

Upute za morfosintaktičko označavanje

Za morfosintaktičko označavanje koristimo se MTE oznakama dostupnima na stranici <http://nl.ijs.si/ME/V5/msd/html/msd-hr.html>.

U ovom su dokumentu sadržane upute za označavanje specifičnih konstrukcija i elemenata koji u specifikaciji MTE oznaka nisu jasno razgraničeni.

Imenice

- **imenice kojima su gramatički rod i spol različiti**

Kod imenica koje su gramatički ženskoga roda, ali muškoga spola, gramatički rod označavamo prema spolu.

<i>izbjeglica</i>	Ncmsn
<i>papa</i>	Ncmsn
<i>kolega</i>	Ncmsn
<i>pederčina</i>	Ncmsn

- **akronimi**

1. akronimi koji se odnose na institucije, stranke, odbore, udruge ili tvrtke označavaju se kao vlastita imena

<i>HAZU</i>	Npmsn
<i>HDZ</i>	Npmsn
<i>INA</i>	Npfsn

2. akronimi koji se odnose na predmete, pojave i sl. označavaju se kao opće imenice

<i>DVD</i>	Ncmsn
<i>SF</i>	Ncmsn
<i>USB</i>	Ncmsn

- **imenice stranoga porijekla**

1. imenice za koje imamo jak osjećaj da se učestalo pojavljuju označavamo kao obične opće imenice hrvatskog jezika, s time da rod dobivaju prema deklinaciji; najčešće su muškoga roda

<i>show</i>	Ncmsn
<i>update</i>	Ncmsn
<i>remix</i>	Ncmsn
<i>botox</i>	Ncmsn

(tako te oblike izjednačavamo s grafijski prilagođenima: *šou*, *remiks*, *botoks*)

2. osobna imena, imena bendova, zemljopisna imena i sl. označavamo kao vlastite imenica (**Np**), bez obzira na to što su po vrsti riječi u tom jeziku

<i>Michael Jackson</i>	Npmsn Npmsn
<i>Metallica</i>	Npmsn
<i>Kings of Leon</i>	Npmsn Npmsn
<i>New York</i>	Npmsn Npmsn

3. imena članaka, pjesama, institucija, knjiga, emisija i slično označavamo kao strane riječi (**Xf**), a ako se unutar njih pojavi riječ koja je osobno ili zemljopisno ime, onda samo ta riječ ima oznaku **Np**

Nothing else matters Xf Xf Xf
Le petit prince Xf Xf Xf
The Oprah Winfrey Show -> Xf Npfsn Npfsn Ncmsn (show u skladu s pravilom 1)

4. X vs. Xf

Oznaku X upotrebljavamo za ostatak koji ne ulazi ni u jednu kategoriju (npr. linkovi, nadimci (v. dolje), a oznaku Xf za strane riječi.

- **oblici imenice *dijete***

Imenica *dijete* srednjega je roda. Iako je zbirna imenica *djeca* gramatički ženskoga roda jednine, zbog ujednačenosti s ostalim resursima (i sročnost je množinska *djeca* **su** otišla) označavamo je kao imenicu srednjega roda množine:

djeca Ncnpn

- **bog / Bog**

Kada je napisana malim slovom, imenicu *bog* označavamo kao opću imenicu, a kada je napisana velikim slovom, imenicu *Bog* označavamo kao vlastitu imenicu.

bog Ncmsn
Bog Npmsn

- **Zapad / zapad**

Strane svijeta označavamo kao opće imenice neovisno o tome jesu li pisane malim ili velikim slovom.

<i>Krenuli smo na zapad</i>	Ncmsn.
<i>Na Zapadu</i>	Ncmsn <i>ništa novo.</i>

- **nadimci na društvenim mrežama**

Nadimci na društvenim mrežama koji nemaju veze sa stvarnim imenima označavaju se kao X.

maaaleenica X
palomalo X

- **polusloženice čija je prva sastavnica broj**

Polusloženice čija je prva sastavnica broj označavamo kao obične opće imenice.

20-godišnjak Ncmsn
50-godišnjakinja Ncfsn

- **brojevne imenice tipa dvojica, trojica**

Brojevne imenice tipa dvojica, trojica označavamo kao obične imenice ženskoga roda.

dvojica dužnosnika Ncfsn Ncmpg
trojica čelnika Ncfsn Ncmpg

- **imena država, gradova i sl. kojima je prva riječ pridjev**

Sve dijelove imena država, gradova i sl. kojima je prva riječ pridjev označavamo kao Np, a pridjev lematiziramo u rodu koji je realiziran.

