

Popis jednotlivých pozic značky

Pozice jsou číslovány od 1 do 16 (16. pozice chybí pouze v korpusech [SYN2000](#), [ORWELL](#) a [InterCorp](#)). Každé pozici odpovídá jiná gramatická kategorie:

1. [Slovní druh](#)
2. [Detailní určení slovního druhu](#)
3. [Jmenný rod](#)
4. [Číslo](#)
5. [Pád](#)
6. [Přivlastňovací rod](#)
7. [Přivlastňovací číslo](#)
8. [Osoba](#)
9. [Čas](#)
10. [Stupeň](#)
11. [Negace](#)
12. [Aktivum/pasivum](#)
13. pozice nepoužita
14. pozice nepoužita
15. [Varianta \(stylový příznak\)](#)
16. [Vid](#)

Pozice 1 - Slovní druh

Označuje hlavní slovní druh, víceméně podle obvyklého schématu známého z českých gramatik včetně školních. Přiřazení i těchto hlavních slovních druhů je však řízeno především potřebami konzistentnosti další analýzy přirozeného jazyka. Proto je možné, že v některých případech (zejména tehdy, kdy se gramatiky a slovníky v určení slovního druhu neshodují nebo uvádějí jiné rozdělení na významy slova) nemusí být zařazení zcela „tradiční“.

Značka	Význam
N	substantivum (podstatné jméno)
A	adjektivum (přídavné jméno)
P	pronomen (zájmeno)
C	numerál (číslovka, nebo číselný výraz s číslicemi)
V	verbum (sloveso)
D	adverbium (příslovce)
R	prepozice (předložka)
J	konjunkce (spojka)
T	partikule (částice)
I	interjekce (citoslovce)
X	neznámý, neurčený, neurčitelný slovní druh
Z	interpunkce, hranice věty

Pozice 2 - Detailní určení slovního druhu

Detailní určení slovního druhu slouží především k zachycení dalších relevantních morfolozických kategorií, které jsou uvedeny na dalších pozicích (ne vždy však jednoznačně). Ze znaku použitého pro detailní určení slovního druhu je možné jednoznačně vyvodit hlavní [slovní druh](#).

Značka	Význam	Značka	Význam
!	zkratka jako adverbium	R	předložka, obyčejná
*	slovo „krát“ (slovní druh: spojka)	S	zájmeno přivlastňovací „můj“, „tvůj“, „jeho“ (vč. plurálu)
,	spojka podřadicí (vč. „aby“ a „kdyby“ ve všech tvarech)	T	částice (slovní druh 'T')
.	zkratka jako adjektivum	U	adjektivum přivlastňovací (na “-ův“ i “-in“)
:	interpunkce všeobecně	V	předložka vokalizovaná („ve“, „pode“, „ku“, ...)
;	zkratka jako substantivum	W	zájmena záporná („nic“, „nikdo“, „nijaký“, „žádný“, ...)
=	číslo psané číslicemi (značkováno jako slovní druh: číslovka - 'C')	X	slovní tvar, který byl rozpoznán, ale značka (ve slovníku) chybí
?	číslovka „kolik“	Y	zájmeno „co“ spojené s předložkou („oč“, „nač“, „zač“)
^	spojka souřadící	Z	zájmeno neurčité („nějaký“, „některý“, „číkoli“, „cosí“, ...)
}	číslovka psaná římskými číslicemi	a	číslovka neurčitá („mnoho“, „málo“, „tolik“, „několik“, „kdovíkolik“, ...)
~	zkratka jako sloveso	b	příslovce (bez určení stupně a negace; „pozadu“, „naplocho“, ...)
@	morfolozickou analýzou nerozpoznaný tvar (slovní druh: 'X' = neznámý)	c	kondicionál slovesa být („by“, „bych“, „bys“, „bychom“, „byste“)
0	předložka s připojeným “-ň“ (něj), „proň“, „naň“, atd. (značkováno jako slovní druh: zájmeno - 'P')	d	číslovka druhová, adjektivní skloňování („jedny“, „dvojí“, „desaterý“, ...)
1	vztažné přivlastňovací zájmeno „jehož“, „jejíž“, ...	e	slovesný tvar přechodníku přítomného (“-e“, “-íc“, “-íce“)
2	slovo před pomlčkou	f	slovesný tvar: infinitiv
3	zkratka jako číslovka	g	příslovce (s určením stupně a negace; „velký“, „zajímavý“, ...)
4	vztažné nebo tázací zájmeno s adjektivním skloňováním (obou typů: „jaký“, „který“, „čí“, ...)	h	číslovky druhové „jedny“ a „nejedny“
5	zájmeno „on“ ve tvarech po předložce (tj. „n-“: „něj“, „něho“, ...)	i	slovesný tvar rozkazovacího způsobu
6	reflexivní zájmeno „se“ v dlouhých tvarech („sebe“, „sobě“, „sebou“)	j	číslovka druhová ≥ 4 , substantivní postavení („čtvero“, „desatero“, ...)
7	reflexivní zájmeno „se“, „si“ pouze v těchto tvarech, a dále „ses“, „sis“	k	číslovka druhová ≥ 4 , adjektivní postavení, krátký tvar („čtvery“, ...)

