

FORMALIZACIJA JEDNOG DELA SRPSKOHrvatskog JEZIKA - Prikaz

Marica D. Prešić

Abstrakt: A survey of formalisation of a part of Serbo-croat is given. The corresponding formal language Serb is a categorial language which consists of nouns, all sorts of pronouns, adjectives, adverbs, verbs, quantifiers and in addition variables and case operations. The main purpose of such a formalisation is to preserve the various roles which noun cases play in Serbo-croat and other languages with inflection. The language Serb, its connection with natural language expressions and a corresponding set - theoretical interpretation are considered in all detail in [4].

Formalizacija engleskog jezika tipa Montague-a [2], [3] prenesena na jezike sa izrazitom fleksijom, kakav je srpskohrvatski, deluje prilično grubo. Jedan od glavnih razloga je što se padeži i njihova mnogostruka uloga na takav način potpuno zaobilaze. U formalizaciji koju ovde izlažemo padeži zauzimaju bitno mesto.

Budući da su padeži samo različiti oblici neke imenice (zamenice, imeničkog ili zameničkog izraza), koje karakterišu različite uloge koje ta imenica može imati u rečenici ili nekom drugom složenom jezičkom izrazu, pri matematičko-logičkom pristupu prirodnim jezicima u vezi sa svakim padežom prirodno se pojavljuje po jedna operacija dužine jedan. Razmotrimo po bliže takav pristup padežima u slučaju srpskohrvatskog jezika. Kao osnovne padeže uzimamo:

(1) nominativ, genitiv, dativ, akuzativ, instrumental i lokativ.

Vokativ isključujemo iz dva razloga. Prvi je što vokativ ima posebnu

Uprkos potrebnoj ulozi koja je ograničena na obraćanje i dozivanje.

U slučaju što su svi padežni oblici neke imenice, izuzev vokativa, u istom licu, čak je vokativ u drugom, pa bi se za strogo gramatičko-logičke razmatranje svih padežnih oblika morali uključiti i indeksi što bi nas dovele u intenzionalnu logiku i indeksitane skupovne interpretacije.

Kao oznake za padežne oblike (1) koristimo gornje indekse<sup>1)</sup>

(2)  $i, 2, 3, 4, 5, 6$

Maže, ako je a neka imenica onda njene osnovne padežne oblike označavamo po redu

(3)  $a^1, a^2, a^3, a^4, a^5, a^6$

Što je poznato pomoću predloga se grade složeni ili izvedeni padežni oblici. Pri tome predlog smatramo određenim tek ako je pratećim uz koji osnovni padež ide. Iako su

u (sa genitivom), u (sa akuzativom), u (sa lokativom)

tri različita predloga, čak su na primer,

u kuću, u kuću, u kući

tri različita izvedena padežna oblika imenice kuća koje u skladu sa uvedenim oznakama (3) zapisujemo i ovako:

(4)  $u\ kuća^2, u\ kuću^4, u\ kući^6$

Pretpostavimo sada da su svi padeži, kako osnovni tako izvedeni poredjani u niz tako da prvo dođu osnovni, zatim oni izvedeni iz nominativa (poredjani na neki dogovorom način) potom oni izvedeni iz genitiva i najzad padeži izvedeni iz lokativa. Koliko je ukupan broj svih padeži, k. onda kao njihove oznake koristimo gornje indekse

$1, 2, 3, 4, 5, 6, 7, 8, \dots, k_0$

U slučaju dogovora u svi predlozi koji idu uz nominativ, onda je

$a^7$

grupa oznaka za ovaj izvedeni predlog

od  $a^1$

imenice a.

Već smo pomenuli da se padeži, kako osnovni tako izvedeni mogu matematički shvatiti kao operacije dužine jedan što podrobnije uvodno narednom definicijom.

Definicija 1. Padeži su operacije dužine jedan za koje po redu koristimo oznake

$$(6) \quad I, II, III, IV, V, VI, VII, \dots, K_0$$

i kojima se od imenice a prelazi po redu na njene padežne oblike

$$(7) \quad a^1, a^2, a^3, a^4, a^5, a^6, a^7, \dots, a^{K_0}$$

Dakle, s-ti padež zadovoljava jednakost

$$(8) \quad \bar{s}(a) = a^s \quad (s = 1, 2, \dots, K_0)$$

Prethodna definicija se može shvatiti na više načina. Jedan je da se kao polazni pojmovi uzmu imenica a i padežne operacije (6); padežni oblici (7) su tada rezultati primena padežnih operacija na imenicu a. Drugi način je da se kao polazni uzmu imenica a i padežni oblici (7) a da se padežne operacije uvedu jednakostima (8).