<i>Crna Gora</i> [leme: Crna Gora]	Npfsn Npfsn
<i>Velika Britanija</i> [leme: Velika Britanija]	Npfsn Npfsn
<i>Slavonski Brod</i> [leme: Slavonski Brod]	Npmsg Npmsg

Pridjevi

- **neodređeni vs. određeni pridjevi**

Je li pridjev neodređenog ili određenog lika, vidi se samo u muškom rodu jednine i tu imamo razlike u oznakama:

velik Agpmsnn
veliki Agpmsny

Posvojni pridjevi na –ov, -ev, -in u nominativu jednine uvijek su neodređeni!

<i>Markov prijatelj</i>	<i>Aspmsnn</i>	<i>Ncmsn</i>
<i>Jakovljevi prijatelji</i>	<i>Aspmsnn</i>	<i>Ncmsn</i>
<i>Ivanin prijatelj</i>	<i>Aspmsnn</i>	<i>Ncmsn</i>

- **pridjev (Agp) vs. glagolski particip (App)**

Iako to nije teorijski najsretnije rješenje, radi usuglašavanja anotatora uveden je sljedeći kriterij: ako se pridjev pojavljuje kao zasebna natuknica u rječniku hrvatskoga jezika na Hrvatskom jezičnom portalu (hjp.znanje.hr), onda se označava kao **Agp**, odnosno kao pridjev. Ako se pak pridjev pojavljuje kao dio glagolske paradigme, onda se označava kao **App**.

u svjetlu učestalih pritužbi Agpfpgy Ncfpg (učestao = zasebna natuknica)
On je zadužen za vanjsku politiku. Appmsnn (zadužen nije zasebna natuknica; paziti na neodređenost!)

- **sav i sve**

Kako su zamjenički pridjev i neodređena zamjenica u nekim svojim oblicima isti, uspostavljen je sljedeći kriterij: kada taj oblik stoji uz imenicu, označava se kao pridjev i lematizira na **sav**, a kada stoji samostalno, označava se kao neodređena zamjenica i lematizira na **sve**.

...podnosi ostavku na sve funkcije Agpfpay Ncfpa
...ispunila je sve uvjete... Agpmpay Ncmpa
...pozivaju sve koji žele isporučiti pomoć... Pi-mpa
...po svemu sudeći... Pi-nsi

NAPOMENA!

Oblik sve može biti i čestica (**Qo**):

...svjedoci smo sve većeg pritiska... Qo
...dokaz sve boljih bilateralnih odnosa... Qo

- **roza i rozi**

Pridjev roza nepromjenjiv je, a pridjev rozi promjenjiv. Ako se u rečenici pojavi promjenjiv oblik, lematizira se na rozi, a ako se pojavi nepromjenjiv oblik, lematizira se na roza. MSD se uvijek pridaje prema imenici uz koji pridjev stoji.

roza automobil lema: roza MSD: Agpmsny
rozi automobil lema: rozi MSD: Agpmsny
roza haljine lema: roza MSD: Agpfsgy
roze haljine lema: rozi MSD: Agpfsgy

- **x-godišnji**

Polusloženice sastavljene od broja i pridjeva godišnji označavamo kao pridjev.

Stradao je 53-godišnji motociklist. Agpmsny

Prilozi vs. čestice

Kako je ponekad teško odrediti je li nešto prilog ili čestica, vrstu riječi određujemo prema sljedećoj dogovorenoj specifikaciji:

- **prilozi (Rgp):** zbilja, zapravo, inače, možda, samo, međutim, naime, ipak, kamoli, također
- **čestice (Qo):** baš, čak, bar, barem, tek, evo, eto

Čestice

- **li / da li**

Upitnu česticu *li* označavamo kao *Qq*. Isto je slučaj sa svezom *da li*, koju označavamo kao dvije upitne čestice: *Qq Qq*.

Jeste li znali? Qq

Da li je došao? Qq Qq

NAPOMENA!

Treba paziti na *li* u službi veznika, kada se označava kao veznik zavisnosložene rečenice: **Cs**.