8	privlastňovací zájmeno „svůj“	l	číslovky základní 1-4, „půl“, ...; sto a tisíc v nesubstantivním skloňování
9	vztažné zájmeno „jenž“, „již“, ... po předložce („n-“: „něhož“, „níž“, ...)	m	slovesný tvar přechodníku minulého, příp. (zastarale) přechodník přítomný dokonavý
A	adjektivum obyčejné	n	číslovky základní ≥ 5
B	sloveso, tvar přítomného nebo budoucího času	o	číslovky násobné neurčité (“-krát“: „mnohokrát“, „tolikrát“, ...)
C	adjektivum, jmenný tvar	p	slovesné tvary minulého aktivního přičestí (včetně přidaného “-s“)
D	zájmeno ukazovací („ten“, „onen“, ...)	q	archaické slovesné tvary minulého aktivního přičestí (zakončení “-t“)
E	vztažné zájmeno „což“	r	číslovky řadové
F	součást předložky, která nikdy nestojí samostatně („nehledě“, „vzhledem“, ...)	s	slovesné tvary pasivního přičestí (vč. přidaného “-s“)
G	přídavné jméno odvozené od slovesného tvaru přítomného přechodníku	t	archaické slovesné tvary přítomného a budoucího času (zakončení “-t“)
H	krátké tvary osobních zájmen („mě“, „mi“, „ti“, „mu“, ...)	u	číslovka tázací násobná „kolikrát“
I	citoslovce (značkováno jako slovní druh: citoslovce - 'I')	v	číslovky násobné (“-krát“: „pětkrát“, „poprvé“ ...)
J	vztažné zájmeno „jenž“ („již“, ...), bez předložky	w	číslovky neurčité s adjektivním skloňováním („nejeden“, „tolikátý“, „několikátý“ ...)
K	zájmeno tázací nebo vztažné „kdo“, vč. tvarů s “-ž“ a “-s“	x	zkratka, slovní druh neurčen/neznámý
L	zájmeno neurčité „všechn“, „sám“	y	zlomky zakončené na “-ina“ (značkováno jako slovní druh: číslovka - 'C')
M	přídavné jméno odvozené od slovesného tvaru minulého přechodníku	z	číslovka tázací řadová „kolikátý“
N	substantivum, obyčejné		
O	samostatně stojící zájmena „svůj“, „nesvůj“, „tentam“		
P	osobní zájmena (vč. tvaru „tys“)		
Q	zájmeno tázací/vztažné „co“, „copak“, „cožpak“		

Pozice 3 - Jmenný rod

Značka

- neurčuje se

Význam

F	femininum (ženský rod)
H	femininum nebo neutrum (tedy nikoli maskulinum)*
I	maskulinum inanimatum (rod mužský neživotný)
M	maskulinum animatum (rod mužský životný)
N	neutrum (střední rod)
Q	femininum singuláru nebo neutrum plurálu (pouze u přičestí a jmenných adjektiv)*
T	masculinum inanimatum nebo femininum (jen plurál u přičestí a jmenných adjektiv)*
X	libovolný rod (F/M/I/N)
Y	masculinum (animatum nebo inanimatum)*
Z	'nikoli femininum' (tj. M/I/N; především u příslovci)*

* Tato značka je k dispozici pouze v korpusech: [SYN2006PUB](#), [SYN2005](#), [SYN2000](#), [ORWELL](#).

Pozice 4 - Číslo

Značka	Význam
-	neurčuje se
D	duál (pouze 7. pád feminin)
P	plurál (množné číslo)
S	singulár (jednotné číslo)
W	pouze v kombinaci s jmenným rodem 'Q' (singulár pro feminina, plurál pro neutra)*
X	libovolné číslo (P/S/D)

* Tato značka je k dispozici pouze v korpusech: [SYN2006PUB](#), [SYN2005](#), [SYN2000](#), [ORWELL](#).

Pozice 5 - Pád

Značka	Význam
-	neurčuje se
1	nominativ (1. pád)
2	genitiv (2. pád)
3	dativ (3. pád)
4	akuzativ (4. pád)
5	vokativ (5. pád)
6	lokál (6. pád)
7	instrumentál (7. pád)
X	libovolný pád (1/2/3/4/5/6/7)*

* Tato značka je k dispozici pouze v korpusech: [SYN2006PUB](#), [SYN2005](#), [SYN2000](#), [ORWELL](#).

