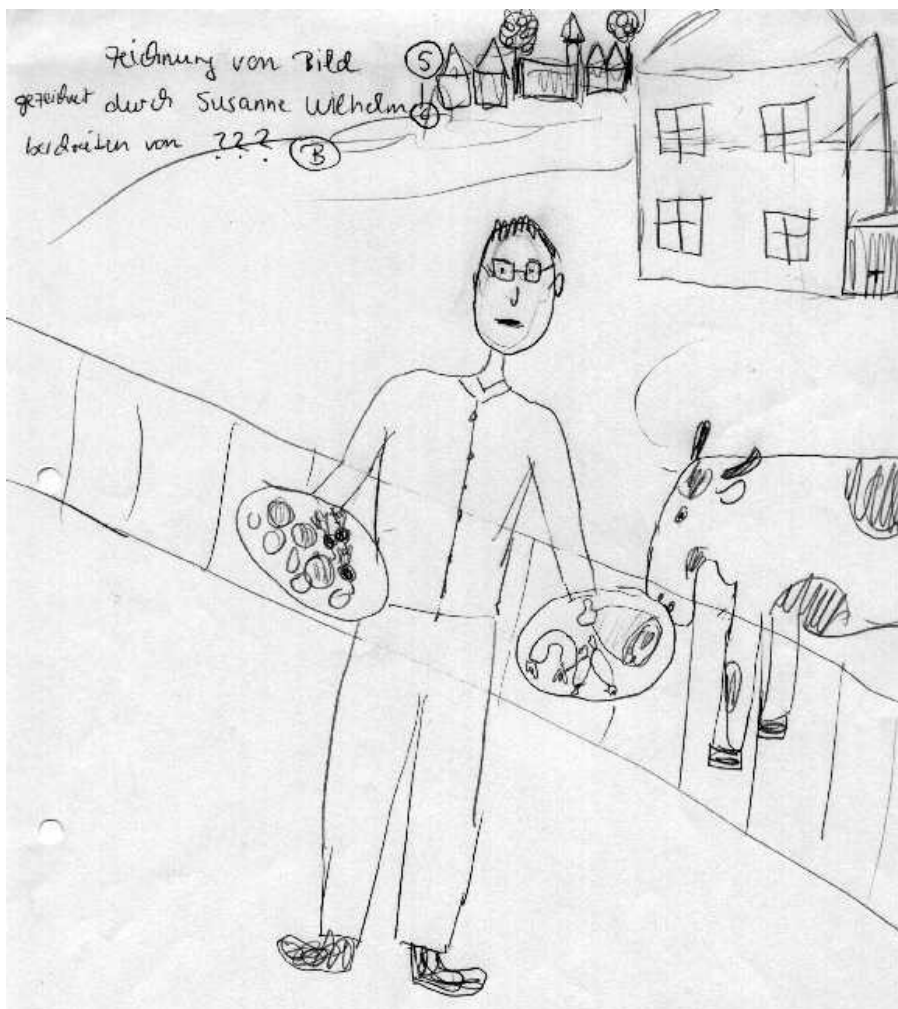
Lernziel: Konkretes Verständnis dafür, wie das Referieren funktioniert, welche Funktion dabei der definite und der indefinite Artikel hat und wie Koreferenz in Texten hergestellt wird.

1. Was ist referentielle Semantik?

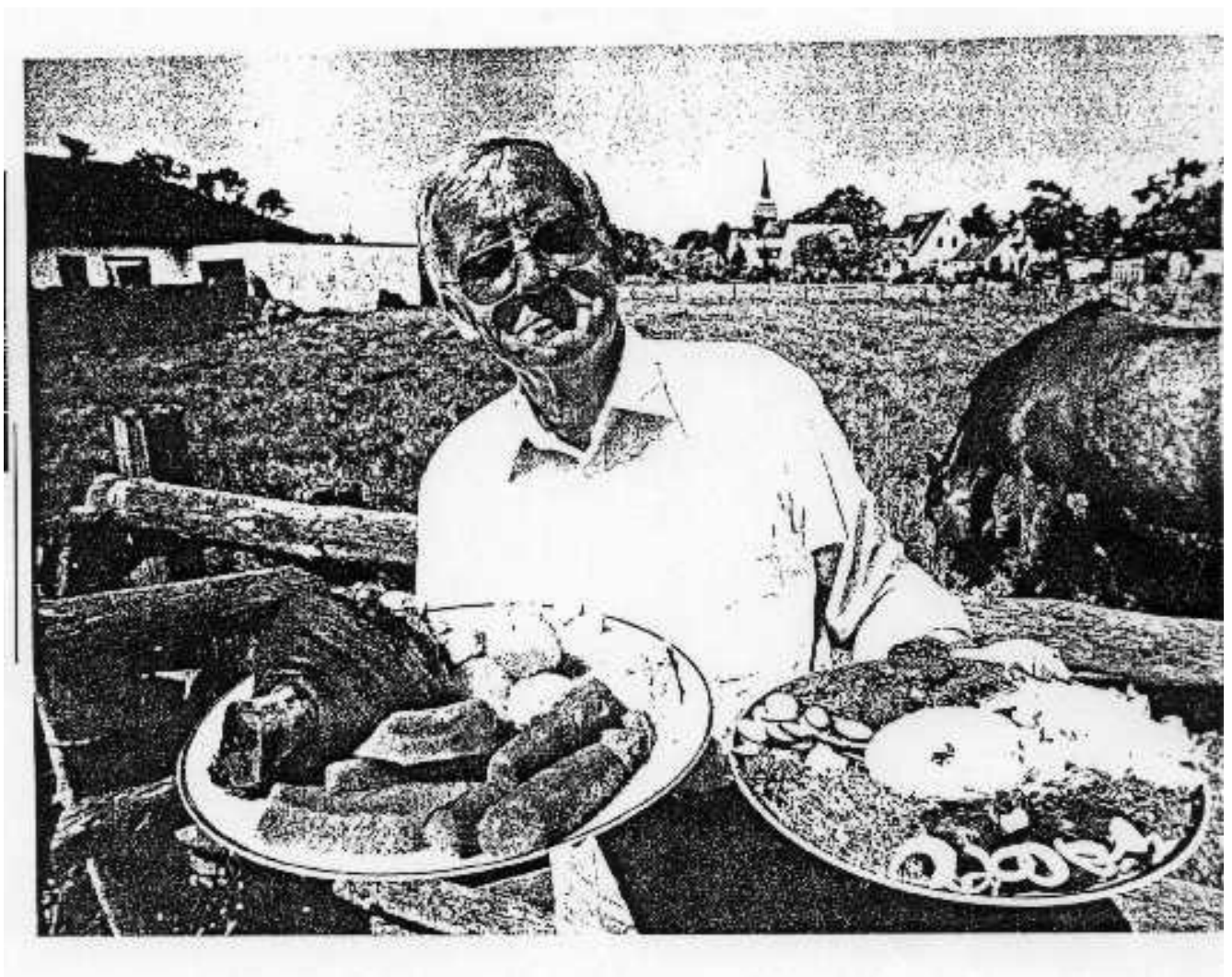
Hier ist eine Bildbeschreibung:

In der Mitte des Bildes ein hagerer älterer Mann mit Brille. Frontansicht, weißes Hemd. Er hält zwei runde Platten in den Händen. Auf der einen Platte (rechte Hand) verschiedenen Gemüsesorten in Salatform, identifizierbar: Radieschen und Zwiebeln. Auf der anderen Platte Schlachterzeugnisse - Schweinshaxe, Würste, etc. Aufgenommen ist das Bild von rechts oben, so dass die Platten zentrale Wirkung haben. Die Platte in der linken Hand ist in der rechten unteren Ecke des Bildes. Der Mann lehnt an oder sitzt auf einem hölzernen, aus zwei Balken bestehenden Weidezaun. Der Zaun verläuft auf dem Bild leicht schräg, vom unterem Drittel des rechten Rands zur Mitte des linken Bildrands. Aus dem rechten Bildrand ragt hinter dem Zaun die vordere Hälfte einer grasenden Kuh. Im oberen rechten Eck des Bildes, dem Horizont nach vorgelagert, sieht man die Hälfte eines Hauses mit Blick auf die Seitenwand, an die rechts eine Garage angebaut ist. Am Horizont die Silhouette eines Dorfes mit Kirchturm in der Mitte, Bäumen, einige Häusern.

Hier ist eine Zeichnung, die allein anhand der Bildbeschreibung angefertigt wurde:



Und hier ist das Originalbild, zu dem die Bildbeschreibung gemacht worden war und das der Zeichner nicht gesehen hatte:



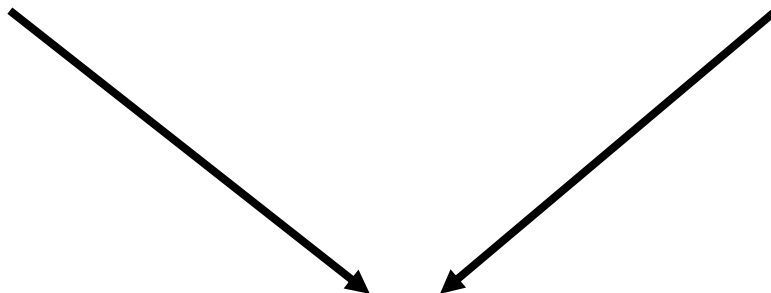
Offenbar kann zwischen einer bildlichen Darstellung (einschließlich der Wahrnehmung einer Szene mit den Augen) und der sprachliche Darstellung hin und her übersetzt werden.

Analoge (u.U. digitale) Darstellung

Propositionale Darstellung

Bildpunkte auf Netzhaut, Film, Graphik

referierende Ausdrücke, Prädikate



Gegenstände, Zusammenhänge

in der Realität

im Kopf

Referenz, Referieren = sich mit sprachlichen Ausdrücken auf Gegenstände beziehen dazu dann

Prädikation, Prädizieren = Gegenständen Eigenschaften und Beziehungen zusprechen

Nebenbei

Prädikate referieren nicht!

Würden sie selbst auf die zugesprochenen Beziehungen und Eigenschaften referieren, käme man in einen infiniten Regress:

Das Schloss ist rot – die Röte haftet am Schloss – das Haften ist zwischen Röte und Schloss usw.

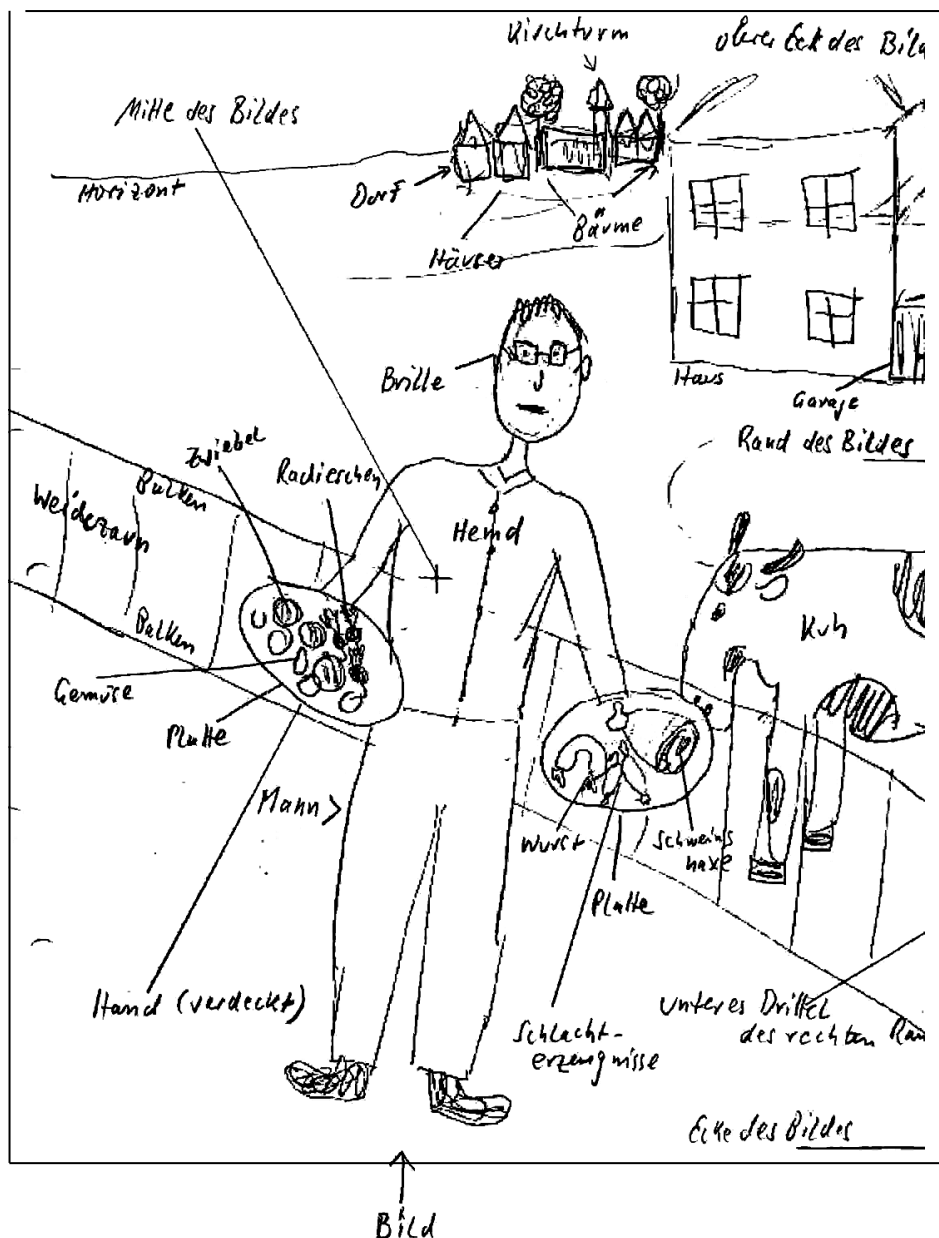
Wohl kann man auf Eigenschaften und Sachverhalte referieren:

Das Schloss ist rot. Das / dieser Sachverhalt ist bekannt. Daß das Schloss rot ist, ist bekannt.

Vgl. Tugendhat: Einführung in die sprachanalytische Philosophie, 1976

Die Zeichnung mit den im Text vorkommenden Bezeichnungen versehen:

Wer das kann, beherrscht die referentielle Semantik!



Kurzer Exkurs in die lexikalische Semantik:

Einige Sinnrelationen der referierende Ausdrücke im Text der Bildbeschreibung (vgl. S 7.1):

Balken	Synonym	Stange
Baum	Hyperonym	Pflanze
Baum	Meronym	Krone
Baum	Meronym	Stamm
Bild	Synonym	Szene
Bild	Hyperonym	Darstellung
Bild	Meronym	Bildbereich
Bildrand	Hyperonym	Bildbereich
Blick	Synonym	Sicht
Brille	Hyperonym	Zubehör
Brille	Meronym	Bügel
Brille	Meronym	Glas
Dorf	Hyperonym	Ort
Dorf	Meronym	Haus
Dorf	Meronym	Kirche
Dorf	Antonym	Stadt
Drittel	Hyperonym	Teil
Eck	Synonym	Ecke
Ecke	Hyperonym	Bereich
Ecke	Antonym	Mitte
Frontansicht	Antonym	Rückansicht
Garage	Hyperonym	Gebäude
Garage	Meronym	Tor
Gemüsesorte	Hyperonym	Nahrungsmittel
Hälfte	Hyperonym	Teil
Hälfte	Antonym	Ganzes
Hand	Hyperonym	Körperteil
Hand	Meronym	Finger
Hand	Holonym	Arm
Haus	Hyperonym	Gebäude
Haus	Meronym	Dach
Hemd	Hyperonym	Kleidungsstück
Hintergrund	Antonym	Vordergrund
Horizont	Synonym	Hintergrund
Horizont	Hyperonym	Bereich
Kirche	Meronym	Kirchturm

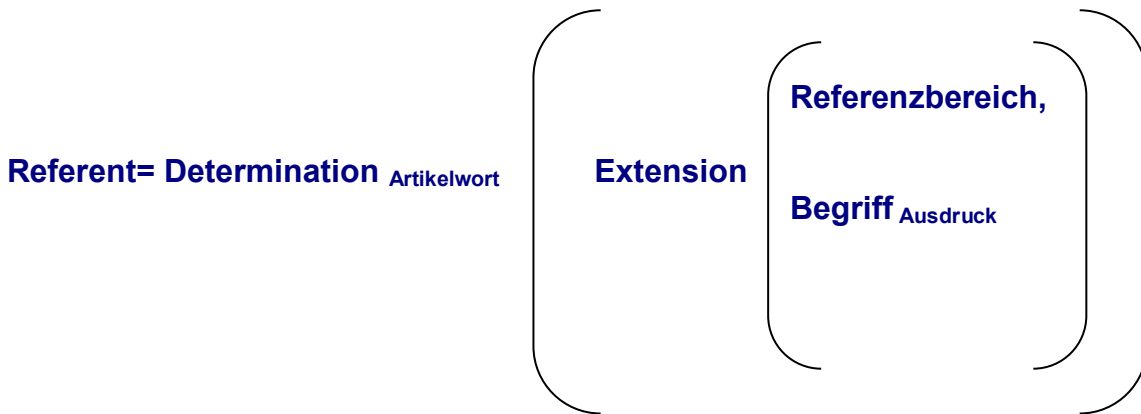
Kirchturm	Hyperonym	Gebäude
Kirchturm	Meronym	Kirchturmspitze
Kuh	Hyperonym	Rind
Mann	Synonym	Herr
Mann	Hyperonym	Person
Mann	Antonym	Frau
Mitte	Synonym	Zentrum
Mitte	Hyperonym	Bereich
Mitte	Antonym	Peripherie
Person	Meronym	Arm
Platte	Synonym	Teller
Platte	Hyperonym	Geschirr
Radieschen	Hyperonym	Gemüsesorte
Rand	Hyperonym	Bereich
Rand	Antonym	Zentrum
Rind	Hyperonym	Säugetier
Salatform	Synonym	Salatart
Säugetier	Meronym	Maul
Schlachterzeugnis	Hyperonym	Nahrungsmittel
Schweinshaxe	Hyperonym	Schlachterzeugnis
Silhouette	Synonym	Umriß
Silhouette	Hyperonym	Form
Weidezaun	Hyperonym	Zaun
Weidezaun	Meronym	Stange
Wirkung	Synonym	Eindruck
Wurst	Hyperonym	Schlachterzeugnis
Zaun	Hyperonym	Einfriedung
Zwiebel	Hyperonym	Gemüsesorte

Stellen Sie sich die in Sinnrelation gesetzten Begriffe als Hintergrundbezeichnung vor. Sie könnten dann mit diesen Begriffen ebenfalls auf Teile des Bildes referieren. Dies geschieht auch in Texten. Referentielle und lexikalische Semantik wirken Hand in Hand.

Siehe Stunde **S 8.2: Grundlagen der Textkonstitution und des Textverstehens.**

2. Der Mechanismus des Referierens

Formel zur Referenz:



Es gibt eine Funktion „Extension“ mit zwei Argumenten: einem Begriff, der mit dem vorliegenden sprachlichen Ausdruck verbunden ist und einem pragmatisch vorliegenden Bereich in der Welt. Die Funktion gibt die Gegenstände aus, welche in dem Referenzbereich unter den Begriff fallen.

Es gibt weiter eine spezielle Funktion „Determination“ zu jedem Artikelwort, deren Argument die Ausgabe der Extensionsfunktion ist (nämlich die Gegenstände, die im Referenzbereich unter den Begriff fallen). Die Determinationsfunktion gibt den Gegenstand in der Welt aus, auf den der Ausdruck referiert, den Referent des Ausdrucks.

Die Determinationsfunktionen werden in vielen Sprachen durch Artikelwörter repräsentiert, engl. *Determiners*. Die wichtigsten sind die Funktionen des bestimmten und unbestimmten Artikels sowie des weggelassenen Artikels (Null-Artikel):

Determination *der, die, das* (inkludierend)

Referent ist genau das, was im gegebenen Referenzbereich in die Extension des Begriffs fällt.

Determination *ein, eine, 0* (exkludierend)

Referent ist etwas, was im gegebenen Referenzbereich in die Extension des Begriffs fällt. Normalerweise aber nicht alles, was in der Extension ist, weil dann der definite Artikel genommen werden muss. Es bleibt etwas ausgeschlossen.

Andere Determinatoren mit jeweils anderer Auswahl aus der Extension des Begriffs im Referenzbereich:

jeder, alle, manche, viele, einige, die Mehrzahl der ...

Erinnerung an die Prädikatenlogik (S 6.2):

Semantik:

ein sog. Modell $M=(D,F)$

D ist eine nicht-leere Menge von Objekten (die Domäne).

F ist eine Funktion, die

- jeder Individuenkonstante ein Element aus D zugeordnet,
- jedem n-stelligen Prädikatensymbol eine n-stellige Relation über D zuordnet

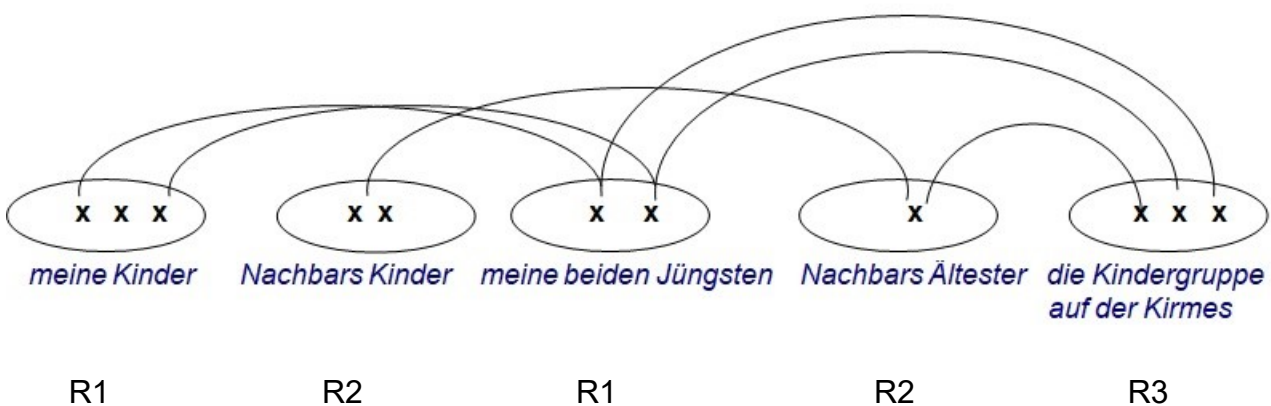
- Die Domäne D ist dasselbe wie der Referenzbereich in unserer Formel. Aber der steht in der natürlichen Kommunikation nicht fest, sondern ist hochgradig variabel, dynamisch, situationsabhängig.

Beispiel:

*Ich habe drei Kinder, mein Nachbar hat zwei. Neulich bin ich mit meinen beiden Jüngsten und dem ältesten Sohn des Nachbarn auf der Kirmes gewesen. Zum Abschluss bekamen **alle Kinder** ein Eis.*

Wer sind **alle Kinder** im letzten Satz??

weder alle meine Kinder, noch alle Kinder des Nachbarn, noch alle Kinder auf der Kirmes, noch alle Kinder in der Welt ...



Referenzbereiche im Text: R1 = meine Familie, R2 = die Familie des Nachbarn, R3 der Kirmesbesuch. Alle Kinder bezieht sich auf R3 und die darin befindlichen Kinder.

- Referenzbereiche sind in verschiedenem Grad offen, Referenten anaphorisch in verschiedenem Grad erreichbar

So wie der Text fortschreitet, gibt es neue Referenzbereiche, die von da ab zur Extensionalisierung benutzt werden können.

Canakkale (dpa) - Ein unter kambodschanischer Flagge fahrender **Frachter** ist am Abend in den **Dardanellen** mit einem türkischen **Tanker** zusammengestoßen und zwei Stunden später gesunken. Die 17 Mann Besatzung konnte rechtzeitig in Sicherheit gebracht werden, berichtete eine türkische Nachrichtenagentur. Warum die beiden Schiffe kollidierten, ist noch unklar. Die 65 Kilometer lange Meeresstraße verbindet das Marmarameer mit der Ägäis. 12.04.2002 23:21

MEZ

fett= Referenzbereiche

Frachter, Tanker, Dardanellen

unterstrichen= extensionalisiert im vorangehenden Referenzbereich

Frachter – 17 Mann Besatzung,

Frachter und Tanker – die beiden Schiffe,

Dardanellen – die . Meeresstraße

Koreferenz

Zwei Ausdrücke in einem Text oder Dialog sind koreferent, wenn sich der Sprecher mit ihnen auf dieselben Gegenstände bezieht.

m.a.W.

Die Determination auf Basis der Extension der beiden Ausdrücke führt zu derselben Menge von Referenten.

Sprachliche Mittel der Koreferenz:

- Wiederholung

*Ein unter kambodschanischer Flagge fahrender **Frachter** ist am Abend mit einem türkischen **Tanker** zusammengestoßen. Der Frachter ist innerhalb von zwei Stunden später gesunken. Der Tanker konnte in den Hafen von Istanbul geschleppt werden.*

- Hyperonym

*Ein unter kambodschanischer Flagge fahrender **Frachter** ist am Abend mit einem türkischen Tanker zusammengestoßen. Warum das Schiff die Route des Tankers kreuzte ist noch unklar.*

- Pronomen

*Ein unter kambodschanischer Flagge fahrender **Frachter** ist am Abend mit einem türkischen Tanker zusammengestoßen. Er ist innerhalb von zwei Stunden gesunken.*

- Ellipse

*Ein unter kambodschanischer Flagge fahrender **Frachter** ist am Abend mit einem türkischen **Tanker** zusammengestoßen und zwei Stunden später gesunken. Die 17 Mann Besatzung [Ellipse des Frachters] konnte rechtzeitig in Sicherheit gebracht werden.*

Korrelierte Referenz, Kontiguität, Bridging

Aufgrund unserer Kenntnis von Prototypen haben wir Vorstellungen von zusammen vorkommenden Objekten und von ihrem wirklichen Zusammenhang. Deshalb sehen wir auch in Texten Zusammenhänge zwischen den entsprechenden Bezeichnungen.

Jack bought a **new car**. The engine needed tuning.
 I went into **the room**. The window was open.
 Jane has a **new house**. The front door is blue.
 Susan sent **her meal** back. The steak was underdone.
 John **went walking** at noon. The park was beautiful.
 Mary **dressed** the baby. The clothes were made of pink wool.

In der Welt gibt es einen Bezug zwischen *car - engine, room - window, house - door, meal - steak, walking - park, dressing - cloths*. Ohne Kenntnis dieser Beziehungen, würden man auch kaum die Objekte wiedererkennen, die ein Begriff bezeichnen kann. Eher vortheoretisch, intuitiv wird dieser Bereich **Kontiguität** genannt. Die sinnrelationale und protoypischen lexikalische Semantik, versucht diese Beziehungen zu systematisieren.

Zusammenhänge wie in den obigen Beispielen werden in der Textlinguistik als **Bridging** bezeichnet. Meine Deutung:

Aufgrund des prototypischen Wissens, das ein Ausdruck im Text aufruft, öffnet sich ein neuer Referenzbereich, z.B. bei **car** alles, was über Autos weiß. In diesem Referenzbereich ist **engine** dann leicht zu extensionalisieren.

nochmal Beispiel Bildbeschreibung

grün = Neueinführung von Referenten (normal mit indefinitem Artikel)

In der Mitte des Bildes ein hagerer älterer Mann mit Brille. Frontansicht, weißes Hemd. Er hält zwei runde Platten in den Händen. Auf der einen Platte (rechte Hand) verschiedenen Gemüsesorten in Salatform, identifizierbar: Radieschen und Zwiebeln. Auf der anderen Platte Schlachterzeugnisse - Schwinshaxe, Würste, etc. Aufgenommen ist das Bild von rechts oben, so dass die Platten zentrale Wirkung haben. Die Platte in der linken Hand ist in der rechten unteren Ecke des Bildes. Der Mann lehnt an oder sitzt auf einem hölzernen, aus zwei Balken bestehenden Weidezaun. Der Zaun verläuft auf dem Bild leicht schräg, vom unterem Drittel des rechten Rands zur Mitte des linken Bildrands. Aus dem rechten Bildrand ragt hinter dem Zaun die vordere Hälfte einer grasenden Kuh. Im oberen rechten Eck des Bildes, dem Horizont nach vorgelagert, sieht man die Hälfte eines Hauses mit Blick auf die Seitenwand, an die rechts eine Garage angebaut ist. Am Horizont die Silhouette eines Dorfes mit Kirchturm in der Mitte, Bäumen, einige Häusern.

Koreferenz und Bridging

1. Pragmatisch vorgeben: *das Bild* - die Mitte des Bildes - das Bild - die rechte untere Ecke des Bildes - auf dem Bild - das untere Drittel des rechten Randes - die Mitte des linken Bildrands - das Bild - der obere rechte Eck des Bildes
2. *ein Mann* - er - die Hände - rechte Hand - die linke Hand - der Mann
3. *zwei Platten* - die eine Platte - die andere Platte - die Platten - die Platte in der linken Hand
4. *Gemüsesorten* - Radieschen – Zwiebeln
5. *Schlachterzeugnisse* - Schweinshaxe – Würstel
6. *ein Weidezaun* - zwei Balken - der Zaun - der Zaun
7. *eine Kuh* - die vordere Hälfte
8. *der Horizont* - am Horizont
9. *ein Haus* - die Hälfte - die Seitenwand - eine Garage
10. *ein Dorf* - die Silhouette - Kirchturm - Bäume - einige Häuser

S 7.3 Satzsemantik

Lernziel: Einsicht in einige satzsemantische Phänomene

1. Abgrenzung

Die Satzsemantik untersucht den Teil der Bedeutung der Sprache, der sich aus dem **Aufbau** des Satzes ergibt (also der Syntagmatik), anders als die referentielle Semantik, die untersucht, was die einzelnen Wörter bezeichnen können.

✓ Man kann sagen: Satzsemantik ist Semantik minus referentieller Bedeutung

Abstraktion von der referentiellen Semantik - was bleibt?

All borogoves were mimsy.

The mome raths are borogoves.

Hence: The mome raths were mimsy.

Dass die Wahrheitswerte von Prämissen und Konklusion abhängig voneinander sind, liegt nicht an der referentiellen Semantik, d.h. an den Gegenständen und dem prototypischen Wissen darüber, sondern nur an der Struktur der Sätze und an Wörtern wie *all, the, be (were, are)*

„Truth Conditional Semantics“

Logik als Lehre was man unter welchen Bedingungen sagen kann/muss = eine Sprachwissenschaft, jedenfalls ein wichtiger Teil der Gebrauchsbedingungen für Sätze:

Welche Sätze folgen auseinander, welche sind äquivalent?

Welche präsupponieren sich, welche bilden einen Widerspruch?

Die logische Form der natürlichen Sprache

Hintergrund: Die Bedeutung der Sprache ist ihr Gebrauch (Wittgenstein)

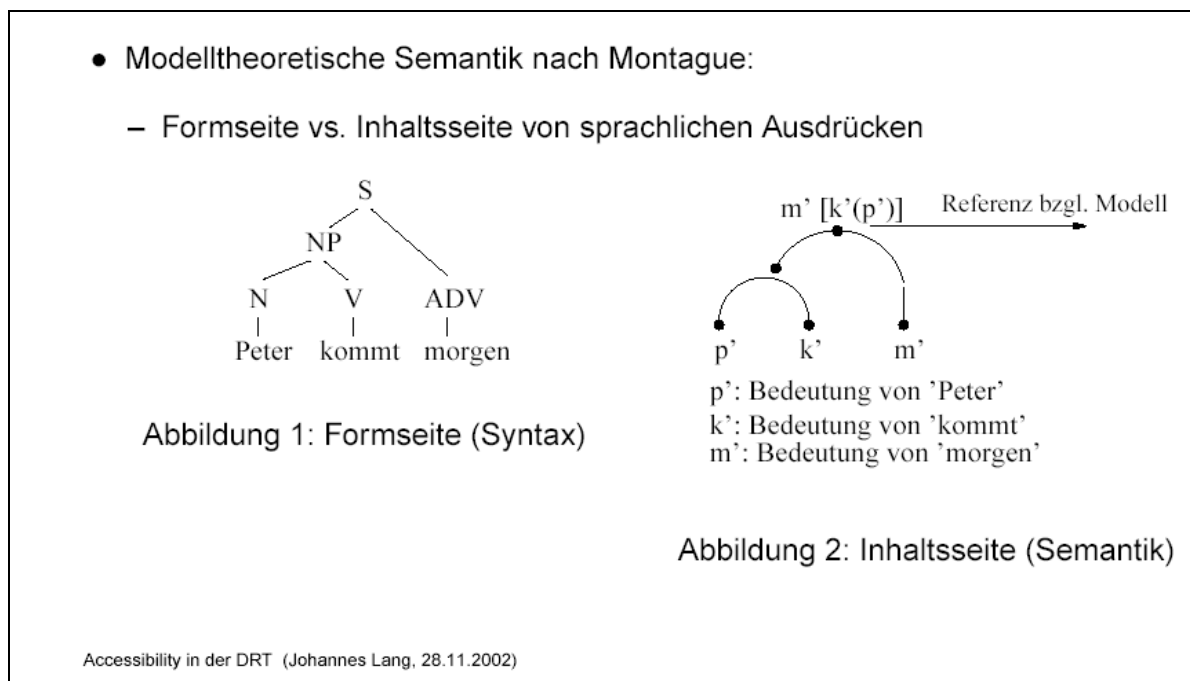
✓ Bedeutung beschreiben heißt demnach Gebrauchsbedingungen angeben.

Es gibt **zwei Herangehensweisen**:

(1) Trennung von Form und Inhalt

- Die Syntax beschäftigt sich ausschließlich mit der sog. **Oberflächenstruktur**. Welche Zeichenketten bilden wohlgeformte Sätze der Sprache? Eine generative Grammatik akzeptiert genau diese und weist alle anderen zurück.
- Für die Satzsemantik wird eine eigene Repräsentation entwickelt, an der man alle Bedeutungsbeziehungen zwischen Sätzen ablesen kann. Im Rahmen der generativen Grammatik war dies die sog. **Tiefenstruktur**. Oder es wurden eigene **semantische Repräsentationen** entwickelt - meist Versuche, Aussagenlogik und Prädikatenlogik so auszubauen, dass damit auch satzsemantische Phänomene der natürlichen Sprache gelöst werden konnten.
- Schließlich wird eine Abbildung von der Oberflächenstruktur in die Tiefenstruktur definiert.