Postavlja se pitanje šta je sama imenica a. Za sada se kao odgovor vide dve mogućnosti. Prva je da se a shvati kao neki opšti padežni oblik, zvaćemo ga i rečnički, koji se sreće prilikom nabiranja, u spiskovima, enciklopedijama, rečnicima. Druga mogućnost, dualna prethodnoj jeste da se a shvati kao neki opšti padež u koji su uključeni svi ostali padeži, i osnovni i izvedeni. U prilog ovoj drugoj mogućnosti ide činjenica da se svaki pojam prirodno raslojava na čitav niz podpojmova i da se jedno takvo raslojavanje može dovesti u vezu sa padežima. Tako, pojam čovjek se raslojava na:

- čoveka nekog od koga ne može nešto uraditi, ostvariti, završiti od on, kao on i sl.
- čoveka od koga nešto potiče, od koga neko nešto uzima, očekuje, dobija, iz koga neko nešto čupa, u koga su vredne ruke i sl.
- čoveka kome neko nešto naučuje, kome je nešto namenjeno, upućeno,

ka kome se neko kreće i sl.

- čoveka koqa neko posmatra, očekuje, voli, u koga neko ima povere-  
nja, po koga je neko dušao, pred koga je neko stao i sl.
- čoveka kojim se neko poslužio za neke ciljeve, pred kojim neko sto-  
ji, za kojim neko žali, s kojim neko radi,
- čoveka o kome se puno govori, u kome se nešto kuva, prema kome se  
svi ravnaju, po kome se može i gaziti i sl.

U skladu sa tim, imenica čovek se može shvatiti kao uređjena  $k_0$ -  
torka svih svojih padežnih oblika tj., u skladu sa usvojenim označa-  
vanjem, kao

(čovek<sup>1</sup>, čovek<sup>2</sup>, čovek<sup>3</sup>, ... , čovek<sup>k<sub>0</sub></sup>)

Oba pristupa su u tesnoj vezi jedan s drugim ali zbog izvesnih pre-  
dnosti pri skupovnoj interpretaciji opredeljujemo se za prvi pristup.  
Budući da se po padažima pored imenica menjaju i zamenice, pridevi,  
brojevi kao i imenički, zamenički, pridevski i brojevni izrazi, to  
se definicija padežnih operacija (8) može na prirodan način produ-  
žiti i na te vrste reči.

Glavna uloga padeža je u rečenici. Ograničavamo se na tzv. elementa-  
rne rečenice, tj. one nastale primenom izvesnog glagolskog izraza  
(predikata) na imeničke ili zameničke izraze, dakle na rečenice u  
kojima ne učestvuju rečenični veznici. Kod fleksivnih jezika u osn-  
ovne gramatičke odrednice glagola (jednostavnosti radi ograničavamo  
se na glagole trećeg lica sadašnjeg vremena izjavnog načina u akti-  
vu) dolaze pored dužine i padeži u kojima izrazi na koje se glagol  
primenjuje moraju biti. Zbog toga u vezi sa svakim glagolom uvodimo  
tzv. naznaku koja je uređjena n-torka oblika

(8)  $\{i_1, i_2, \dots, i_n\}$

gde je n dužina glagola, a  $i_j$  utvorenjeni elementi skupa

(1L)  $\{1, 2, 3, \dots, k_0\}$

- skupa sačinjenog od prirodnih brojeva koji odgovaraju padežima  
(8). Tako svi neprelazni glagoli imaju naznaku

(1)

budući da se takvi glagoli primenjuju na imeničke ili zameničke izraze u nominativu. Dakle, na primer glagolima:

radi (neko nešto), piše (neko nekome nešto)

vadi (neko nešto iz nečega), prenosi (neko nešto iz nečega u nešto drugo)

odgovaraju po redu oznake

(1, 4), (1, 3, 4), (1, 4, iz (2)), (1, 4, iz (2), u (4))