Pozice 6 - Přivlastňovací rod

Vyjadřuje rod subjektu nebo objektu, jemuž zájmeno nebo adjektivum přivlastňuje: *jeho/její* (*kniha*), kde *jeho* je tagováno jako přivlastňovací maskulinum a *její* jako přivlastňovací femininum, analogicky u adjektiv *otcův/matčín* (*bratr*) je *otcův* značkován jako maskulinum a *matčín* jako femininum.

Rody mužský neživotný a střední se nikdy jako samostatná kategorie nevyskytují. Rod mužský životný (M) se může vyskytnout jen u přivlastňovacích adjektiv.

Značka	Význam
-	neurčuje se
F	femininum (ženský rod)
M	maskulinum animatum (rod mužský životný)
X	libovolný rod (F/M/I/N)
Z	'nikoli femininum' (tj. M/I/N; u přivlastňovacích adjektiv)*

* Tato značka je k dispozici pouze v korpusech: [SYN2006PUB](#), [SYN2005](#), [SYN2000](#), [ORWELL](#).

Pozice 7 - Přivlastňovací číslo

Kategorie, která se uplatňuje u zájmen: *můj, tvůj, jeho* atp. představují přivlastňovací singulár, protože přivlastňují subjektu nebo objektu v singuláru, zatímco *náš, váš, jejich* jsou označeny jako přivlastňovací plurál.

Značka	Význam
-	neurčuje se
P	plurál (množné číslo)
S	singulár (jednotné číslo)

Pozice 8 - Osoba

Značka	Význam
-	neurčuje se
1	1. osoba
2	2. osoba
3	3. osoba
X	libovolná osoba (1/2/3)*

* Tato značka je k dispozici pouze v korpusech: [SYN2006PUB](#), [SYN2005](#), [SYN2000](#), [ORWELL](#).

Pozice 9 - Čas

Značka	Význam
-	neurčuje se

F	futurum (budoucí čas)
H	minulost nebo přítomnost (P/R)*
P	prézens (přítomný čas)
R	minulý čas
X	libovolný čas (F/R/P)*

* Tato značka je k dispozici pouze v korpusech: [SYN2006PUB](#), [SYN2005](#), [SYN2000](#), [ORWELL](#).

Pozice 10 - Stupeň

Značka Význam

-	neurčuje se
1	1. stupeň
2	2. stupeň
3	3. stupeň

Pozice 11 - Negace

Značka Význam

-	neurčuje se
A	afirmativ (bez negativní předpony „ne-“)
N	negace (tvar s negativní předponou „ne-“)

Pozice 12 - Aktivum/pasivum

Značka Význam

-	neurčuje se
A	aktivum nebo 'nikoli pasívum'
P	pasívum

Pozice 13 - Nepoužito

Značka Význam

-	neurčuje se
---	-------------

Pozice 14 - Nepoužito

Značka Význam

-	neurčuje se
---	-------------

Pozice 15 - Varianta (stylový příznak)

Značkování stylového příznaku je z podstaty věci značně subjektivní a vázané na širší kontext, který při automatickém zpracování nelze brát v úvahu. Hodnoty přisuzované jednotlivým jevům jsou založeny na hodnocení, které se objevují v tradičních jazykovědných příručkách, můžou tedy být zastaralé nebo zcela mylné. Doporučujeme proto údaje na této pozici buď zcela ignorovat nebo je brát s velkou rezervou.

Značka	Význam
-	neurčuje se („základní“ tvar pro kategorie v pozicích 1-14)
1	varianta, víceméně rovnocenná („méně častá“)
2	řidká, archaická nebo knižní varianta
3	velmi archaický tvar, též hovorový
4	velmi archaický nebo knižní tvar, pouze spisovný (ve své době)
5	hovorový tvar, ale v zásadě tolerovaný ve veřejných projevech
6	hovorový tvar (koncovka standardní obecné češtiny)
7	hovorový tvar (koncovka standardní obecné češtiny), varianta k '6'
8	zkratky
9	speciální použití (tvary zájmen po předložkách apod.)

Pozice 16 - Vid

Tato pozice byla k původní sadě doplněna na základě slovníku morfologické analýzy. Tato pozice není k dispozici v korpusech [SYN2000](#) a [ORWELL](#).

Značka	Význam
-	neurčuje se
P	perfektivum (dokonavé sloveso)
I	imperfektivum (nedokonavé sloveso)
B	obouvidé sloveso

— *Jan Hajič, Václav Cvrček*