Pokažu li se predviđanja točnima, očekuje se pomlađivanje momčadi. Cs

Prilozi

- **glagolski prilog prošli i glagolski prilog sadašnji**

I glagolski prilog prošli i glagolski prilog sadašnji označavaju se kao *Rr*:

izražavajući Rr

označivši Rr

- **pol/pola, posto, put/puta**

Pol/pola, posto i put/puta označavamo kao prilog (**Rgp**). *Pol* i *pola* lematiziramo na *pola*, a *put* i *puta* na *put*.

...godinu i pol... Rgp

...na pola puta... Rgp

...imaju 65 posto manju šansu... Rgp

...četiri puta tjedno... Rgp

...prvi put... Rgp

Veznici

- **jednostavni veznici**

- **dakle** označavamo kao Cc

Dakle pokriveni ste kako god okrenete. Cc
Živi s mamom koja ima oko 80, dakle od života nema ništa. Cc

- **kao** označavamo kao Cs, osim kad je nedvosmisleno čestica

izgleda kao odlična ideja Cs
kroz prijevare kao što je izdavanje lažnih računa Cs Cs
Kao prvi susjedi, želimo izgraditi dobre odnose. Cs

On je kao lud. Qo

- **složeni veznici**

Svaku od sastavnica složenih veznika označavamo kao Cs, odnosno Cc, ovisno o tome je li rečenica zavisnosložena ili nezavisnosložena. Složeni veznici nezavisnosloženih rečenica: *tek što, jedino što, samo što*. Ostali su složeni veznici zavisnosloženi.

Prije nego što odem, reći ću ti još jedno. Cs Cs Cs
Budući da si napisao zadaću, možeš na igranje. Cs Cs
Ići ćemo u kino nakon što napišemo zadaću. Cs Cs
Zatvoreni su zato što nisu plaćali porez. Cs Cs

- **s obzirom da**

Svezu *s obzirom da* ne označavamo kao složeni veznik nego kao kombinaciju prijedložno-padežnog izraza i veznika: **Si Ncmsi Cs**.

...s obzirom da je odbio pristati na dogovor s vladom. Si Ncmsi Cs

- **ne samo... ..nego**

Svezu *ne samo... ..nego* ne označavamo kao složeni veznik nego kao kombinaciju čestice i priloga i veznika: **Qz Rgp Cc**.

...koji utječu ne samo na državne institucije, nego i na ljudska prava. Qz Rgp Cc

Prijedlozi

- **prema**

Prijedlog *prema* uvijek dolazi uz lokativ.

obveze prema vjerovnicima SI Ncmpl
sve ide prema planu SI Ncmsl

- **prijedlozi s i sa**

Bez obzira na to je li prijedlog ostvaren kao *s* ili *sa*, lema je uvijek **sa**.

Nastavi s radom. s sa S Si (token, type, POS, MSD)
jagode sa šlagom sa sa S Si (token, type, POS, MSD)

Brojevi i brojevnice riječi

- **sati 19:00 ili 0:36**

Sate izražene brojevima označavamo kao **Mds**.

...koji je na rasporedu u 21:45 sati. Mds Ncmprg
Radimo od 10:00 do 16:00 h. Mds Mds Y
ukrcaj putnika u 7,30 Mds

- **70-ih**

Godine u obliku dekada označavamo kao **Mds**.

početkom 70-ih godina prošlog stoljeća Mds
...kada ste solo u svojim 30-ima... Mds

- **30-ak, 50-ak**

Brojevnice riječi koje označavaju približnu količinu označavamo kao **Mds**.

Odsvirali su 30-ak koncerata. Mds
...predložili su još 50-ak svjedoka. Mds

- **nedeklinirani oblici brojeva od dva do četiri**

Kada brojevi koji se inače dekliniraju dolaze u nedekliniranom obliku, označavamo ih kao **Mlc**.

od strane dvije autorice Mlc
prije tri godine Mlc

- **telefonski brojevi, datumi**

Kada imamo cijele telefonske brojeve i datume, označavamo ih kao **Mds**.

njen broj je 098 450 525
njen broj je 098450450
U petak, 26.03.2010. (...)

Mds Mds Mds (jer su zbog razmaka tri odvojene pojavnice)
Mds (ako nema razmaka, onda je jedna pojavnica)
Mds