Beispiel:



(2) Symbiose von Form und Funktion

Die Beschreibung der Syntagmatik (der Verbindung zwischen Sprachelementen) berücksichtigt von Anfang an sowohl die Form wie die Funktion (oder Rolle) der syntaktischen Einheiten. Indem untersucht wird, welche Rolle die Konstituenten spielen, wird gleichzeitig deren Wohlgeformtheit definiert. Dies ist das Design von funktionalen Grammatiken (hauptsächlich Dependenzgrammatiken). Die traditionelle Sprachwissenschaft mit ihren Kategorien Subjekt, Prädikat, Attribut, Adverb usw. geht auch so vor. Wir haben oben bei der Syntaxanalyse schon diese Haltung eingenommen (vgl. S 5.1).

2. Ausgewählte semantische Beziehungen aufgrund von Satzstrukturen

- **Paraphrasen** (logisch Äquivalenz)

(Sternchen * bedeutet, dass der Satz nicht korrekt ist, d.h. dass die Paraphrase nicht möglich ist.)

Aktiv- und der Passivsatz

Er kann Lydia leicht überreden

Lydia kann von ihm leicht überredet werden

Der Vorfall überraschte/betraf/berührte ihn kaum

Er wurde von dem Vorfall kaum überrascht/betroffen/berührt

Der Vorfall kümmerte/ärgerte/entsetzte ihn kaum

** Er wurde von dem Vorfall kaum gekümmert/geärgert/entsetzt*

- ✓ Das passivfähige Objekt muss eine andere Rollenkatgorie bekommen als das nicht-passivfähige.

„ist zu“ – „kann werden“ - „läßt sich“

Lydia ist leicht zu überzeugen/zu täuschen/zu beschwichtigen

Lydia kann leicht überzeugt/getäuscht/beschwichtigt werden

Lydia lässt sich leicht überzeugen/täuschen/beschwichtigen

Lydia ist begierig/geeignet/unfähig zu überzeugen

** Lydia kann begierig/geeignet/unfähig überzeugt werden*

** Lydia lässt sich begierig/geeignet/unfähig überzeugen*

- ✓ Die beiden Ausgangssätze sehen nur an der Oberfläche ähnlich aus. In der ersten Gruppe steht **sein + Verb im Infinitiv-mit-zu**. Hierfür gilt die Paraphrase. In der zweiten Gruppe steht **sein + Prädikatsadjektiv**. Dafür gibt es keine Paraphrase. Der Infinitiv hängt im zweiten Satz vom Adjektiv ab.

Lydia ist leicht/schwer/unmöglich zu überzeugen .

Es ist leicht/schwer/unmöglich, Lydia zu überreden.

Lydia ist schnell/hoffentlich/schwerlich/morgen zu überreden .

** Es ist schnell/hoffentlich/schwerlich/morgen, Lydia zu überreden.*

- ✓ *Unterschiedliche Rollen der Adverbien in den beiden Gruppen*

„dass“-Satz - Infinitiv mit zu

Ich hoffe/freue mich/bin sicher, dass ich an der Tagung teilnehmen kann

Ich hoffe/ freue mich/bin sicher, an der Tagung teilnehmen zu können

Ich weiß, dass ich an der Tagung teilnehmen kann

** Ich weiß, an der Tagung teilnehmen zu können*

- **Implikation, Implikatur**

Peter ist mit dem Fahrrad gekommen impliziert *Peter ist gekommen.*

W	→	W
F	→	W/F

Ich habe versucht , Hans gestern anzurufen Implikatur: *Ich habe Hans nicht erreicht*

- **Präsupposition**

Der König von Frankreich ist kahlköpfig.

Der König von Frankreich ist nicht kahlköpfig

} Präsupposition:
*Es gibt gegenwärtig einen
König in Frankreich.*

Mayer hat aufgehört, seine Frau zu schlagen.

*Mayer hat nicht aufgehört, seine Frau zu schlagen
geschlagen*

} Präsupposition:
Mayer hat bisher seine Frau

- ✓ Die Präsupposition muß wahr sein, damit der Satz, der sie voraussetzt, überhaupt wahr oder falsch sein kann. Ist die Präsupposition falsch, ist der Satz sinnlos.

S 8.1: Grundlagen der Pragmatik und Sprechakttheorie

Lernziel: Begreifen, dass Sprechen (oder Schreiben) eine von anderen sozialen Handlungen ist und dass man ohne diesen Aspekt Sprache nicht richtig beschreibt.

1. Sprechen als Handeln

Sprache hat eine soziale Funktion. Sie ist ein Mittel sozialer Organisation.

Auch das Denken ist - über die Sprache - sozial geprägt.

Definitionen:

Pragmatik - Kunst, richtig zu handeln
 - wissenschaftliche Untersuchung des Handelns

Handlung ziel- und zweckgerichtete Tätigkeit,
 Verhalten, das als absichtsvoll interpretiert werden kann

Interaktion zwischenmenschliches Handeln,
 Handlungen, die Beziehungen zwischen mindestens zwei Personen
 herstellen, bewahren, ändern, abbrechen

Handlungen: Auto fahren, Kochen, Schwimmen, Straße überqueren
 ebenso Sprech- und Schreibhandlungen

Sprachhandlungen können nicht weniger wirksam sein als andere Handlungen.

Ein falsches Wort kann u.U. ähnlich schwere Folgen haben wie falsches Überqueren der Straße. Man kann ins Gefängnis kommen, andere ins Gefängnis bringen, gar andere ums Leben bringen, aus Versehen oder absichtlich.

Handeln ist ein Verhalten, das als absichtsvoll interpretiert werden kann, d.h. eine Handlung hat ein **Intention**.

Interaktion: ein Sonderfall von Handlungen, nämlich zwischenmenschliches Handeln.

Handlungen, die Beziehungen zwischen mindestens 2 Personen herstellen, bewahren, ändern, abbrechen.

➤ **Sprechen/Schreiben ist eine spezielle Form der Interaktion.**

Welche sprachlichen Interaktionen gibt es?

In einer früheren Hausaufgabe sollten Verben für Sprechakte aus dem Wörterbuch herausgesucht und dazu ein Beispiel erfunden werden. Wer hätte gedacht, dass es so viele sind?

<i>beanspruchen</i>	"Das gehört mir!"
<i>beanstanden</i>	"Das Gerät funktioniert nicht"
<i>beantragen</i>	"Wir beantragen einen Parkausweis für die Altstadt"
<i>beauftragen</i>	"Bitte kaufen Sie 100 Stück davon!"
<i>bedanken</i>	"Vielen Dank für das schöne Geschenk!"
<i>befehlen</i>	"Sei still!"
<i>befragen</i>	"Was halten Sie davon?"
<i>befürworten</i>	„Das sollte man versuchen.“
<i>beglaubigen</i>	"Die Unterschrift stimmt mit dem Original überein".
<i>beglückwünschen</i>	"Alles Gute zum Geburtstag!"
<i>begnadigen</i>	"Der Angeklagte wird begnadigt!"
<i>begründen</i>	"Ich kann nicht kommen, weil ich Krank bin."
<i>bezeugen</i>	"Mit Sicherheit habe ich gesehen, daß der Unfallgegner ein rotes Auto gefahren hat."
<i>darlegen</i>	"Zu dem Angeklagten stehe ich in keinem besonderen Verhältnis. Ich sah ihn ein oder zwei mal auf einer Party; lange unterhalten haben wir uns aber nie. Er kam mir auf Anhieb sehr eigenartig vor."
<i>dedizieren</i>	"Für Julia"
<i>deklarieren</i>	"Ich möchte vier Flaschen Rum und 25 Zigarren verzollen."
<i>dementieren</i>	"Der in der Presse erwähnte Zwischenfall hat nie stattgefunden."
<i>diskreditieren</i>	"Diesem Herrn Olaf solltest Du kein Wort glauben. Er ist ein Hochstapler. Um 400 Mark hat er mich betrogen!"
<i>diskutieren</i>	..."Das menschliche Gebiß zeigt Merkmale, die sich auch bei den Gebissen anderer Fleischesser finden lassen. Also ist scheint es doch nur natürlich, daß wir Fleisch essen." "Der Mensch kann aber ohne Schwierigkeiten überleben, ohne Fleisch zu verzehren. Ganz abgesehen davon kann ein physiologisches Merkmal doch keine Berechtigung zum Töten sein."...
<i>dokumentieren</i>	"Die Funktion find_last sucht das alphabetisch kleinste Element der Liste und gibt anschließend die Adresse des Elements zurück."
<i>dolmetschen</i>	"Sie sagt, das sei zu teuer. In Deutschland seien die Hemden billiger, sagt sie."
<i>drohen</i>	"Wenn Sie meiner Tochter noch einmal in die Nähe kommen, können Sie `was erleben!"

<i>durchsagen</i>	"An Gleis 5 bitte einsteigen. Die Türen schließen selbsttätig. Vorsicht bei der Abfahrt!"
<i>duzen</i>	"Kannst du mir mal helfen, Hans?"
<i>ehelichen</i>	Pfarrer: „Willst Du xy zum Mann/zur Frau nehmen?“ Er/sie: „Ja, ich will!“
<i>einbestellen</i>	„Kommen Sie morgen 14 Uhr in mein Büro.“
<i>einfordern</i>	„Gib mir mein Lineal zurück“
<i>einklagen</i>	„Herr Richter, ich fordere die Einlösung des Vertrages.“
<i>einladen</i>	„Hast Du Lust mich morgen zu besuchen?“
<i>einlenken</i>	„Ich gebe zu, in diesem Punkt hatte ich Unrecht“
<i>einreden auf jmd.</i>	„Ach bitte, bitte... und schau doch mal... und sieh doch mal... und willst du es dir nicht noch mal überlegen?“
<i>einschärfen</i>	„Und wehe, Du vergißt die Blumen zu gießen!“
<i>einweisen</i>	„Dies ist Ihr neuer Arbeitsplatz.“
<i>einwilligen</i>	"Mit Ihrem Projekt bin ich einverstanden."
<i>erklären</i>	„...und danach ziehen sie die Lasche durch den Nippel....“
<i>fachsimpeln</i>	"Die neuen Dieselmotoren sind ja wirklich sparsam und auch durchzugskräftig und schau mal auf die Elastizität der Drehzahlen."
<i>fantasieren</i>	"Ich könnte mir gut vorstellen, König von Deutschland zu sein."
<i>feilschen</i>	"Das kostet 100 DM? Ich gebe ihnen maximal 60 DM dafür"
<i>festlegen</i>	„wir definieren hiermit die ...“
<i>feststellen -</i>	„Ich stelle hiermit fest, das <i>Feststellen</i> eine sprachlich geartete Handlung vorraussetzen kann.“
<i>folgern</i>	"Im Wetterbericht wurde gesagt, daß es morgen 3 Grad C werden, das bedeutet wohl, das es kalt wird"
<i>formulieren</i>	„ich moechte das gerne mal so sagen...“
<i>fortfahren</i>	„Wo war ich denn noch stehen geblieben? Also ein Sprechakt verlangt“
<i>freihalten</i>	"Dieser Platz ist leider schon besetzt."
<i>freisprechen</i>	„Wir kommen zur Urteilsverkündigung. Der Angeklagte wird freigesprochen.“
<i>gratulieren</i>	“happy birthday to you...”
<i>hänseln</i>	'Max, du siehst aus wie Pumuckl, mit deinen roten Haaren! Pumuckl, Pumuckl, Pumuckl!'
<i>handeln</i>	'Also zehn Mark ist mir zu viel dafür... aber für acht würde ich es nehmen...?!'

<i>herausreden</i>	"Ich wollte mit dir darüber sprechen, was mir Sorgen bereitet"
<i>herausstreichen</i>	'Ich muss besonders hervorheben, dass der Kollege Meier dieses Jahr hervorragende Leistungen erbracht hat.'
<i>herrufen</i>	'Komm mal bitte her!'
<i>heucheln</i>	'Oooch, das habe ich gar nicht gewusst, das tut mir aber leid...'
<i>heulen</i>	'huhuhuuuuuu'
<i>hinausposaunen</i>	"Hast du die Nachricht noch nicht gehört? Ich erzähle dir alles gleich!"
<i>hinauswerfen</i>	'Sie sind fristlos entlassen!'
<i>hochjubeln</i>	'Dieser neue Film - fantastisch!! Also so etwas geniales hätte ich nicht für möglich gehalten! Daran ist echt alles perfekt!'
<i>huldigen</i>	: 'Oh großer Koenig, ich verehere dich sehr.'
<i>informieren</i>	"Die Versammlung findet am 10.01. statt"
<i>instruieren</i>	"Wenn sie die Akten sortiert haben, bringen Sie sie bitte in mein Büro!"
<i>interviewen</i>	"Sagen Sie bitte Ihre Meinung zu diesem Problem"
<i>jammern</i>	"Och, tun mir heute die Füße weh."
<i>jauchzen</i>	"Hach, die Welt ist so schön!"
<i>jaulen</i>	"Mami, wann sind wir endlich da?...wie lange dauert es denn noch?"
<i>jm etw verraten</i>	"Oliver will eine CD für Weihnachten, aber normalerweise, müsstest du das nicht wissen. OK?"
<i>jodeln</i>	"Hollarüdüdio"
<i>johlen</i>	" Oleeeoleole..."
<i>jubeln</i>	"Ja, ich habe die Prüfung doch bestanden!"
<i>jubilieren</i>	"Juhuu! Es klappt!"
<i>klagen</i>	„Das Wasser steht mir bis zum Hals.“
<i>klarmachen</i>	„Jeder wird dir doch bestätigen, dass Bier nach Wein das absolute Ende ist.“
<i>klatschen</i>	"Weisst du, warum die ihren Freund verlassen hat?"
<i>konjugieren</i>	"ich gehe, du gehst, er/sie/es geht..."
<i>korrigieren</i>	"Es heißt <i>der Hund</i> und nicht <i>das Hund</i> ."
<i>kritisieren</i>	"Dein Aufsatz ist etwas unzusammenhängend."
<i>(herum)kritisieren</i>	"Du hast wieder alles falsch gemacht!"

<i>labern</i>	"Die alkoholische Gärung ist die Gärung des Alkohols, welcher alkoholisch ist und daher als Produkt der alkoholischen Gärung Alkohol produziert..."
<i>lästern</i>	"Der hat vielleicht einen hässlichen Pullover an!"
<i>lehren</i>	"Merke dir: Wer andern eine Grube gräbt, fällt selbst hinein."
<i>loben</i>	"Das hast du gut gemacht."
<i>lügen</i>	"I did not have any sexual relationship with that person Monica Lewinsky."
<i>mahnen</i>	"Sie müssen Ihre Rechnung bezahlen"
<i>moralisieren</i>	"Ihr müsst immer verantwortlich und selbstbewusst sein"
<i>nachbestellen</i>	"Hiermit möchte ich die fehlenden Ausgaben des Time-Magazins(Jahrgang 1999) nachbestellen."
<i>nachfragen</i>	"Können Sie mir bitte zeigen, wo man dieses verfl... Ding einschaltet?"
<i>nachhaken</i>	"Bitte erkläre das doch mal genauer."
<i>nachsagen</i>	"Er ist wirklich ein höchst krimineller Mensch."
<i>nennen</i>	"Ich nenne Dich jetzt einfach Ding-ohne-Namen."
<i>passen</i>	"Ich gebe auf."
<i>patentieren</i>	"Laut § xyz gehören Ihnen sämtliche Rechte an dieser Erfindung."
<i>petzen</i>	"Der war' s!"
<i>plädieren</i>	"Die Verteidigung fordert Freispruch für den Angeklagten"
<i>plaudern</i>	"Schönes Wetter heute, nicht?"
<i>prahlen</i>	"Mein Haus, mein Auto, meine Yacht."
<i>predigen</i>	"...denn keiner von uns ist ohne Sünde..."
<i>proklamieren</i>	"Die neue Verfassung hat von heute an Gültigkeit."
<i>prophezeien</i>	"Y2K wird sämtliche Computer lahmlegen"
<i>quengeln</i>	"Ich will aber die neue "Action-Man"-Figur haben, Mami!!!"
<i>radebrechen</i>	"Ich ... Ähh ... Ich wollte doch nur ... also ... ich ..."
<i>raten</i>	"Schauen sie sich auf jeden Fall das neue Gebäude von Heidelberg Druck an, wenn sie am Bahnhof vorbeikommen."
<i>raten</i>	"Ich denke, das Streichholz ist in deiner linken Hand"
<i>rechten</i>	...A: "Die Ästhetik des 3. Reiches war die konsequente Weiterführung der Ideen der Bauhaus Künstler" B: "Ach was, das war doch eine totale Pervertierung von deren Gedanken."

<i>reden</i>	„Meine Damen und Herren, es ist mir eine Freude“
<i>reimen</i>	Ja ich weiß, woher ich stamme, Ungesättigt gleich der Flamme, Glühe und verzehr ich mich
<i>reklamieren</i>	"Herr Ober, das Essen ist kalt."
<i>reservieren</i>	"Könnten sie einen Tisch für 5 Personen auf halb acht reservieren?"
<i>rufen</i>	"Riiicola"
<i>schildern</i>	"Ich schildere Ihnen nun den exakten Ablauf des Geschehens.."
<i>schwindeln</i>	„Ich habe schon als Baby sofort aus der Tasse getrunken.“
<i>schwören</i>	"Ich schwöre bei meiner Großmutter: Ich werde nie wieder Nachbars Katze in unserer Waschmaschine schleudern!"
<i>sich geloben</i>	"Ich trinke nie in meinem Leben Alkohol!"
<i>sich verabreden</i>	"-Treffen wir uns um 2 Uhr im Cafe? -Ja, OK.
<i>sich versprechen</i>	"Der Gewinner des Rennens ist Schuhmacher...eh...Haekinnen."
<i>suspendieren</i>	"Packen Sie Ihre Sachen und verschwinden Sie auf der Stelle! Hiermit sind Sie für die nächsten die nächsten fünf Wochen suspendiert."
<i>tadeln</i>	„Du sollst nicht immer die Kleidung sauber in die Wäsche legen!“
<i>taufen</i>	„Ich taufe Dich auf den Namen Hans.“
<i>verbieten</i>	"Du darfst hier nicht rauchen!"
<i>verdeutlichen</i>	"Ich werde versuchen, das dir ein bisschen besser zu erklären : eine Lautmalerei ist wie "Pluf!", "Doing!" oder"Ring!"."
<i>vereinbaren</i>	"Dann können wir uns heute Abend im Kinosehen."
<i>verlangen</i>	"Gib' mir meine 2 Mark zurück!"
<i>verlautbaren</i>	"Der Minister werde erst in einer Woche seine Arbeit wieder aufnehmen."
<i>vermerken</i>	"Bitte die folgende Adresse ändern!"
<i>voraussagen</i>	„Morgen wird es regnen.“
<i>weitererzählen</i>	"Du musst diese Information mit ins Grab nehmen."
<i>wetten</i>	"Falls das passiert, bekommst du von mir eine Flasche Wein."
<i>widerrufen</i>	"Hiermit erkläre ich unseren Beschluss für ungültig."
<i>widersprechen</i>	"Hiermit erheben wir Einspruch gegen Ihre Entscheidung."
<i>widerstreiten</i>	"Da bin ich aber einer ganz anderen Meinung!"
<i>wispern</i>	"Sprechen Sie bitte ganz leise, damit uns niemand hört."
<i>zählen</i>	Eins, zwei, drei..."

<i>zanken</i>	"So kann es nicht weitergehen!"
<i>zensieren</i>	"Sie dürfen dieses Bild nicht in unserer Werbung verwenden!"
<i>zerreden</i>	(Sehr lange Kombination von Sprechakten)
<i>zitieren</i>	"Wie der große Goethe gesagt hat, ..."
<i>zusagen</i>	"Gut, ich nehme die Einladung an und komme heute Abend."

Schema, Konvention, Regel

- Für die Ausführung von Handlungen gibt es **Schemata, Muster**
eine Schleife binden, Kravatte binden, Gabel benutzen, Stäbchen benutzen ...
Erst muss man das Muster lernen, dann wird es **internalisiert**.
- Schemata von Interaktionen sind oft **Konvention**
zur Begrüßung die Hand schütteln, Grüß Gott sagen , Nase aneinander reiben
Man könnte es auch anders machen, aber es ist nun mal so Sitte:
Eine Konvention gilt in einer sozialen Gruppe. Auch manche nicht sprachliche Handlungen sind Konvention
im Straßenverkehr rechts fahren
- Konventionelle Muster heißen **Regel**
Regelbefolgung, Regelabweichung – mit je sozialer Wirkung
Sprachregeln, wir haben schon viele kennengelernt (S 1.1 – S 7.3)

Sprechakttheorie – die Lehre von mündlichen und schriftlichen Sprachhandlungen

Verschiedene Sprechakte:

Fragesatz: "Gehst Du in die Stadt?"

Imperativ: „Geh in die Stadt!“

Aussage: „Du gehst in die Stadt“

„Entweder du gehst in die Stadt oder du gehst nicht in die Stadt“

Aber auch sehr unterschiedliche Sprechakte mit demselben Satz möglich

„Du hast in München zuletzt getankt.“

je nach Situation: Behauptung, Warnung, Erklärung, Vorwurf ... (s.u.)

Und derselbe Sprechakt ist mit sehr unterschiedlichen Sätzen realisierbar:

„Monika, mach das Fenster zu!“

„Monika, würdest du bitte mal das Fenster zumachen?“

„Monika, du könntest mal das Fenster zumachen.“

„Monika, warum machst du nicht das Fenster zu?“

„Monika, das Fenster steht auf!“

„Monika, es zieht!“

„Monika, merkst du denn gar nicht, dass es zieht?“

(Aus Funkkolleg Sprache)

Wörtliche Bedeutung wäre bei manchen Beispielen falsch (wenn auch absichtliches Missverständnis ein beliebter Trick ist):

„Monika, warum machst du nicht das Fenster zu?“	<i>weil es mir nicht kalt ist</i>
„Monika, das Fenster steht auf!“	<i>fein</i>
„Monika, es zieht!“	<i>oh je, das ist aber ungesund</i>
„Monika, merkst du denn gar nicht, dass es zieht?“	<i>doch, ist mir auch schon aufgefallen</i>

- Also die Bedeutung eines Satzes ergibt sich nicht allein aus der Form, sondern erst aus dem Handlungszusammenhang. Deshalb kann man über Sätze in Isolation nichts Endgültiges sagen.

Systematische Untersuchung dieser Phänomene in der sog. Sprechakttheorie

Austin, Searle, sprachanalytische Philosophie

Zerlegung, Analyse, Komponenten, Teilaktivitäten

2. Sprechaktanalyse

Jeder Sprechakt resultiert aus dem gleichzeitigen Befolgen mehrerer zusammenwirkender Regeln:

<u>Lokution</u>	das geordnete Hervorbringen von Ausdrücken; die äußere Form; die Äußerung
<u>Propositor</u>	Bezugnahme auf Dinge und Sachverhalte; die wörtliche Bedeutung, ohne den Satzwert
<u>Illokution</u>	wie der Sprechakt als Interaktion gemeint ist; was für eine Interaktion vorliegt
<u>Perlokution</u>	der Zweck, der mit dem Sprechakt erreicht werden soll; das Ziel des Sprechers, die Intention des Sprechers

Beispiel

Beifahrer sagt zum Fahrer: *Du hast in München zuletzt getankt*

Lokution	"Du hast in München zuletzt getankt" * "Du getankt in hast München zuletzt" (wäre inadäquate Lokution)
Propositor	daß der Angesprochene in München getankt hat (Referenz, Prädikation)
Illokution	Warnung / Aufforderung zum Tanken / Erklärung für Liegenbleiben (je nach Situation) / Vorwurf
Perlokution	z.B. bei Illokution Warnung: der Beifahrer möchte nicht mit dem Auto auf der Straße liegen bleiben

In Wirklichkeit ist der Ablauf umgekehrt von der Perlokution zur Lokution.

Perlokution: S ist auf H nicht gut zu sprechen und will ihn verletzen aus aktueller Wut oder aus weiter zurückliegenden Gründen

Illokution: S sucht einen passenden Sprechakt aus, z.B. eine Beleidigung.

Propositon: S sucht den passenden Inhalt aus, z.B. eine ehrenrührige Behauptung, dass H typischerweise das Tanken vergisst.

Lokution: S verbalisiert diesen Inhalt

Das sieht Dir mal wieder ähnlich! Du merkst wohl gar nicht, das wir mal wieder tanken müssen!

Glücken und Gelingen von Sprechakten

Glücken Hörer fasst einen Sprechakt als diejenige Illokution auf, die der Sprechakt wirklich hat

Gelingen Sprecher erreicht seine Perlokution

Weiteres Beispiel:

(Aus Maas/Wunderlich: Pragmatik und sprachliches Handeln. 1972)

<p>1. ER: <i>Die Kleine hat in die Windeln gemacht.</i></p>	<p>Proposition, daß die Kleine in die Hose gemacht hat.</p> <p>Illokution: Mitteilung</p> <p>Perlokution?</p> <p>Gewisse Folgen aus der Sachlage: es muss was getan werden. Möglich:</p> <p style="padding-left: 40px;">a) Kind soll trocken liegen</p> <p style="padding-left: 40px;">b) er möchte aber nichts dafür tun,</p> <p>also möglicherweise Illokution indirekt eine Aufforderung.</p>
--	--

<p>2. SIE: -</p>	<p>Kein Sprechakt, keine Reaktion.</p> <p>Wenn 1 eine Aufforderung war, wäre eine Reaktion nötig. Bewusste Nicht-Reaktion = indirekte Ablehnung der Aufforderung.</p> <p>Möglich weil 1 nicht deutlich die Illokution zum Ausdruck brachte.</p> <p>Perlokution? a) er soll deutlich machen, dass er Kleine nicht trocken legen will b) sie will es nicht tun weil sie es ungern macht, weil er auch was machen soll</p>
<p>3. ER: <i>Ich würde ihr eine neue Windel geben</i></p>	<p>Folgerung aus 1: was wäre zu tun? Vgl. "<i>Sie müsste eine neue Windel haben</i>"</p> <p>Form aber als Ratschlag eines Nichtbetroffenen. "<i>ich würde an Deiner Stelle</i>" Illokution noch nicht explizit, aber deutlich, dass sie es tun soll.</p>
<p>4. SIE: <i>Das freut mich.</i></p>	<p>Sie übergeht "<i>an deiner Stelle</i>" von 3</p> <p>Perlokution: Abweisung der Aufforderung durch absichtliches Nicht-Verstehen</p> <p>Illokution: Provokation</p>
<p>5. ER: <i>Kannst DU ihr nicht die Windel geben?</i></p>	<p>Proposition deutlich auf sie bezogen.</p> <p>Illokution äußerlich eine Frage – in Wirklichkeit eine Aufforderung (nicht + Frage = Bitte)</p> <p>Perlokution: Tut so, als läge Missverständnis vor, geht also nicht auf Provokation ein.</p>
<p>6. SIE: <i>Du kannst es doch auch.</i></p>	<p>Proposition über das Können, was eine Voraussetzung für eine Aufforderung ist.</p> <p>Perlokution: Gegenangriff, Abweisung der Aufforderung durch Behauptung von Alternativen, Beginn einer Argumentation Illokution implizit Aufforderung an ihn.</p>
<p>7. ER: <i>Ich muss gleich weg. Kannst DU ihr nicht eine geben?</i></p>	<p>1. Illokution: direkt Begründung, warum er nicht infrage kommt, indirekt: Ausflucht, Abblocken der Diskussion</p> <p>2. Form Imperativ, explizite Illokution: Befehl.</p>

Fazit: ER wird gezwungen immer expliziter die **Illokution** zu markieren, den Befehlscharakter anzuzeigen. SIE setzt verschiedene Taktiken ein: Missverstehen, Alternativen diskutieren, vermeidet ebenfalls direkte Konfrontation.

Reden über Sprechakte

(auch als heuristischer Test einsetzbar)

1. Lokution, Äußerung	Was sagte er/sie? Er sagte: „Die Kleine hat in die Windeln gemacht.“
2. Proposition	Worüber sprach er/sie? Er sprach über den Zustand der Windel der Kleinen
3. Illokution	Was TAT er/sie ?
primär	Er forderte seine Frau dazu auf , die Windel der Kleinen zu wechseln
sekundär	Er informierte seine Frau darüber, dass die Kleine in die Windel gemacht habe
4. Perlokution, Intention	Was wollte er/sie erreichen? Er versuchte (vergeblich) zu erreichen , dass seine Frau die Windel wechseln
5. Glücken	Hat der Hörer verstanden?
6. Gelingen	Hat der Sprecher erreicht, was er wollte?

Schema:

[] Leerstelle in die etwas eingesetzt wird.

Sprecher erreichte/versuchte zu erreichen, dass [Perlokution], dadurch dass er Hörer [primäre Illokution], indem er Hörer [sekundäre Illokution], dass [Proposition], mit den Worten [Lokution]

Schema ausgefüllt:

ER versuchte zu erreichen, dass [SIE die Windel wechselte], dadurch dass er sie dazu [aufforderte, die Windel zu wechseln], indem er sie [informierte], [dass die Kleine in die Windel voll hatte], mit den Worten [„Die Kleine hat in die Windeln gemacht].“

Noch ein Beispiel, an dem man das Analysieren von Sprechakten üben kann:

(Hausmeister H und Schüler S auf dem Schulhof)

H: Würdest Du bitte die Papiertüte dort aufheben?

S: Ich habe sie nicht hingeworfen.

H: Das ist egal. Heb sofort die Tüte auf!

S: Sie haben mir nichts zu befehlen.

H: Wenn Du nicht sofort die Tüte aufhebst, bringe ich Dich zum Direktor

S: Tun Sie's doch!

H: (Haut dem Schüler eine runter.)

S: Das laß ich mir nicht gefallen.

(H. und S. gehen zum Direktor.)

Performative

Die Illokution ist meist implizit. Sie ergibt sich in der Situation.

Man kann die Illokution aber auch explizit machen. Dies geschieht durch

"performatorischer Ausdrücke"

ich verspreche dir, ich warne dich, ich bitte dich, bitte, hiermit fordere ich Sie auf, ...
auch Fragesatzstellung, Imperativ, Konjunktiv ...

Sprechaktsequenzen

Eine Folge interaktiver Sprechakte kann man als **Sprachspiele (games)** betrachten.

Ferdinand De Saussure, Ludwig Wittgenstein : Sprache ist wie ein Spiel, z.B. ein Schachspiel

eine fruchtbare Analogie:

- Ein Spiel wird definiert durch konventionelle Regeln, ebenso die Sprache
- Züge: wer zieht gerade, was kann er tun, was der andere?
- Gewinn, Verlust, Glücken, Gelingen, Ziel erreichen, verfehlen ...

Dies führt zu einer pragmatischen Semantiktheorie:

- Die Bedeutung der Sprachelemente ist ihr möglicher Gebrauch im Sprachspiel.

3. Pragmatische Bedingungen

➤ soziale und faktische Bedingungen, die erfüllt sein müssen, damit der Sprechakt glückt

Beispiel Pragmatische Bedingungen für das Versprechen (nach John R. Searle)

Wenn ein Sprecher S einen Satz T in Gegenwart eines Hörers H äußert, dann verspricht S in einer aufrichtigen (und vollständigen) Äußerung von T dem Hörer H, daß p, wenn und nur wenn folgende Bedingungen erfüllt sind:

- (1) Normale Eingabe- und Ausgabebedingungen
- (2) In der Äußerung von T drückt S aus, dass p.
- (3) Indem S ausdrückt, dass p, sagt S einen künftigen Akt A von S voraus.
- (4) H sieht lieber die Ausführung als die Unterlassung von A durch S, und S glaubt, daß H lieber seine Ausführung als seine Unterlassung von A sieht.
- (5) Für S und auch für H ist es nicht offensichtlich, dass S bei einem normalen Verlauf der Ereignisse A tun wird.
- (6) S beabsichtigt, A zu tun.
- (7) S hat die Absicht, sich zur Ausführung von A zu verpflichten, wenn er T äußert.

Die pragmatischen Bedingungen sind für jeden Sprechakt anders und müssten für jede Illokution (siehe oben die vielen Verben!) erforscht werden.

Dabei sollte man wieder strukturalistisch vorgehen. Es muss wenigstens eine andere Bedingung geben, durch die sich zwei Sprechakte unterscheiden, wenn sie nicht dieselbe Illokution haben.

Streit über Sprechakte und ihre pragmatischen Bedingungen

Wer darf was sagen? Political Corectness.

Wer muss sich bei wem für was entschuldigen?

usw.

Man kann auch andere symbolische Handlungen mit dem Instrumentarium der Sprechakttheorie analysieren.

Was ist z.B. die Lokution, Proposition, Illokution, Perlokution dieser Denkmäler?



S 8.2: Grundlagen der Textkonstitution und des Textverstehens

Lernziel: Eine klare Vorstellung von Kohäsion und Kohärenz von Texten

1. Textkonstitution

Was unterscheidet einen Text von einer willkürlichen Ansammlung von Sätzen?

TEXTPUZZLE

1. Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch.	4. Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.
2. Die CDU/CSU lehnte die Vorlage ab.	5. Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen.
3. Die Novelle zielt auf eine Verkürzung des Rechtsweges ab.	

Was ist die richtige Reihenfolge der Sätze in diesem Beispiel ?

ORIGINALTEXT

Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen.

Die Novelle zielt auf eine Verkürzung des Rechtsweges ab.

Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch.

Die CDU/CSU lehnte die Vorlage ab.

Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.

Nicht nur die Reihenfolge der Sätze spielt eine Rolle, auch die Wortstellung in den Sätzen.

WORTFOLGE IM SATZ GEÄNDERT

Der Bundestag hat gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber unter dem Druck der erneut gestiegenen Zahl der Asylbewerber beschlossen.

Auf eine Verkürzung des Rechtsweges zielt die Novelle ab.

