Positional Tags: Quick Reference (Czech "HM" Morphology)

Author: Jan Hajic, 2000

Categories

No.	Name	Description
1	POS	Part of Speech
2	SUBPOS	Detailed Part of Speech
3	GENDER	Gender
4	NUMBER	Number
5	CASE	Case
6	POSSGENDER	Possessor's Gender
7	POSSNUMBER	Possessor's Number
8	PERSON	Person
9	TENSE	Tense
10	GRADE	Degree of comparison
11	NEGATION	Negation
12	VOICE	Voice
13	RESERVE1	Unused
14	RESERVE2	Unused
15	VAR	Variant, Style, Register, Special Usage

Values and Description by Category

POS

Value	Description
A	Adjective
C	Numeral
D	Adverb
I	Interjection
J	Conjunction
N	Noun
P	Pronoun
V	Verb
R	Preposition
T	Particle
X	Unknown, Not Determined, Unclassifiable
Z	Punctuation (also used for the *Sentence Boundary* token)

SUBPOS

Value	Description
!	Abbreviation used as an adverb (now obsolete)
#	Sentence boundary (for the *virtual* word `###`)
*	Word `krát` (lit.: *times*) (`POS`: `C`, numeral)
,	Conjunction subordinate (incl. `aby`, `kdyby` in all forms)
.	Abbreviation used as an adjective (now obsolete)
0	Preposition with attached `-ň` (pronoun `něj`, lit. *him*); `proň, naň, ...` (`POS`: `P`, pronoun)
1	Relative possessive pronoun `jehož`, `jejíž`, ... (lit. *whose* in subordinate relative clause)
2	Hyphen (always as a separate token)
3	Abbreviation used as a numeral (now obsolete)

4	Relative/interrogative pronoun with adjectival declension of both types (*soft* and *hard*) (`jaký, který, čí, ...`, lit. *what*, *which*, *whose*, ...)
5	The pronoun *he* in forms requested after any preposition (with prefix `n-`: `něj, něho, ...`, lit. *him* in various cases)
6	Reflexive pronoun `se` in long forms (`sebe, sobě, sebou`, lit. *myself* / *yourself* / *herself* / *himself* in various cases; `se` is personless)
7	Reflexive pronouns `se` (`CASE` = `4`), `si` (`CASE` = `3`), plus the same two forms with contracted `-s`: `ses, sis` (distinguished by `PERSON` = `2`; also number is singular only)
8	Possessive reflexive pronoun `svůj` (lit. *my*/*your*/*her*/*his* when the possessor is the subject of the sentence)
9	Relative pronoun `jenž, již, ...` after a preposition (`n-`: `něhož, niž, ...`, lit. *who*)
:	Punctuation (except for the virtual sentence boundary word `###`, which uses the `SUBPOS` `#`)
;	Abbreviation used as a noun (now obsolete)
=	Number written using digits (`POS`: `C`, numeral)
?	Numeral `kolik` (lit. *how many*/*how much*)
@	Unrecognized word form (`POS`: `X`, unknown)
A	Adjective, general
B	Verb, present or future form
C	Adjective, nominal (short, participial) form `rád, schopen, ...`
D	Pronoun, demonstrative (`ten, onen, ...`, lit. *this*, *that*, *that ... over there*, ...)
E	Relative pronoun `což` (corresponding to English *which* in subordinate clauses referring to a part of the preceding text)
F	Preposition, part of; never appears isolated, always in a phrase (`nehledě (na), vzhledem (k), ...`, lit. *regardless*, *because of*)
G	Adjective derived from present transgressive form of a verb
H	Personal pronoun, clitical (short) form (`mě, mi, ti, mu, ...`); these forms are used in the second position in a clause (lit. *me*, *you*, *her*, *him*), even though some of them (`mě`) might be regularly used anywhere as well
I	Interjections (`POS`: `I`)
J	Relative pronoun `jenž, již, ...` not after a preposition (lit. *who*, *whom*)
K	Relative/interrogative pronoun `kdo` (lit. *who*), incl. forms with affixes `-I` and `-s` (affixes are