Pri tome smo sa iz (2), u (4) označili prirodne brojeve koji odgovaraju složenim padežima izvedenim pomoću predloga iz koji ide sa genitivom i predloga u koji ide sa akuzativom. Slično označavanje koristimo i u donjem izlaganju. Napominjemo da ćemo glagol smatrati određenim tek kad je precizirano koja mu je naznaka, budući da jednom glagolu prirodnog jezika može odgovarati više naznaka. Tako, u vezi sa različitim ulogama glagola piše<sup>1)</sup> smatrano da su

piše (neko)<sup>2)</sup>

piše (neko nešto)

piše (neko nekome nešto)

piše (neko nešto zbog nečega)

piše (neko nešto nekuda zbog nečega)

medjusobno različiti glagoli čije su naznake po redu

(1), (1, 4), (1, 3, 4), (1, 4, zbog (2)), (1, 4 u(4)zbog (2))

ostale uloge padeža zaobilazimo u ovom izlaganju. Napominjemo samo da se uloga padeža odnosno padežnih oblika imenice (zamenice, imeničkog ili zameničkog izreza) u sintagmama može matematički podvesti pod pojam operacije: i imeničke, pridevske, priloške, glagolske. Posebno raspravljanje svih vrsta sintagmi zahteva, međutim, mnogo prostora i vremena. Principa radi ističemo samo da se široka klasa

1) taj primer potiče sa ilge išecke - Tomić

2) tj. bavi se pisanjem

padežnih sintagmi može obuhvatiti jednom vrstom složenih imeničkih, odnosno zameničkih izraza u kojima padežni oblici imaju ulogu imeničkih, odnosno zameničkih operacija dužine jedan. Odgovarajuća definicija glasi:

Definicija 2. Neka je  $a^S$  imenički (ili zamenički) izraz proizvoljnog padeža  $s$  a  $b^i$  izraz datog padeža  $i$ . Primenom izraza  $b^i$ , snvaćenog kao imenička (zamenička) operacija dužine jedan, na izraz  $a^S$  nastaje nov imenički (zamenički) izraz istog padeža  $s$ . Umesto uobičajene matematičke oznake

$$(11) \quad a^i (a^S)$$

za rezultat primene operacije  $b^i$  na izraz  $a^S$  koristimo i oznaku:

$$(12) \quad (a^S b^i)$$

koja je bliže pisanju u prirodnom jeziku. Izgrađe, kojih u prirodnom jeziku nema, koristimo iz čistu matematičkih razloga: da bismo obezbedili jednoznačnost izgradnje složenih izraza pomoću njegovih podizraza.

Brisanjem zagrada iz (12) nastaju uobičajeni izrazi prirodnog jezika kakvi su na primer:

žena<sup>S</sup> plavih očiju, sto<sup>S</sup> sa drvenim nogama, dete<sup>S</sup> bez  
roditelja, visoka<sup>S</sup> kuća<sup>S</sup> bez prozora

kao i složeniji izrazi:

(žena<sup>S</sup> sa detetom) bez oca, žena<sup>S</sup> sa (detetom bez oca)  
čovek<sup>S</sup> sa (detetom koga vidiš), (tkanina<sup>S</sup> od vune) sa  
čvorovima

Kod kojih se brisanjem dolazi do višeznačnih izraza, čija se značenje u prirodnom jeziku obično precizira pisanjem ili ne pisanjem izvesnih zagrada.

Da bi se izbegle višesmislenosti prirodnog jezika, koje najčešće nastaju zbog višeznačne uloge koju reči mogu imati, pored zagrada koje su često opasnosne, koriste se i vrste (kategorije, tipovi) -

pomoćni znaci koji se pridružuju svakoj reči i služe za preciziranje njene uloge. Pri tome je čest slučaj da jednoj reči, u skladu sa raznim ulogama koje ona može imati, odgovara više vrsta.

U skladu sa prethodnim izlaganjem o padežima nameće se kao prirodno da se za polazne vrste izaberu znaci

$$(13) \quad \emptyset, 0, 1, 2, 3, 4, \dots, k_0$$

koji će služiti kao vrste redom za:

formule, rečnički padež, nominativ, genitiv, ... , padež  $k_0$

Polazeći od znakova (13) skup vrsta - označimo ga sa Vrs - uvodimo narednom definicijom

Definicija 3.

