

UDK 811.163.6:681.3

Jure Zupan

Kemijski inštitut, Ljubljana

PROBLEMI IN NEKAJ REŠITEV RAČUNALNIŠKIH OBDELAV SLOVENSКИH BESEDIL

Opisan je problem avtomatske besedne analize slovenskih besedil. Posebej je izpostavljen problem različnosti končnic pri sklanjatvah in spregatvah. Navedeni so vsi nizi različnih računalniških končnic, ki so bili najdeni pri pregledu in klasifikaciji približno 156.000 slovenskih besed. Predstavljeni računalniški slovar, katerega vsako geslo je sestavljeno iz *računalniške osnove* besede in pripadajočega niza *računalniških končnic* vsebuje približno 150.000 pregibnih in okrog 6.000 nepregibnih gesel. Opisani slovar omogoča dokajšno stopnjo avtomatizacije pri besedni analizi. Navedena sta tudi dva primera uporabe.

A problem of automatic word analysis of Slovene texts is described and discussed. As the essential issue the variety of *computational endings* in Slovenian declination is exposed. All sets of different endings identified during the classification of about 156.000 Slovenian words are listed. The dictionary in which each entry is composed of the *computational root* and the set of *computational endings* consists of approximately 150.000 and about 6,000 flexible and non-flexible words. The described dictionary enables a variety of applications with two of them shown as examples.

Prva naloga, ki jo je pri obširnejši obdelavi besedila treba opraviti, je podrobna besedna analiza. Besedna analiza je izhodišče za nadaljnje raziskave, kot so npr.: primerjave med različnimi avtorji oziroma načini njihovih pisanj, študije naraščanja besednih zakladov v odvisnosti o obdobja pisanja in »zorenja« pisateljev, iskanje vplivov posameznih literarnih smeri v opusih sodobnikov, iskanje konkordanc v kontekstih itd. Žal mora vsak raziskovalec besedno analizo v pretežni meri še vedno opraviti »ročno«. Povsem računalniško vodena (brez posredovanja strokovnjaka) besedna analiza je zaradi mnogoterosti slovenskih pregibnih oblik in variacij tako otežkočena, da doslej še ni uspelo izdelati računalniškega programa, ki bi lahko posej zadovoljivo opravil. Beseda *zadovoljivo* je v tem kontekstu mišljena tako, da bi morala računalniška besedna analiza zadostiti standardom, ki sta jih postavila Jakopin in Bizjak (Jakopin 1995; Jakopin, Bizjak 1997) in biti hkrati narejena nekaj velikostnih razredov hitreje kot ročna. Posebej je treba poudariti, da se vsi, ki na tem področju delamo, zavedamo, da bo določen delež ročne analize nujen še precej časa. Hkrati vemo, da vsak napredek v tej smeri precej olajša ta najbolj utrujajoč in za marsikoga tudi najmanj zanimiv del raziskav. Avtomatična besedna analiza bo precej pripomogla k bolj natančnim in hitreje dobljenim končnim rezultatom kompleksnejših raziskav.

Zavedamo se tudi, da je prednost slovenščine pred drugimi jeziki, ki imajo bistveno revnejše pregibne možnosti, predvsem pred angleščino, prav bogastvo in razvejanost pregibnih oblik. Natančno poznavanje in sistemizacija vseh možnosti pregibnih oblik nudi izhodišče za izboljšave avtomatske (računalniško vodene) analize besedil. Hkrati z novimi dosežki narašča upanje, da bomo povsem (ali vsaj v veliki meri) avtomatizirano besedno analizo lahko nekoč tudi res imeli. Pri jezikih

s skromno pregibno raznolikostjo različne besedne vrste pogosto nastopajo v isti formalni (črkovni) obliki. Zato jim lahko šele iz konteksta določimo pravilno formalno opredelitev. Primer je beseda *iron*, ki v angleščini pomeni v zvezah *to iron*, *the iron in the iron (lady)* bodisi *likati* (glagol), *železo* (samostalnik) ali pa *odločno* (*železno*) žensko (pridevnik). Njihov pomen je potemtakem določen glede na položaj v stavku (kontekst) in mnogokrat tudi glede na ločila. Obdelava konteksta je tudi pri besedni analizi za take jezike neizogibna. Za razliko od angleščine (Popovič, Willet 1990; Zupan 1999) se da pri slovenščini v bistveno večjem številu primerov že iz same besedne forme (črkovne oblike = osnova + končnica) ugotoviti, za katero besedno vrsto gre. Večinoma lahko hkrati ugotovimo še sklon, spol, število ali čas. Natančna analiza možnih končnic pri sklanjatvi ali spregatvi nam ob upoštevanju pravil, ki določajo povezave med končnicami besed in predlogi, zaimki ali pomožnimi glagoli, omogoča precejšno mero avtomatizma.

