

Computerlinguistik

E04: Morphologie

Morphologie

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

Begriffe: Wort - Morphem - Allomorph

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

Flexion und Wortbildung

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

Sprachtypologie anhand der Wortstruktur

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

Analytische vs. generative Morphologie

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

Morphologische Analyse

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

Grundmodelle der generativen Morphologie

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

Morphologie mit endlichen Automaten

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

Flexionsparadigmen

WARTEN:

Ich	wart-e	wart-et-e
Du	wart-est	wart-et-est
Es	wart-et	wart-et- e
Wir	wart-en	wart-et-en
Ihr	wart-et	wart-et-et
Sie	wart-en	wart-et-en

Partizip: ge-wart-et

HALTEN:

Ich	halt-e	hielt- ø
Du	h ä lt-st	hielt-est
Es	h ä lt- ø	hielt- ø
Wir	halt-en	hielt-en
Ihr	halt-et	hielt-et
Sie	halt-en	hielt-en

Partizip: ge-halt-**en**

Morphologie mit Default-Vererbungsnetzen

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

Wissensrepräsentation in DATR

WARTEN:

<form präs sg eins> == warte

<form präs sg zwei> == wartest

<form präs sg drei> == wartet

(...)

.

DATR: Resolutionspfade

WARTEN:

<form> == <wurzel> <endung>

<wurzel> == wart

<endung präs sg eins> == e

<endung präs sg zwei> == est

<endung präs sg drei> == et

<endung präs pl eins> == en

<endung präs pl zwei> == et

<endung präs pl drei> == en

.

DATR: Redundanzen

WARTEN:

<form> == <wurzel> <endung>

<wurzel> == wart

<endung **präs** sg eins> == e

<endung **präs** sg zwei> == est

<endung **präs** sg drei> == et

<endung **präs** pl eins> == en

<endung **präs** pl zwei> == et

<endung **präs** pl drei> == en

.

DATR: Defaults und Neutralisierung

WARTEN:

<form>	==	<wurzel> <endung>
<wurzel>	==	wart
<endung sg>	==	e
<endung sg zwei>	==	est
<endung präs sg drei>	==	et
<endung pl>	==	en
<endung pl zwei>	==	et
<endung präs>	==	<endung>
<endung prät>	==	et <endung>

.

DATR: Klassenbildung

VERB_SCHWACH_T:

<form> == <wurzel> <endung>

<endung sg> == e

<endung sg zwei> == est

<endung präs sg drei> == et

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

.

WARTEN:

<> == VERB_SCHWACH_T

<wurzel> == wart

.

DATR: Nichtlokale Abhängigkeiten

VERB_SCHWACH_T:

<form> == <wurzel> <endung>

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

<endung sg> == e

<endung sg zwei> == est

<endung präs sg drei> == et

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

.

DATR: Nichtkonkative Phänomene

VERB_STARK:

<form> == <wurzel> <endung>
<wurzel> == <w_anf> <w_vok> <w_end>
<endung sing> == Ø
<endung präs sing eins> == e
<endung sing zwei> == st
<endung plur> == en
<endung plur zwei> == et

.

HALTEN:

<> == VERB_STARK
<w_anf> == h
<w_end> == lt
<w_vok> == a
<w_vok prät> == ie <prät>
<v_vok präs sg zwei> == ä <präs sg zwei>
<v_vok präs sg drei> == ä <präs sg drei>

.

Literatur / Hausaufgabe

➤ **Zur Nachbereitung:**

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

➤ **Zur Vorbereitung:**

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