Noch bis zu acht Jahren nimmt das Anerkennungsverfahren bisher in Anspruch.

Die Vorlage lehnte die CDU/CSU ab.

Alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern schöpfe das neue Gesetz noch nicht aus.

THESE: Es gibt immer gleichzeitig zwei Arten von Zusammenhang in einem Text:

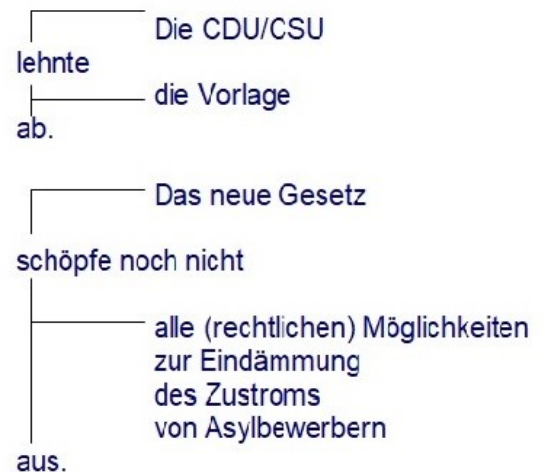
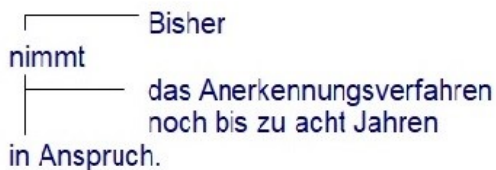
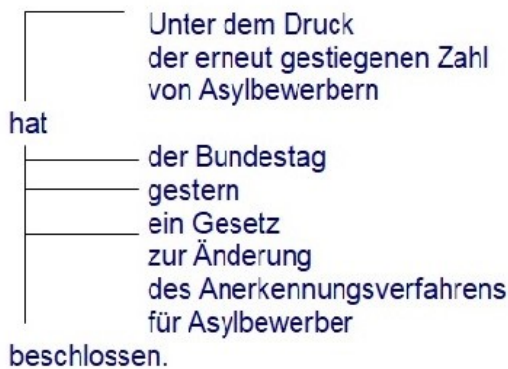
die "**Kohäsion**" und die "**Kohärenz**".

Die Textstruktur hat also zwei Dimensionen.

Kohäsion

Die Kohäsion wird durch die verwendeten sprachlichen Mittel erreicht; vor allem

- durch Prädikate und Ergänzungen in der **Syntax** (siehe S 5)



- durch Referenz, Koreferenz und prototypischen Zusammenhängen (Kontiguität, Bridging) im **Wortschatz** (siehe S 7)

Der Bundestag beschließt Gesetze.

Gesetze regeln den Rechtsweg zur Durchsetzung von Ansprüchen.

Das Anerkennungsverfahren für Asylbewerber ist der Rechtsweg zur Durchsetzung der Anerkennung als Asylantrag.

Der Bundestag besteht aus Vertretern der Parteien.

Die CDU/CSU ist eine Partei im Bundestag.

Der Zustrom von Asylbewerbern führt zu einer steigenden Zahl von Asylbewerbern.

Druck durch Zustrom verlangt Eindämmung des Zustroms.

Eine Verkürzung ist eine Änderung.

Nach einer Änderung ist etwas anders als bisher.

Kohärenz

Kohärenz ist gegeben, wenn die Aufeinanderfolge von Sprechakten pragmatisch einen Sinn macht.

➤ Kohärenz ist also eine Eigenschaft der Sprechaktsequentialisierung.

Dialoge mit alternativen, doch jeweils kohärenten Reaktionen:

Inspektor: *Sie sagen nicht die Wahrheit, Madam!*

Mrs. B. *Ich muß doch sehr bitten!*

Doch. Ich kann es beweisen.

Frage: *Wann haben Sie sich verletzt?*

Antwort: *Am 11. September*

Ich weiß es nicht mehr.

Das habe ich Ihnen doch schon gesagt

2. Sonderfall monologische Texte

Kohäsion allein reicht nicht

Am Zaun steht ein Fahrrad. Das Fahrrad gehört meinem Nachbarn. Der Nachbar arbeitet bei einer Gärtnerei. Die Gärtnerei baut Gemüse an. Gemüse ist gesund. Gesunde Ernährung verhindert Krebs. Krebs ist eine Volkskrankheit ...

Der obige Text ist sehr kohäsiv. **Aber er "macht irgendwie keinen Sinn".**

Es fehlt also Kohärenz. Aber was soll das sein, wo es sich bei einem Text doch um einen einzigen Sprechakt zu handeln scheint? Allenfalls um eine Aneinanderreihung von Sprechakten des gleichen Typs: **lauter Aussagen.**

Wir nehmen an, dass jeder Aussage eine Frage zugrundeliegt. Wenn der Sprecher nicht den Eindruck hätte, dass dem Hörer etwas fraglich ist (oder fraglich sein sollte), würde er die Aussage nicht machen. Das ist der Schlüssel zur Kohärenz monologischer Texte.

- **Für jeden Aussagesatz muss es eine im Kontext der vorausgegangenen Sätze eine sinnvolle Fragestellung geben, auf die der Satz eine Antwort gibt. Die kann durch eine explizite Frage realisiert sein. oder aber nur als implizite Voraussetzung existieren.**

Zentrale Analysemethode zur Explizitmachung kohärenzschaffender Fragen:

Versuche zu jedem Satz eine Frage $F_i \rightarrow j$ zu formulieren,

1. auf die der Satz j eine Antwort ist und
2. die sich aus einem der vorangehenden Sätze i sinnvoll ableiten läßt

(F 0 → 1)	Was gibt es Neues? Was wird gegen die steigende Zahl von Asylbewerbern unternommen?
1	Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen.
(F 1 → 2)	Was enthält dieses Gesetz? Inwiefern trägt es zur Lösung des Problems der Asylbewerber bei?
2	Die Novelle zielt auf eine Verkürzung des Rechtsweges ab.
(F 2 → 3)	Wieso wäre eine Verkürzung des Rechtsweges ein Fortschritt? Wie lange dauert das Verfahren denn bisher?
3	Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch.
(F 1 → 4)	Auf welche Weise kam das Gesetz zustande? Wie verhielt sich die Oppositionsfraktion CDU/CSU?
4	Die CDU/CSU lehnte die Vorlage ab.
(F 4 → 5)	Warum lehnte die CDU/CSU die Vorlage ab?
5	Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.

Definitionen

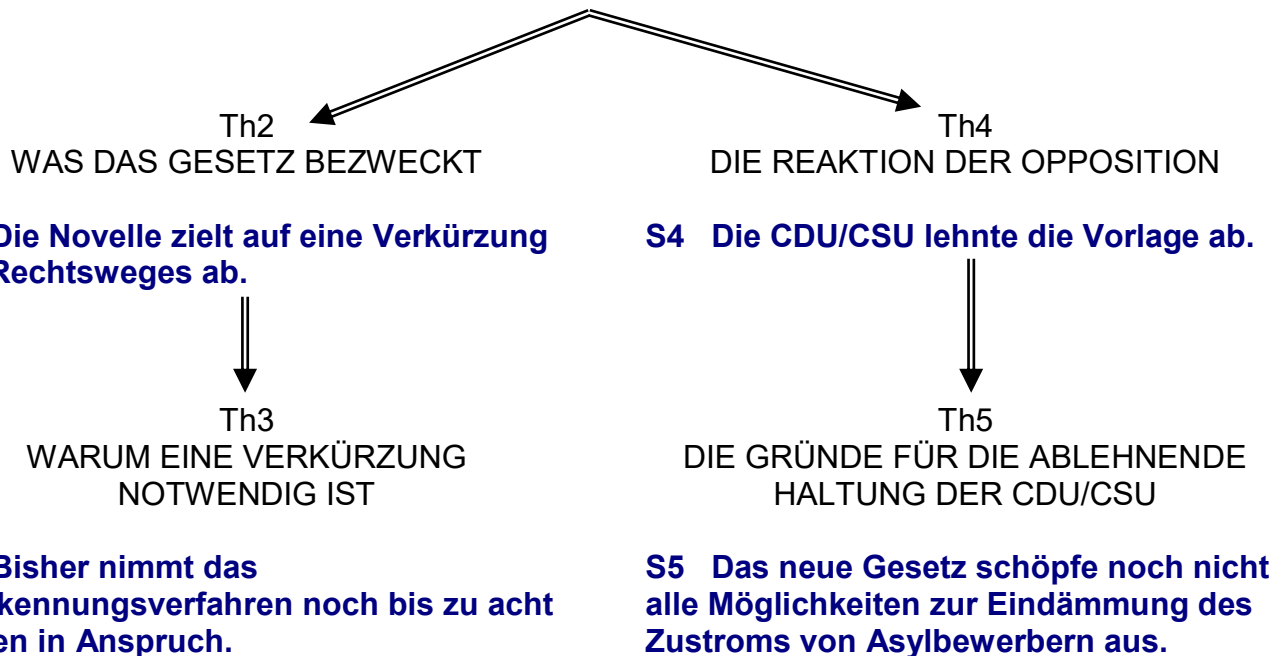
Das Fragliche, das durch eine satzverbindende Frage $F_i \rightarrow j$ explizit gemacht. ist **das Thema** des folgenden Satzes. Der Satz ist **die Exposition** zum Thema.

Themen stehen oft als Überschriften im Text.

KOHÄRENZBAUM (mit Themen Th und Expositionen S)

Th1
NEUES AUS DEM BUNDESTAG
WIE DIE ZAHL DER ASYLBEWERBER EINGEDÄMMT
WERDEN SOLL

S1 Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen.

**Progression**

- Zur Formalisierung der Kohäsion ist das Modell einer Hierarchie von Konstituenten unbrauchbar. Eine Textgrammatik muß auf den Operationen des Anknüpfens und Fortschreitens basieren. Die Grundlage der ersten Dimension der Textstruktur ist **die "referentielle Progression"**.
- Die Kohärenzstruktur ist nicht als Teil-Ganzes-Hierarchie darstellbar, da sich nachgeordnete Themen nicht ohne Rückgriff auf vorangehende Expositionen einordnen lassen. Auch hier ist ein dynamisches Modell adäquater: Die Basis der zweiten Dimension der Textstruktur ist **die "thematische Progression"**.

Wie oben beobachtet, hängt die Wortfolge im Satz von seiner Rolle im Text ab.

These: Der Satz ist die Schnittstelle zwischen den beiden Progressionen des Textes. Die aktuelle Satzstruktur ist das Ergebnis einer zweifachen Kontextualisierung.

Die Wortfolge in Sätzen monologischer Texte scheint ebenso von der impliziten Frage abzuhängen, wie die Wortfolge der Antwort von der expliziten Frage in Dialogen.

Offene Frage - als Antwort dient eine ganze Aussage, mit neuem Prädikat und neuen Referenzobjekten.

Was macht Ihr im Urlaub?

Wir machen eine Kreuzfahrt

Wir renovieren unseren Wintergarten.

Geschlossene Frage - die Frage gibt bereits einen propositionalen Rahmen vor und richtet den Fokus auf einen Aspekt.

die Antwort füllt nur noch die offenen Stellen darin aus.

Wohin fahrt ihr im Urlaub?

Wir fahren nach Mallorca

Wortstellungstests in Frage und Antwort (in rot = holprig bis inakzeptabel)

Wie oft geht ihr ins Kino?

Ins Kino gehen wir jeden Tag.

** Wir gehen jeden Tag ins Kino*

Was macht ihr im Urlaub?

Wir gehen jeden Tag ins Kino

**Ins Kino gehen wir jeden Tag.*

Wo sind die Bücher?

Die Bücher sind in der Kiste.

**Die Kiste enthält lauter Bücher*

Was ist in der Kiste?

Die Kiste enthält lauter Bücher

**Die Bücher sind in der Kiste.*

Einteilung des Satzes (Thema - Rhema)

Thema: Wie oft geht ihr ins Kino?

Supposition: Ins Kino gehen wir ...

Rhema: jeden Tag.

Thema: Was ist in der Kiste

Supposition: In der Kiste sind ...

Rhema: lauter Bücher.

Normal steht die Supposition am Anfang des Satzes und das Rhema am Ende.

Genau genommen ist dies aber eine Sache der **Intonation**. Das Rhema muss den letzten fallenden Ton bekommen. Thema, bzw. Supposition haben steigenden Ton..

Weitere Beispiele:

Thema, thematisch: am Satzanfang, Tonhöhenbewegung nach oben

Rhema, rhematisch: Tonhöhenbewegung nach unten, letzter Tonbruch im Satz

Supposition: (phrasierende Tonhöhenbewegungen, steigend und fallend, nicht letzter Tonbruch)

Thematische Satzanfänge

Den Fernsehzuschauern versuchte Reagan weiszumachen, *schuld daran sei die Sturheit der Demokraten.*

Militärregimes haben wir unter dem *Grundtyp Autoritarismus eingeordnet.* Wir unterscheiden traditionelle und neue Militärregimes. *Den Fall Kuba* können wir mit diesen Haupttypen *nicht erfassen.*

Rhematische Satzanfänge

Keinen Gedenktag, sondern einen Park forderten die Fans von John Lennon.

Noch ohne konkrete Ergebnisse blieb gestern die fieberhafte Fahndung nach dem Einbrecher, der Mittwoch früh einen 33-jährigen türkischen Gastarbeiter auf brutale Weise niedergeschossen und lebensgefährlich verletzt hat.

4. Textverarbeitung

Zu einem Text oder aus einem Text wird ein anderer gemacht..

Dies zeigt Beziehungen zwischen Texten auf.

Unterscheidung zwischen

Objekttext = ein Text, der sich auf die Objekte in der Welt bezieht.

Metatext = ein Text, der sich auf einen anderen Text bezieht (Rezension, Abstract, Titel)

Volltext

Kondensat = verkürzter Text

Texte, die aus anderen Texten (maschinell?) hergestellt werden::

Zusammenfassung = Objekttext kondensiert	Abstract = Metatext kondensiert
Schlagzeile = stark kondensierter Objekttext	Titel = stark kondensierter Metatext

S 8.3 Wissensrepräsentationen

Lernziel: Übersicht über Darstellungsformen von Wissen verschiedener Art und eine nüchterne Beurteilung der Reichweite von wissensbasierten Systemen.

1. Grundlagen

Zum Verständnis eines Satzes ist viel **Hintergrundwissen** nötig.

- Die Bauern verkauften ihre Kühe, denn **sie** gaben keine Milch mehr

- Die Bauern verkauften ihre Kühe, denn **sie** brauchten Geld

Worauf bezieht sich "sie"? Wieso weiß man das?

Wissensbasierte Systeme

Computerprogramme mit Hintergrundwissen über den Anwendungsbereich. Sie verfügen über eine sog. Wissensbasis.

Einsatz von Wissen zum Zwecke der Problemlösung

Wissensbasierte Systeme haben neben der Wissensbasis eine Problemlösungskomponente.

Natural language processing systems (NLP) = sprachverarbeitenden Systeme mit

Wissensbasis: und Problemlösungskomponente:

maschinelle Übersetzung

automatische Fragebeantwortung,

Mensch-Maschine-Kommunikation ...

u.a.

Arten von Wissen

Wissen über

- **Objekte**

In erster Linie versteht man unter Wissen die Kenntnis von Fakten über Objekte in der Welt, die uns umgibt:

"Vögel haben Flügel. Schwalben sind Vögel. Schnee ist weiß".

Wir müssen daher Objekte, Klassen von Objekten, Eigenschaften von Objekten, und Beziehungen zwischen Objekten repräsentieren können.

- **Ereignisse**

Zusätzlich zu Objekten gibt es Vorgänge und Ereignisse in der Welt.

"Der Ätna ist ausgebrochen. Das World Trade Center wurde zerstört."

Neben der Darstellung der Ereignisse selbst, muß der zeitliche Ablauf von Ereignissen und die zwischen ihnen bestehenden Ursache-Wirkungs-Beziehungen erfasst werden können.

- **Handlungswissen**

Eine spezielle Art von Wissen ist das, wie eine bestimmte Handlungen (vgl. S 8.1) auszuführen ist, z.B. wie man eine Schleife bindet. Auch Sprechakte verlangen Handlungswissen.

- **Metawissen**

Das ist Wissen über das Wissen, z.B. über die Verlässlichkeit oder Wichtigkeit bestimmter Information, oder das Bewußtsein über Objekt-, Ereignis- und Handlungswissen zu verfügen oder auch nicht.

Wissensrepräsentation

Wissensrepräsentation ist die symbolische Darstellung von Wissen über einen Gegenstandsbereich. Zur Wissensrepräsentation gehört

- **Die Syntax**, d.i. die Menge der Symbole, die verwendet werden können, sowie die mögliche Anordnung der Symbole
- **Die Semantik**, d.i. eine Interpretation (eine Bedeutung) der Symbole. Als Interpretation kann gelten, welche Operationen mit den Symbolen vorgenommen werden können. .Vgl. unsere Definition von "Rechnen" als Symbolmanipulation nach einer Vorschrift.

Drei Aufgaben für Systementwickler:

- Wissensrepräsentation (*knowledge representation*)
- Wissensakquisition (*knowledge acquisition*)
- Wissensverarbeitung (*knowledge processing*)

Trennung dieser drei Bereiche; die Wissensbasis soll nicht im Programm versteckt sein, sondern explizit repräsentiert sein. Die Wissensbasis soll so repräsentiert sein, dass auch andere Programme darauf zurückgreifen können.

2. Repräsentationsformate

Der wichtigste Unterschied in unserem Zusammenhang (vgl. S 7.3):

- **analoge Repräsentationen:**

Bilder, Graphen, Modelle

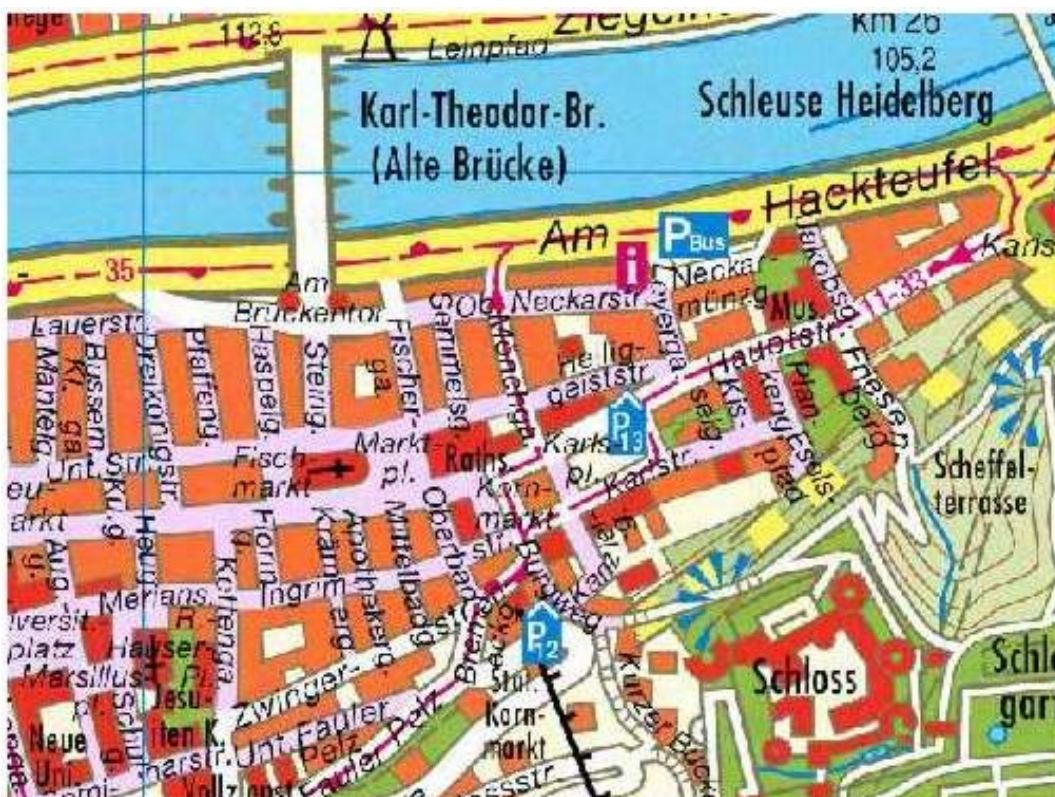
Ablezen der Zusammenhänge am Modell aufgrund von Homomorphie mit dem Objekt

- **symbolische (propositionale) Repräsentationen:**

natürliche Sprache, Kunstsprachen, Logikkalkül, mathematische Formeln

Erschließen der Zusammenhänge über Symbolmanipulation nach Regeln

analog



symbolisch propositional

Die Altstadt von Heidelberg liegt auf dem südlichen Ufer des Neckar. Am Südufer führt die Straße „Am Hackteufel“ entlang. Am Nordufer führt ein Leinpfad entlang. Der Neckar wird überspannt von der Karl-Theodor-Brücke (Alte Brücke). Die Brücke mündet auf der Heidelberger Seite im Brückenturm. In Verlängerung der Brücke, verbunden mit ihr durch die Steingasse, steht eine Kirche. Östlich an diese schließt sich der Marktplatz an und jenseits des Marktplatzes ist das Rathaus. Heidelberg wird im Südosten dominiert vom Schloss. (usw.)

Unterschiedliche Verrbeitung aber

- Repräsentationen sind teilweise ineinander übersetzbar, soweit sie dieselben Gegenstände und Sachverhalte abbilden.

Im Zusammenhang mit Computermodellierung unterscheide:

Deklarative Wissensrepräsentation

"Eine Nominalphrase besteht aus einem Artikel und einem Substantiv"

NP = Art + Subs

Prozedurale Wissensrepräsentation

"Verknüpfe ein Artikelwort mit einem Substantiv, dann hast Du eine Nominalphrase"

NP = CONCAT(Art, Subs)

Einige Formate

- **Graphen, Bäume, Netze**

bestehen aus **Knoten, Kanten, Beschriftungen**

ungerichtete Graphen = die Kanten zwischen den Knoten entsprechen einer symmetrischen Relation

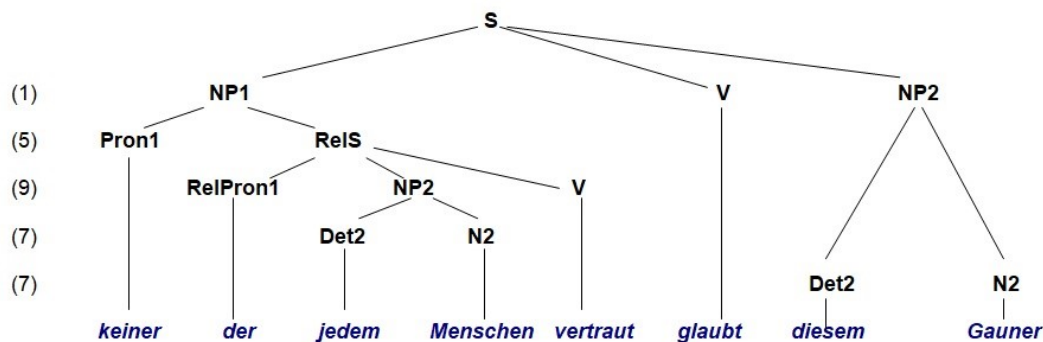
gerichtete Graphen = die Kanten zwischen den Knoten entsprechen einer asymmetrischen Relation

Baumgraphen

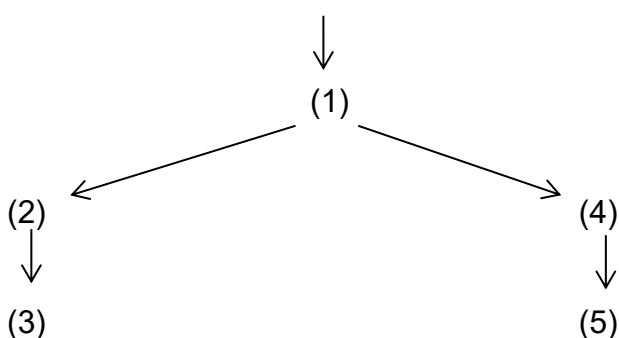
- gerichtet
- alle Knoten miteinander über Kanten verbunden
- genau ein Knoten, von dem (mittelbar) alle Kanten ausgehen
- keine Kante führt zum Knoten, von dem sie ausgeht
- zu jedem Knoten führt genau eine Kante

Verarbeitung: Durchsuchen vom obersten Knoten über die Kanten zu allen anderen Knoten (effiziente Algorithmen verfügbar)

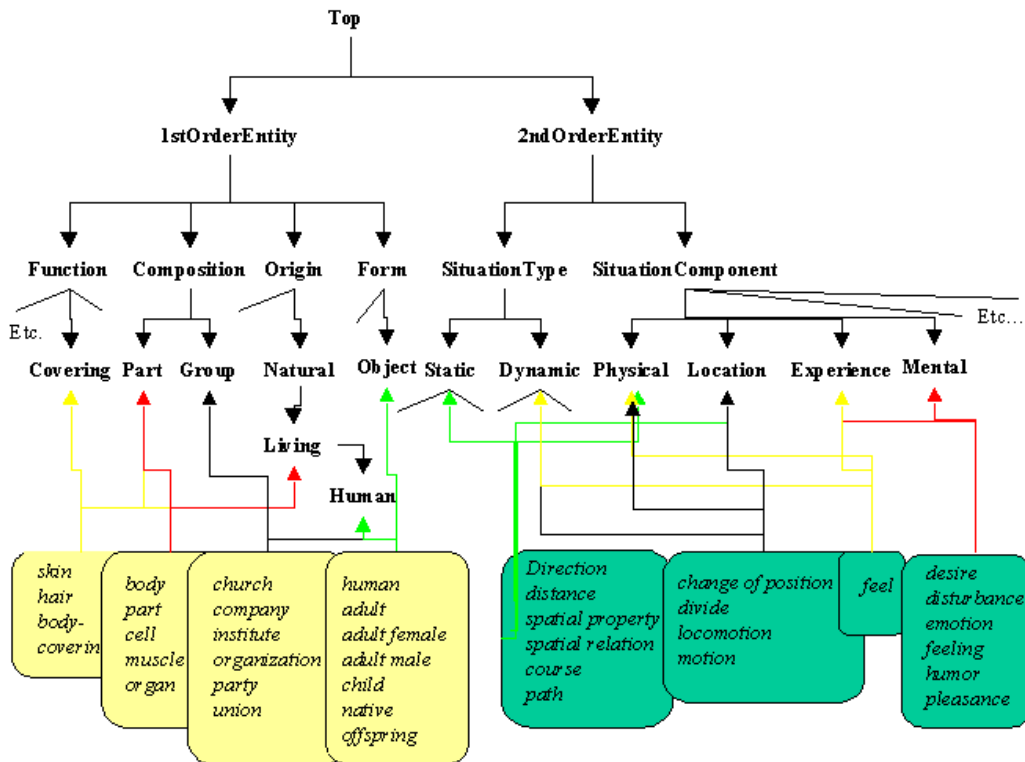
Verwendung von Baumgraphen in der Syntax (vgl. S5.2)



Kohärenzbaum in der Textlinguistik (vgl.S 8.2)



Baumgraphen in der lexikalischen Semantik für sog. „Ontologien“

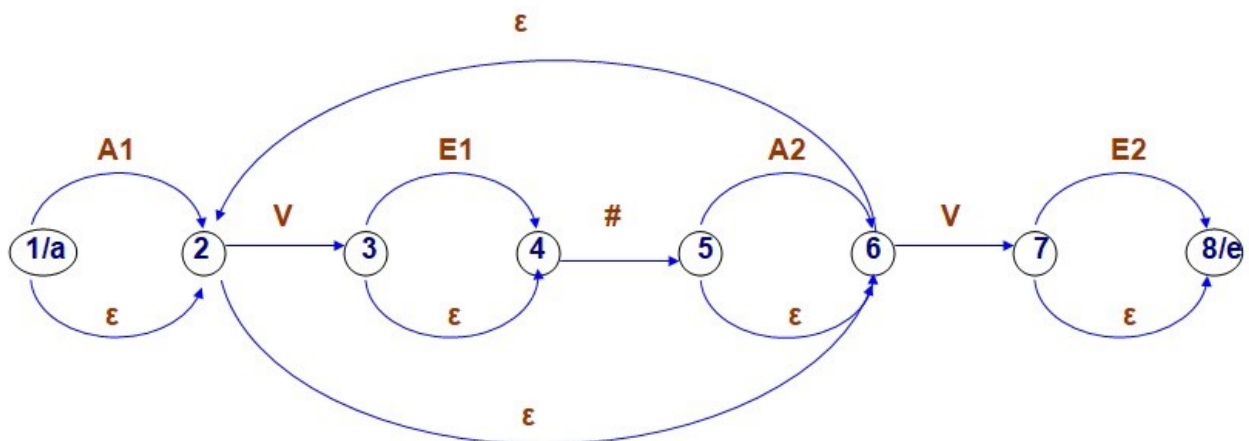


aus Eurowordnet

- **Netze**

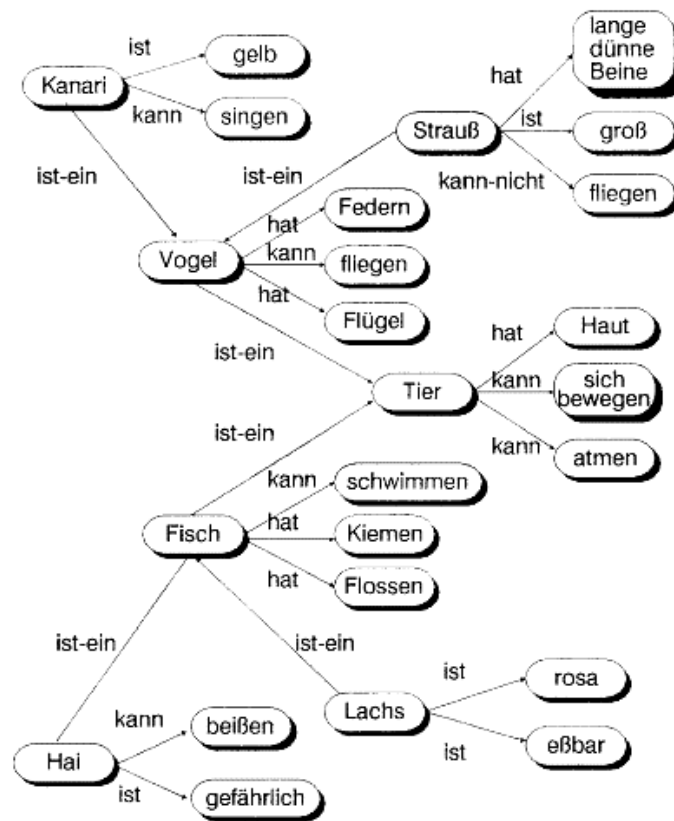
- normalerweise gerichtet
- alle Knoten miteinander über Kanten verbunden
- ein oder mehrere Knoten als Eingang, ein oder mehrere Knoten als Ausgang
- Kanten dürfen auch zu dem Knoten zurückführen, von dem sie ausgehen

Übergangnetzwerke in der Phonetik, Morphologie und Syntax (vgl.S 5.32)



/a Eingang, /e Ausgang, # bedeutet Silbengrenze

Netze in der lexikalische Semantik: Prototypenbeschreibung, Begriffsnetz



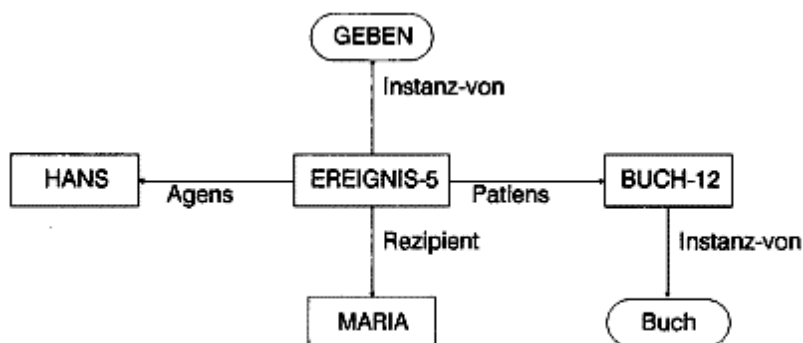
Exkurs: "Konzepte" - Begriffe

Konzeptklasse = mehrere Gegenstände fallen unter den Begriff (*Extension* > 1)

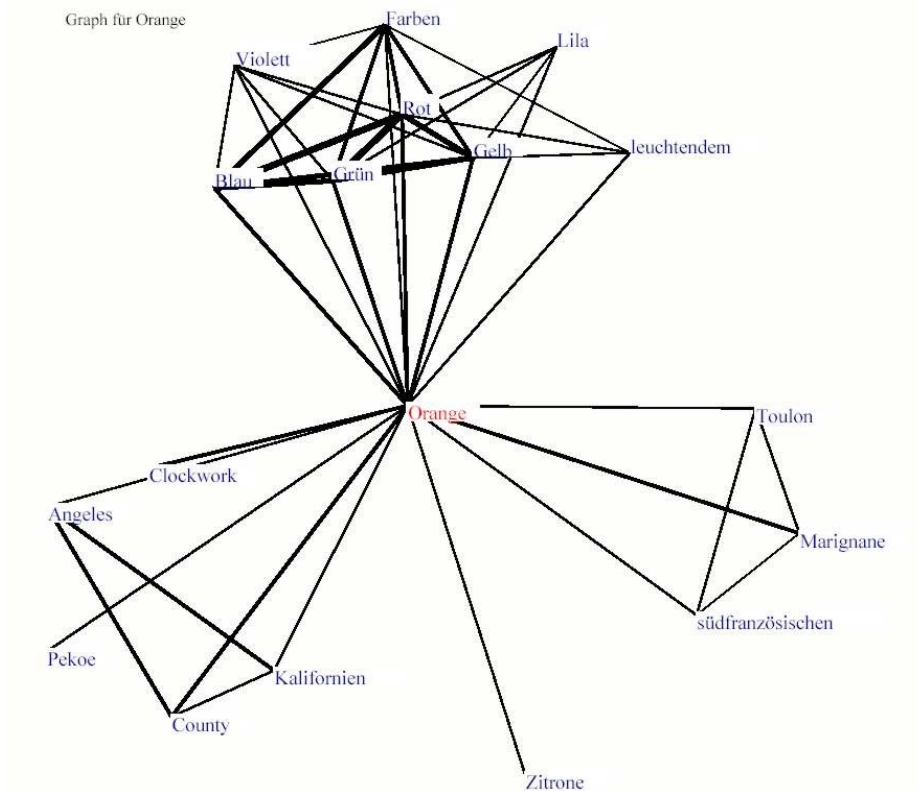
Individuumkonzept = bezeichnet einen konkreten Gegenstand (*Extension* = 1)

Verarbeitung: Durchsuchen des Netzes von jedem Knoten aus möglich.