1<sup>o</sup> Znaci (13) jesu vrste

2<sup>o</sup> Ako su  $t, t_1, \dots, t_n$  vrste onda je vrsta i znak <sup>1)</sup>

$$(14) \quad (t, (t_1, \dots, t_n))$$

3<sup>o</sup> Vrste, tj. članovi skupa Vrs se dobijaju jedino konačnom primenom prethodnih dvaju pravila.

U slučaju drugih fleksivnih jezika definicija skupa Vrs će se razlikovati samo u dužini polaznog niza (13).

Deo prirodnog jezika indeksiran skupom Vrs nazivamo vrstovni jezik.

U daljem izlažemo vrstovni jezik Srps koji se sastoji od izvesnih reči srpskohrvatskog jezika i dodatnih znakova promenljivih, zagrada i zareza. Jednostavnosti radi ograničavamo na jedninu, na stvarne imerice i to samo na vlastite i zajedničke i na aktivno i redikalne glasove u trećem licu sadašnjeg vremena izjavnog načina u aktivu.

Jezik Srps čine naredne reči:

1) u slučaju  $n=1$  koristimo i kraće pisanje  $(t, t_1)$ , ukoliko ne postoji mogućnost pojave dvosmislenosti.

(i) Sve stvarne vlastite i zajedničke imenice srpskonrvatskog jezika nabrojane prema nekom standardnom rečniku pri čemu imenici padeža i odgovara vrsta

(15)  $\quad i$

Svakoj imenici a dodeljene su odrednice za rod, živost i brojnost koje označavano po redu:

Rod (a), Živ (a), Broj (a)

Pored osnovne uloge da označavaju bića stvarne imenice mogu, prema Definiciji 2, imati i ulogu imeničkih operacija, a takodje i pridevskih i priloških operacija. U prvom slučaju imenici padeža i odgovara vrsta

(16)  $\quad (i, i)$

a u drugom i trećem

(17)  $\quad (t, t)$

gde je t vrsta prideva, odnosno priloga (videti odeljke (iv) i (x)).

(ii) Za svaku vrstu  $i \in \underline{Vrs}$  po jedan prebrojiv niz imeničkih prome-nljivih

(18)  $\quad x_1^i, x_2^i, x_3^i, x_4^i, \dots$

(iii) Sve lične zamenice za jedninu, pri čemu je zamenica padeža i vrste (13)

(iv) Svi pridevi, izuzev kontekstnih kakvi su recimo:

ovdašnji, ondašnji, sadašnji, budući, prethodni, naredni

Pri tome je pridev datog roda, živnosti i padeža i ili vrste

(19)  $\quad (i, i)$

ili vrste

(20)  $\quad ((\emptyset, i), (\emptyset, i))$

Prva vrsta se odnosi na pridev u atributskoj a druga na pridev u predikatskoj ulozi.

Prvevrste uključuje o i sve komparative i superlative (ukoliko postoje). Pri tom komparativi pored vrsta (19), (20) mogu biti i vrste



(21) (i, (i, od (2)))

što se odnosi na ulogu komparativa u izrazima kao:

devojka lepša od vile, letilica brža od zvuka, drvo  
više od kuće, plastika tvrdja od lima

(v) Sve prisvojne zamenice u jednini, tj. zamenice

(22) moja<sup>i</sup> (moja<sup>i</sup>, moje<sup>i</sup>), tvoj<sup>i</sup>, njegov<sup>i</sup>, njen<sup>i</sup>  
( $i \in \{0, 1, 2, \dots, k_0\}$ )

koje kao i pridevi, mogu biti vrste (19) ili (20) u zavisnosti od toga da li imaju atributsku ili predikatsku ulogu.

(vi) Sve pokazne imeničke, odnosno pridevske zamenice:

(23) ovaj<sup>i</sup> (ova<sup>i</sup>, ovo<sup>i</sup>), taj<sup>i</sup>, onaj<sup>i</sup>

(24) ovakav<sup>i</sup> (ovakva<sup>i</sup>, ovakvo<sup>i</sup>), takav<sup>i</sup>, onakav<sup>i</sup>

(25) ovoliki<sup>i</sup> (ovolika<sup>i</sup>, ovoliko<sup>i</sup>), toliki<sup>i</sup>, onlikiki<sup>i</sup>  
( $i \in \{0, 1, 2, \dots, k_0\}$ )

koje su sve vrste (19) a zamenice (23) su još i vrste (1b), dok su zamenice (24) i (25) još i vrste (20).