	distinguished by the category VAR (for -ĺ) and PERSON (for -s))
L	Pronoun, indefinite vąechnen, sám (lit. *all*, *alone*)
M	Adjective derived from verbal past transgressive form
N	Noun (general)
O	Pronoun svůj, nesvůj, tentam alone (lit. *own self*, *not-in-mood*, *gone*)
P	Personal pronoun já, ty, on (lit. *I*, *you*, *he*) (incl. forms with the enclitic -s, e.g. tys, lit. *you're*); gender position is used for third person to distinguish on/ona/ono (lit. *he/she/it*), and number for all three persons
Q	Pronoun relative/interrogative co, copak, coľpak (lit. *what*, *isn't-it-true-that*)
R	Preposition (general, without vocalization)
S	Pronoun possessive můj, tvůj, jeho (lit. *my*, *your*, *his*); gender position used for third person to distinguish jeho, její, jeho (lit. *his, her, its*), and number for all three pronouns
T	Particle (POS: T, particle)
U	Adjective possessive (with the masculine ending -ův as well as feminine -in)
V	Preposition (with vocalization -e or -u): (ve, pode, ku, ..., lit. *in*, *under*, *to*)
W	Pronoun negative (nic, nikdo, nijaký, ľádný, ..., lit. *nothing*, *nobody*, *not-worth-mentioning*, *no/none*)
X	(temporary) Word form recognized, but tag is missing in dictionary due to delays in (asynchronous) dictionary creation
Y	Pronoun relative/interrogative co as an enclitic (after a preposition) (oč, nač, zač, lit. *about what*, *on/onto what*, *after/for what*)
Z	Pronoun indefinite (nějaký, některý, čikoli, cosi, ..., lit. *some*, *some*, *anybody's*, *something*)
^	Conjunction (connecting main clauses, not subordinate)
a	Numeral, indefinite (mnoho, málo, tolik, několik, kdovíkolik, ..., lit. *much/many*, *little/few*, *that much/many*, *some (number of)*, *who-knows-how-much/many*)
b	Adverb (without a possibility to form negation and degrees of comparison, e.g. pozadu, naplocho, ..., lit. *behind*, *flatly*); i.e. both the NEGATION as well as the GRADE attributes in the same tag are marked by - (Not applicable)
c	Conditional (of the verb být (lit. *to be*) only) (by, bych, bys, bychom, byste, lit. *would*)
d	Numeral, generic with adjectival declension (dvojí, desaterý, ..., lit. *two-kinds/...*, *ten-...*)

e	Verb, transgressive present (endings `-e/-ě, -íc, -ície`)
f	Verb, infinitive
g	Adverb (forming negation (`NEGATION` set to `A/N`) and degrees of comparison `GRADE` set to `1/2/3` (comparative/superlative), e.g. `velký`, `za\-jí\-ma\-vý`, ..., lit. *big*, *interesting*
h	Numeral, generic; only `jedny` and `nejedny` (lit. *one-kind/sort-of*, *not-only-one-kind/sort-of*)
i	Verb, imperative form
j	Numeral, generic greater than or equal to 4 used as a syntactic noun (`čtvero, desatero, ...`, lit. *four-kinds/sorts-of*, *ten-...*)
k	Numeral, generic greater than or equal to 4 used as a syntactic adjective, short form (`čtvery`, ..., lit. *four-kinds/sorts-of*)
l	Numeral, cardinal `jeden`, `dva`, `tři`, `čtyři`, `půl`, `...` (lit. *one*, *two*, *three*, *four*); also `sto` and `tisíc` (lit. *hundred*, *thousand*) if noun declension is not used
m	Verb, past transgressive; also archaic present transgressive of perfective verbs (ex.: `udělav`, lit. *(he-)having-done*; arch. also `udělaje` (`VAR` = `4`), lit. *(he-)having-done*)
n	Numeral, cardinal greater than or equal to 5
o	Numeral, multiplicative indefinite (`-krát`, lit. (*times*): `mnohokrát`, `tolikrát`, `...`, lit. *many times*, *that many times*)
p	Verb, past participle, active (including forms with the enclitic `-s`, lit. *'re* (*are*))
q	Verb, past participle, active, with the enclitic `-»`, lit. (perhaps) *-could-you-imagine-that?* or *but-because-* (both archaic)
r	Numeral, ordinal (adjective declension without degrees of comparison)
s	Verb, past participle, passive (including forms with the enclitic `-s`, lit. *'re* (*are*))
t	Verb, present or future tense, with the enclitic `-»`, lit. (perhaps) *-could-you-imagine-that?* or *but-because-* (both archaic)
u	Numeral, interrogative `kolikrát`, lit. *how many times?*
v	Numeral, multiplicative, definite (`-krát`, lit. *times*: `pětkrát, ...`, lit. *five times*)
w	Numeral, indefinite, adjectival declension (`nejeden`, `tolikátý`, ..., lit. *not-only-one*, *so-many-times-repeated*)
x	Abbreviation, part of speech unknown/indeterminable (now obsolete)
y	Numeral, fraction ending at `-ina` (`POS`: `C`, numeral); used as a noun (`pětina`, lit. *one-fifth*)
z	Numeral, interrogative `kolikátý`, lit. *what* (*at-what-position-place-in-a-sequence*)