Vererbung der Eigenschaften von der Klasse auf das Individuum



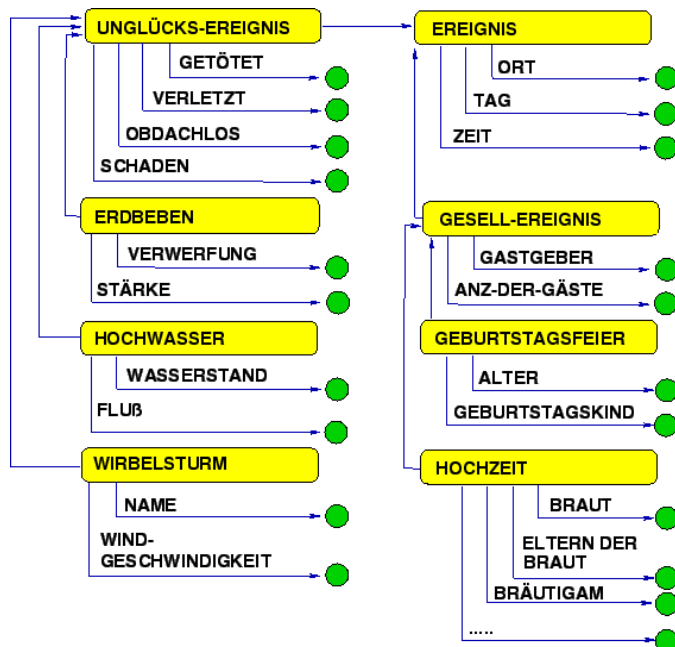
Kollokationsnetz (Dicke der Linien = Häufigkeit):



Aus: Schmidt, Fabian. Automatische Ermittlung semantischer Zusammenhänge lexikalischer Einheiten und deren graphische Darstellung. Leipzig 1999

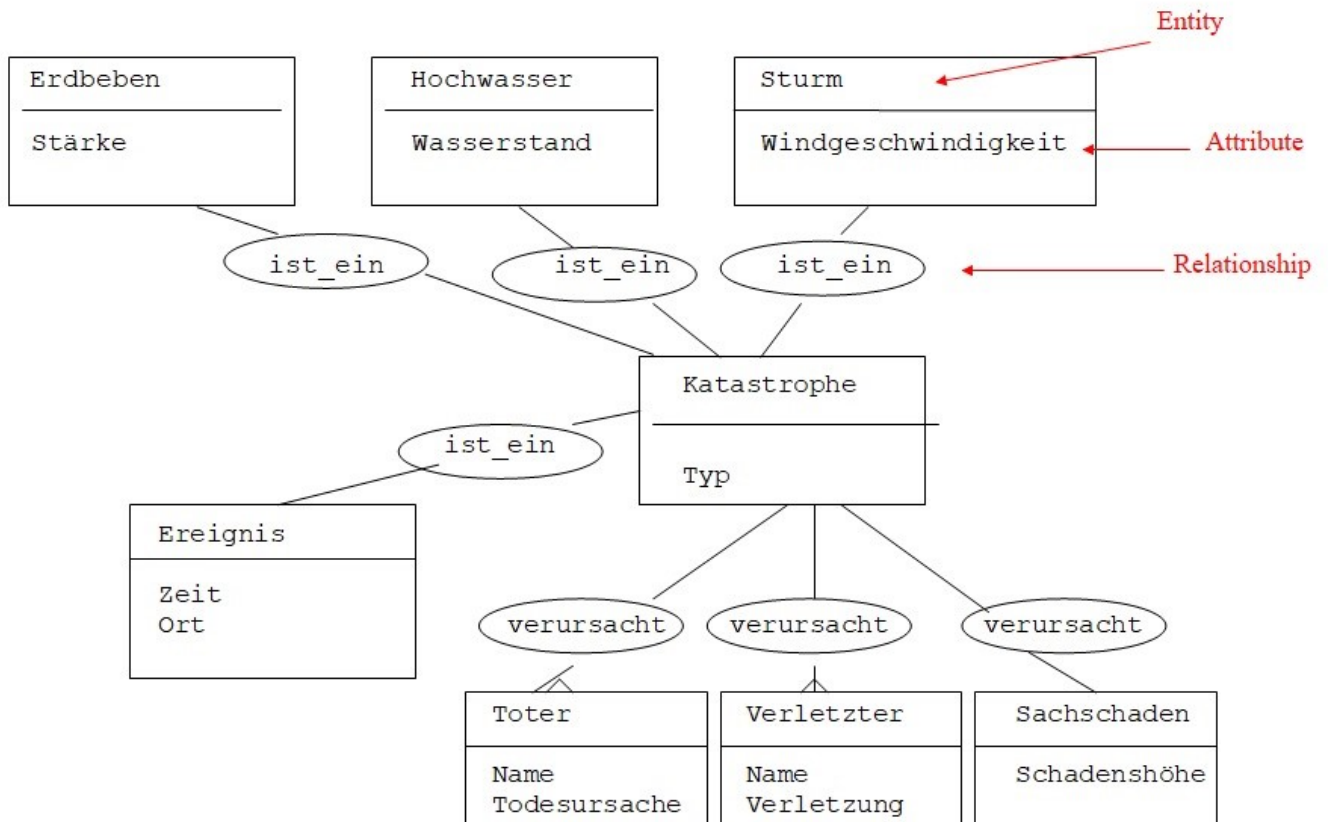
- **Frames**

Datenstruktur mit Kategorien ("*Rollen*") und "*Slots*" d.s. Leerstellen zum Ausfüllen, Prinzip des Fragebogens. Beispiel mit verknüpften Frames:



- **Datenbanken**

Auf der konzeptuellen Ebene wird das : **Entity, Attribute, Relationship** - Modell verwendet.
(eng verwandt mit Frames)



Auf der Implementierungsebene sind Datenbanken Datensätze (z.B. Tabellen mit Spalten und Zeilen) , wobei jeder Eintrag einen Schlüssel hat. Über die Schlüssel können Einträge in Relation gesetzt werden. Zum Beispiel:

Ereignisse

ID	Zeit	Ort	Typ
#1	12.8.00	San Fransico	Erdbeben
#2	10.11.00	Koblenz	Hochwasser
#3	26.11.01	Feldberg	Sturm
#4	04.05.02	Udine	Erdbeben
#5	22.08.02	Dresden	Hochwasser

Erdbeben

Ereignis_ID	Stärke
#1	6,3
#4	7,0

Hochwasser

Ereignis_ID	Wasserstand
#2	9,3 m
#5	8,0 m

Sturm

Ereignis_ID	Geschwindigkeit
#3	190 kmh

Katastrophen

Ereignis_ID
#1
#2
#3
#4
#5

Verletzte

Ereignis_ID	Name	Verletzung
#1	Miller	Beinbruch
#1	Smith	Gehirnerschütter.
#4	de Lima	Gehirnerschütter
#5	Braun	Armbruch
#5	Liebherr	Beckenbruch
#5	Meyer	Beinbruch

Sachschaden

Ereignis_ID	Höhe
#1	2 Mio \$
#4	7 Mio DM

- **Regeln**

Regeln sind bedingte Vorschriften für Operationen, d.h. sie spezifizieren direkt Algorithmen. Allgemein besteht eine Regel aus

BEDINGUNGSTEIL	AKTIONSTEIL
----------------	-------------

(vgl. IF-THEN oder DO WHILE in Programmiersprachen)

Der Bedingungsteil stellt normalerweise ein **Muster** für einen Anwendungsfall dar. Er kann Konstante und Variable enthalten. Im Aktionsteil kann auf Elemente, die im Bedingungsteil identifiziert wurden, Bezug genommen werden.

Regeln können auf immer neue Einzelfälle zutreffen und angewandt werden, denn

- ✓ Regeln sind Generalisierungen aus bisherigen Einzelfällen
- ✓ Regeln sind Vorhersagen auf künftige Einzelfälle

(im Unterschied zu Datenbanken, die nur bisherige Einzelfälle enthalten)

- **Formalismus**

Eine Menge von Regeln mit einer festgelegten Syntax, welche zusammen einen bestimmten Problembereich abdecken.

Zum Beispiel eine Menge regulärer Ausdrücke, welche bestimmte Zeichenketten in Texten identifizieren. Oder ein Grammatikformalismus mit Ersetzungsregeln, der jedem Satz einen Baumgraphen zuordnet, welcher die syntaktische Struktur zeigt. Oder ein Logikformalismus, der die Gültigkeit einer Aussage auf der Basis einer Aussagenmenge berechnet.

Vor allem Programmiersprachen sind Beispiele für Formalismen, mit denen nahezu jede Symbolmanipulation beschrieben und sodann mit dem Computer ausgeführt werden kann.

A 1.1 Grundlagen der Sprachübersetzung und die Arbeitswelt des Übersetzers und Dolmetschers

Lernziel: Eine Vorstellung von den Schwierigkeiten der Übersetzung und eine kurze Einführung in das Übersetzungswesen

Wenn jemandem eine Übersetzung auffällt, dann meist nur dann, wenn sie schlecht ist - gute Übersetzungen nehmen wir normalerweise nicht als Übersetzungen wahr.

Übersetzungsprobleme

Mehrdeutige Lexeme

Die Firma stellte einen neuen Mitarbeiter ein.

Der Mechaniker stellte den Vergaser ein.

Die Gäste stellten das Rauchen ein.

Ich stelle mich auf Schwierigkeiten ein.

Falsche Freunde

engl. *actual* - dtsch. * *aktuell*

engl. *sensible* - dtsch * *sensibel*

engl. *become* - dtsch * *bekommen*

ndl. *bellen* - dtsch * *bellen*

lexikalische Kollokationen

starker Raucher - **strong smoker* - *heavy smoker*

starker Regen - *heavy rain*

tief schlafen - * *to sleep deeply* - *to sleep soundly*

the man eats - *the dog eats*

der Mann isst - **der Hund isst* - *der Hund frisst*

Diskontinuierliche Lexeme, Mehrwortlexeme versus Einwort Lexeme

Die Geschichte fängt mit einer Explosion an.

**The history catches with an explosion on*

History begins with an explosion

Die Sonne geht im Osten auf und im Westen unter

**The sun goes in the east up and in the west down.*

The sun rises in the east and sets in the west

Unterschiedliche Wortwahl

in the street - auf der Straße - on the road

auf dem Postamt - at the post office

on the wall - an der Wand

on the table - auf dem Tisch

Unterschiedliche Wortarten

Die Entwicklung nahm ihren Anfang mit ...

The development began with ...

Er studiert gern Physik

He likes to study physics

Er studierte lieber Physik

He preferred to study physics

Er sprach weiter

He continued to talk

I suppose I must have fainted

Ich muß wohl ohnmächtig geworden sein

Übersetzungswissenschaft

Die Übersetzungswissenschaft steht zwischen Sprachwissenschaft und Kommunikationswissenschaft. Ihr Ziel:

theoretisch-methodische Grundlagen liefern, die es erlauben,

- den Übersetzungsvorgang adäquat zu beschreiben
- Möglichkeiten und Grenzen der Übersetzbarkeit festzustellen
- dem Übersetzer Arbeitshypothesen und Übersetzungsstrategien zu geben, die ihm erlauben, sich in die verschiedenen Bereiche einer zweisprachigen Kommunikation einzuarbeiten

Der Übersetzungsvorgang

Der Übersetzer trifft zunächst einmal auf ein Textvorkommen, d. h. auf eine in einer bestimmten Sprache vorliegende Information, die von ihm **erfasst** werden muss, um dann ausgehend vom Resultat des Verstehensprozesses in der Zielsprache neu **gefasst** zu werden.

Wichtige Unterscheidungen

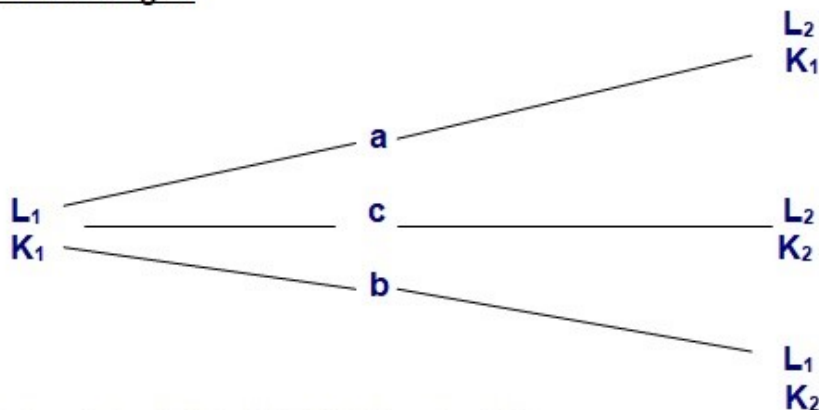
Kommunikationsgemeinschaft K:

Eine Anzahl von Menschen, bei denen das Bedürfnis und die Notwendigkeit besteht, miteinander in einen Informationsaustausch zu treten.

Sprachgemeinschaft L:

Eine Anzahl von Menschen, die dasselbe Sprachsystem verwenden.

Wechselbeziehungen



- a) Wechsel der L bei gleichbleibender K
- b) Wechsel der K bei gleichbleibender L
- c) Wechsel der L und Wechsel der K

Übersetzbarkeit von Sprachen

Wie sagtman?

Überschreiten der Gleise verboten - **No crossing** - It is forbidden to cross the tracks

Das Rauchen in der Eingangshalle ist verboten - **No smoking**- It is forbidden to smoke in the lobby

Allgemeines Vorurteil, der Übersetzungsprozess sei ein reiner Umkodierungsprozess

Kulturelle Unterschiede. Der Wortschatz einer Sprache repräsentiert eine bestimmte Organisation der Wirklichkeit. In anderen Sprachen ist die Wirklichkeit anders gegliedert.

Allgemein bekannt ist die Aufteilung der Farbenskala, zum Beispiel:

kymrisch	gwyrdd	glas		llwyd		
engl.	green	blue	grey	brown		
Navaho	icii	ico		doot'iz		
engl.	red	orange	yellow	green	blue	purple

Viele Sprachen haben kein Äquivalent zu *Onkel* sondern unterscheiden *Bruder von Vater* und *Bruder von Mutter* und trennen dies von *Ehemann von blutsverwandter Tante* ab.

.Übersetzungsäquivalenz

Wechselseitige Austauschbarkeit von einem Text in einer Sprache L2 (ZS-Text) und einem Text in einer Sprache L1 (AS-Text) bezogen auf Situationen und Kontexte.

Äquivalenz ist schwächer als Bedeutungsgleichheit. Wörter mit völlig gleicher Bedeutung gibt es in verschiedenen Sprachen nur wenige.

„ Man kann daher für die Äquivalenz nur fordern, daß äquivalente AS-Ausdrücke und ZS-Ausdrücke überlappende Bedeutungen haben, wobei der Durchschnitt zwischen beiden die in dem betreffenden Kontext aktualisierte(n)Bedeutung(en) enthält. Das "(en)" ist z.B. für die Übersetzung von Wortspielen etc. relevant, die eben oft nicht übersetzbar sind.“
(Jürgen Kunze)

Wortspiele, meist nicht übersetzbar:

Dem Junggesellen fehlt zum Glück die Frau
Sie fahren mit Abstand am besten.

- **Letztes Kriterium ist die Pragmatik, die Perlokution des Textes in der Zielgruppe**
 - Für wen wird übersetzt?
 - Zu welchem Zweck wird übersetzt?

Je nachdem ist „ the smallest translatable unit“ etwas anderes.

Das Berufsprofil

Übersetzer, Urkundenübersetzer, Medienübersetzer, Konferenzübersetzer, Redaktion
maschinellem Übersetzungen und literarisches Übersetzen

Dolmetscher, Gesprächs- und Flüsterdolmetscher, Gerichtsdolmetscher und
Mediendolmetscher

Angestellte oder Freiberufler

Gesetzlich geschützte akademische Grade: Diplom-Dolmetscher, Diplom-Übersetzer, Diplom-
Fachübersetzer, Diplom-Technikübersetzer und Diplom-Sprachmittler.

Sprachen	Selbständige Übersetzer %	Angestellte Übersetzer %
Englisch / Amerikanisch	68,6	74,5
Französisch	45,2	44,1
Spanisch	14,3	<u>21,6</u>
Italienisch	13,8	11,8
Andere romanische Sprachen	5,2	7,8
Russisch	<u>7,6</u>	4,9
Andere slawische Sprachen	<u>9,0</u>	1,0
Niederländisch / skand. Sprachen	<u>8,6</u>	2,9
Sonstige Sprachen	10,0	6,9

Art der übersetzten Texte	Selbständige Übersetzer %	Angestellte Übersetzer %
Bücher	<u>81,0</u>	1,0
Zeitschriften-, Zeitungsartikel	<u>44,8</u>	9,8
Verträge, Urkunden, Dokumente, Korrespondenzen	29,5	<u>85,3</u>
Bühnenstücke, Drehbücher, TV-/Film- Manuskripte	<u>25,2</u>	1,0
Werbetexte, Gebrauchsanweisungen, Bedienungsanleitungen, Produktbeschreibungen	21,4	<u>60,8</u>

Ausbildung

Tabstopp links		
Formalabschluss zum Übersetzer	Selbständige Übersetzer %	Angestellte Übersetzer %
Keinerlei Abschluss oder Zertifikat	61,4	17,6
Staatliche Prüfung als Übersetzer	19,5	20,6
Prüfung bei öffentlichen Einrichtungen	<u>5,7</u>	<u>14,7</u>
Prüfung bei privaten Einrichtungen	8,6	<u>25,5</u>
Keine Angaben	4,8	21,6
Allgemeinausbildung und sprachqualifizierte Erfahrungen	Selbständige Übersetzer	Angestellte Übersetzer
Sprachenschule	19,0	<u>51,9</u>
Fachhochschule	3,3	3,9
Universität	51,9	19,6
Fern- und Selbststudium	<u>31,4</u>	<u>36,3</u>
Mehrsprachige Erziehung	16,7	10,8
Langjährige Auslandsaufenthalte	<u>47,6</u>	20,6
Sonstige sprachnahe Erfahrungen	5,7	6,9

Nur 7,1 % der selbständigen Übersetzer sehen ihren Beruf durch technische Entwicklungen bedroht.

Typen von Übersetzungen

- Interlinearversion
- Wörtliche Übersetzung
- Philologische Übersetzung
- Kommunikative Übersetzung
- Bearbeitende Übersetzung

Software für Übersetzer

"Computer aided Translation" - CAT

- Textverarbeitung, Alignment
- Rechtschreib- und Grammatikkorrektur
- [Wörterbuch, Synonymenwörterbuch, Kollokationswörterbuch](#)
- [Terminologie](#)
- Sachwissen, Enzyklopädie
- Vergleichstexte, Stil
- [Translation Memory](#)

Synonymwörterbuch

Eingebaut in Microsoft Word, Rechtsklick auf ein Wort im Text ergibt eine Liste der Bedeutungen des Wortes, Auswahl einer Bedeutung ergibt eine Synonymenliste



Online-Suche im Kollokationswörterbuch

Häcki Buhofer, Annelies, Marcel Dräger, Stefanie Meier & Tobias Roth (2014): Feste Wortverbindungen des Deutschen. Kollokationenwörterbuch für den Alltag. Tübingen: Francke.

einstellen

in ein Arbeitsverhältnis nehmen, anstellen

ADJEKTIVE/ADVERBIEN/NOMEN

dauerhaft • unbefristet • fest • sofort • vorübergehend

justieren, regulieren

ADJEKTIVE/ADVERBIEN/NOMEN

genau • exakt • präzise •
korrekt • richtig • optimal •
scharf

neu • manuell • variabel •

stufenlos
aufeinander • medikamentös

niedrig • hoch • falsch

vorübergehend nicht fortsetzen, aufhören

ADJEKTIVE/ADVERBIEN/NOMEN

frühzeitig • vorübergehend •
vorläufig • vorerst

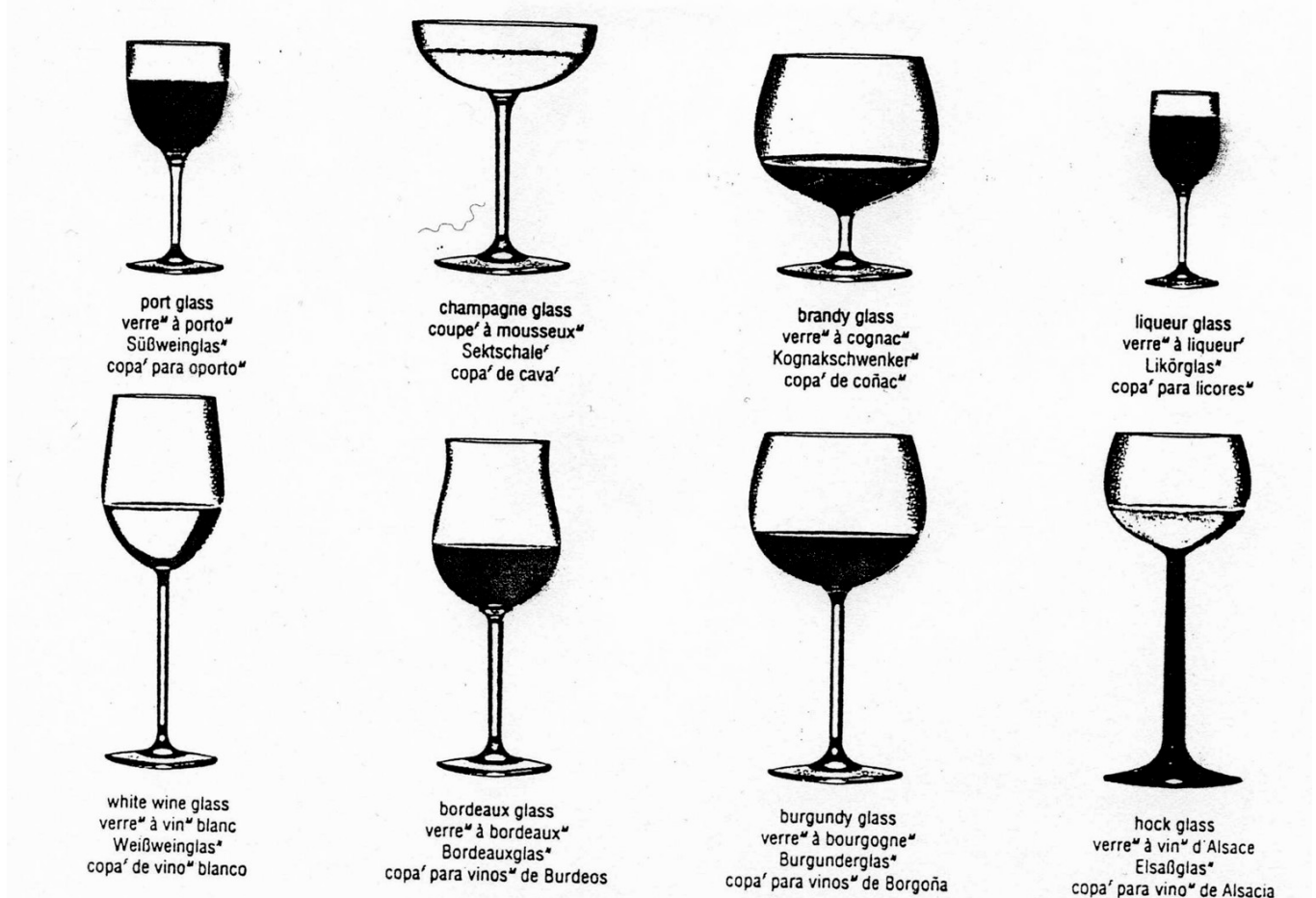
unverzüglich • sofort • Wirkungsdauerhaft • endgültig •
gänzlich • vollständig

weitere Verbindungen (alphabetisch)

NOMEN

Aktivität • Anzeige • Arbeit • Arbeiterin Arbeiter
• Arbeitgeberin Arbeitgeber • Bau • Betrieb •
Dienst • Erfolg • Feuer • Hilfe • Kamera •
Kampf • Kanal • Kraft • Leben • Lehrling •
Leute • Lieferung • Linie • Produktion • Projekt
• Situation • Sprache • Tätigkeit • Verfahren •

Terminologie (Computerunterstützung besonders wichtig)



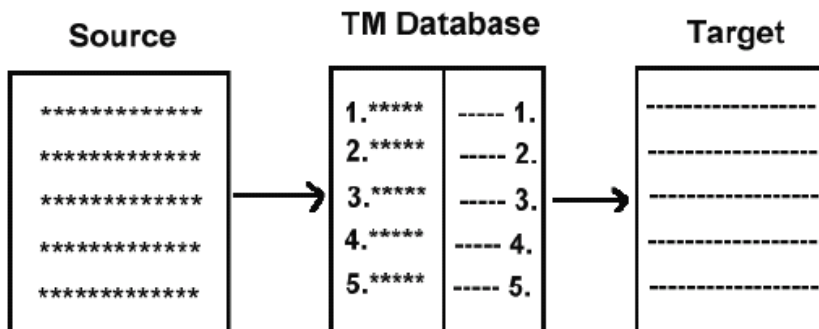
Translation Memory

ein Textarchiv mit parallel angeordneten ("aligned") Texten in verschiedenen Sprachen

speziell erstellt oder einfach durch Aufheben früherer Übersetzungen

eine Datenbank mit Textstücken (Sätzen, Teilsätzen oder auch größeren Abschnitten) in zwei Sprachen

bevor der Übersetzer etwas neu übersetzt, wird alles automatisch ersetzt, was im TM vorhanden ist



© Lynn E. Webb

Genauere Übereinstimmung von Textstellen bis auf den i-Tüpfelchen sind selten. Daher geht es bei der TM-Software vor allem um das Auffinden ähnlicher Textstellen ("fuzzy matching")

A.1.2 Maschinelle Übersetzungssysteme

Was hier 2004 in meiner Vorlesung stand, ist heute (2021) weitgehend veraltet. Es wird deshalb hier nicht mehr vorgetragen. Damals fanden wir bei Tests folgende Fehler

	Fehlertyp:	Was tun?
[1]	Von mehreren Bedeutungen eines Wortes die falsche herausgegriffen	Differenzieren der Wortbedeutungen im Lexikon und zu jeder Bedeutung Erkennungsmerkmale eintragen
[2]	Feste, idiomatische Redewendung nicht erkannt und wörtlich übersetzt	Lexikon und Grammatik müssen mit Mehrwort-Wendungen fertig werden
[3]	Terminologisierte Begriffe nicht erkannt und mit falschen Varianten übersetzt	Lexikon verbessern
[4]	Wort ganz unbekannt	Lexikon verbessern
[5]	morphologische (Komposita!) oder syntaktische Analyse falsch	Parser verbessern
[6]	falsche Pronomenauflösung	Vorgängersatz einbeziehen und ein gutes Modul für Pronomenauflösung entwickeln
[7]	komplett schief gegangen; Grund unbekannt	
[8]	Fehler beim Generieren der Zielsprache	besseren Generierer entwickeln

Heute tritt beim Übersetzen der damaligen Testsätze keiner der Fehler mehr auf. Auf der Homepage des aktuell benutzten Programms **babelfish.de** heißt es (21.9.2021):

"Wie funktioniert babelfish.de?"

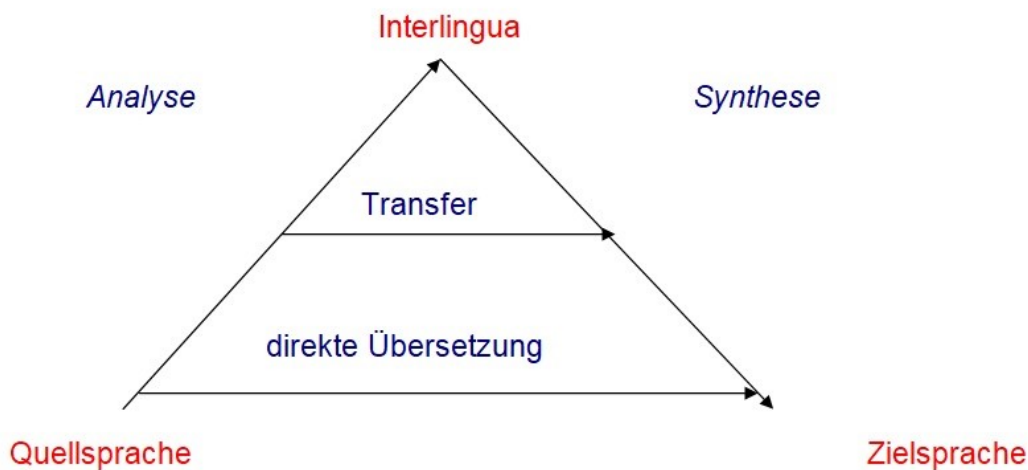
babelfish.de basiert auf einem speziellen Suchalgorithmus, der für jeden Ausgangstext die besten Resultate aus einem mehrsprachigen Wörterbuch, einer automatischen Textübersetzung sowie aus Millionen professionell übersetzter Dokumente ermittelt.

Woher stammen die Übersetzungen?

Neben mehrsprachigen Webseiten werden für babelfish.de beispielsweise auch Dokumente der Europäischen Union sowie der Vereinten Nationen indiziert. Nutzer haben zudem die Möglichkeit, eigene Übersetzungen vorzuschlagen."

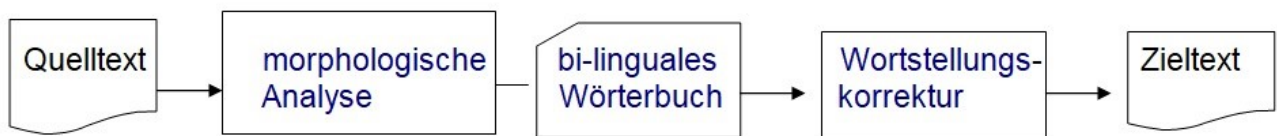
Lange diente zur Klassifizierung von maschinellen Übersetzungssystemen das folgende Schema:

MÜ-Dreieck (nach Vauquois)



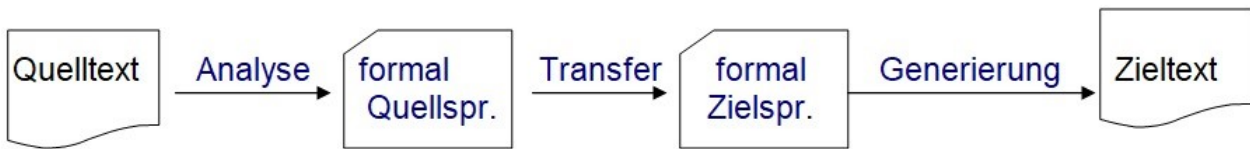
Dies ergibt drei Grundtypen von Übersetzungssystemen

Direkte Übersetzung



Direkte Übersetzung ist oberflächennah, berücksichtigt ad-hoc Besonderheiten für Übergang von einer Sprache zur anderen. Je nach Erfordernis sind die Übersetzungseinheiten (**smallest translatable unit**) verschieden: [einzelne Wörter](#), [Wortgruppen](#), [Idiome](#), [Homographen](#), [Komposita](#), [Satzteile](#), bis hin zu ganzen (kurzen) Sätzen!

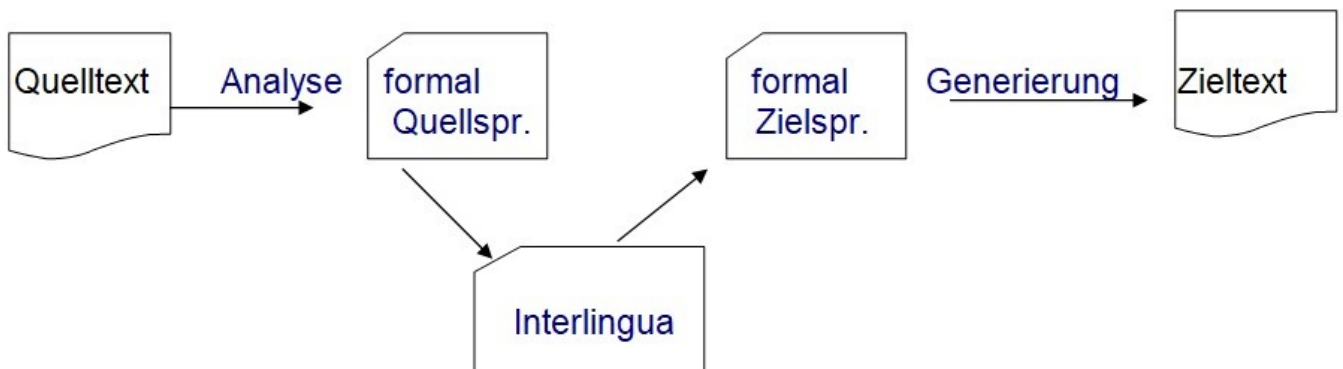
Transfermodell - bilingual



Transfermodell mit kompletter tiefer Analyse, formaler Repräsentation L1, Transfer, formale Repräsentation L2, Generierung

Nachteil: $n - 1 \times n$ viele Transfermodule bei n Sprachen nötig.

Interlingua-Modell - multilingual



Interlingua-Modell multilingual mit Transfer in die interlingua im Anschluß an jeweilige Analyse
Nur $n - 1 + n$ viele Transfermodule, aber Schwierigkeit eine universale Interlingua zu entwerfen.

Zu dieser regelbasierten Übersetzung sind heute statistische Verfahren zur Erschließung riesiger Corpora von Vergleichstexten hinzu gekommen.

