

Introduction to Tags

The Universal Dependencies framework establishes a comprehensive and universal set of tags for parts of speech (POS), morphological features and syntactic dependencies that can be adopted in the treebanks of individual languages, or supplemented with new morphological features or derivations of core relations when necessary. For the Slovene language data, this includes the adoption of all 17 parts of speech (see Table 1), 22 morphological features spanning 62 distinct values (see Table 2), and 35 types of dependency relations (see Table 3).

Tag	Description
ADJ	adjective
ADP	adposition
ADV	adverb
AUX	auxiliary
CCONJ	coordinating conjunction
DET	determiner
INTJ	interjection
NOUN	noun
NUM	numeral
PART	particle
PRON	pronoun
PROPN	proper noun
PUNCT	punctuation
SCONJ	subordinating conjunction
SYM	symbol
VERB	verb
X	other

Table 1: Part-of-speech tags used in Slovene texts.

Feature	Value	Description
Abbr	Yes	abbreviation

Feature	Value	Description
Animacy	Anim, Inanim	animacy
Aspect	Imp, Perf	aspect
Case	Nom, Gen, Dat, Acc, Loc, Ins	case
Definite	Ind, Def	definiteness or state
Degree	Pos, Cmp, Sup	degree
Foreign	Yes	is this a foreign word?
Gender	Masc, Fem, Neut	gender
Gender[psor]	Masc. Fem, Neut	possessor's gender
Mood	Ind, Imp, Cnd	mood
Number	Sing, Dual, Plur	number
Number[psor]	Sing, Dual, Plur	possessor's number
NumForm	Word, Digit, Roman	numeral form
NumType	Card, Ord, Mult, Sets	numeral type
Person	1, 2, 3	person
Polarity	Neg, Pos	polarity
Poss	Yes	possessive
PronType	Prs, Int, Rel, Dem, Tot, Neg, Ind	pronominal type
Reflex	Yes	reflexive
Tense	Pres, Fut	tense
Variant	Bound, Short	alternative form of word
VerbForm	Fin, Inf, Sup, Part, Conv	form of verb or deverbative

Table 2: Tags for morphological features used in Slovene texts. In the corpus, these are listed in the form of feature and value pairs (e.g., Tense=Pres).

Tag	Description
acl	clausal modifier of noun
advcl	adverbial clause modifier
advmod	adverbial modifier
amod	adjectival modifier
appos	appositional modifier

Tag	Description
aux	auxiliary verb
case	case marking preposition
cc	coordinating conjunction
ccomp	clausal complement
conj	conjunct
cop	copula verb
csbj	clausal subject
dep	unspecified dependency
det	determiner
discourse	discourse element
dislocated	dislocated element
expl	expletive
fixed	fixed multi-word expression
flat	flat multi word-expression
goeswith	disjointed token
iobj	indirect object
list	list
mark	marker (subordinating conjunction)
nmod	nominal modifier
nsbj	nominal subject
nummod	numeric modifier
obj	(direct) object
obl	oblique nominal (adjunct)
orphan	dependent of missing parent
parataxis	parataxis
punct	punctuation symbol
reparandum	overriden disfluency
root	root element
vocative	vocative
xcomp	open clausal complement

Table 3: Tags for syntactic dependency relations (without subtypes) used in Slovene texts.

Revision #3
Created 13 November 2023 14:21:55 by Tina Munda
Updated 23 January 2024 12:52:31 by Tina Munda