}	Numeral, written using Roman numerals (`XIV`)
~	Abbreviation used as a verb (now obsolete)

GENDER

Value	Description
-	Not applicable
F	Feminine
H	Feminine or Neuter
I	Masculine inanimate
M	Masculine animate
N	Neuter
Q	Feminine (with singular only) or Neuter (with plural only); used only with participles and nominal forms of adjectives
T	Masculine inanimate or Feminine (plural only); used only with participles and nominal forms of adjectives
X	Any of the basic four genders
Y	Masculine (either animate or inanimate)
Z	Not fenimine (i.e., Masculine animate/inanimate or Neuter); only for (some) pronoun forms and certain numerals

NUMBER

Value	Description
-	Not applicable
D	Dual
P	Plural
S	Singular
W	Singular for feminine gender, plural with neuter; can only appear in participle or nominal adjective form with gender value `Q`
X	Any

CASE

Value	Description
-	Not applicable
1	Nominative
2	Genitive
3	Dative
4	Accusative
5	Vocative
6	Locative
7	Instrumental
X	Any

POSSGENDER

Value	Description
-	Not applicable
F	Feminine possessor
M	Masculine animate possessor (adjectives only)
X	Any gender
Z	Not feminine (both masculine or neuter)

POSSNUMBER

Value	Description
-	Not applicable
P	Plural (possessor)
S	Singular (possessor)

PERSON

Value	Description
-	Not applicable
1	1st person

2	2nd person
3	3rd person
X	Any person

TENSE

Value	Description
-	Not applicable
F	Future
H	Past or Present
P	Present
R	Past
X	Any (Past, Present, or Future)

GRADE

Value	Description
-	Not applicable
1	Positive
2	Comparative
3	Superlative

NEGATION

Value	Description
-	Not applicable
A	Affirmative (not negated)
N	Negated

VOICE

Value	Description
-	Not applicable
A	Active

P	Passive

RESERVE1

Value	Description
-	Not applicable

RESERVE2

Value	Description
-	Not applicable

VAR

Value	Description
-	Not applicable (basic variant, standard contemporary style; also used for standard forms allowed for use in writing by the Czech Standard Orthography Rules despite being marked there as colloquial)
1	Variant, second most used (*less frequent*), still standard
2	Variant, rarely used, bookish, or archaic
3	Very archaic, also archaic + colloquial
4	Very archaic or bookish, but standard at the time
5	Colloquial, but (almost) tolerated even in public
6	Colloquial (standard in spoken Czech)
7	Colloquial (standard in spoken Czech), less frequent variant
8	Abbreviations
9	Special uses, e.g. personal pronouns after prepositions etc.