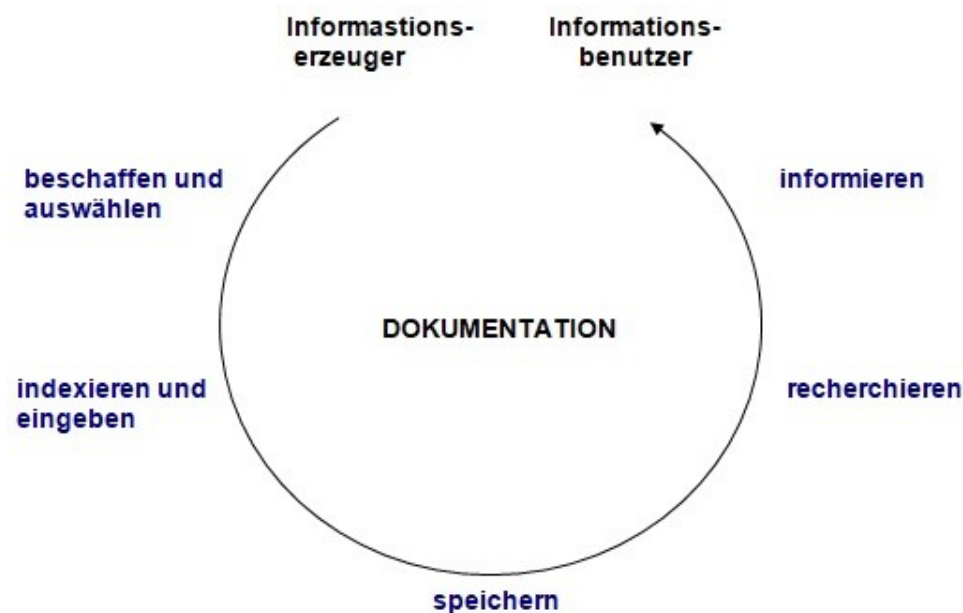
A.2.1 Grundlagen des Dokumentationswesens und des Information Retrieval

Lernziel: Einsicht in das Anwendungsfeld

1. Dokumentationswesen

Information und Dokumentation (IuD)

Vermittlung zwischen Informationsangebot und Informationsnachfrage



Objekte der Dokumentation

dokumentarische Bezugseinheit DBE (=das Original)

Texte, Kunstwerke, Photos, Filme, Musikstücke usw.

insbesondere Objekte des Publikationswesens: Artikel, Bücher, Zeitschriften

Dokumentationsmittel

Dokumentationseinheit DE (= z.B. ein Eintrag im Katalog)

Kataloge, Bibliographien, Referatenorgane, Datenbanken

Institutionen

Bibliotheken (Library of Congress!), Dokumentationszentren, das Internet

Berufsbilder

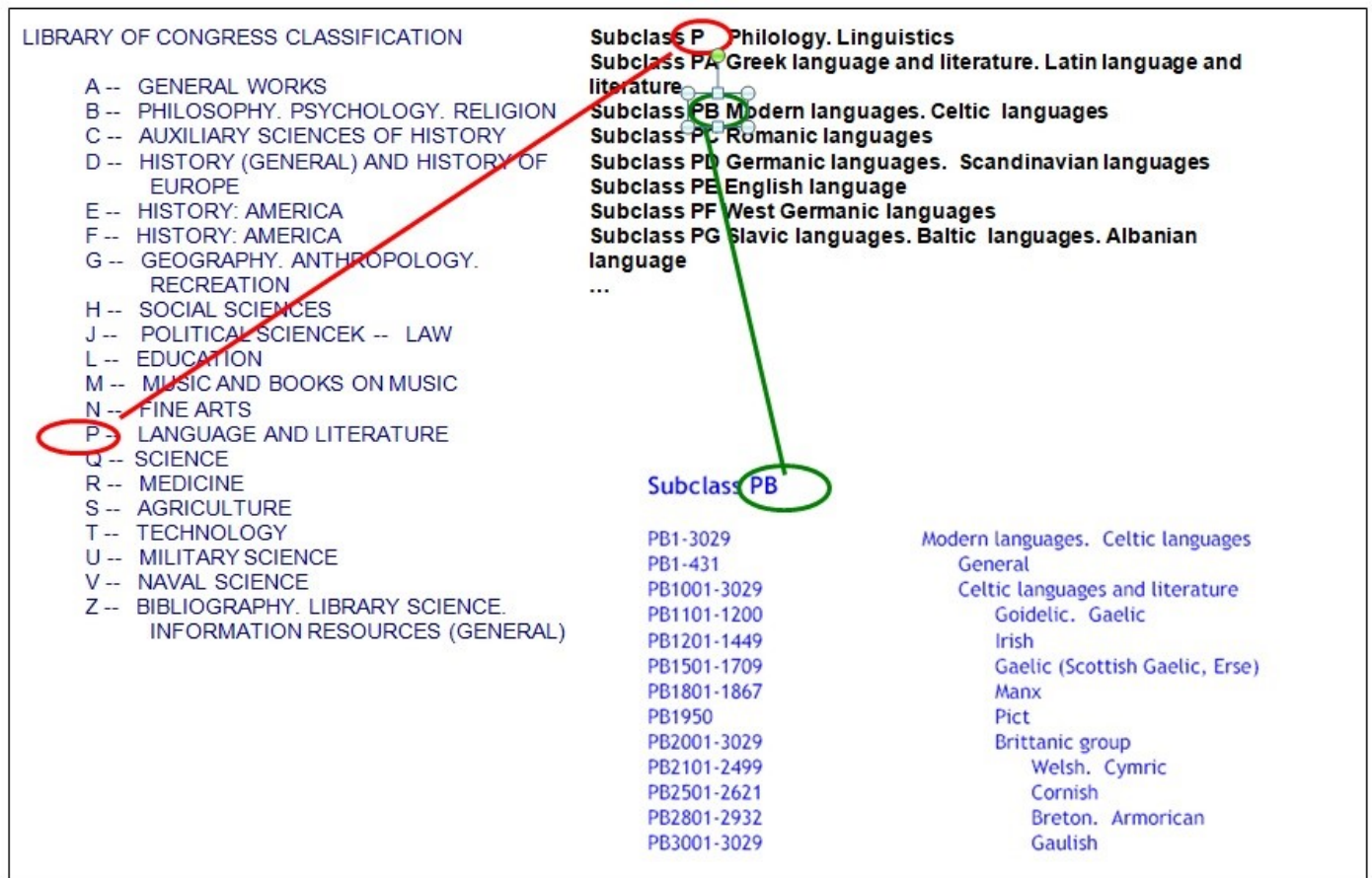
Bibliothekar, Dokumentar, Information Broker
 Informationswissenschaftler

Der Klassischer Dokumentationsprozess

Beschaffen, Erfassen: Verfasser, Titel, Erscheinungsort, Jahr, Verlag, Signatur
 inhaltlich Erschließen: Klassifikation, Indexieren, Abstract

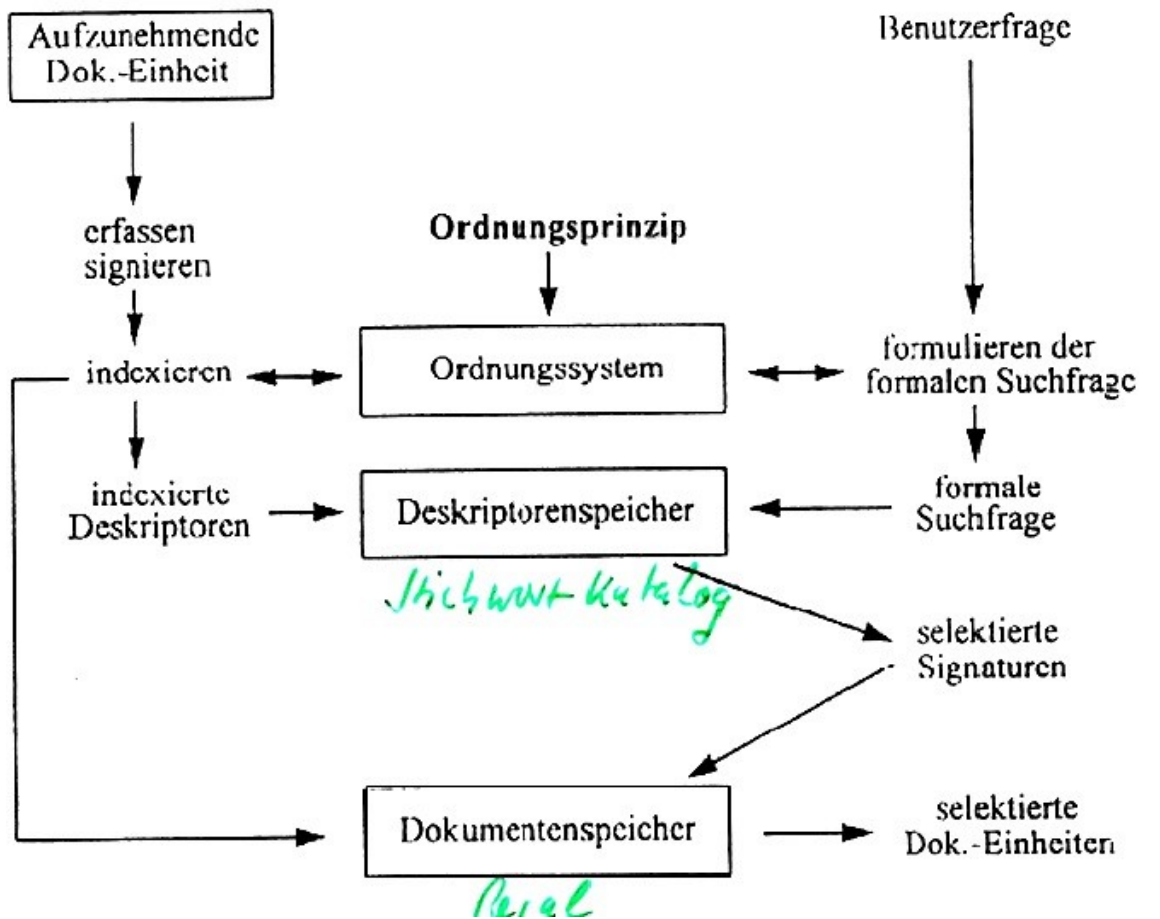
Klassifikation

hierarchisch



Indexierung

- **Stichwort** aus dem Text selbst, nicht beschränkt
- **Schlagwort** Deskriptor zugeordnet, kontrolliert, aus einem Thesaurus



Sachgruppen als Deskriptoren der Deutschen Bibliothek (Auszug)

- | | |
|--------------------------------|------------------------------|
| Nutzungsrecht | Pathologie |
| - Natürliche Ressourcen 7.5c | - allgemein 27.4 |
| Obstbau 32.4 | Personalpolitik 10.11h |
| Öffentliche Aufgaben 10.3a | Pflanzenanatomie 24.2a |
| Öffentliche Bautätigkeit 10.8a | Pflanzengeographie 24.2b |
| Öffentliche Verwaltung 7.15a | - Nerv 27.3d |
| - Geschichte 7.15b | - Pflanzen 24.2a |
| Öffentliche Wirtschaft 10.3b | - Sinnesorgan 27.3d |
| Öffentlicher Dienst 7.15a | - Stoffwechsel 27.3c |
| Öffentlichkeitsarbeit | - Tiere 25.2a |
| - allgemein 10.15 | Physiologische Chemie 27.3a |
| Onkologie 27.4 | Physische Geographie 19.1b |
| Operations research 29 | Plasmaphysik 21.5 |
| Optik 21.3 | Plastik 13.2 |
| Organisationssoziologie 9.3c | Politik |
| Organische Chemie 22.5 | - allgemein 8.1 |
| Orthopädie 27.10 | Politiker 16.5p |
| Ortsgebundenes Bauwerk 31.3ab | Politische Organisation 8.2b |
| Ostkirche 3.6c | Politische Theorie 8.1 |
| Paläontologie 19.4d | Polizeirecht 7.6a |
| Papierherstellung 31.14 | Praktische Theologie 3.5 |
| Parapsychologie 5.6 | Presse 2.3 |

Referieren, Zusammenfassung, Abstract

Das sind Texte über Texte! mehr darüber in A 2.2

2. Dokument Retrieval mit Computer

Volltextsuche (vgl. Internet-Suchmaschinen)

basiert auf Stichwörtern, ggf. auf alle Wörter im Text

schnelle Aufsuche durch Volltext-Index (sog. "invertierte Datei", Register)

Stop-Wörter, Lemmatisierung

starker Filtereffekt durch geschickte Kombination von Stichwörtern

mit Beschränkungen auf Bereiche

Suche mit Deskriptoren (Schlagwörtern)

Deskriptoren werden dem Dokument in der Datenbank zugeordnet

Jedes Dokument D wird gekennzeichnet durch eine Menge von Deskriptoren T

	T1	T2	T3	T4	Tn
D1	1	0	0	1	1
D2	0	1	1	0	0
D3	1	0	1	0	0
...

Suche mit Booleschen Verknüpfungen (UND, ODER, NICHT) von Deskriptoren

(T1 UND NICHT T2) UND T3

$T1 = \{D1, D3\}$

$\text{NICHT } T2 = \{D1, D3\}$

$(T1 \text{ UND NICHT } T2) = \{D1, D3\}$

$T3 = \{D2, D3\}$

$(T1 \text{ UND NICHT } T2) \text{ UND } T3 = D3$

Erweiterungen der Suche mit dem Thesaurus

Praktisch steht die gesamte Datenbanktechnologie zur Verfügung

Vector Space Retrieval Modell

Grundidee: Modellierung von Dokumenten wie Anfragen durch einen n-dimensionalen Vektorraum bei n unterschiedlichen Termen in der Indexierungssprache.

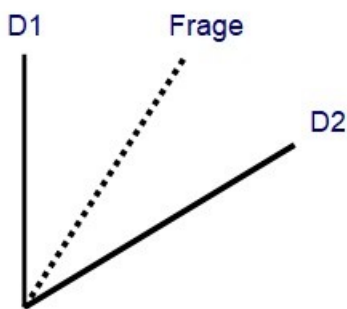
Jeder Deskriptor stellt eine Dimension im Raum dar.

Jedes Dokument steht aufgrund seiner relevanten Deskriptoren an bestimmter Stelle im Raum.

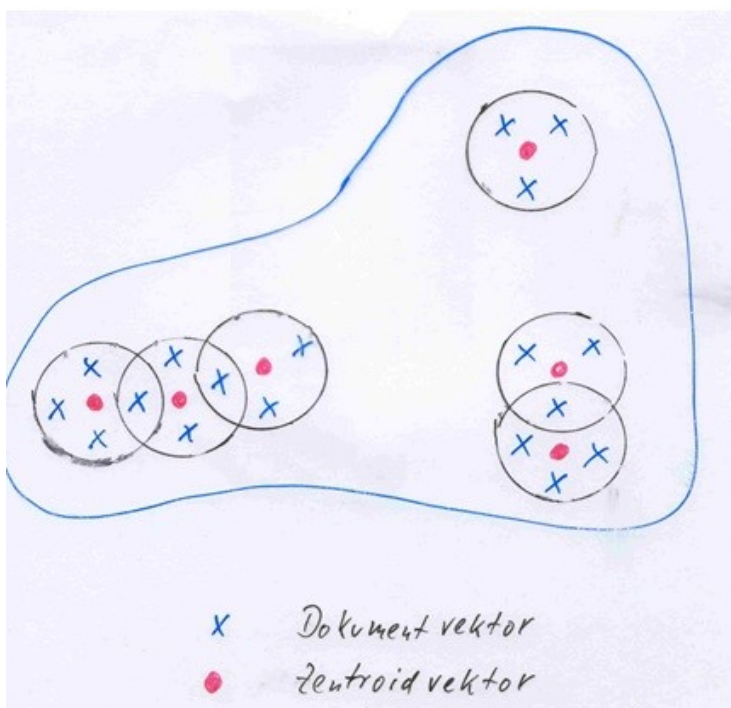
Dasselbe gilt für die Anfrage.

Die Ähnlichkeit zwischen den Vektoren der Dokumente und dem Vektor der Frage wird berechnet, z.B. als Winkel, nicht absolut, sondern relativ

Gewichtung der Ausgabe nach der Nähe der Dokumente zur Anfrage



Clustering im Vector Space



Maschinelle Ermittlung der Deskriptoren (Maschinelle Indexierung)

- linguistische Verfahren

Einsatz von Methoden der lexikalischen Semantik bis hin zur Textlinguistik

- statistische Verfahren

Wieso funktionieren **statistische Verfahren**?

Wortbedeutung muß Wortbedeutung im Text treffen, nicht unbedingt selbst bekannt sein

Grundidee statistischen Indexing:

Kontrast zwischen spezieller Häufigkeit von Wörtern im Dokument - in einer Gruppe von Dokumenten - im Vergleich zur Verteilung in der Alltagssprache. Was häufig in einem Dokument vorkommt, ist möglicherweise relevanter für das Dokument als der Rest.

$\text{relative Häufigkeit eines Wortes} = \frac{\text{absolute Häufigkeit des Wortes im Text}}{\text{durch Zahl aller Wörter im Text}}$
$\text{Auffälligkeitswert eines Wortes (weirdness)} = \frac{\text{relative Häufigkeit des Wortes im Text}}{\text{relative Häufigkeit des Wortes in der Alltagssprache}}$
<p>d.h. Wort kommt hier x-mal häufiger vor als im Durchschnitt</p>

Erfolgsmaße: recall und precision

- **recall** der Anteil der gefundenen Fälle unter allen relevanten im Suchraum
- **precision** der Anteil der relevanten Fälle unter den gefundenen

Wieviel Schrott ist dabei?

	relevant	nicht relevant
gefunden	g (gut) - z.B. 8	s (Schrott) - z.B. 72
nicht gefunden	f (fehlt) - z.B. 2	a (ausgeschlossen)

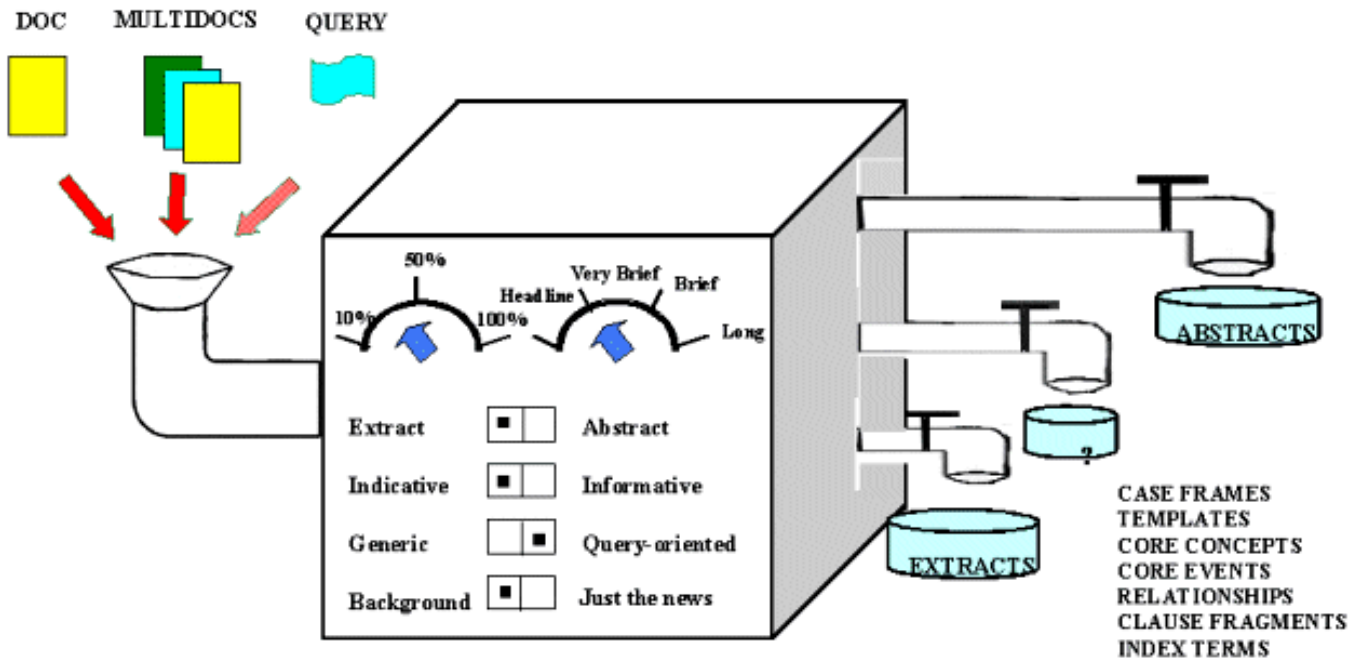
recall = $g / g+f$ z.B. $8/10 = 0.8 = 80\%$

precision = $g / g+s$ z.B. $8/80 = 0.1 = 10\%$

A 2.2 Automatische Texterschließung und Textzusammenfassung

Lernziel: Einiges Verständnis dafür, wie Textzusammenfassung durch einen menschlichen Bearbeiter funktioniert und von da aus eine Bewertung maschineller Verfahren.

A Summarization Machine



© Eduard Hovy and Daniel Marcu

Eingabe

- Einzeldokument
- mehrere Dokumente, z.B. minütlich aufeinander folgende Nachrichten einer Presseagentur
(evtl. auch mehrere Dokumente in verschiedenen Sprachen)
- Benutzeranfrage mit bestimmtem Informationsinteresse

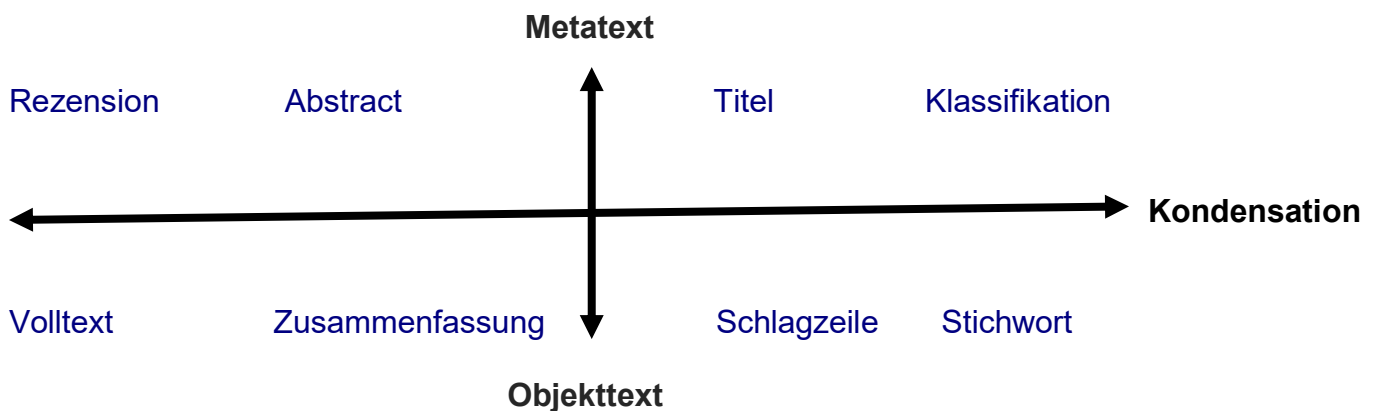
Ausgabe

- **Extrakt** = einzelne Passagen oder Sätze oder Wörter werden einem Text entnommen: Zitate, „Zettelkasten“, vor allem auch Stichwörter, die später Deskriptoren zugeordnet werden können
Extrakte sind unvollständige Texte, herausgerissen aus dem Kontext, oft nicht kohärent und kohärent

- **Zusammenfassung** = ein Originaltext wird auf das Wichtigste verkürzt, kondensiert
Pragmatische Funktion: Z. steht häufig am Ende eines Buches, das Wichtigste soll noch einmal genannt werden, damit es nicht vergessen wird (*informative summary*)
- **Abstract** = ein Metatext, der beschreibt, worum es im Originaltext geht und der in der Regel ebenfalls kurz ist
Pragmatische Funktion: A. steht häufig am Anfang eines Artikels, Entscheidungshilfe für den Leser, ob er den Originaltext überhaupt lesen soll (*indicative summary*)

➤ Linguistisch geht es um semantische Zusammenhänge zwischen verschiedenen Texten.

Zwei Dimensionen des Text-zu-Text-Zusammenhanges:



Objekttext = ein Text über Fakten und Ereignisse in einem Objektbereich

z.B. Nachrichten, Erzählungen, wissenschaftliche Artikel, Gebrauchsanweisungen, usw., usw.

Metatext = ein Text über einen anderen Text

z.B. Rezension, Abstract, Titel, Inhaltsverzeichnis, Bibliographie

Schlagzeile = eine extrem verkürzte Zusammenfassung

Titel = ein extrem verkürztes Abstract

Grundsätzliche Herangehensweise

- **Top-Down Texterschließung - nachfrageorientiert**
Der Benutzer formuliert eine Frage. Diese wird mit den vorhandenen Texten verglichen. Alles Irrelevante wird ignoriert. Zur Frage Passendes wird ausgegeben. (Siehe document retrieval A 2.1 oder fact retrieval A 3.2)
- **Bottom-Up Texterschließung - angebotsorientiert**
Alles, was (nach dem Urteil des Bearbeiters) wichtig ist, wird den vorhandenen Texten entnommen und dem Benutzer dargeboten

Textzusammenfassung - mindestens drei Aufgaben:

- ANALYSE - Verstehen des Originaltextes, Kohäsion und Kohärenz
- AUSWAHL - Beurteilung, welche Information darin mehr und welche weniger wichtig ist
- GENERIEREN - Wiedergabe der wichtigsten Information in stark komprimierter Form

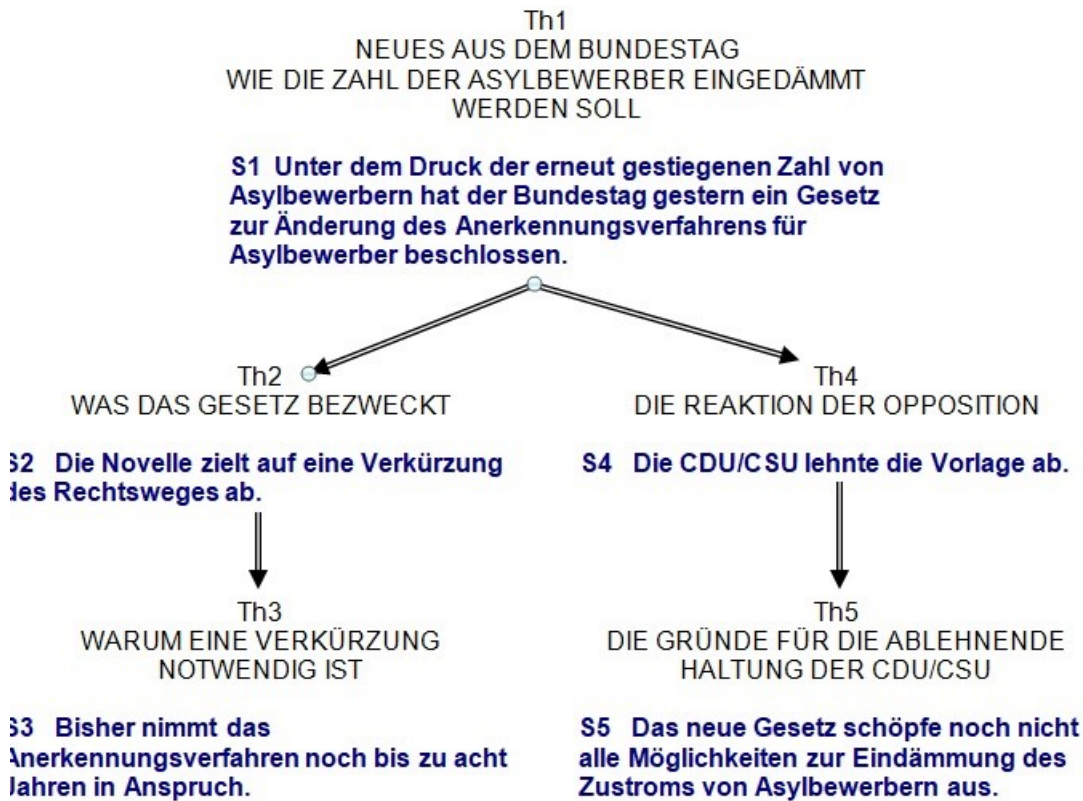
Anmerkung: Die überwiegende Zahl gegenwärtiger Programme zur Zusammenfassung von Texten beschränkt sich auf die AUSWAHL. Mit verschiedenen - teils statistischen, teils linguistischen - Methoden wird jedem Satz ein Maß an Informationshaltigkeit zugeordnet. Anschließen werden je nach gewünschter Länge des Outputs die Sätze mit dem höchsten Wert schlicht extrahiert und ausgegeben. Dass so erzeugte Zusammenfassungen oft sowohl in Bezug auf Kohäsion wie Kohärenz defekt sind, ist nicht erstaunlich.

Beispiel

Originaltext

- (1) Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen.
- (2) Die Novelle zielt auf eine Verkürzung des Rechtsweges ab.
- (3) Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch.
- (4) Die CDU/CSU lehnte die Vorlage ab.
- (5) Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.

Kohärenzbaum



Aufgabe: Komprimierung des Textes auf x Prozent Länge -

Zwei Vorgehensweisen

- "Von außen nach innen" – den vorliegenden Text immer kürzer machen durch Eliminieren von weniger wichtigen Sätzen, bis x erreicht ist.
- "Von innen nach außen" – Verstehen des ganzen Textes, Wissensrepräsentation erstellen, daraus das Wichtigste zuerst sagen, Text immer länger machen durch Hinzufügen von weiteren Sätzen mit abnehmender Wichtigkeit, bis x erreicht ist.

Operationen beim Zusammenfassen von außen nach innen

- Weglassen
- Integrieren (Fusionieren)
- Generalisieren

dabei von den Blättern im Kohärenzbaum zur Wurzel hin vorgehen.

Ergebnisse eines Experiments

manuelle Zusammenfassung

- **Proband 1**

Der Bundestag hat gestern ein Gesetz zur Verkürzung des Anerkennungsverfahrens für Asylbewerber beschlossen. Die CDU/CSU lehnt die Vorlage ab, da sie nicht ausreicht, den Zustrom von Asylbewerbern einzudämmen. (50%)

Satz 1 (Teile weggelassen) + Satz 2 (integriert) - (Satz 3 weggelassen) - + Satz 4 + Satz 5 (i

Neues Asylbewerbergesetz verabschiedet

Schlagzeile: Teile von Satz 1

- **Proband 2**

Wegen der erneut gestiegenen Anzahl von Asylbewerbern hat der Bundestag ein Gesetz zur Änderung des Anerkennungsverfahrens beschlossen. Die Novelle zielt auf eine Verkürzung des Rechtsweges ab. Die CDU/CSU lehnte die Vorlage ab.

Satz1 + Satz 2 + Satz 4 (nur Weglassen von Satz3 und Satz5)

- **Proband 3**

Der Bundestag hat gestern ein Gesetz zur Verkürzung des bisher noch acht Jahre dauernden Anerkennungsverfahrens für Asylbewerber beschlossen. Die CDU/CSU lehnte die Vorlage ab.

Satz1 + Satz 2 + Satz 3 (zu einem Satz integriert) + Satz 4 (Satz 5 weggelassen)

- **Proband 4**

Der Bundestag hat ein Gesetz zur Verkürzung des Anerkennungsverfahrens für Asylbewerber beschlossen, welches von der CDU/CSU abgelehnt wurde.

Teile von Satz 1 + Satz 3 integriert

- **Proband 5**

Gestern beschloß der Bundestag trotz Ablehnung durch die CDU/SCU ein Gesetz zur Eindämmung des Asylbewerberzustroms.

Teile von Satz 1 + Satz 3 integriert

Beispiel für automatische Textzusammenfassung

MS Word 25% (nur das rot gedruckte kommt in die Zusammenfassung)

Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen. Die Novelle zielt auf eine Verkürzung des Rechtsweges ab. Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch. Die CDU/CSU lehnte die Vorlage ab. Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.

MS Word 40%

Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen. Die Novelle zielt auf eine Verkürzung des Rechtsweges ab. Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch. Die CDU/CSU lehnte die Vorlage ab. Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.

MS Word 60%

Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen. Die Novelle zielt auf eine Verkürzung des Rechtsweges ab. Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch. Die CDU/CSU lehnte die Vorlage ab. Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.

MS Word 75%

Unter dem Druck der erneut gestiegenen Zahl von Asylbewerbern hat der Bundestag gestern ein Gesetz zur Änderung des Anerkennungsverfahrens für Asylbewerber beschlossen. Die Novelle zielt auf eine Verkürzung des Rechtsweges ab. Bisher nimmt das Anerkennungsverfahren noch bis zu acht Jahren in Anspruch. Die CDU/CSU lehnte die Vorlage ab. Das neue Gesetz schöpfe noch nicht alle Möglichkeiten zur Eindämmung des Zustroms von Asylbewerbern aus.

Die wichtigsten Verfahren zur maschinellen Textzusammenfassung

Da bisher kein vollständiges Textverstehen möglich ist (bzw. zu aufwändig wäre) sind die meisten Systeme **hybrid**, d.h. sie setzen verschiedene Techniken für punktuelle Einzelphänomene ein. Anschließend "Übereinanderlegen" der Ergebnisse, so dass eine einigermaßen gute Zusammenfassung entsteht.

- **Statistische Methoden zur Extraktion**

- **Methoden zum Erkennen der Kohärenzstruktur** (*topic detection, rhetorical structure detection*)

- Lexikalische Ketten (lexical chains) spiegeln Kontiguität
- Kohärenzstruktur (rhetorical structure detection) u.a. mittels "cue phrases "

- **Methoden zur Auswahl des Wichtigsten**

Gewichtung je nach Position (optimal position policy), Sätze oder Stichwörtern am Textanfang/Ende, am Anfang und Ende eines Absatzes, Position des Absatzes

- Stopwörter und VIP Wörter, Thesaurus, Templates

- **Generierung**

Es gibt bisher nur wenig über die Fusion von Sätzen im Rahmen von Integration und Generalisierung.

