

Textgenerierung und Neuronale Netze

Pluralbildung von
Substantiven

Morphologie

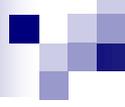
- Formenlehre
- untersucht die systematischen Beziehungen zwischen Wörtern und Wortformen bzw. die Regeln nach denen Wörter/Wortformen gebildet werden
- formale Umsetzung solcher Regeln in der Computerlinguistik dient dazu, Vollformlexika zu ergänzen oder zu ersetzen

Grundbegriffe der Morphologie

- **Wort** : (=Lexem) abstrakte Einheit, die verschiedenen Formen zugrunde liegt und dem Eintrag eines Wörterbuchs entspricht
- **Wortform**: verschiedene Formen eines Lexems
- **Morpheme**: minimale Phonem/Graphem-Ketten mit einer festgelegten Bedeutung (freie/gebundene Morpheme, Grundmorpheme/periphere Morpheme)

Grundbegriffe der Morphologie

- **Allomorphe:** realisieren dieselbe Bedeutung in verschiedenen morphologischen Kontexten (z.B. wird Plural in *Kind-er* durch *-er* und in *Wind-e* aber durch *-e* ausgedrückt)
- **Morphophonologie:** Wenn der Unterschied zwischen verschiedenen Allomorphen nur minimal ist, spricht man von **partieller Allomorphie** (Umlaut : Mutter-Mütter, Vater-Väter)



Grundbegriffe der Morphologie

- **Nichtkonkatenative Morphologie:** bedeutet, dass eine Kategorie wie Plural nicht durch ein Morphem, sondern durch andere Mittel wie z.B. den Umlautungsprozess ausgedrückt wird

Probleme der morphologischen Beschreibung

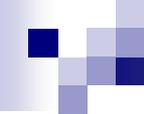
- **Neutralisierung:** Wörter/Wortformen spiegeln ihre Bedeutung oder Funktion nur teilweise wieder z.B. *Haken* kann Pl. oder Sg. sein
- **Nichtkonkatenativität**
- **Regularitäten und Ausnahmen:** Bestimmte Formen und Affixe sind regelmäßig, andere ganz oder teilweise unregelmäßig
- **Allomorphie und Phonologie**

Pluralbildung im Deutschen

	GENUS		
<u>Pl.-Allomorph</u>	<u>Maskulinum</u>	<u>Femininum</u>	<u>Neutrum</u>
-e	Fisch/Fische	Kenntnis/-nisse	Jahr/Jahre
-(e) n	Bauer/Bauern	Tür/Türen	Auge/Augen
-er	Geist/Geister	-	Kind/Kinder
-s	Park/Parks	Mutti/Muttis	Auto/Autos
-ø	Adler/Adler	-	Fenster/Fenster
Umlaut	Bruder/Brüder	Tochter/Töchter	Kloster/Klöster
Umlaut + -e	Sohn/Söhne	Kuh/Kühe	Floß/Flöße
Umlaut + -er	Wald/Wälder	-	Volk/Völker

Pluralbildung

- Für die Zuweisung der aufgelisteten Pluralmarkierungen zu den nominalen Lexikoneinträgen existieren keine generellen Regeln
- Pluralmarkierung muss als Teil eines jeden Nomens mitgelernt werden
- Zuweisung und Auswahl einer Pluralmarkierung für ein spezifisches nominales Lexem aus einer finiten Anzahl von Möglichkeiten ist arbiträr



*Textgenerierung: Umwandlung der internen
Repräsentation in natürliche Sprache*

Kommunikation

Maschine \rightarrow Mensch

Maschine \leftrightarrow Mensch (Dialogsysteme)

Wie soll die Textgenerierung vor sich gehen?

- Symbolisch → generative Linguistik
 - System: Regeln + Ausnahmen

oder

- Konnektionistisch → Konnektionismus
 - System: lernt
 - Kein Grammatikregeln zugänglich

Beispiele für Textgenerierung

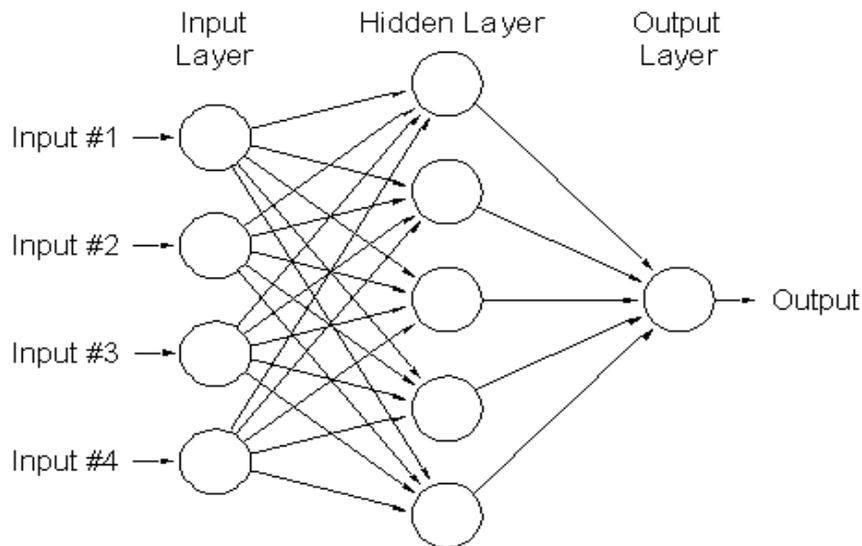
- Ansage- und Auskunftsdienste
(Zeitansage, Wetterdienst, Fahrplanauskunft, ...)
- Bestell- und Buchungssysteme
(Versandhäuser, Verlage, Reisebüros, ...)
- akustische Ausgabe als Ergänzung zum Bildschirm
(Multimedia-Präsentationen, Computer-unterstützter Unterricht, ...)

