

Adjektiv-Schema

Deutsche Nomen und Adjektive haben wesentlich mehr Paradigmen als unterschiedliche lexikalische Formen. Für die Endungsformen -en und -e ergeben sich 26 bzw. 11 Paradigmen. Eine Darstellung der Funktion von Paradigmen auf Formen (wie oben) ist deshalb komplexer als die umgekehrte Funktion von Formen zu möglichen Paradigmen, wie folgt. Zudem lassen sich durch systematische *Unterspezifikation* diese Formen ohne Informationsverlust auf Muster abbilden. Die 26 Paradigmen der Form -en lassen sich beispielsweise durch nur 10 unterspezifizierte Muster abbilden. [\[Pollard & Sag 94: 64-7\]](#)

| FORM | KASUS | NUMERUS | GENUS | DEKLIN. | MUSTER/PARA |
|------|-------|---------|--------|---------|-------------|
| -en | Gen | (Sg) | | Schw | 10 / 26 |
| | Gen | (Sg) | | Gem | |
| | Dat | (Sg) | | Schw | |
| | Dat | (Sg) | | Gem | |
| | | Pl | | Schw | |
| | | Pl | | Gem | |
| | Akk | Sg | Masc | | |
| | Gen | Sg | Masc | (Stark) | |
| | Gen | Sg | Ntr | (Stark) | |
| | Dat | Pl | | (Stark) | |
| -e | Nom | Sg | Fem | | 7 / 11 |
| | Akk | Sg | Fem | | |
| | Nom | Sg | Ntr | Schw | |
| | Akk | Sg | Ntr | Schw | |
| | Nom | Sg | (Masc) | Schw | |
| | Nom | Pl | | Stark | |
| | Akk | Pl | | Stark | |
| -er | Nom | Sg | Masc | Gem | 5 / 5 |
| | Nom | Sg | Masc | Stark | |
| | Gen | (Sg) | Fem | Stark | |
| | Dat | Sg | Fem | Stark | |
| | Gen | Pl | | Stark | |
| -es | Nom | Sg | Ntr | Gem | 4 / 4 |
| | Akk | Sg | Ntr | Gem | |
| | Nom | Sg | Ntr | Stark | |
| | Akk | Sg | Ntr | Stark | |
| -em | Dat | Sg | Masc | Stark | 2 / 2 |
| | Dat | Sg | Ntr | Stark | |

Hierzu [ein kleines Beispielprogramm in Prolog](#).

Verb-Konjugation im Deutschen

Die deutschen Verben bieten auf den ersten Blick kaum Probleme für die automatische Lemmatisierung. Sie werden eingeteilt in starke (unregelmässige), schwache (regelmässige) und gemischte (nur Vokalwechsel, sonst regelmässig; z.B. nennen, senden) Konjugation. Bei starken Verben erhält jeder Stamm einen separaten Lexikoneintrag.

Verbleibende Probleme:

- Bei Verben, deren Stamm auf -s, -x oder -z endet, wird in der 2. Person die Endung verschmolzen (z.B. wachsen -> wächs+st -> wächst)
- Bei Verben, deren Stamm auf -n oder -m endet, wird in der 2. Person ein -e- eingeschoben (z.B. atmen -> atm+st -> atmest)
- Bei Verben mit dem festen Präfix ver- oder be- wird das Partizip Perfekt ohne ge- gebildet (z.B. verlieren -> verloren; behindern -> behindert; vergessen -> vergessen)
- Bei Verben mit einem abtrennbaren Präfix wird das Partizip Perfekt durch einen Einschub von ge- und der starke Infinitiv durch einen Einschub von zu- gebildet (z.B. abfahren -> abgefahren, abzufahren; umfahren -> umgefahren, umzufahren)
- Einige Verben haben starke und schwache Formen und davon abhängig (leicht) unterschiedliche Bedeutungen (z.B. erschrecken, schaffen, wachsen, weichen).
- ...

Das Flexionsparadigma

| | Starke Konjugation | Schwache Konjugation |
|---------------|---|---|
| Präsens | schwimme schwimmst schwimmt schwimmen schwimmt schwimmen | zeige zeigst zeigt zeigen zeigt zeigen |
| Präteritum | schwamm schwammst schwamm schwammen schwammt schwammen | zeigte zeigtest zeigte zeigten zeigtet zeigten |
| Konjunktiv I | schwimme schwimmest schwimme schwimmen schwimmet schwimmen | zeige zeigest zeige zeigen zeiget zeigen |
| Konjunktiv II | schwämme schwämmest schwämme schwämmen schwämmet schwämmen | zeigte zeigtest zeigte zeigten zeigtet zeigten |
| andere | schwimm(e) schwimmt geschwommen schwimmend | zeig(e) zeigt gezeigt zeigend |
| | | |

Zugriffstabellen auf morphol. Information bei Verben

Um bei der Analyse einer Verbform die Geschwindigkeit des Zugriffs auf die morphologische Information zu erhöhen, können die Verbendungen in Zugriffstabellen (auch Indextabellen genannt) angeordnet werden. Diese Tabellen können durch die Hinzunahme eines Endungsbaums noch effizienter genutzt werden.