Lexical chains

(herauszufinden mittels lexikalischer Datenbank, Prototypensemantik, Pronomenauflösung):

Auf dem Bild befindet sich eine Maus auf einer Glasplatte.
Man sieht sie von unten?
Ihre Haltung ist horizontal.
Ihre Schnauze berührt den linken Bildrand exakt in der Mitte.
Körper und Schwanz sind ausgestreckt und reichen insgesamt zu drei Vierteln über das Bild zum rechten Bildrand .
Auf der rechten Seite des Bildes befindet sich ein Mann , von dem man nur Kopf und Brust sieht.
Sein Gesicht ist ein wenig faltig.
Er hat kurze Haare und einen Seitenscheitel auf der linken Kopfseite.
Sein Blick ist auf den Mäuseschwanz gerichtet.
Seine Hand macht eine greifende Bewegung und befindet sich (auf dem Bild) genau unterhalb des Mäuseschwanzes.
Er trägt eine Krawatte.
Sein Kopf ist im rechten oberen Bildsechstel , sein Kinn reicht gerade über die verlängerte Mausschwanzlinie, die das Bild in zwei etwa gleich große Hälften teilt.
Der Mann greift mit Daumen, Zeige- und Mittelfinger, die anderen beiden Finger sind an die Handfläche angewinkelt.
Die Krawatte des Mannes ist von rechts oben nach links unten gestreift, die Streifen sind auf dem Bild ca. 1 cm breit.
Man sieht drei Streifen.
Der Mann trägt ein weißes Hemd.
Auch die Maus ist weiß.

Cue words or phrases = Wörter oder Konstruktionen, die auf Metatext-Passagen hindeuten und so Anzeichen für die Kohärenzstruktur sind.

+

Indicator Phrase	Quality Score	argumentative or rhetorical roles
we argued	2	BACKGROUND TOPIC RELATED WORK PURPOSE/PROBLEM SOLUTION RESULTS CONCLUSIONS/CLAIMS, (nach Teufel und Moens)
we have argued	1	
we have argued that	1	
we will argue	1	
what I have argued is	1	
what we have argued is	1	
This article	3	
in this article	3	
is an attempt to	1	
I attempt to	2	
I have attempted	2	
I have attempted to	2	
our work attempts	2	
the present paper is an attempt to	2	
supported by grant	-1	

Ein letztes Beispiel für Textzusammenfassung:

Mrs. Coolidge: "What did the preacher preach about?"

Coolidge: "Sin."

Mrs. Coolidge: "What did he say?"

Coolidge: "He's against it."

A 3.1 Grundlagen sogenannter Expertensysteme

Lernziel: Der Unterschied zwischen wissensbasierten Systemen und konventioneller Programmierung soll klar werden.

Arten von Informationssuche und Informationsverarbeitung

- Dokumentation, "document retrieval", automatische Indexierung und Klassifikation (A2.1)
- Texterschließung und Textbearbeitung, Zusammenfassung, Abstract-Erzeugung (A 2.2)
- Fragebeantwortung, Problemlösung, "fact retrieval" (A 3.1 und A 3.2)

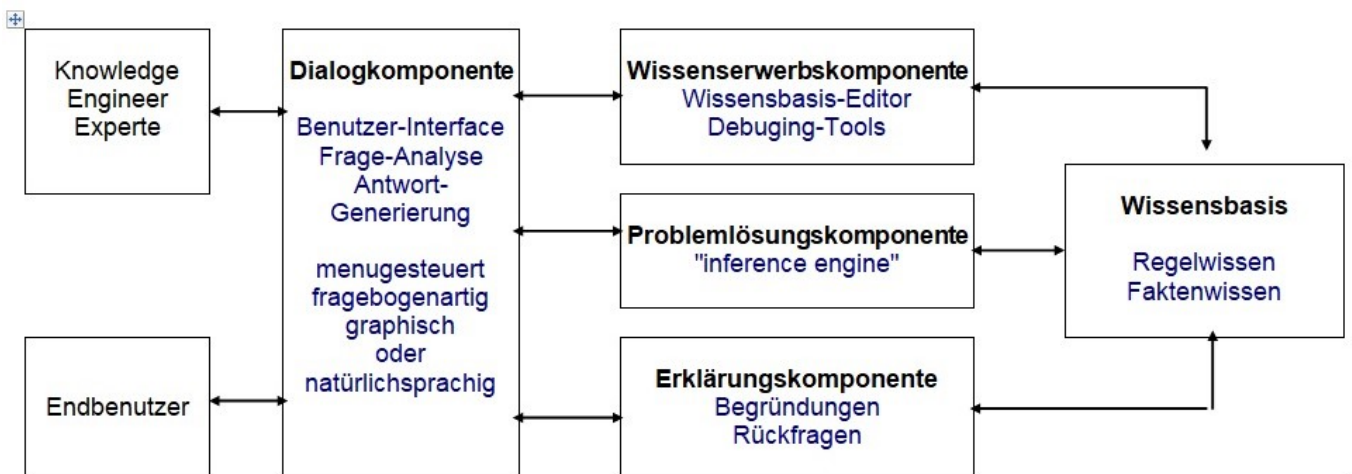
Expertensystem

Ein Programm, mit dem das Spezialwissen und die Schlussfolgerungsfähigkeit von Experten zur Lösung von Problemen auf einem bestimmten Arbeitsgebiet nachgebildet werden sollen.

Das ist etwas ganz anderes als eine Datenbank.

Expertensysteme zählen zum Kernbereich der KI (Künstliche Intelligenz-Forschung)

Typische Bestandteile eines Expertensystems



"Shell" = die Komponenten und ihre Kommunikation noch ohne Daten - inhaltsleer
man kann verschiedene konkrete Expertensysteme daraus machen

Hauptcharakteristikum wissensbasiert

Der Problemlösungsweg ist nicht durch den Algorithmus des Programms fest vorgegeben, sondern hängt von den Daten ab, die als Regelwissen in der Wissensbasis gespeichert sind. Das System findet den Lösungsweg selbst.

Ein **wissensbasiertes System** ist ein Softwaresystem, bei dem das Wissen über die Lösung eines Problems von den üblichen Kontrollstrukturen algorithmischer Programmiersprachen (Sequenz, Selektion und Iteration) getrennt ist.

Ein Expertensystem gibt nicht nur feststehende Fakten aus, sondern kann neue Fakten erschließen:

- z.B. Routenplaner, der die Route anhand einer Karte berechnet,
- ein System, das aufgrund von Patientendaten eine Diagnose erstellt und eine Therapie vorschlägt
- ein System das anhand von meteorologischer Daten eine Wettervorhersage erstellt.

Zum größten Teil wird logisches Schlussfolgern (WENN-DANN Regeln) eingesetzt. Aber auch analoge Verfahren, wie Graphiken, Landkarten, sonstige Modelle

Experte

Jemand, der auf einem Spezialgebiet Fachwissen in Form von Fakten und Regeln hat, vor allem aber auch jemand der *Erfahrung* hat und daher nicht stur alle Fakten und alle Regeln durchgehen muss.

Erfahrungswissen besteht aus:

- Heuristiken
- Assoziationsvermögen
- Urteilsfähigkeit auf Grundlage persönlicher Entscheidungskriterien
- experimentelle Erfahrung

Expertensysteme enthalten ebenfalls Heuristiken, zum einen wo das Wissen diffus ist, z.B. bei Krankheitssymptomen, oder wo die algorithmische Lösung zu umfangreich wird, z.B. bei Schach.

Erklärungskomponente

Expertensysteme können Auskunft geben

- welche Hypothesen sie gerade verfolgen,
- warum sie einen eingeschlagenen Lösungsweg gewählt haben,
- zu welchen Schlussfolgerungen sie bereits gelangt sind ,
- warum sie gerade zu diesen Schlussfolgerung gekommen sind.

Expertensysteme können fehlendes Wissen durch gezielte Rückfragen an den Benutzer selbst anfordern.

Wissenserwerbskomponente

Knowledge Engineering = das Wissen aus Wissensquellen wird abgebildet auf die Wissensbasis des Expertensystems

Wissensquellen sind

- menschliche Experten,
- Fachbücher, Fachartikel,
- Messdaten, usw

In diesem Zusammenhang auch Bemühungen, *Lernen* zu simulieren.

Beispiel

MYCIN - medizinisches Beratungssystem für Ärzte

Mitte der 70er an der Stanford University entwickelt

Beispiel: MYCIN, Sitzungsablauf

- Daten über einen Patienten werden erhoben: Name, Alter, Geschlecht, etc., Ergebnisse aus bereits durchgeführten Untersuchungen und Befunde.
- Mycin generiert Annahmen über mögliche Erreger. Dazu werden vom Arzt weitere Informationen oder Einschätzungen verlangt.
- Mycin erstellt eine Diagnose (nach ca. 30-90 Fragen).
- Mycin erstellt einen Therapievorschlag, der die Risiken für den Patienten minimiert.

Mögliche Antworten des Arztes:

<i>unb</i>	-	unbekannt
<i>?</i>	-	gib mir Beispiele für die hier möglichen Antworten
<i>??</i>	-	nenne mir alle möglichen Antworten
<i>Regel</i>	-	nenne mir die Regel, die verwendet wurde
<i>FA</i>	-	ich möchte nun beliebige Fragen stellen und deshalb in den Frage - Antwort Modus gehen
<i>Warum</i>	-	erkläre mir, warum die letzte Frage an mich gestellt wurde
<i>zurück x</i>	-	laß uns bei Frage x noch einmal beginnen
<i>Stop</i>	-	die Konsultation ist sofort abzubrechen

Beispieldialog

Bitte beantworten Sie die folgenden Fragen:

1. Name

Robert Mustermann

2. Geschlecht?

männlich

3. Alter?

21

Ich beziehe mich jetzt auf die aktuelle Kultur, für die Sie einen therapeutischen Rat wollen und nenne sie ab jetzt: KULTUR_1

4. Von wo wurde der Abstrich KULTUR_1 genommen?

Blutt

=Blut Fehlertoleranz!

5. Vor wievielen Tagen wurde die Blutkultur angelegt?

2

Ich werde mich jetzt auf den ersten Organismus dieser KULTUR_1 beziehen und ihn nennen: ORGANISMUS_1

6. Um welchen Organismus handelt es sich bei ORGANISMUS_1?

unb

7) Wie ist die Gramfärbung des ORGANISMUS_1?

?

Erwartete Antworten sind: säurebeständig grampos gramneg

gramnegativ

8. Handelt es sich bei diesem ORGANISMUS_1 um Stäbchen - oder Coccusbakterien ?

Stäbchen

9. Wurden noch andere signifikante Erreger (die aber nicht mehr bekämpft werden müssen) innerhalb der letzten ca.

30 Tagen erkannt? ja.

19. Von woher sind die Erreger vermutlich an diesen Ort gekommen?

Regel

Regel 121 Wenn die Stelle, von der die Kultur genommen wurde, normalerweise steril ist und die Stelle, an der der Erreger eingedrungen ist, dieselbe Stelle ist, an der vorher ein Erreger eingedrungen war und dieser vorige Erreger dieselbe Identität wie der jetzige Erreger hat dann ist anzunehmen (0.3), daß alle Erreger, die gefunden wurden, dieselbe Identität haben.

26. Wurden an Patient_1 irgendwelche chirurgischen Eingriffe oder andere Eingriffe am Blasen-Trakt vorgenommen?

Warum

Ich versuche, die Identität des ORGANISMUS_1 zu finden. Ich weiß, dass

- die Stelle an der der Erreger eingedrungen ist: obere Harnwege,
- die Morphologie des Organismus: Stäbchenbakterie,
- die Gramfärbung des Organismus: gramnegativ,
- die Kultur wurde angelegt: vom Blut.

Deshalb, wenn der Patient keine Eingriffe an den Harnwegen hatte und der Patient nicht wegen einer Harnwege -Infektion

behandelt wurde, dann ist anzunehmen (0.6), daß die Identität des Organismus E.coli ist. (Regel 156)

Ziele beim Einsatz von Expertensystemen

- Erhalt von Expertenwissen
- Unterstützung angelernter Laien und/oder von Experten
- Verringern des Bedarfs, „teure“ Experten zuziehen zu müssen
- Standardisierung von Abläufen

Oft ist ein Expertensystem letztlich eher ein Assistentensystem oder ein Tutorsystem.

Charakteristik der Probleme, bei denen typischerweise Expertensysteme eingesetzt werden:

- Erfolgsrate < 100% ist akzeptabel
(=> Mensch spielt trotz allem eine wichtige Rolle im Lösungsprozess)
- Probleme, deren Randbedingungen sich schnell ändern
(=> Systeme sollen flexibel bleiben)

Was kann die Computerlinguistik beitragen?

Dialog: und Erklärung:

Momentane Systeme können noch nicht (gut)

- vorherige Erklärungen weiter ausführen,
- Missverständnisse erkennen und die Erklärung an der richtigen Stelle wieder aufnehmen,
- auf Folgefragen im Dialogkontext antworten, sie nicht als unabhängige Fragen nehmen

Wissensaquisition:

- Wissen aus Texten, Analyse, Repräsentation

Problemlösung:

- begriffliche Gliederung der Welt im Wortschatz (z.B. Raumausdrücke bei GIS)
- flexible Repräsentation jeglichen Wissens (Folgern in natürlicher Sprache)

A 3.2 Natürlichsprachige Frage-Antwort-Systeme (FAS)

Lernziel: Ein paar Eindrücke

Allgemein besteht Mensch-Computer-Interaktion aus vielerlei Ein- und Ausgabe (Meldungen, Eingabeaufforderungen usw.) . Das Programm kann an bestimmten Stellen natürlichsprachige Ausgaben erzeugen oder auch bestimmte Eingaben erwarten, z.B. bestimmte Befehle. Dies ist auch in gesprochener Sprache möglich, wenn ein Spracherkenner zwischengeschaltet wird.

All dies dies ist sozusagen "eingemachter Text" (canned text) und damit kein Problem.

Engere Definition

Ein natürlichsprachiges FAS ist ein Expertensystem mit natürlicher Sprache als Wissensrepräsentation

Pro

- unübertroffene Ausdrucksfähigkeit natürlicher Sprache bei komplizierten Sachverhalten
- konzeptuell umfassend und flexibel
- ökonomisch

Contra

- zu vage und mehrdeutig
- viel zu schwierig und daher zu zeitaufwendig zu verarbeiten

Meine Antwort

Im Experimentierfeld Frage-Beantwortung können viele semantische Zusammenhänge der Sprachverwendung studiert werden. Ähnlich wie die automatische Übersetzung sind FAS daher schon allein aus theoretischen Gründen interessant.

Beispiel SHRDLU

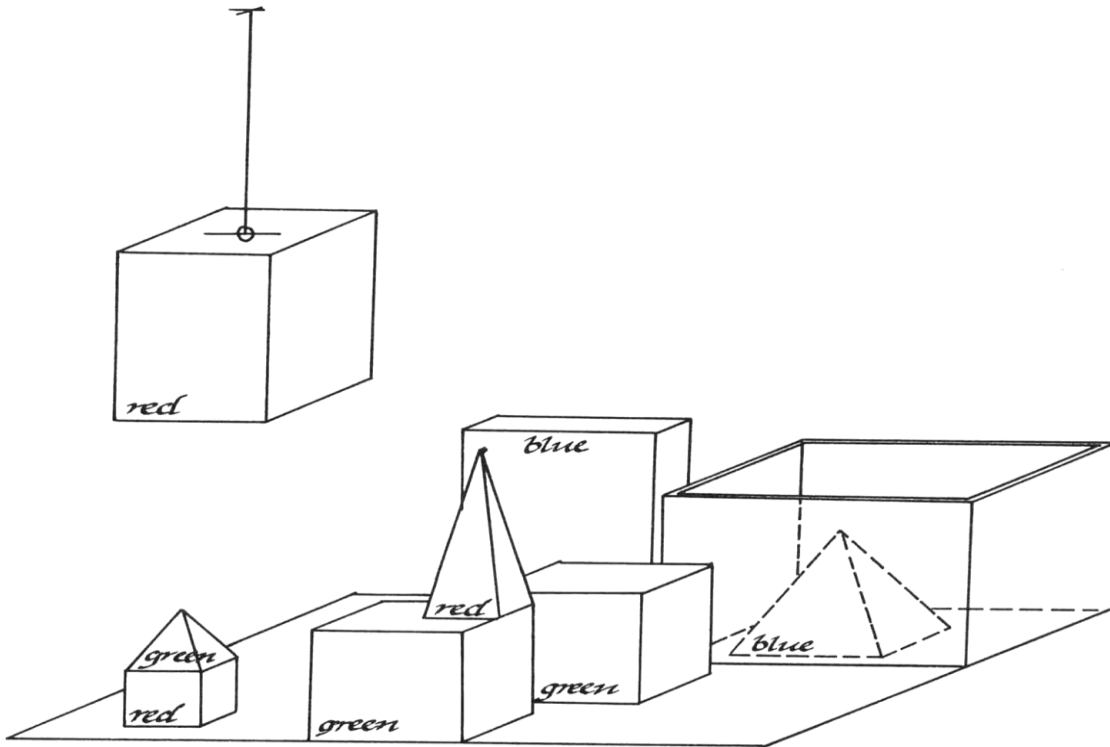
Das Robotersystem von Terry Winograd, Massachusetts Institute of Technology (MIT), 1971
Es hat den Boom der Künstlichen Intelligenz in den Siebzigern und Achtzigern inspiriert.

Gegenstandsbereich: Tischoberfläche mit Klötzchen verschiedener Farbe, Größe und Form
Szene

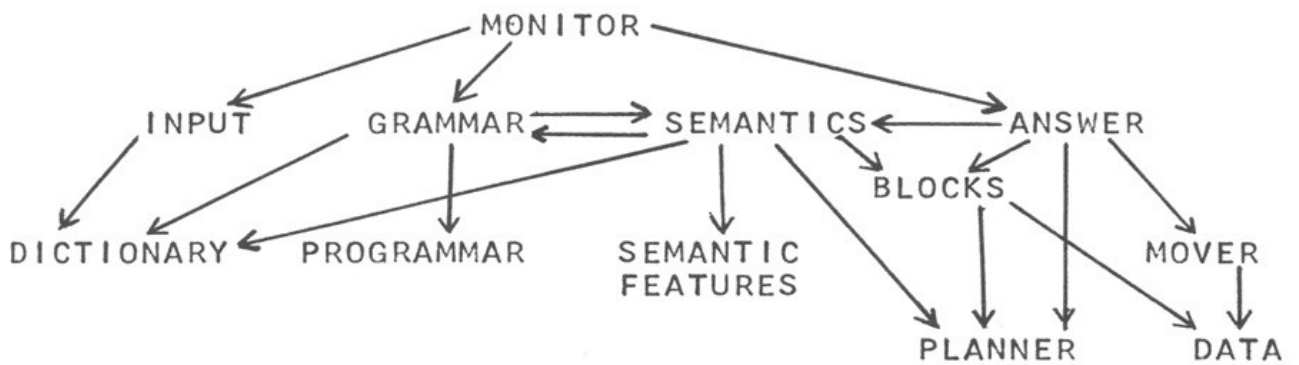
:

Pick up a big red block.

OK.



Programm-Architektur:



MONITOR koordiniert alle Komponenten.

INPUT liest Eingabe, schaut im **DICTIONARY** nach.

GRAMMAR analysiert Eingabe grammatisch mittel der Grammatik im Formalismus **PROGRAMMAR**

SEMANTICS weist Elementen in der Eingabe **SEMANTIC FEATURES** zu und setzt sie in Beziehung zur Blockwelt in **BLOCKS**, ruft die Deduktionskomponente **PLANNER** zur Problemlösung

BLOCKS enthält das physikalische Wissen über die Klötzchenwelt in Regelform

DATA enthält Fakten über die aktuelle Konstellation der Klötzchen

ANSWER erzeugt in Zusammenarbeit mit **PLANNER** auf die Blockwelt **BLOCKS** bezogene Antworten, wobei ein Dialoggedächtnis eingesetzt wird

PLANNER erzeugt Folgerungen mithilfe der Regeln in **BLOCKS**

MOVER führt Befehle aus, wodurch sich **DATA** ändert.

BEISPIELDIALOG

[...]

--How many blocks are not in the box?

Four of them.

--Is at least one of them narrower than the one which I told you to pick up?

Yes, the red cube.

[...]

--Can the table pick up blocks?

No.

--Can a pyramid be supported by a block?

Yes.

--Can a pyramid support a pyramid?

I don't know.

--Stack up two pyramids.

[The robot tries and fails.]

I can't.

[...]

--Does a steeple

Sorry, I don't understand the word 'steeple'.

[Shrdlu interrupts as soon as he encounters a word that he doesn't know.]

--A 'steeple' is a stack which contains two green cubes and a pyramid.

I understand.

--Are there any steeples now?

No.

--Will you please stack up both of the red blocks and either a green cube or a pyramid.

OK.

[...]

--When did you touch a pyramid?

When I was stacking up a red cube, a large green cube and a large red block.

--Why?

To get rid of it.

--Why did you do that?

To clear off the red cube.

--Why did you clear off that cube?

To put it on a large green cube.

--Why did you do that?

To stack up a red cube, a large green cube and a large red block.

--Why did you do that?

Because you asked me to.

Ein anderes berühmtes natürlichsprachiges Dialogsystem:

ELIZA von Josef Weizenbaum

Online ausprobieren unter <http://www-ai.ijs.si/eliza/eliza.html>

Dialog

HELLO. I AM ELIZA. HOW CAN I HELP YOU?

Men are all alike.

IN WHAT WAY?

They're always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE? ?

Well, my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME

He says I'm depressed much of the time.

I AM SORRY TO HEAR YOU ARE

...

Funktionsweise von ELIZA

- Eingabe wird nach Schlüsselwörtern und minimalem Kontext durchsucht
(z.B. *I - am, you - are*)
- mit Schlüsselwörtern sind Regeln verbunden, welche einen Teil der Eingabe verwenden und eine Ausgabe erzeugen
- wird kein Schlüsselwort gefunden, wird eine in jedem Kontext verwendbare Äußerung ausgegeben oder es wird auf vorherige Aussagen zurückgegriffen
- Schlüsselwörter sind gewichtet; das höherwertige wird genommen; Gewichtung kann sich dynamisch ändern.
- Komma oder Punkt führen dazu, dass der Text rechts oder links davon ignoriert wird, je nachdem welche Seite kein oder ein geringer gewichtetes Schlüsselwort enthält; so wird immer nur ein Satz zur Ausgabe umgeformt
- Die Schlüsselwörter plus Regeln bilden ein sog. Skript, was ELIZA interpretiert; man kann das Skript also austauschen, so dass ELIZA anders reagiert oder eine andere Sprache spricht.
- Das Skript kann (könnte?) im Gespräch fortgeschrieben werden, so dass ELIZA zu "lernen" scheint.

A.4.1 Computer in der Lehre und Ausbildung, "e-Learning"

CAI: Computer-aided Instruction

Lernziel: Aufmerksamkeit für eine der möglichen Anwendungen computerlinguistischer Technologie, zugleich Offenheit dafür, selbst elektronische Medien beim Studieren zu benutzen

Einteilung der CAI-Software:

Eine Systematik von Bertram C. Bruce und James A. Levin von der University of Illinois

A. Media for Inquiry

1. *Theory building--technology as media for thinking*

- Model exploration and simulation toolkits
- Visualization software
- Virtual reality environments
- Data modelling--defining categories, relations, representations
- Procedural models
- Mathematical models
- Knowledge representation: semantic network, outline tools, etc.
- Knowledge integration

2. *Data access--connecting to the world of texts, video, data*

- Hypertext and hypermedia environments
- Library access and ordering
- Digital libraries
- Databases
- Music, voice, images, graphics, video, data tables, graphs, text

3. *Data collection--using technology to extend the senses*

- Remote scientific instruments accessible via networks
- Microcomputer-based laboratories, with sensors for temperature, motion, heart rate, etc.
- Survey makers for student-run surveys and interviews
- Video and sound recording

4. *Data analysis*

- Exploratory data analysis
- Statistical analysis
- Environments for inquiry
- Image processing
- Spreadsheets
- Programs to make tables and graphs
- Problem-solving programs

B. Media for Communication

1. Document preparation

- Word processing
- Outlining
- Graphics
- Spelling, grammar, usage, and style aids
- Symbolic expressions
- Desktop publishing
- Presentation graphics

2. Communication--with other students, teachers, experts in various fields, and people around the world

- Electronic mail
- Asynchronous computer conferencing
- Synchronous computer conferencing (text, audio, video, etc.)
- Distributed information servers like the World-wide Web
- Student-created hypertext environments

3. Collaborative Media

- Collaborative data environments
- Group decision support systems
- Shared document preparation
- Social spreadsheets

4. Teaching Media

- Tutoring systems
- Instructional simulations
- Drill and practice systems
- Telementoring

C. Media for Construction

- Control systems--using technology to affect the physical world
- Robotics
- Control of equipment
- Computer-aided design
- Construction of graphs and charts

D. Media for Expression

- Drawing and painting programs
- Music making and accompaniment
- Music composing and editing
- Interactive video and hypermedia
- Animation software
- Multimedia composition

Klassifikation von CAI Programmen

(nach T.O'Shea, J. Self. Learning and teaching with computers, 1987)

Lineare	Einfachste Programme, predeterminierter Programmablauf
Verzweigte (branching)	Mehrere Varianten des Programmablaufs, die an die Antworten von Studenten angepasst werden.
Generative	Programm generiert die Aufgaben und macht die Musterlösungen, die mit Lösungen vom Studenten verglichen werden
Simulationen	Modellierung realer Situationen, Computer als Labor, interaktive Graphiken
Spiele	Spielform, audiovisuelle Effekte
Dialogsysteme	Komplexe Wissensrepräsentationen, Interaction in der natürlichen Sprache, Computer als Tutor und Expertensystem

Vorteile von Lernprogrammen im Vergleich mit anderen Lernverfahren

- Planung des Stoffes, Zerlegung in Schritte
- Anpassung an den einzelnen Schüler, eigene Lerngeschwindigkeit, eigene Aufgaben, gezielte Wiederholung
- hält Schüler an, aktiv zu sein (Passivität wie im Frontalunterricht unmöglich)
- feed back, Kontrolle des Schülers über den eigenen Fortschritt
- Erfolgskontrolle für den Lehrer, leichtere Verwaltung der Lehre
- Vernetzung, leichte Verfügbarkeit der Materialien

Nachteile

- Programme gehen höchstens scheinbar auf den einzelnen ein, in Wirklichkeit Entpersönlichung
- mangelnde Robustheit der Programme

A 4.2 Sprachlernsoftware

CALL: Computer-aided Language Learning

Verschiedene Typen von Sprachlernsoftware

- Drill-Programme (multiple choice, Lückentexte, Vokabelabfrage)
- Simulation von Sprachgebrauch in Kommunikationssituationen
- Mensch-Maschine-Interaktion, der Computer als Tutor (ICALL = Intelligent CALL)
- sonstige?

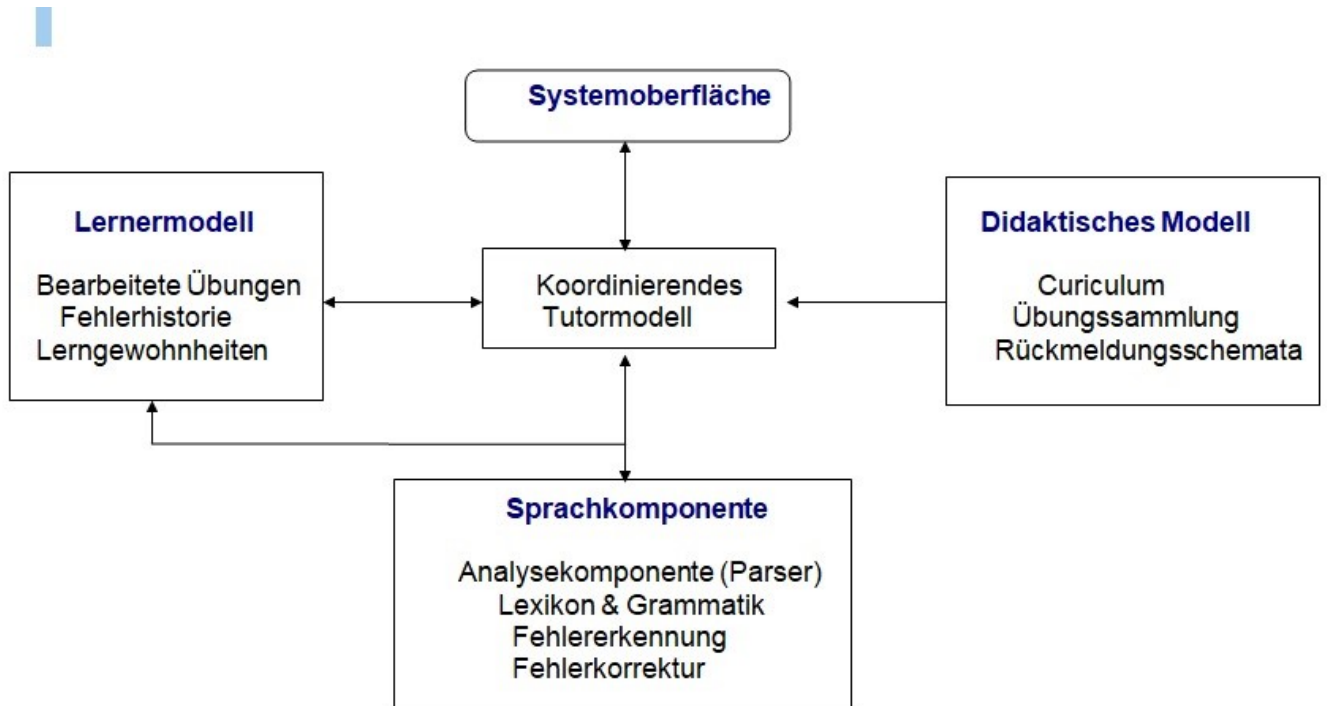
Voraussetzungen für CALL-Programme

- CAI nebst ihren Hilfswissenschaften
- Angewandte Linguistik, Psycholinguistik: Theorie des Spracherwerbs (Zweitsprache, Fremdsprache), Sprachdidaktik
- Text- und Dialoglinguistik, Gesprächsanalyse
- Computerlinguistik: Einsatz von Programmen zur Verarbeitung der verschiedenen Komponenten des Sprachsystems (NLP - natural language processing), bis hin zum kompletten Sprachverstehen

Probleme

- vor allem die Prüfung und Korrektur der Schülereingabe bei komplexeren Aufgaben (reines Auflisten möglicher Antworten unpraktikabel)
daher Einsatz von computerlinguistischen Techniken
- neben korrekten Eingaben auch Fehler erkennen (theoretisch fast ein Widerspruch), Fehlertypen unterscheiden, Fehler korrigieren, Fehler erklären

Architektur eines CALL-Systems



K 1 Wissenschaftstheorie und Methodologie der CL

Lernziel: Reflexion darüber, wie wissenschaftliches Herangehen an computerlinguistische Aufgaben aussehen sollte.

Computerlinguistik = eine empirische, theoretische und anwendungsorientierte Wissenschaft

Wissenschaftstheorie

Untersuchung der Voraussetzungen, Methoden und Ziele der Wissenschaften

Vorab Erinnerung an die Syllogismen

Schlussfigur

wenn a dann b

a

also b



MAJOR-PRÄMISSE

MINOR-PRÄMISSE

KONKLUSION

Die Major-Prämisse ist sozusagen eine (Mini) Theorie, die erlaubt von a auf b zu schließen.

Deduktion

Deduktion ist die Anwendung einer Theorie, nämlich aufgrund der Majorprämisse *wenn a dann b* aus vorliegenden Untersuchungen, *dass a der Fall ist*, zu prognostizieren, *dass b der Fall sein* wird.

Induktion

Induktion heißt, an empirischen Daten eine Gesetzmäßigkeit zu erkennen und diese zu beschreiben, also im einfachsten Fall eine Majorprämisse aufzustellen. Die Beschreibung von Gesetzmäßigkeiten nennt man dann eine Theorie.

Der Zugang zur Welt

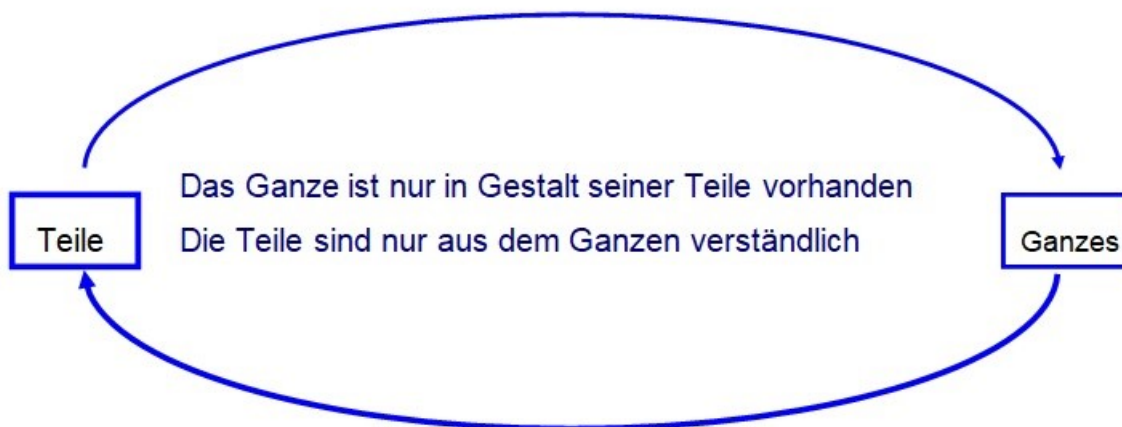
Positivismus versus Hermeneutik ?

Hermeneutik = Kunst der Auslegung, der Interpretation

- Erkenntnisform, die auf **Erfassung von Sinn** aus ist.
- Der Zugang zum Erkenntnisobjekt ist **das Verstehen**.
- Postulat einer spezifisch geisteswissenschaftlichen Methode (im Gegensatz zu den Naturwissenschaften)
- Betonung der Geschichtlichkeit des Menschen in seiner Lebenswelt
- ganzheitliche Sicht, Verstehen von (Lebens-) Äußerungen des Menschen (etwa in der Kunst) im Ganzen seines (Welt) Horizontes ("Weltanschauung")

Schleiermacher: *"Die Natur erklären wir, das Seelenleben verstehen wir."*

Der hermeneutische Zirkel:




Positivismus = Wissenschaft soll sich auf das Positive, Tatsächliche, Wirkliche und Zweifellose beschränken.