(vii) Pokazne priloške zamenice

(26) ovako, tako, onako

koje su, već prema tome koju ulogu imaju jedne od vrsta

(27) ((i, i), (i, i))

(28) (((∅, 1), (∅, 1)), ((∅, 1), (∅, 1)))

(29) (t, t) gde je t vrsta nekog glagola

(30) (∅, ∅)

(viii) Sve odnosne pridevske zamenice

(31) koji<sup>i</sup> (koja<sup>i</sup>, koje<sup>i</sup>), kakav<sup>i</sup>, koliki<sup>i</sup>, čiji<sup>i</sup>  
( $i \in \{0, 1, \dots, k_0\}$ )

kao i složen simbol

(32) takav<sup>i</sup> da (takva<sup>i</sup> da, takvo<sup>i</sup> da)

pri čemu su i (31) i (32) vrste

(33) ((i, ∅), i)

(ix) Sve odnosne priloške zamenice (tj. zamenički prilozi):

(24) Kako, koliko, kad, otkad, dokad, kuda, otkud, kamo, gde, odakle i sl.

koji su svi vrste oblika

(35)  $(\langle t, \mathcal{V} \rangle, t)$

gde je  $t$  vrsta nekog priloga (višeći naredni odeljak)

(x) Svi prilozi, kojima kao vrste odgovaraju (19), (27), (28), (29), (30) ili

(36)  $(\langle \emptyset, \emptyset \rangle, 0)$

paš samo jedino prilozi za količinu mogu biti vrste (19), oonus (36).

(xi) Svi redni brojevi; redni broj padeža i je vrste (9)

(xii) Padežne operacije (6), koje se pored imeničkih i zameničkih izraza mogu primenjivati i na prideve i obrnovo te su stoga jedne od vrsta

(37)  $(i, 0)$

(vrste padežne operacije I)

(38)  $(\langle i, i \rangle, \langle 0, 0 \rangle)$

(xiii) Svi kolikovnici, čiji je spisak inače poduži [4]. Tu spadaju izrazi oblika

$C a$

gde je  $a$  neki opšti imenički izraz dok  $C$  može biti neka količni imenica kao:

većina, manjina, mnoštvo, obilje

prilog ili priloški izraz za količinu, glavni broj, pokazna priloška zamenica, neodređena pridevska zamenica, pridev ili pridevski izraz za količinu kao

množi, brojan, sav, ceo, skoro sav

Pored navedenih u kolikovnike spadaju i neodređene imeničke zamenice svako, neko i broj obadva kao i razni složeni izrazi za količinu kakvi su na primer:

postoji broj ljudi, izvesta. broj ljudi, postoji sasvim malo vode, svaka približno ovalika količina vode, skoro svaki čovak i sl.

(xiv) Svi glagolci prvog stupnja, pri čemu glagolci naznake (9) odgovaraju vrsti

$$(39) \quad (Z, (i_1, i_2, \dots, i_n))$$

(xv) Svi glagolci višeg stupnja, pri čemu glagolci stupnja  $n+1$  i naznake

$$(i_1, \dots, i_{r-1}, v, i_{r+1}, \dots, i_n) \quad (v \text{ je naznaka nekog glagola stupnja } n)$$

odgovaraju vrsti

$$(40) \quad (Z, (i_1, \dots, i_{r-1}, t, i_{r+1}, \dots, i_n))$$

gde je  $t$  vrsta koja odgovara naznaci  $v$ .

(xvi) Svi rečenični veznici pri čemu su veznici dužine jedan vrste

(30) a veznici dužine dva vrste

$$(41) \quad (Z, (\emptyset, \emptyset))$$

i uopšte veznici dužine  $n$  ( $n=1, 2, 3, \dots$ ) su vrste

$$(42) \quad (Z, (\emptyset, \emptyset, \dots, \emptyset))$$

n-puta

(xvii) Svi pridevski veznici, pri čemu su veznici dužine jedan vrste

(27), odnosno (28), dok su veznici dužine dva vrste

$$(43) \quad ((i, i), ((i, i), (i, i)))$$

odnosno vrste

$$(44) \quad ((Z, 1), (Z, 1)), ((\emptyset, 1), (Z, 1)), ((Z, 1), (Z, 1))$$

(xviii) Svi priloški veznici su vrste (31) a veznici dužine jedan vrste

$$(45) \quad (t, t)$$

dok su veznici dužine dva vrste

$$(46) \quad (t, (t, t))$$

gde je  $t$  neka vrsta priloga.