Einsatz: Die (potentielle) Endung einer Verbform wird abgetrennt, und es wird mit Hilfe des Stammformenlexikons geprüft, ob der verbleibende Rest ein möglicher Verbstamm ist. Falls ja, dann ist in diesem Lexikon vermerkt, zu welcher Konjugation das Verb gehört und gegebenenfalls von welchem Typ der Stamm ist. Mit diesen 3 Angaben (Konjugation, Endung, Typ des Stamms) kann in einer der folgenden Tabellen die morphologische Information ermittelt werden.

Indextabelle für starke Konjugation

| Knoten | Endung | Präsens-Stamm | Präteritum-Stamm | Konjunktiv-Stamm |
|--------|--------|---|-------------------------|---------------------------|
| 1 | -t | 3. Sg und 2. Pl Präsens; Imperativ Pl | 2. Pl Prät. | |
| 2 | -et | 2. Pl KonjI | | 2.Pl KonjII |
| 4 | -st | 2. Sg Präsens | 2. Sg Prät. | |
| 5 | -est | 2. Sg KonjI | | 2.Sg KonjII |
| 7 | -e | 1. Sg Präsens; 1. Sg und 3. Sg KonjI; Imperativ Sg | | 1. Sg und 3. Sg KonjII |
| 9 | -en | 1. Pl und 3. Pl Präsens; 1. Pl und 3. Pl KonjI; Infinitiv | 1. Pl und 3. Pl Prät | 1. Pl und 3. Pl KonjII |
| 11 | -end | Partizip Präsens | | |
| 12 | - | | 1. Sg und 3. Sg Prät | |

Indextabelle für schwache Konjugation

| Knoten | Endung | |
|---------------|---------------|---|
| 1 | -t | 3. Sg und 2. Pl Präsens; Imperativ Pl; (Partizip Perfekt) |
| 2 | -et | 2. Pl KonjI |
| 3 | -tet | 2. Pl Prät.; 2.Pl KonjII |
| 4 | -st | 2. Sg Präsens |
| 5 | -est | 2. Sg KonjI |
| 6 | -test | 2. Sg Prät.; 2.Sg KonjII |
| 7 | -e | 1. Sg Präsens; 1. Sg und 3. Sg KonjI; Imperativ Sg |
| 8 | -te | 1. Sg und 3. Sg Prät; 1. Sg und 3. Sg KonjII |
| 9 | -en | 1. Pl und 3. Pl Präsens; 1. Pl und 3. Pl KonjI; Infinitiv |
| 10 | -ten | 1. Pl und 3. Pl Prät; 1. Pl und 3. Pl KonjII |
| 11 | -end | Partizip Präsens |

Zusammenfassung

1. Die Morphologiesysteme und -phänomene natürlicher Sprachen können in die Klassen isolierend, agglutinierend, inflektional und polysynthetisch eingeteilt werden.
2. Morpheme können eingeteilt werden in Kern-, Partikel-, Derivations- und Flexionsmorpheme.
3. Zur Beschreibung der offenen Wortarten (Adjektive, Substantive, Verben) in einem Stammformenlexikon werden diese aufgrund von formalen Merkmalen in Unterklassen eingeteilt.
4. Um ein Stammformenlexikon mit möglichst geringem Aufwand erweitern zu können, sollte man die Merkmale in Form einer Entscheidungstabelle organisieren. [s. z.B. das [Klassifikationsschema für deutsche Substantive](#)]
5. Um die Kombinationsvielfalt der Merkmale einzudämmen, kann man Unterspezifikation verwenden. D.h., wenn eine Wortform in allen Ausprägungen (d.h. Werten) eines Merkmals identisch ist, braucht dieses Merkmal nicht spezifiziert werden. [s. z.B. das [Adjektiv-Schema](#)]
6. Um den Zugriff zu morphologischer Information möglichst effizient durchführen zu können, setzt man Endungsbäume und Indextabellen ein. [s. z.B. [Zugriffstabellen auf morphologische Information bei Verben](#)]

Gerold Schneider, Martin Volk

www.ifi.unizh.ch/CL/gschneid/LexMorphVorl/Lexikon03.html