- Ziel ist die **Erklärung** des Gegenstandsbereichs durch Aufstellung von Theorien, Gesetzen und Hypothesen
- Als Basis für wissenschaftliche Erkenntnis sind nur Tatsachen zugelassen.
- Tatsachen müssen sich objektiv erkennen oder beweisen lassen.
- **Verifizierbarkeitsprinzip:** Eine Aussage ist sinnvoll, wenn und nur wenn man sie prinzipiell durch Erfahrung als wahr oder falsch erweisen kann.
- Metaphysische Aussagen sind (per definitionem) theoretisch unmöglich und praktisch Unsinn.

Sog. Positivismusstreit zwischen Theodor W. Adorno und Karl Popper (1961)

Ich mag unrecht
haben und Du magst recht haben; und so
wann wir uns bemühen, dann können wir
zusammen vielleicht der Wahrheit etwas
näher kommen.



Karl Popper 1902 - 1994

Kritische Rationalismus (Popper) :

- Allgemeine Theorien sind nicht logisch zwingend aus Erfahrungen abzuleiten. Jede Induktion gilt nur vorläufig bis zum Beweis des Gegenteils.
- Eine Theorie kann aber durch Erfahrungen widerlegt werden. Prinzip der Falsifikation (im Gegensatz zu dem Verifikationsprinzip der Positivisten).
- Eine Aussage, die sich prinzipiell nicht falsifizieren läßt, ist nicht sinnlos, aber unwissenschaftlich.
- Also Verzicht auf einen absoluten Wahrheitsanspruch.

Es gibt Wahrheit, aber der Mensch kann sich nie sicher sein, ob er sie gefunden hat. Erfahrung ist subjektiv, möglicherweise ungenügend. Fehlersuche wird zum Prinzip erhoben. Damit erreichen wir zwar nie endgültige Wahrheiten - denn wir werden auch die neue bzw. modifizierte Theorie zu widerlegen suchen - aber wir nähern uns der Wahrheit an.

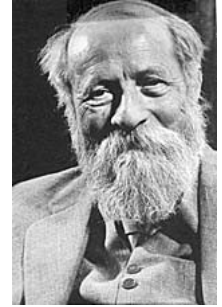
- ✓ Die Falsifikationstheorie ist eine konsequent antidogmatische und antikonservative Strategie..

Berechtigung sowohl des hermeneutischen wie rationalistischen Vorgehens .

"Ich und Du" von Martin Buber, Erstauflage 1923

"DEN MENSCHEN, zu dem ich Du sage, erfahre ich nicht. Aber ich stehe in der Beziehung zu ihm, im heiligen Grundwort Ich-Du. Erst wenn ich daraus trete, erfahre ich ihn wieder. Erfahrung ist Du-Ferne. (...)"

Martin Buber 1878-1965



Nach Buber gibt es zwei unterschiedlichen Arten, der Welt gegenüberzutreten

- **Die Ich-Es Haltung** vergegenständlicht die Welt
Ich sehe einen Baum als soundsoviel Kubikmeter Holz
Ich sehe in einem Menschen Haut und Pickel.
- **Die Ich-Du Haltung** tritt mit dem Gegenüber in eine Beziehung ein, die auf das Wesen gerichtet ist.
Der Baum macht auf mich als Ganzes einen erhebenden oder erschreckenden Eindruck.
Ich begegne einer Person als attraktiver Frau und Freundin.

„Das einzelne Du muß, nach Ablauf des Beziehungsvorgangs, zu einem Es werden. Das einzelne Es kann, durch Eintritt in den Beziehungsvorgang, zu einem Du werden.“ S. 37:

Beides unverzichtbar.

Der Verstand ist nicht alles , das Leben besteht nicht nur aus Wissenschaft. Aber in bestimmten Situationen soll man sich besser auf den kritischen Verstand verlassen, als auf Gefühl.

Theorie und Praxis

werden gerne gegeneinander ausgepielt. Aber

Gibt es überhaupt Praxis ohne Theorie? (oder nur Praxis mit unreflektierter Theorie?)

Und gibt es wirklich zweckfreie Wissenschaften?

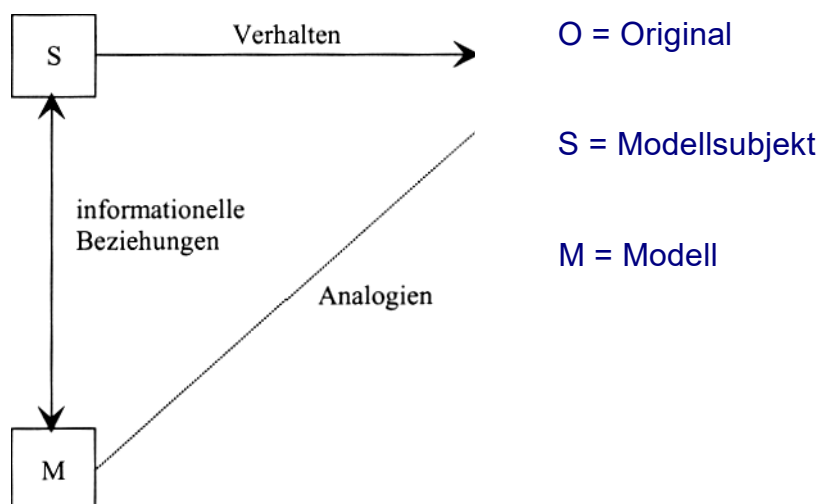
"Nothing is so practical as a good theory, and there are few better theoretical challenges than a precise practical problem, empirically unsolvable by mere practicalities."

Hans Karlgren

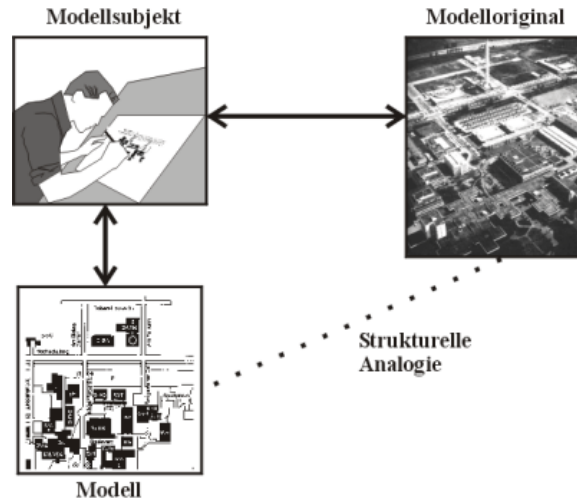
Wissenschaftliche Methodik

- **Fragestellung**, Interesse
- **Empirie**, Beobachtung, Beschreibung der Phänomene
- **Abstraktion**, nur die in Bezug auf die Fragestellung relevanten Eigenschaften werden betrachtet
- **Reduktionismus**, minimale aber noch ausreichende Parameter
- **Generalisierung**, Regel, die auf alle Einzelfälle paßt (dabei Variablen, Klassenbildung) Vorhersage (Hypothese)
- **Überprüfung**, Falsifizierung, Revision
Experiment (Empirie auf der Basis von schon vorhandenen Hypothesen)
- Generalisierung auf mehreren Stufen, Verknüpfung von Hypothesen, allgemeinere Prinzipien, wieder Reduktionismus, die "Weltformel"

Die Modellmethode



Auch Modell im Kopf (M)
versus Wirklichkeit (O)



Karl-Heinz Wagner, Uni Bremen, Grundkurs Sprachwissenschaft

Definition Modell

Ein MODELL, ist ein Objekt (ein Gegenstand, ein materielles oder symbolisches System), das auf der Grundlage einer Struktur-, Funktions- oder Verhaltensanalogie zu einem entsprechenden Original von einem Subjekt (z.B. einem Menschen) eingesetzt und genutzt wird, um eine bestimmte Aufgabe lösen zu können, deren Durchführung mittels direkter Operationen am Original zunächst oder überhaupt nicht möglich bzw. unter gegebenen Bedingungen zu aufwendig ist.

Die Funktion des Modells ergibt sich im Rahmen eines aus Subjekt (S), Original (O) und Modell (M) bestehenden MODELLSYSTEMS in Abhängigkeit von der gegebenen Zielstellung des Subjekts (z.B. den Erkenntnisinteressen eines Forschers)...

(nach KLAUS/BUHR 1971, s.v. Modell)

Beispiele für Modelle

Streckennetzplan der Bahn

Linien und Punkte als Modell für die Bahnverbindungen zwischen Orten in der Wirklichkeit

Fahrplan

Zahlentabellen als Modell für die Abfahrts- und Ankunftszeiten von Zügen in der Wirklichkeit

Straßenkarte

Linien unterschiedlicher Farbe als Modell für Straßen unterschiedlicher Art in der Wirklichkeit

geologische Karte

Farben als Modell für die Bodenarten in der Wirklichkeit

Schaltplan der elektrischen Anlage eines Autos

Symbole und Linien als Modell für elektrische Apparate und die Leitungen zwischen ihnen in der Wirklichkeit

Planetarium

Apparatur mit verkleinerten Nachbildungen von wirklichen Himmelskörpern als Modell ihrer Bewegungen

Grundriss eines Hauses

Symbole und Linien als Modell für ein Haus in der Wirklichkeit

➤ **Ein Modell ist keineswegs ein vollständiges Abbild der Wirklichkeit.**

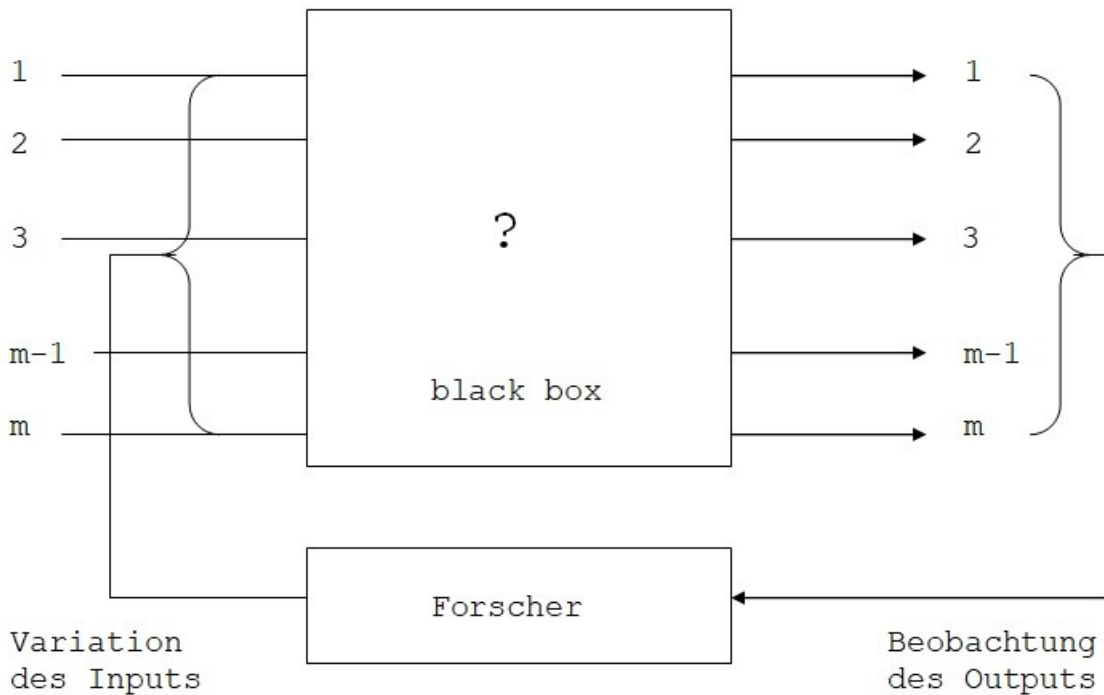
Ohne Abstraktion, ohne Beschränkung der Analogie auf die Fragestellung, wäre das Modell in der Praxis unbrauchbar. Viele nicht interessierende Daten würden den Blick verstellen.

➤ **Achtung! Die Analogiefalle.**

Manchmal werden Objekte aus der Wirklichkeit als Modelle für andere Objekte aus der Wirklichkeit genommen. Das kann fruchtbar sein. Ist aber auch gefährlich, weil unkontrollierte, möglicherweise nicht zutreffende Eigenschaften unversehens von einem auf das andere Objekt übertragen werden.

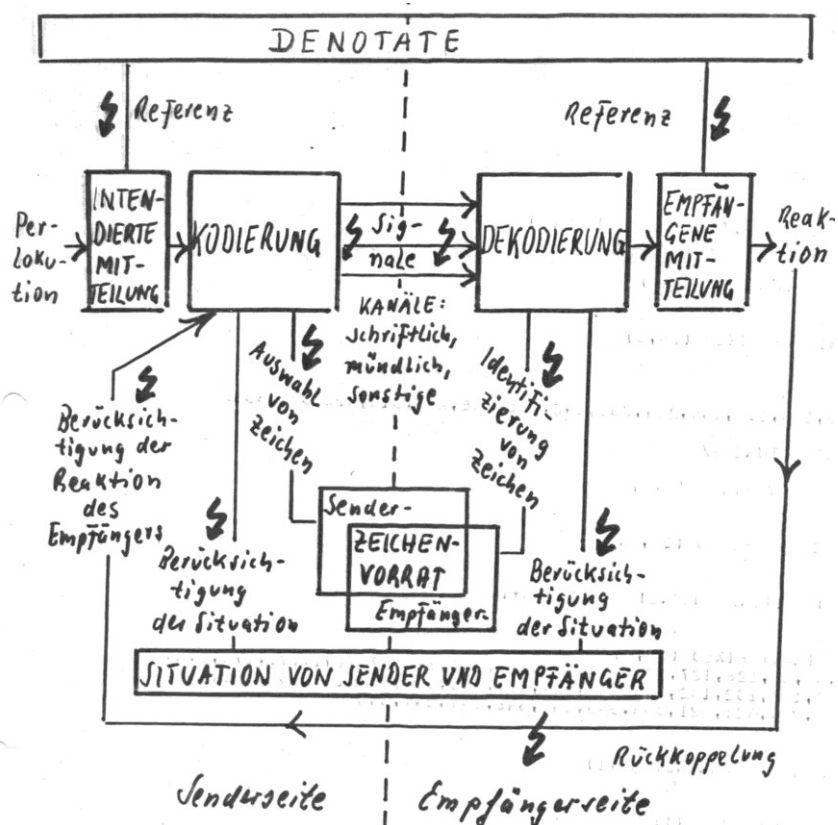
Vorsicht. Hinkende Vergleiche können auch zur Manipulation eingesetzt werden.

Black-box Methode



Es kann nur das Verhältnis von Einwirkung und Reaktion eines Untersuchungsobjekts beobachtet werden. Solche Modelle sind im besten Fall **beobachtungsadäquat** und **ergebnisgleich**. Das reicht, um in der Praxis Prognosen zu machen. Sie sind nicht **erklärungsadäquat** und **ablaufgleich**. Das heißt, sie führen nicht zum wirklichen Verstehen des Objekts.

Ein Kommunikationsmodell. Beispiel für Zerlegung in kleinere Boxen und Input-Output-Relationen zwischen ihnen:



Gültigkeit von Theorien

subjektiv - objektiv?

- Subjektivität lässt sich nicht eliminieren
- das Erkenntnisinteresse ist subjektiv
- auch schon die Datenaufnahme geschieht durch eine bestimmte Brille, selbst sog. objektive Messungen

Konsequenz: objektiv heißt intersubjektiv

- rekonstruierbare Operationen (vgl. Minimalpaaranalyse, heuristische Tests in der Syntax)
- Untersuchung so, daß man Gewährsleute befragen könnte, oder wirklich fragt
- operationale Definitionen aller Erhebungsdaten (Instruktionsbuch)

Bewertungskriterien für Theorien:

Explizitheit

Ist die Theorie so explizit formuliert, daß sie nachprüfbar und widerlegbar ist?

(Achtung: Immunisierungsstrategie durch verschwommene Sprache oder undefinierte Fremdwörter)

Widerspruchsfreiheit

Lässt sich ein Widerspruch aufdecken?

Adäquatheit

beobachtungsadäquat - sind die Daten korrekt und (bezüglich der Fragestellung) vollständig und machen die Regeln korrekte Vorhersagen über das Objekt?

(Objekt ist eine black box)

erklärungsadäquat - entsprechen die Regeln den inneren Zusammenhängen des Objekts und machen so das Objekt verständlich? (Objekt ist eine white box)

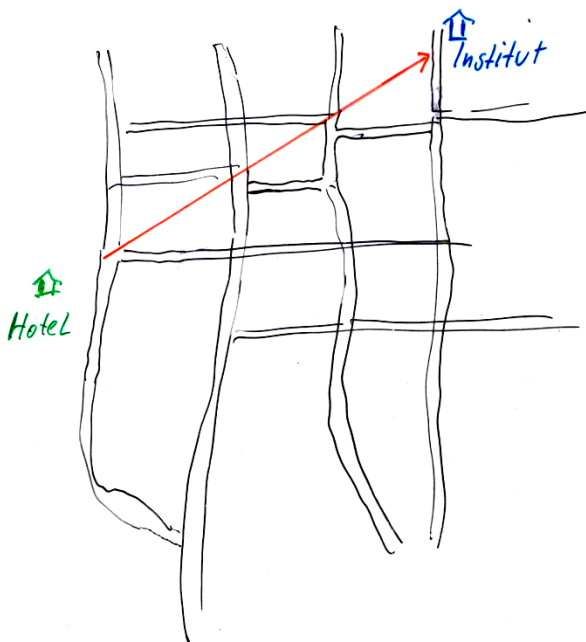
Einfachheit

"Unser Denken geht vom Naiven über das Komplizierte zum Einfachen" (A. St. Exupéry)

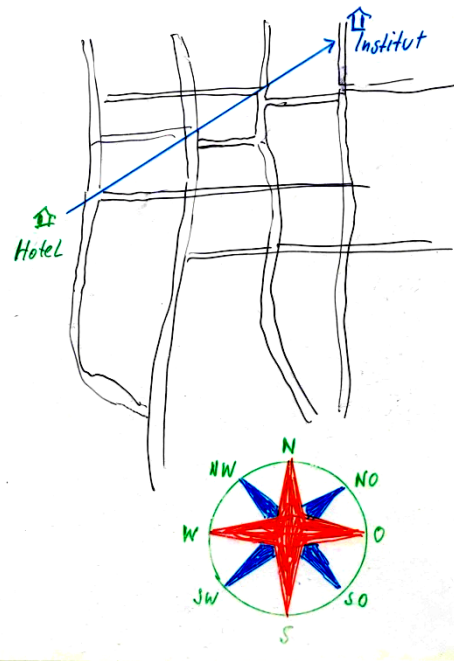
Ab einem bestimmten Stadium verbessert man eine Theorie, indem man wieder etwas wegnimmt. (Aber Achtung: Gefahr durch übermäßigen Reduktionismus)

Gibt es als Ziel die eine "wahre" umfassende Theorie, an der wir alle arbeiten, oder viele gleichberechtigte Theorien?

Theorie 1: Diagonale



Theorie 2: Richtung NO



Wie kommen wir vom Hotel Duomo zum Instituto di linguistica computazionale in Pisa?
Zwei Theorien: 1. Wir müssen diagonal zum Straßenverlauf gehen, d.h. immer rechts, links, rechts, links. 2. Wir müssen nach Nordosten gehen.

Welche Theorie ist im Moment praktischer? Welche global verwendbarer?

Die Blinden und der Elefant

Es war einmal, so erzählt Buddha, ein König von Benares, der rief zu seiner Zerstreung etliche Bettler zusammen, die von Geburt an blind waren und setzte einen Preis aus für denjenigen, der ihm die beste Beschreibung eines Elefanten geben würde. Zufällig geriet der erste Bettler, der den Elefanten untersuchte, an dessen Bein, und er berichtete, daß der Elefant ein Baumstamm sei. Der zweite, der den Schwanz erfaßte, erklärte, der Elefant sei wie ein Seil. Ein anderer, welcher ein Ohr griff, beteuerte, daß der Elefant einem Palmenblatt gleiche und so fort. Die Bettler begannen untereinander zu streiten, und der König war überaus belustigt

⁶Aus: Neumüller, G. und Niel, F. Gott und Gottesbilder, Reihe Konzepte 2 (Materialien für den Religionsunterricht in der Sekundarstufe 2), Verlag Moritz Diesterweg u. Köselverlag, 1977, S. 1.

Methodologie der Computerlinguistik

Besonderheiten bei Sprache als Gegenstand

Dichotomien Ferdinand de Saussure/Noam Chomsky:

langue/competence = potentiell, Sprache als abstraktes System bzw. als internalisiertes Wissen

parole/performance = aktuell, individuelles Vorkommen sprachlicher Ausdrücke



Noam Chomsky

Wie gewinnt man Daten?

- parole - belegorientiert, philologisch, Korpora, große Textmengen, Statistik
- langue - durch Introspektion, kognitiv, sein Sprachbewußtsein in Regeln umsetzen

Ein Gegensatz?

Die einen analysieren riesige Mengen von Sprachdaten die anderen schreiben Regeln. Mich interessiert nicht die riesige Menge der Daten, sondern darunter nur die Fälle, für die ich noch keine Regel geschrieben habe. Ich glaube, dass es bei Kleinkindern genauso ist. Aber man mag bei ausschließlich introspektivem Vorgehen vieles übersehen, besonders durch mangelnde Imagination von Sprechakt-Situationen. Also Sprachdatensammlungen (sog. Corpora) sind schon gut.

Besondere Klasse unter den Modellen:

Formale Modelle

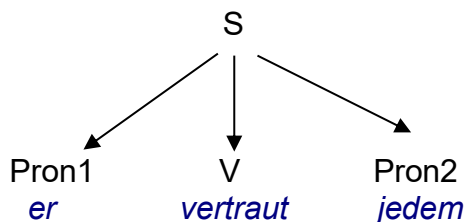
- Inventar von Symbolen
- Regeln = Symbolmanipulation nach Vorschrift (Rechnen)

Die natürliche Sprache besteht selbst aus Symbolen. Eine Kopie davon kann in ein formales Modell der Sprache inkorporiert werden, nämlich als sog. terminale Symbole..

So kann man mit dem Modell das Modellobjekt direkt ausrechnen! Rechnen mit Sprache.

Beispiel aus S 5.2

S	->	Pron1 + V + Pron2
Pron1	->	er
V	->	vertraut
Pron2	->	jedem

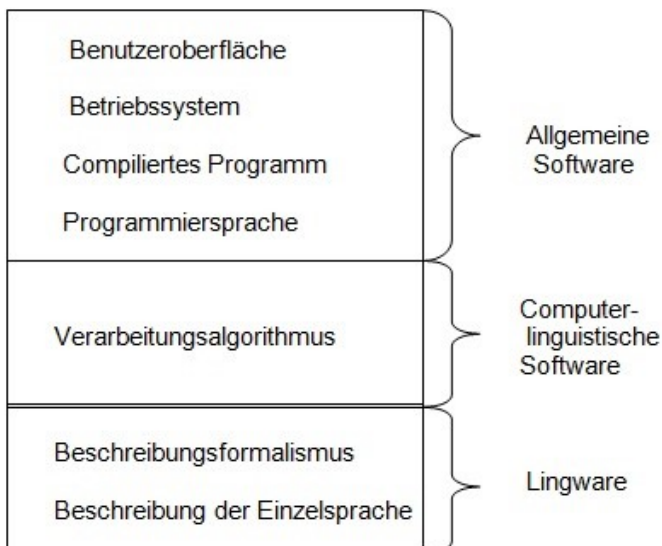


Gibt man die Symbole an den Endknoten des Strukturgraphen aus, bekommt man den objektsprachlichen Satz: *er vertraut jedem*, zu dem der Graph ein Modell ist.

Dadurch können formale Theorien bei uns extrem rigoros bewertet werden:

100% explizit, widerspruchsfrei, adäquat - zusätzlich möglichst einfach.

Hierarchie der Symbolsysteme:



Konkrete Arbeit

Zunächst Eindruck vom **Stand der Wissenschaft** gewinnen.

Literatur und Internetrecherche

aber nicht als Selbstzweck sondern **mehr als Anregung zum selber Forschen!**

Empirische Untersuchung der Phänomene, "Feldforschung"

ist ganz wichtig und macht Spaß

Segmentieren, Klassifizieren, Opposition, Distribution, Experiment

umfangreiche Beschreibung der Daten, Beispiele systematisch angeordnet, kommentiert

Regeln aufstellen

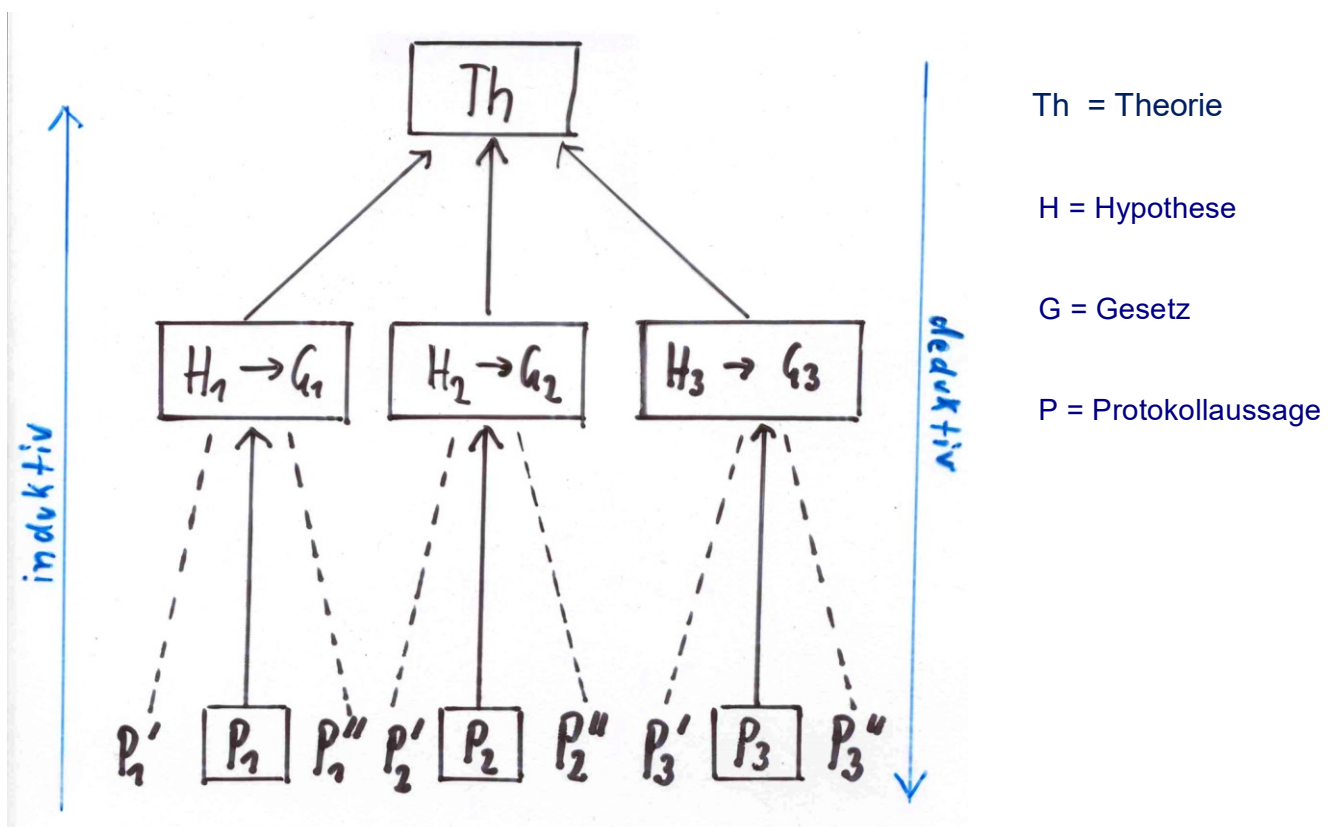
Von den Daten zu geeigneten Hypothesen zu gelangen ein kreativer Akt.

Erforderliches Talent: **Strukturen sehen!**

Testen

Weitere Daten heranziehen. Sind sie regelkonform?

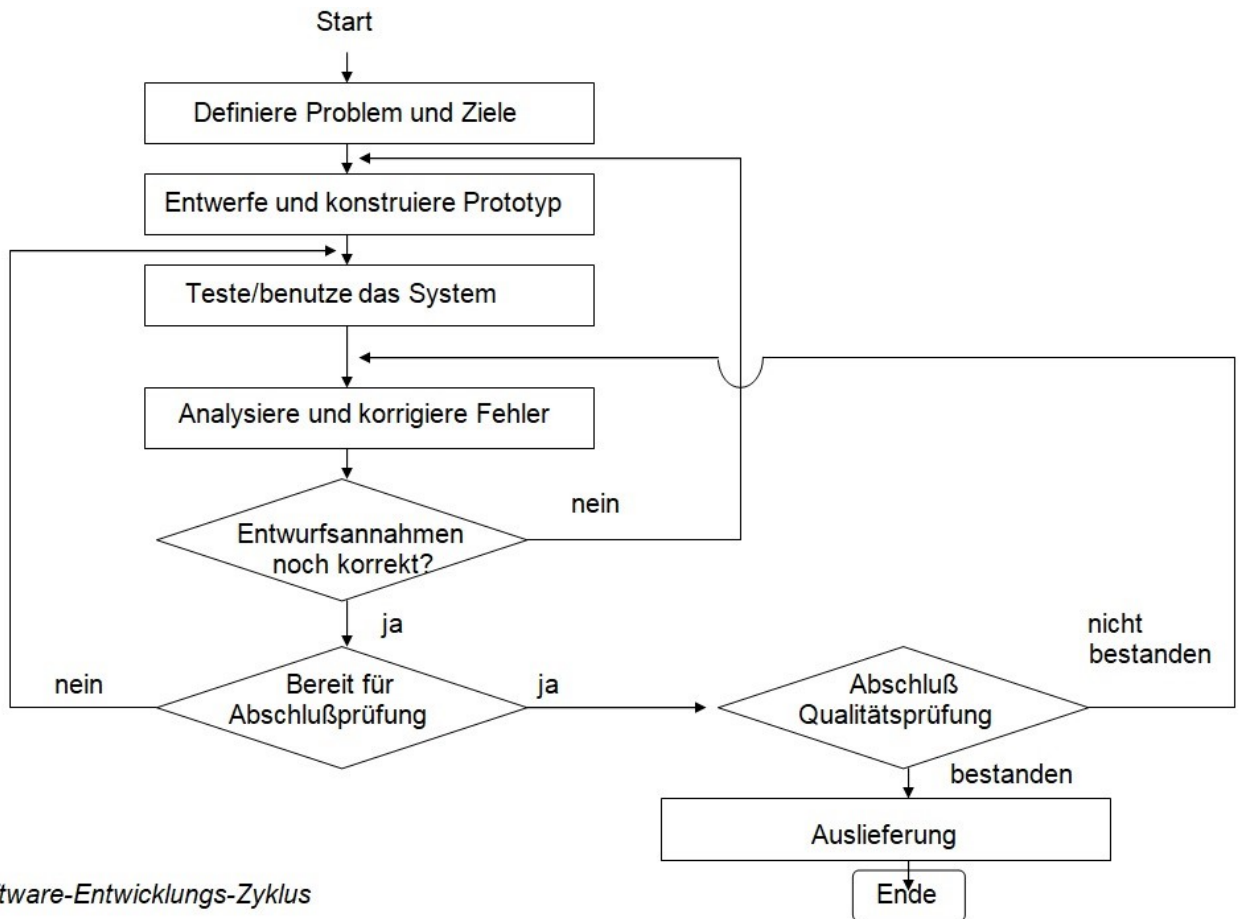
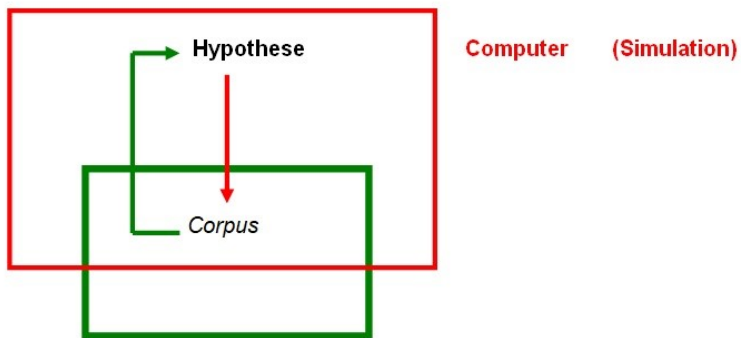
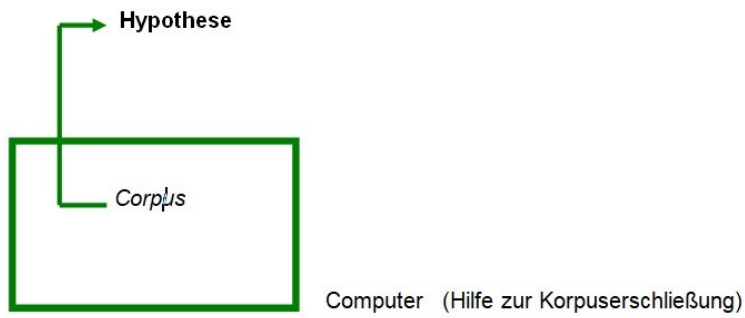
Wenn nicht, die Regeln revidieren.



nach I.M. Bochenski

Einsatz des Computers

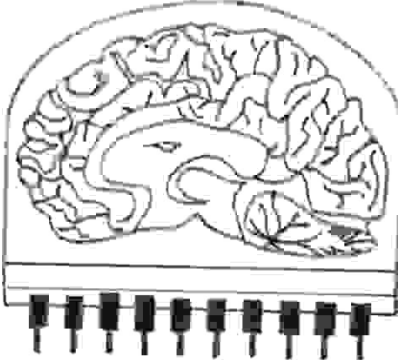
bei der empirischen Untersuchung und der Hypothesenüberprüfung



K 2 Computer und Denken, Künstliche Intelligenz

Lernziel: Situierung der Computerlinguistik im größeren Zusammenhang der Kognitionswissenschaften

GWAI 85



Künstliche Intelligenz

"Hätte man früher angefangen, dann gäbe es Künstliche Intelligenz seit 1500 Jahren, und jeder von uns wäre heute in der Lage, eine Sicherheitskopie von sich selbst zu machen. Können Sie sich Ihr Sprachverarbeitungssystem in - sagen wir nur - hundert Jahren vorstellen? Kann man Ihr Programm dann in einen Neurochip im Gehirn eines Kindes laden, und es hat alle Wortbedeutungen und Grammatiken dieser Welt verinnerlicht?"