(vix) Poročni znaci

(vii)  $\{ \dots \}$ ,

ei = srpski jezik Serb završava.

Izrazi jezika Serb se dalje definišu slično kao u svakom vrstovnom jeziku. Glavni dio te definicije glasi:

Ako je  $t$  izraz vrste  $(t, (t_1, \dots, t_n))$  a  $B_1, \dots, B_n$  su po redu izrazi vrsta  $t_1, \dots, t_n$  onda je

$$(47) \quad (t(B_1, \dots, B_n))$$

izraz vrste  $t$ .

Na to se još često zantavaju i razni uslovi o rodu, broju, živosti. Tako, ako je  $A$  pridev ili pridevski izraz a  $B_1$  imenički izraz, onda odgovarajući uslovi glase:

$$(48) \quad \text{Rod}(A) = \text{Rod}(B_1), \text{ Broj}(A) = \text{Broj}(B_1), \\ \text{Živ}(A) = \text{Živ}(B_1)$$

Od vrstovnih izraza jezika Serb lako se prelazi na izraze prirodno srpskohrvatskog jezika. Najčešće je dovoljno obrisati zagrada i kad-kad još treba izvršiti i izvesno prerasporedjivanje polaznih izraza, kao u slučaju elementarne formule

(piše (Jovan, Nadi, pismo) )

kojoj odgovara ova rečenica (sa uobičajenim, prirodnim redosledom reči):

Jovan piše Nadi pismo

Da jezik Serb se može izgraditi odgovarajuća skupovna interpretacija, koja je funkcijska, slučajno kao što je to slučaj sa drugim poznatim jezicima takvog tipa. Evo ističemo samo neke glavne karakteristike te interpretacije:

se polazni delovi uzimaju se izvesni neprazni skupovi

$$(49) \quad T_1, T_2, \dots, T_k$$

koji odgovaraju polaznim vrsta  $a$  (43) pri čemu

$$(50) \quad T_1 = \dots,$$

valje da za svaki član  $s$  iz  $\Sigma$  postoji po jedan ili en  $\tau_s$  naredna rekurzivno definicija

definicija 4. Ako je  $s$  polazni znak iz  $\Sigma_S = \Sigma_0$  a ako je  $s$  culita  $(s, t_1, \dots, t_n)$  onda

$$(51) \quad \tau_s = \{ \tau_{t_i} : t_1 \dots t_n \rightarrow s \}$$

tj.

$$(52) \quad \tau_s = \tau_{t_1} \times \dots \times \tau_{t_n}$$

Kada toj definiciji, na pri ep, badožna operaci a  $\tau$  vrste (57) se tumači kao funkcija koja prevodi  $\Sigma_0$  u  $\Sigma_1$ , tj. kao član skupa

$$(53) \quad \Sigma_1^{-10}$$

lično se odnosi i na ostale izraze jezika Ser, a cin što neki od njih imaju ulogu pronenjivih a drugi ulogu znakova konstanta na čemu u ovom pregledu nismo insistirali a što se u strogo i podrobnoc izlažanju mora uzeti u obzir (Videti detalje u [4]).

## LITERATURA

- [1] Chomsky, A. S., Grammars as theories: the case for axiomatic grammar I, Theoret. Linguist. 1 (1974), no. 1/2, 33-115, 1, 3(1976), no. 1/2, 1-96
- [2] Montague, R., English as a formal language I, Linguaggi nella Società e nella Tecnica (E. Loescher editore), Milano 1970
- [3] Montague, R., The proper treatment of quantification in ordinary English, Lectures to natural language (L. Lindblom, E. G. Trautson, E. Suppes editors), Dordrecht 1973
- [4] Lesić, M. S., Logička gramatika (manuskript predložio pripremljen za štampu), Zagreb 1982