

Morphosyntaktische Etikettierung

Die folgenden Etikettierungsrichtlinien erläutern Abkürzungen und Besonderheiten bei der morphosyntaktischen Etikettierung mit Hilfe unseres Etikettierungsformulars¹. Der folgende Screenshot vermittelt einen Einblick in das Vorgehen bei der Etikettierung. Die Etikettierung erfolgt Wort für Wort anhand der in den Dropdown-Menüs auszuwählenden Etiketten.

	am	'vaījtər	'flajən	də	gə'drɛçt	'bladər	an
Satz	11	11	11	11	11	11	11
Funktion	Atemp	Atemp	Vfin	ENom	ENom	ENom	Alok
Wortart	praep	subst	verb	art	adj	subst	praep
Spezif.	lok temp mod det	normal	voll		Gebrauch attr subst	normal	lok temp mod
Modus			ind				
Person			3				
Numerus		sg	pl	pl	pl	pl	
Kasus		dat		nom	nom	nom	
Tempus			praesens				
Genus		m		n	n	n	
Lemma	im_1	Winter_2	fliegen_3	der_4	trocknen_5	Blatt_6	in_7
Konzept							
speichern							

S A T Z

1: Satz 1

11: Hauptsatz 1

12: Hauptsatz 2

13: Hauptsatz 3

¹ Der Zugang zur Etikettierung ist passwortgeschützt. Bei Interesse wenden Sie sich bitte an Stephan Lücke oder Thomas Krefeld: luecke@lmu.de / thomas.krefeld@lmu.de

2: Satz 2

21: Hauptsatz 1

22: ...

Bsp.:

{ {Der gute alte Mann ist mit dem Pferd durchs Eis gebrochen}11 {und in das kalte Wasser gefallen.}12 }1

Bei falsch gesetzten Intervallgrenzen (Satzanfang gehört noch zum vorherigen Intervall), wird der Satzanfang als Satz 0 behandelt.

N E B E N S A T Z

kaus: kausal

temp: temporal

kond: konditional

adv: adverbial

konz: konzessiv

mod: modal

rel: relativ

fin: final

kons: konsekutiv

obj: Objektsatz/Nebensatzergänzung (Bsp.: Ich glaube, dass das sehr klug ist.)

ergHS: Objektsatz/Hauptsatzergänzung (Bsp.: Ich glaube, das ist sehr klug.)

F U N K T I O N

VK: Verbkomplex

Vfin: finites Verb

Vinfin: infinites Verb

VKteil: Verbalkomplextteil

VPart: Verbpartikel

praedEl: prädikatives Element

Erg: Ergänzungen

ENom: Nominativergänzung

EGen: Genitivergänzung

EDat: Dativergänzung

EAcc: Akkusativergänzung

EPraep: präpositionale Ergänzung

EAdv: adverbiale Ergänzung

EAdj: Adjektivergänzung

Ang: Angaben

Alok: lokale Angabe

Atemp: temporale Angabe

Amod: modale Angabe

frDat: freier Dativ

Attr: Attribut

App: Apposition

W O R T A R T

Flekt=1: Flektierbare Wortarten

subst: Substantiv (normal, Name)

adj: Adjektiv (Gebrauch: attributiv, substantiviert, adverbial; Komparation: Komparativ, Superlativ; Zahl; unflektiert; Partizip Präsens, Partizip Präteritum)

pron: Pronomen (Personal-, Reflexiv-, Possessiv-, Demonstrativ-, Relativ-, Interrogativ-, Indefinit-)

art: Artikel (determiniert, indeterminiert)

verb: Verb (Voll-, Modal-, Hilfs-)

Flekt=0: Unflektierbare Wortarten

adv: Adverb (Pronominaladverb, Konjunkionaladverb)

praep: Präposition (lokal, temporal, modal, determiniert) > Kombination möglich, z.B. „im“ lok+det bzw. temp+det

junkt: Junktion (Konjunktion, Subjunktion)

mod: Modalwort

prtkl: Partikel (Gradpartikel, Fokus-, Negations-, Abtönungs-, Gesprächs-, Interjektions-)

infprtkl: Infinitivpartikel („zu“, det: „zum“)

L E M M A

Bei der Lemmavergabe orientieren wir uns am Duden. Grundsätzlich gilt nom.sg.m. als Grundform. Einige Besonderheiten gibt es bei Artikeln und Pronomina (z.B. der, die, das > der; es > es; wir > wir; ihr > ihr, s. Duden).

Wenn ein sächsisches Wort sich nicht auf ein hochdeutsches Lemma der gleichen Wurzel zurückführen lässt (z.B. 'Rech' vs. 'Hügel'), wird das sächsische Wort als Lemma verwendet und mit Ausrufezeichen gekennzeichnet. Die hochdeutsche Entsprechung wird als Konzept eingetragen:

!Rech

Hügel

Lemma-Nummerierung:

Im hochdeutschen Basissatz sind die Lemmata fortlaufend ganzzahlig nummeriert. Die Nummern werden mit Unterstrich an das Lemma angehängt, z.B. 'Kuchen_7'.

Für den Dialektsatz gelten folgende Regeln:

- Wenn im Dialektsatz ein Wort steht, das im Basissatz keine Entsprechung hat, wird ein neues Lemma ohne Nummer vergeben.

- Wenn unterschiedliche Lexeme sich entsprechen, wird die gleiche Nummer mit unterschiedlichen Lemmata vergeben.
- Wenn zwei Lemmata aus dem Muster im Dialekt zu einem zusammengezogen werden, werden beide Nummern mit Pluszeichen verbunden: 'durch_4' 'das_5' > 'durchs_4+5'
- Im umgekehrten Fall so: 'durch_4.1' 'das_4.2'
- Bei Wiederholungen im Dialektsatz wird die Nummer aus dem Mustersatz mehrfach vergeben (1:n-Beziehung).

Lehnübersetzungen werden durch ein vorangestelltes Sprachkürzel gekennzeichnet. Mehrwortlexien erhalten zusätzlich eine Gruppennummer. Die Zählung beginnt in jedem Intervall neu.

und	sagten	die	Prüfung
und_22	R1_sagen_23	R1_der_24	R1_Prüfung_2
	ablegen		

KONZEPT

- Wenn kein Konzept angegeben wird, wird beim Abspeichern das Lemma als Konzept übernommen.
- Bei Einzelwörtern wird das Konzept ohne Nummer eingetragen:

Funkich_7
Kuchen

- Setzt sich das Konzept aus mehreren Lemmata zusammen, werden wie bei den Lemmata Nummern vergeben:

hölzern_8.1	Löffel_8.2
Kochlöffel_1	Kochlöffel_2