Marvin Minsky, MIT

GWAI 85 = 9th German Workshop on Artificial Intelligence 1985

Computer und Denken? Können wir Denken mit Computer nachbilden?

Extrempositionen, Sicherheitskopie vom Gehirn, Unsterblichkeit des Menschen ...

Maschinen mit Bewußtsein, Maschinen die lernen, sich selbst schaffen, Menschen werden überflüssig

PRO

- der Mensch ist auch nur eine hochentwickelte Maschine (Materialismus)
- auch das Gehirn läuft nach einem festen Algorithmus ab
- wenn intelligentes Verhalten simuliert werden kann, ist die Maschine intelligent (Behaviorismus)

CONTRA

- es gibt einen fundamentalen Unterschied zwischen Körper und Geist (Dualismus)
- der Mensch ist in seinem Verhalten nicht vorherbestimmt, sondern kreativ (Freiheit statt Determinismus)
- Geist darf nicht mit abstraktem logischen Denken gleichgesetzt werden
- der Mensch hat Bewußtsein, die Maschine nicht

Die Debatte ist immer Janus-köpfig:

- die Maschine kann das auch: kreativ sein, lernen, sozialisiert werden, Bewußtsein haben
- der Mensch ist auch nicht so was besonderes, z.B. kreativ, selbstbestimmt

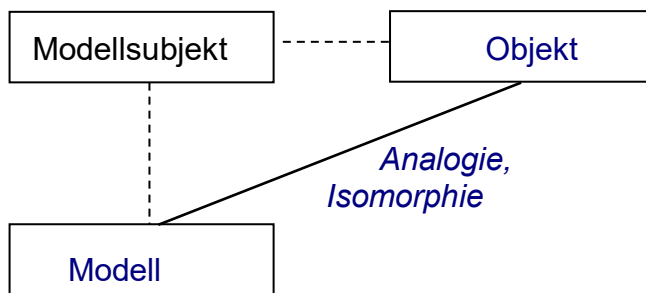
Worum geht es überhaupt

- um Erkenntnis über den Menschen "starke KI" ←
- oder um praktische Programme? "schwache KI" ←

entsprechende Forschungsziele, Erkenntnisinteresse, wissenschaftliche Ansprüche

Mal nüchtern Anknüpfen an Wissenschaftstheorie

Positivismus (= Bereich der Wissenschaft) - Hermeneutik (= der Sinn, die Sinnggebung)



Simulation: was durch was?

Denken durch Computer

Verstand durch Computer?

Geist durch Computer?

Gehirn durch Computer?

Bewußtsein durch Computer?

Intelligenz durch Computer?

Worin besteht die Analogie zwischen Objekt und Modell in diesem Fall?

A. Was ist eigentlich das Modellobjekt bei der künstlichen Intelligenz?

in der Regel schlecht definiert

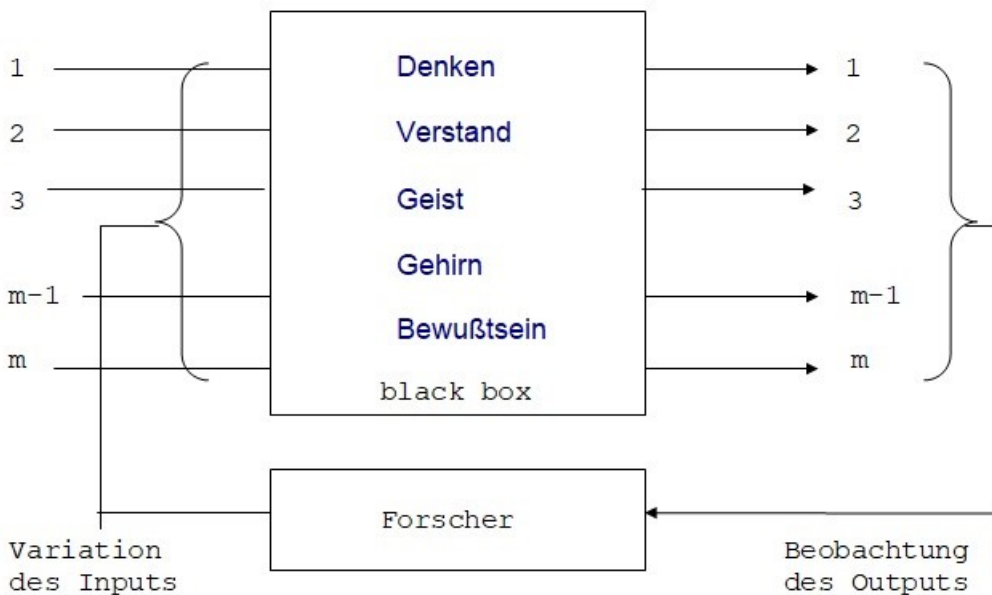
DUDEN "denken - die menschliche Fähigkeit des Erkennens u. Urteilens anwenden, mit dem Verstand arbeiten, überlegen" (d.h. sich ein Modell von der Außenwelt im Kopf bilden und benutzen)

LANGENSCHIED "engl. intelligence - Verstand, Auffassungsgabe, Einsicht, Verständnis, Nachricht, Mitteilung, Auskunft, Nachrichtendienst"

DUDEN 1. Intelligenz = Fähigkeit des Menschen, abstrakt und vernünftig zu denken und daraus zweckvolles Handeln abzuleiten (d.i. Theorien zu entwickeln und anzuwenden)

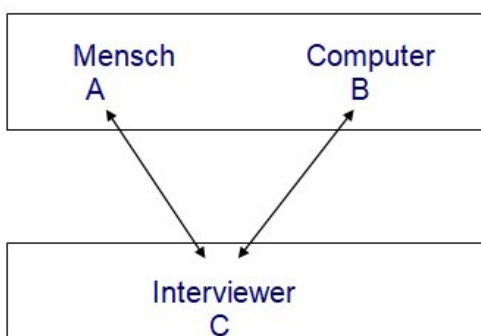
2. Gesamtheit der Intellektuellen, Schicht der wissenschaftlich Gebildeten

B: Was ist das Modell?



Blackbox - was ist das Ziel: nur **ergebnisgleich**? oder **ablaufgleich** mit dem Original ?

Der Turingtest



Versuchsanordnung: Person C in einem Raum im anderen Raum ein Mensch A und ein Computer B C weiß daß er über eine Datenleitung mit einem Menschen und über eine Datenleitung mit einem Computer verbunden ist.

Er soll durch Fragen herausfinden, welches der Mensch und welches der Computer ist.

Der Computer ist intelligent, wenn C die Aufgabe nicht lösen kann.

Turingtest = ergebnisgleich, d.h. der Test setzt Intelligenz mit **intelligenter**

Informationsverarbeitung gleich. Über Intelligenz selbst, über Psyche oder Gehirn, wird dadurch wenig bekannt.

Informationsverarbeitung durch Computer = Symbolmanipulation nach Regeln
 Umkehrschluß: wenn Computer ein adäquates Modell ist,
 dann muß das Objekt erklärbar rein als reine Informationsverarbeitung.

Ist das richtig bei: Denken, Verstand, Geist, Bewußtsein, Intelligenz ?

PRO KI: ja!

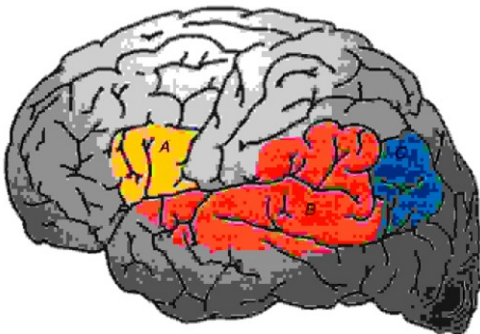
Der Mensch ist auch ein informationsverarbeitendes System (und nichts als das)
Physical Symbol Hypothesis Newell (1957), Psyche ist ein physikalisches Symbolsystem

Was ist die physikalische Repräsentation beim Menschen? **Neuronale Zustände im Gehirn!**

Nächster Schritt also: Neuronen - Denken?

⇒ *Verbindung zur Gehirnforschung*

Sind neuronale Verhältnisse im Gehirn mit symbolischer Repräsentation in Beziehung zu setzen?



Sprachregionen im Gehirn

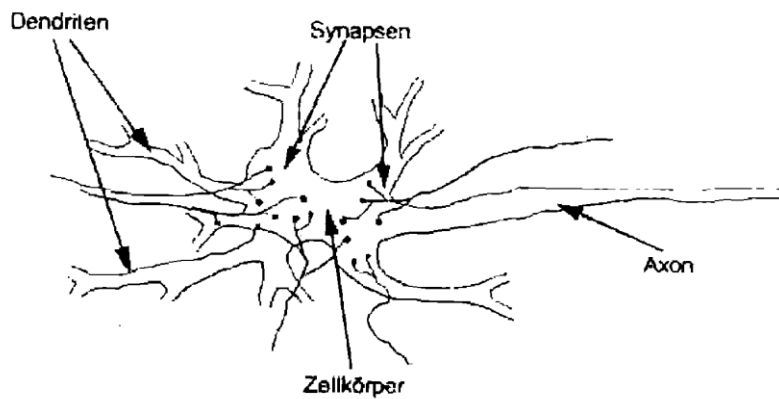
A Broca-Region

B Wernicke-Areal

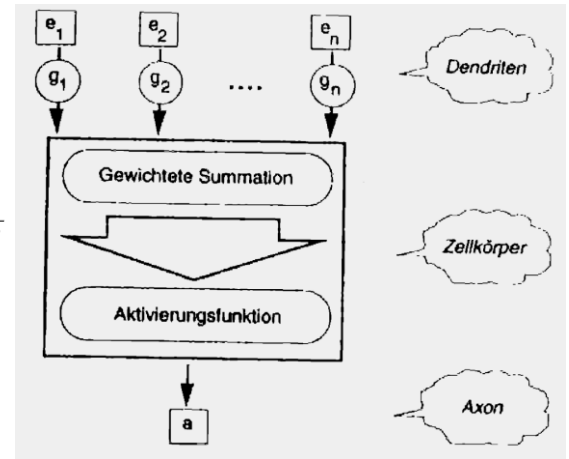
C Gyrus angularis

Aus: Spektrum der Wissenschaft, Dossier
 4/97: Kopf oder Computer

Alle psychischen Funktionen können aufgrund von Störungen im Gehirn selektiv verloren gehen. Durch solche selektiven Ausfälle liefert die Natur gleichsam einen Existenzbeweis dieser Funktionen, denn hiermit wird unmittelbar die Kopplung von Struktur und psychischer Funktion gezeigt (E. Pöppel)



Schematischer Aufbau eines Neurons im Gehirn



Künstliches Neuron in einem neuronalen Netz

Aber was ist jeweils der Input und der Output und sind sie in Objekt und Modell analog?

Searle's chinesisches Zimmer

Ein Mensch, der kein Chinesisch versteht, ist in einem Zimmer. Durch ein Fenster werden ihm chinesische Schriftzeichen hingereicht. Daraufhin reicht er andere chinesische Schriftzeichen hinaus. Der Mensch verfügt über eine Tabelle mit Anweisungen: auf Zeichen A antworte mit Zeichen X.

Versteht dieser Mensch Chinesisch?

Nein?

Irgendwie möchte man die Semantik retten. Der Mann kennt die Bedeutung nicht. Aber was ist das? Bedeutung ist der Gebrauch, sagt Wittgenstein. Der Mensch im chinesischen Zimmer gebraucht die Zeichen korrekt.

Letztlich HERMENEUTISCH, significance, Sinn für mich!

Auch bei der Gebrauchstheorie geht es um Sinn. Was ist pragmatisch sinnvoll?

Aber?

Wir können eigentlich nicht sicher sein, was der andere für Sinn zuschreibt. Letztlich müssen wir uns doch wieder rein auf das Verhalten der anderen als Basis verlassen. Für die Kunden des chinesischen Zimmers versteht der Mann Chinesisch. Es kann ihnen auch egal sein, solange die Antworten, die er herausreicht, korrekt sind.

Kritische Literatur:

Hubert Dreyfus: Die Grenze Künstlicher Intelligenz. Was Computer nicht können

Josef Weizenbaum: Die Macht der Computer und die Ohnmacht der Vernunft

Anmerkungen zur Geschichte der KI

Künstliche Intelligenz aus den Informatikinstitutionen erwachsen, Entdeckung der dynamischen, logischen Verarbeitung

Programmablauf von den Daten her gesteuert, "Wissensbasis", wissensbasierte Systeme

Gebiete

- Sprachverarbeitung, natürlich-sprachige Systeme (NLP)
- Logik, Deduktionssysteme, Theorembeweiser
- Expertensysteme
- Bildverarbeitung
- Robotertechnologie
- Wissensrepräsentation, common sense knowledge
- heuristisches Programmieren
- "real world computing", Simulation von Wirklichkeitsbereichen

Computerlinguistik

aus linguistischen Institutionen entstanden

aber gleiche Fragestellungen auf dem Feld NLP

und auch weitere Überlappung oder Berührungspunkte



Computerlinguistik gehört dazu

(heute ist es eher eine Frage der Institutionen, wie man sich nennt.)

K 3 Technologiefolgenabschätzung

Lernziel: Die Bereitschaft, sich mit den gesellschaftlichen Folgen der Computerisierung auseinanderzusetzen.

Brooks, Rodney:

Menschmaschinen

Wie uns die Zukunftstechnologien neu erschaffen

2002. 280 S. m. 6 Abb. 22 cm

CAMPUS VERLAG

Der Mensch bereitet gerade seinen nächsten Evolutionsschritt vor: die Verschmelzung mit der Maschine. Rodney Brooks nimmt seine Leser mit auf eine packende Reise durch Zukunftslabors, in denen Forscher dem Geheimnis menschlicher Intelligenz auf die Spur kommen. In seinem weltberühmten Institut am MIT entstehen Roboter, die laufen, mit Menschen kommunizieren und sogar Gefühle zeigen. Diese Visionen sind faszinierend und beunruhigend zugleich und werfen die Frage auf, ob die Zukunft uns überhaupt noch braucht. Ja, meint Brooks, denn der Mensch wird die eigene Evolution mit Hilfe der neuen Technologien vorantreiben. Mensch und Maschine verschmelzen, der Mensch erweitert seine Möglichkeiten Schritt für Schritt. Schon heute gibt es elektronische Gehörschnecken und Netzhäute aus Silikon. Bald wird es möglich sein, künstliche Prothesen vom Gehirn aus zu steuern. Vielleicht werden unsere Enkel ihre Hausaufgaben mit implantiertem Internetzugang machen.



Technologie-Schübe und ihre Auswirkungen auf die Entwicklung der Menschheit

- *Sprache* → soziale Organisation
- *Werkzeuge* → verbesserte Überlebenschancen, Eroberung unwirtlicher Räume
- *Ackerbau* → Seßhaftigkeit, Arbeitsteilung, Ermöglichung von Wissenschaft und Philosophie
- *Schrift* → Überlieferung, Akkumulierung von Erfahrung
- *Buchdruck* → Massenverbreitung von Wissen
- *Industrialisierung (Dampfmaschine, Webstuhl, Elektrifizierung, Chemie)* → Massenproduktion
- *Motorisierung (Eisenbahn, Dampfschiffahrt, Auto, Flugzeug)* → Massenmobilität

Welche Auswirkung hat der jüngste Technologie-Schub

- *Computer, Kybernetik, Mikroelektronik, Robotik, Nachrichtentechnik, www, KI, "intelligente" Systeme, Computerlinguistik, "human language technology"*

Anmerkung:

*Sind die Veränderungen wirklich wesentlich oder nur graduell, oberflächlich?
(der Mensch, weiterhin zwischen Geburt und Tod, die großen persönlichen Probleme -
dieselben wie eh und je ?)*

Verheißungen - Befürchtungen

Wissenschaftlicher Fortschritt - bringt Glückseligkeit oder die Katastrophe

Beispiel Atombombe: das soll uns nicht wieder passieren!

Furcht, das etwas außer Kontrolle gerät

MEINUNG 1

"Failing to understand the consequences of our inventions while we are in the rapture of discovery and innovation seems to be a common fault of scientists and technologists."

"Our most powerful 21st-century technologies - robotics, genetic engineering, and nanotech - are threatening to make humans an endangered species."

"These combinations open up the opportunity to completely redesign the world, for better or worse."

"Specifically, robots, engineered organisms, and nanobots share a dangerous amplifying factor: They can self-replicate. A bomb is blown up only once - but one bot can become many, and quickly get out of control."

(aus Bill Joy "Why the future doesn't need us." In Wired Magazine 6/2000)

MEINUNG 2

Rezept für die Bill-Joy-Debatte:

- zwei populistische Forscher und Bestseller-Autoren
- ein alternder Software-Ingenieur
- reichlich Schnaps
- viel Einfalt
- etwas Pest, Terrorismus und Verfolgungswahn
- eine Prise Ahnungslosigkeit

Das Ganze in »Wired« einwickeln und dann für ein paar Monate im Web gar kochen. Zum Schluss mit einem Schluck Zukunftsangst und Technologiefeindlichkeit abschmecken und durchs deutsche Feuilleton drehen. Mahlzeit.

"Kein Mensch weiß, wie Denken funktioniert und wie man Intelligenz schaffen kann. Folgerichtig kann auch kein Mensch ein intelligentes Wesen schaffen. Und deshalb können uns intelligente Roboter auch nicht in den Hintern treten. So einfach ist das."

(Wolf Lotter: So kocht man eine Katastrophe, brandeins Magazin, 2000, Ausgabe 7)

Stell dir vor, es geht, aber keiner kriegt's hin!



(1) Einfluß des Computers auf Wirtschaft und Beschäftigung

Was bringt die Erfindung von "Denkverstärkern"

- Entscheidend ist, daß der Anwender die im Programm enthaltene "Intelligenz" für seine eigene (geistige) Arbeit nutzen kann, ohne daß er die zugrunde liegenden Qualifikationen selbst besitzen muß.
- www - individuelle Verfügbarkeit von Information und Problemlösungen unabhängig von Zeit und Raum.
- Büro- und Verwaltungsbereich, "Trödefaktor PC", der Produktivitätszuwachs betrug von 1970 bis 1980 in der Fertigung 85 %, im Büro nur 4 %.
- Die Angestellten können zu Hause arbeiten, Transport zur Arbeitsstelle unnötig, Auflösung des Betriebes, aber dafür kann man u.a. Arbeit und Familie wieder unter einen Hut zu bekommen.
- Expertensysteme als kollektives Gedächtnis. Das Expertensystem bietet z.B. dem Arzt schnell und zuverlässig eine bedeutend höhere Zahl von Entscheidungsalternativen, als er sie selbst je behalten könnte.
- Aber auch Verlust der Information. Wie früher Papierzerfall, so jetzt Softwareverfall. Angebliche Halbwertszeit von Wissen nur 5 Jahre ???
- Verhältnis von verlorenen Jobs zu neu hinzugekommenen: 7 : 1
- Wenn die Technik so fortschreitet, kann die Arbeitszeit nicht dieselbe bleiben. Was passiert, wenn Arbeit nicht mehr nötig ist, weil Maschinen sie verrichten und höchstens noch eine kleine Elite die Wartung übernehmen muß?

(2) Risiken der Abhängigkeit von Computersystemen

A. Disfunktionalität (bugs)

Es kann nicht garantiert werden, daß ein komplexes Programm auch in unvorhergesehenen Situationen korrekt arbeitet. Murphy's law - "Anything that can go wrong, will."

Es gibt wahrscheinlich Millionenschäden durch Programmabsturz, die nicht publik gemacht werden.

B. Computerkriminalität

Computerviren, Störprogramme, welche von sich selbst Kopien erzeugen und diese dann in ein bestehendes Programm einpflanzen (infizieren). Folgen harmlos bis hin zur Zerstörung von Daten und Programmen, so daß die gesamte Datenverarbeitungsanlage unbrauchbar wird. Binäre Viren enthalten ausführbaren Code; sie können Programmdateien und Boot-Sektoren infizieren.

Makroviren basieren auf der Makro-Programmiersprache, die für die meisten Textverarbeitungs- und Tabellenkalkulationsprogramme verwendet wird. Makros erlauben es dem Benutzer, die Arbeit zu erleichtern. Sobald jedoch ein infiziertes Makros in einer Datei enthalten ist, ist es möglich, bei dem Aufruf dieses Dokumentes das Makro zu aktivieren (Auto-Makros). Dadurch ist es dem Virus möglich, auf Dateien zuzugreifen.

Trojanische Pferde - Programmstücke, welche einem Benutzer aufgrund eines Passworts oder eines Ereignisses Systemfunktionen zur Verfügung stellen, die ihm nicht zustehen. Können über einen langen Zeitraum "schlafen" und erst aufgrund eines Ereignisses ihre (meist negative) Wirkung entfalten.

Würmer - Programme, deren "Lebensbereich" Rechnernetze sind. Ein Wurm kann eine Kopie von sich an andere Rechner verschicken. Hierzu muß er das Protokoll des jeweiligen Netzes kennen. Zudem muß er die Adressenliste, in welchen die einzelnen Knotenrechner des Rechners verzeichnet sind, inspizieren können.

Trap-Doors nennt man das Einrichten von "Hintertüren" durch den Programmierer (Software-Sabotage)

C. Komplexitätsproblem

"As society and the problems that face it become more and more complex and machines become more and more intelligent, people will let machines make more of their decisions for them, simply because machine-made decisions will bring better results than man-made ones. Eventually a stage may be reached at which the decisions necessary to keep the system running will be so complex that human beings will be incapable of making them intelligently. At that stage the machines will be in effective control. People won't be able to just turn the machines off, because they will be so dependent on them that turning them off would amount to suicide." (Bill Joy)

Bei sehr komplizierten System ist der Eingriff des Menschen hochgefährlich.

Und noch etwas: Jedes Computersystem muss gewartet werden. Es ist schwer, sich in Programmcode einzuarbeiten, wenn er nicht außerordentlich gut dokumentiert ist. Haben wir in Zukunft genug Programmierer, die schlau genug sind, die zahlreichen hochkomplexen Systeme zu warten und zu reparieren?

Letzte Meldung

Neuste Entwicklungen in der IT Branche

"Hallo, ich bin Mike und ich begrüße sie alle zu einer neuen Folge von Amazing Discoveries. Heute haben wir einen Gast extra aus England zu uns einfliegen lassen, John, und John hat uns etwas mitgebracht.

Willkommen John!

- Danke Mike.

Was hast du mitgebracht, John?

- Nun Mike, ich hab hier etwas ganz besonderes. Etwas, worauf die ganze Welt gewartet hat.

Unglaublich, sag uns schnell was es ist...

- Ich hab hier ein weißes viereckiges Material und ein Stöckchen. Sie werden sich jetzt fragen, "was ist das nun schon wieder?"

Ja, wir sind schon ganz gespannt John. Spann uns nicht länger auf die Folter...

- Nun Mike, wenn ich dieses Stöckchen über das weiße Material bewege, dann verfärbt sich dieses genau an jenen Stellen, wo das Stöckchen das weiße Material berührt hat.

Das ist ja unglaublich....

- Ja, aber das ist noch nicht alles Mike. Wenn ich mit dem Stöckchen, das man übrigens einen Bleistift nennt, einen Buchstaben, so einen wie man ihn normalerweise auf seinen Computerbildschirm sieht, auf das Material zeichnet.

Unglaublich, langsam begreife ich es. Also wenn man mehrere Buchstaben nebeneinander zeichnet, dann kann man sogar lesen, ohne einen Bildschirm zu benötigen.

- Das ist tatsächlich so Mike, du begreifst es. Das Material nennt man übrigens Papier.

Unglaublich. Also wenn das keine Amazing Discovery ist... Und man benötigt dafür überhaupt keinen Strom?

- Nein, nein Mike, fantastisch, nicht? Man braucht selbst keine Batterien oder Akkus.

Junge, Junge das ist ja unglaublich. Ha John, du hast es zugeklappt. Das kann ich mit meinem Notebook auch.

- Nein Mike, das ist anders, du kannst es so oft falten, wie du willst, bis es das gewünschte Maß hat.

Hey, du hörst ja gar nicht mehr auf zu falten und es wird immer kleiner und kleiner. Aber John, was tust du jetzt??? Nein, das geht doch nicht du zerreißt das Papier in zwei Teile.

- Dieses Material ist so fantastisch, schau Mike. Ich halte die zwei Teile aneinander und man kann es immer noch lesen. Aber das ist noch nicht alles!

Nee?

- Du kannst es überall hin mitnehmen, du kannst es sowohl bei hohen als auch bei niedrigen Temperaturen benutzen. Und wenn du es nicht mehr benötigst, kannst du noch immer deine Nase damit putzen oder auf dem WC...., du verstehst was ich meine, Mike.

Ja John, unglaublich... Sag mal, aber das würde bedeuten, daß wir eines Tages überhaupt keine Computer und Notebooks mehr brauchen. Also John, du hast mich voll überzeugt... Nun sagt mal Leute, ist das nicht fantastisch???

- genau, fantastisch!"

Checkliste - Das Wichtigste aus der Vorlesung

S.1.1 Grundlagen der Phonologie und artikulatorischen Phonetik

Unterscheidung Phonologie - Phonetik

artikulatorische, auditive und akustische Phonetik

Strukturalismus, der Begriff des Systems
Opposition, Distribution, Minimalpaaranalyse

Phon, Phonem, Allophon
distinktives Merkmal
z.B. artikulatorische, distinktive Merkmale der Vokale

in []: phonetisches Element
in / /: phonemisches Element
in < >: graphemisches Element

Wichtig: Phonologie als Beispiel der strukturalistischen Methode

S 2.1 Grundlagen der Graphemik und die Umwandlung von Schrift in Laute

strukturalistische Analyse geschriebener Sprache

Graph, Graphem, Allograph
distinktive Merkmale

Das Verhältnis von Graphemen zu Phonemen

S.4.1 Grundlagen der Morphologie

Gegenstand der Morphologie

syntagmatische und paradigmatische Beziehungen :
Paradigma

Wort und Wortform

Morphem, Morph, Allomorph

Morpheme nach ihrer Funktion:
lexikalische. grammatische Morpheme

morphologische Operationen: Flexion, Derivation, Komposition

Morphologische Phänomene:

Präfixe, Infixe, Suffixe, Transfix

diskontinuierliches Morphem, Null-Morphem, Portmanteau-Morphem

Morpho-syntaktische Kategorien: Wortart, grammatische Merkmale, Kongruenz

S.5.1 Heuristische Grundlagen der Syntaxanalyse

Gegenstand der Syntax

syntagmatische und paradigmatische Relationen:

Syntagma

Tests: Permutation, Kommutation, Weglaßprobe, Implikationstest, Exklusions- und Koordinationsprobe, Paraphraseprobe, semantische Spezialtests

Einige syntaktische Phänomene:

Hierarchie, Abgrenzung von Einheiten auf verschiedenen Ebenen

Dimensionen: lexikalische Bedeutung, morpho-syntaktische Form, grammatische Funktion

Valenz und Rektion

Kongruenz

S 5.2 Formale Syntaxbeschreibung: Phrasenstrukturgrammatik

Chomskys Idee der generativen Grammatik

Konstituentenstruktur - satzorientiert

Teil-Ganzes-Relation, unmittelbare Konstituenten

Konstituentenstruktur-Baum

terminale und nicht-terminale Kategorien

S 5.3 Formale Syntaxbeschreibung: Dependenzgrammatik

Dependenzstruktur - wortorientiert

Wörter, Abhängigkeit

Komplemente und Adjunkte

Dependenzstruktur-Baum

mit komplexen Kategorien

S 5.4 Maschinelle Syntaxanalyse (Parsing)

Aufgaben eines Parsers

Verhältnis von Grammatikformalismus, Grammatik und Parser

Abarbeitung der Regeln (top-down, bottom-up)

Verfahren bei Alternativen: Backtracking, Parallelverarbeitung

S.6.1 Grundlagen der Aussagenlogik

Gebiet der Aussagenlogik

Aussagenlogische Ausdrücke:

Aussagen, Aussageform, Aussagenverknüpfung

Junktoren: Negation, Disjunktion, Konjunktion, Implikation, Äquivalenz

Semantik des Formalismus

Wahrheitswerte und Wahrheitstafeln für Junktoren

tautologische, kontradiktorischen und kontingente Aussageformen

Gesetze der Aussagenlogik (Tautologien)

Benutzen zum Umformen von Aussagenverknüpfungen

Logischer Schluß, Prämissen, Konklusion

Schlußfigur

modus ponens

Übersetzung sprachlicher Aussagen in aussagenlogische Formeln

S.6.2 Grundlagen der Prädikatenlogik

Gebiet der Prädikatenlogik

Prädikatenlogische Ausdrücke:

Prädikator, Argumente - Stelligkeit

Individuenkonstanten, Individuenvariablen

Quantoren, gebundene und freie Variablen

Junktoren

Referenz:

Individuenbereich, Individuenkonstanten (Namen)

durch Quantoren gebundene Individuenvariablen, u.a. Kennzeichnung

Proposition

Schlußfiguren in der Prädikatenlogik (Syllogismen)

Gesetze der Prädikatenlogik

S.7.1 Grundlagen der lexikalische Semantik

Lexikologie: der "Wortschatz" der Sprache

Onomasiologie - Semasiologie

Ausdruck vs. Inhalt (oder besser: Ausdruck und Gebrauch)

Relationen zwischen Ausdruck und Inhalt: n zu n

Ambiguität (Homophonie, Homographie, Homonymie, Polysemie)

Disambiguierung

(1) Die Wortfeldidee:

Abhängigkeit der semantischen Reichweite der Wörter von der Existenz anderer (Jost Trier)

(2) Der strukturalistische Ansatz, Komponentenanalyse:

Opposition, Minimalpaarverfahren, distinktive semantische Merkmale

Bedeutung eines Wortes (Semems) als Menge von Merkmalen (Semen),

Merkmalsmatrix aus Wörter / Merkmalen und +, -, 0

(3) Sinnrelationen:

logisches Kriterium, logische Relationen, z.B. Implikation:

x ist ein Mann -> x ist ein Mensch

x ist erstickt -> x ist umgekommen

(4) Lexikalische Funktionen $LF(x)=y$:

- Synonym
- Hyponym, Hyperonym, Kohyponym
- Antonym
- Komplenym
- Meronym, Holonym

(5) Prototypensemantik:

Nicht-Trennbarkeit von Realdefinitionen und Nominaldefinitionen,

beim Bedeutungslernen etwas Erfahren über Welt und über den sprachlichen Ausdruck zugleich.

Daher besser Prototypen-Beschreibung statt Oppositionen (Fillmore, Putnam)

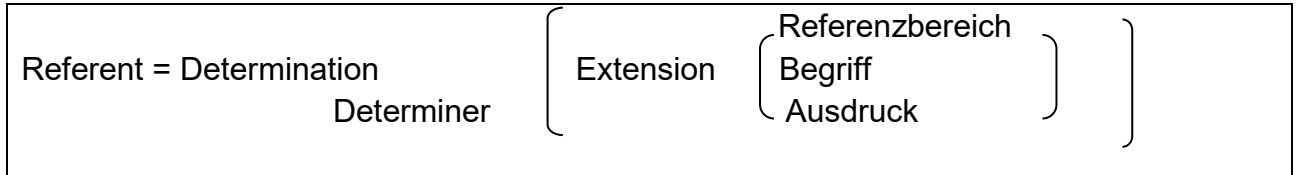
S.7.2 Referentielle Semantik (Weltwissen und Sprache)

Begriff der Referenz versus Sinn/Bedeutung ("*der Morgenstern*" / "*der Abendstern*")
Intension - Extension

Im Satz

- referierende Ausdrücke (bezeichnen Redeobjekte)
- nicht-referierende Ausdrücke - z.B. Prädikate referieren nicht (Tugendhat)

Formel zur Referenz:



Deixis, deiktische Ausdrücke

Anapher:

sprachliche Mittel für Ko-Referenz im Text (welche?)

Probleme der Anapherauflösung

S.7.3 Satzsemantik (und logische Semantik der natürlichen Sprache)

Phänomene der Satzsemantik (mit der Syntax verknüpfte Bedeutung)

Grammatische Rollen (Prädikat, Subjekt, Objekt usw.)

Semantische Beziehungen zwischen Sätzen:

Paraphrase, Implikation, Präsupposition

Schlußfiguren in natürlicher Sprache (Syllogismen)

logische Konstituentenanalyse

S.8.1 Grundlagen der Pragmatik und Sprechakttheorie

Sprache als Handeln (Pragmatik)
 Was kann man sprachlich alles machen?
 (Austin: "How to do things with words")

Beispiele für verschiedene Sprechakte

Bestimmungsstücke von Sprechakten:

- Äußerung
- propositionaler Gehalt
- Illokution
- Perlokution

performatorische Ausdrücke

Glücken eines Sprechakts
 Gelingen eines Sprechakts

pragmatische Bedingungen
 (vgl. Searles Bedingungen für das "Versprechen")

Dialoge, Sprechaktsequenzen ("Sprachspiele")

indirekte Sprechakte

S.8.2 Grundlagen der Textkonstitution und des Textverstehens

Zwei Arten des Textzusammenhanges: Kohäsion und Kohärenz

(1) Kohäsion - syntagmatische Beziehungen

- Prädikate u.a. Relatoren
- Referenz und Korreferenz
- Kontiguität

-> referentielle Progression

(2) Kohärenz - pragmatischer Sinn, Sprechaktsequenzierung

Kohärenz in monologischen Texten
 implizite (teils explizite) Frage-Antwortstruktur
 Fragestellung = Thema

-> thematische Progression

Zusammenhang von Satz- und Textstruktur:

Intonation, Wortfolge
 Topic-Comment