Konnektionismus

- Ein System entsteht aus der Wechselwirkung vieler vernetzter, einfacher Einheiten
Komplex → erst durch Vernetzung
- Methode der Konnektionismus
 - z.B. Künstliche neuronale Netze, (Sensitivitätsmodell)
 - → Die Funktion besteht aus vielen einfachen gleichartigen Teilen.

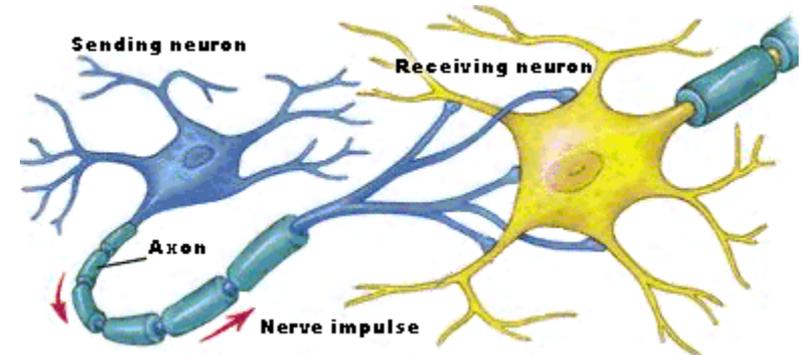
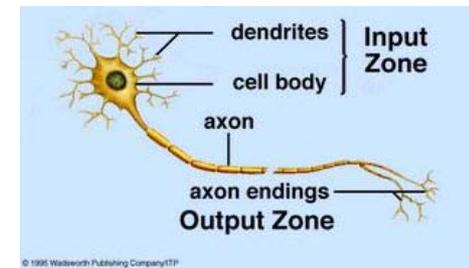
Neuronale Netze

■ Künstliche neuronale Netzwerke



→ Abstraktion von Informationsverarbeitung

■ Biologische neuronale Netzwerke



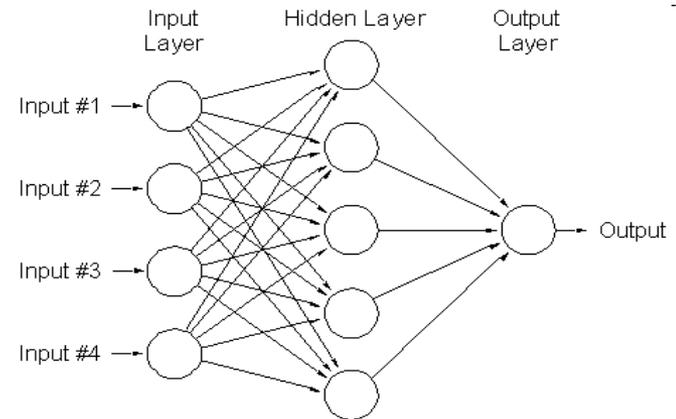
Künstliche Neuronale Netzwerke

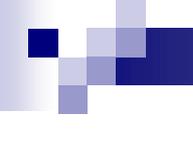
- Konstruktion
- Trainingsphase (Lernen & Adaption)
- Abfrage

Hebb'sche Lernregel

Wenn 2 Neuronen zur gleichen Zeit aktiv sind
wird das Gewicht zwischen ihnen erhöht

→ Grundlage für die meisten Lernsysteme





Vorteile künstlicher neuronaler Netzwerke

- sind anpassungsfähig
- sind lernfähig
- arbeiten auch bei unvollständigen Daten
- sind robust beim Ausfall von Teilen des Systems



Nachteile künstlicher neuronaler Netzwerke

- Ein globales oder nur ein lokales Optimum?
- Überspezifizierung
- Unterspezifizierung
- Ausnahmen

Spracherwerb

KNN können partiell den Spracherwerbsprozess modellieren:

3 Phasen (beim Kind):

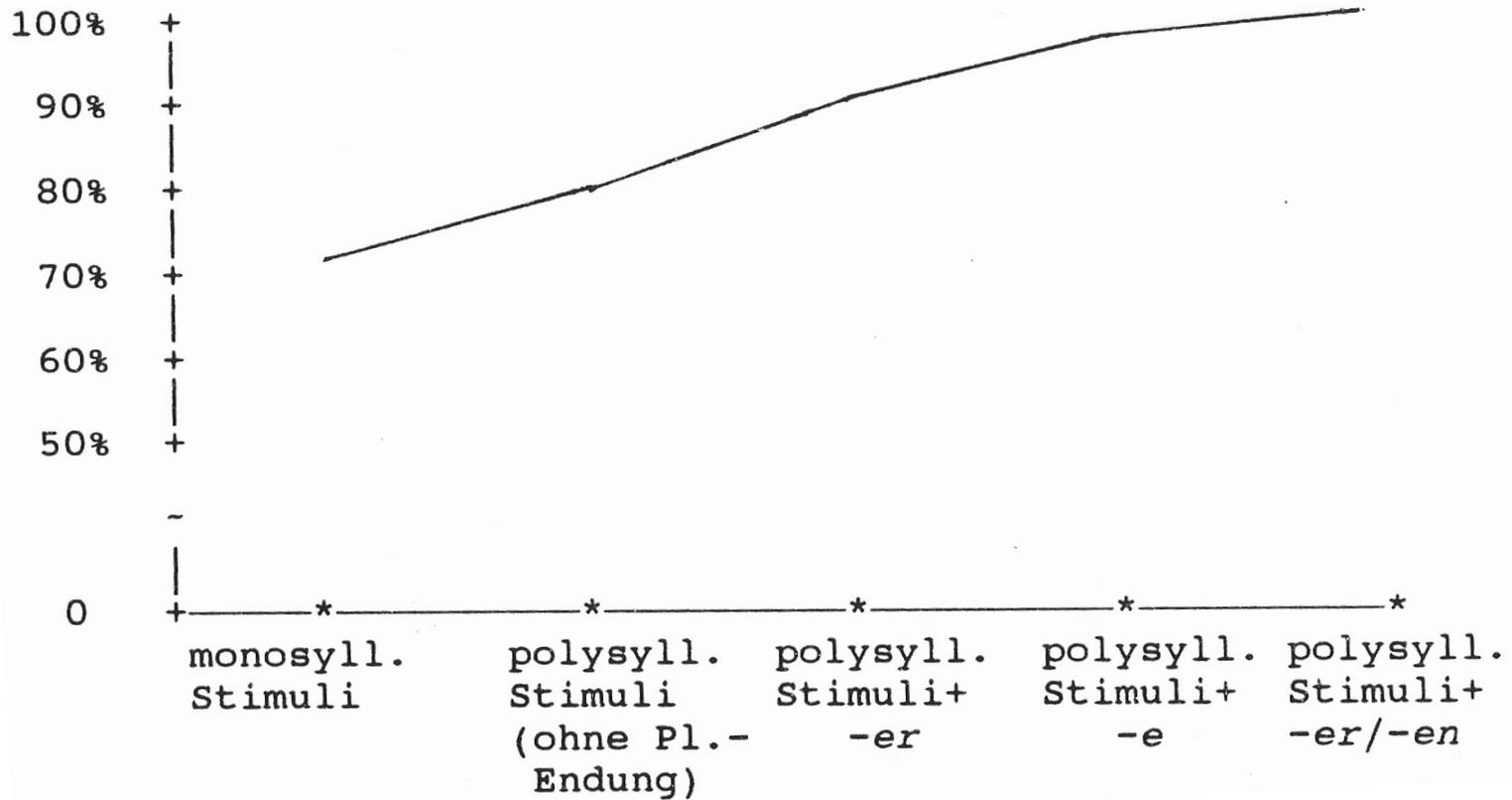
1. *Merken*
2. *Übergeneralisieren*
3. *richtige Balance*

Prinzipien, die beim Pluralerwerb von Bedeutung sind

- rote learning
- Analogielernen
- Regellernen
- Schema-Lernen

Das Kunstwortexperiment von Berko

Stimulus	erwartete Pl.-mark.	Antworten der Kinder					
		-∅	-e	-(e)n	-er	-s	Neubild.
der /šmirl/	-e/-(e)n	13	9		1		2
das /gɛtrydɔ/	-∅	24		1			
die /albər/	-(e)n	23		1		1	
das /tindɔ/	-(e)n	23	1			1	
die /za:ri/*	-s	18		1		5	1
das /gɛlɛk/	-e	18	7				
der /nɛ:bər/	-∅	25					
die /rondatš/	-(e)n	22	3				
das /bros/	-e/-er	17	8/3				
die /arl/	-e/-(e)n	19	5			1	
der /ha:gən/	-∅	25					
die /ne:bɔ/	-(e)n	17		7			1
das /hɛklain/	-∅	21	3			1	
der /fɛndɔ/	-(e)n	24		1			
das /kundər/	-∅	25					
der /farst/	-e	22	1	2			



Verwendung der Nullmarkierung