Na načelni jezikoslovni ravni so pri sklanjanju in spreganju končnice in pravila za uporabo predlogov ter oblik pomožnega glagola dokaj natančno določene. Jezikoslovna pravila, po katerih delimo sklanjatve in spregatve (Toporišič 1975; 1984; 1994; Lenček 1982) in po katerih tvorimo pravilne končnice, mehčamo soglasnike ali izpuščamo polglasnik, so znana, a so za računalniško uporabo preveč raznolika. Večina slovničnih pravil je v etimološkem smislu in iz pedagoško-didaktičnega vidika sicer logična, žal pa ni uporabna pri računalniški logiki v obliki: »če – potem« (*»if–then«*). Primer: pravila živosti za tožilnik ednine samostalnikov moškega spola ne moremo sprogramirati drugače, kot da označimo vse samostalnike moškega spola, da so bodisi »živi« bodisi »neživi«. S samo oznako živosti/neživosti pa še ne povemo, kako se ti samostalniki dejansko sklanjajo. Zato moramo istočasno povedati še, katere končnice pritičejo posameznemu samostalniku pri vseh sklonih. Tak način uporabe logičnih slovničnih pravil je torej neekonomičen in v bistvu ne spremeni dejstva, da je za vsak samostalnik potrebno eksplicitno navesti, kako se le-ta sklanja.

Hipoteza v tem prispevku pravi, da je primerneje pripisati vsaki besedi v slovarju (geslu) vse možne končnice, kot pa iskati, postaviti in prevesti v računalniško logiko vsa slovnična pravila in vse izjeme, ki se pri besedni analizi pojavljajo. V sestavku so navedeni vsi nizi računalniških končnic, ki so bili dobljeni po pregledu in klasifikaciji približno 160.000 slovarskih gesel. Za vse vrste pregibnih besed je naštetih 240 različnih nizov računalniških končnic. Vsaka beseda v slovarju, če je seveda pregibna, ima poleg računalniške osnove navedeno še številko niza računalniških končnic. Niz pa vsebuje vse možne končnice tega gesla. Vsi ugotovljeni nizi računalniških končnic so navedeni v tabelah II do IX. Zbiranje končnic, pregledovanje besedišča in klasifikacija besed je trajalo osem let (Zupan 1991).

Osnovna izhodišča in definicije

Velika večina slovnice je namenjena učenju maternega jezika in zato vedno predpostavlja neko minimalno jezikovno predznanje. Zaradi te predpostavke v slovnice niso navedena vsa slovnična pravila in vse izjeme, ki so potrebna za pravilno sklanjanje in spreganje. Domneva, da bo učenec po nekaj opisanih primerih znal vse

nadaljnje nadgraditi sam, se izkaže za tujca kot precej zahteven pogoj. Kot primer pogledajmo sklanjatve samostalnikov. V slovenskih slovnica (Žagar 1991) navajajo, da so pri samostalnikih moškega in ženskega spola štirje, pri samostalnikih srednjega spola pa trije sklanjatveni vzorci. Dodana so še pravila, ki pojasnjujejo, kako se pri različnih spolih in sklonih spreminjajo naglasi; povedanih je nekaj izjem, kot npr. podaljševanje osnove *z j* (*slikar*, *slikar-j-a*) in potem izjem od izjem (*meter*, *met(e)r-(j)-a*). Niso pa navedena še druga pravila in izjeme, ki so večini učencev s slovenskim maternim jezikom tako samoumevna, da jih je moč izpustiti ali pa so tako izjemna in redka, da praktično ne pridejo v poštev. Npr.: imenovalnik v množini, ki se konča na *-je*, je navadno pojasnjen le s tremi ali štirimi primeri (*brat-je*, *fant-je*, *gost-je*, *golob-je*), kar Slovincu zadostuje. Nikoli pa niso naštet **vsi** taki primeri (*škof-je*, *mož-je*, *ud-je*, *voz-je* in še mnogo drugih tudi tipološko posebnih *mož-je*, *las-je*, *voz-je*, *volc-je zob-je*). Tudi tu niso naštet vsi, ki pa so zbrani v računalniškem slovarju. Za računalniško obdelavo je namreč treba eksplicitno navesti prav vse. Primer za redko izjemo, ki skoraj nikoli ni omenjena v učbenikih, je sklanjanje samostalnika srednjega spola *igo* (*igo*, *ižesa*, *ižesu* ...). Upamo, da rojenemu Slovincu ali Slovenki posluh omogoči spoznanje, da se beseda sklanja podobno kot *uho* ali *oko*. Splošna pravila torej ne dajejo zanesljivih navodil, ki bi veljala za **vse** primere, in zato z njimi ne moremo narediti uporabnih računalniških programov, ker le-ti nimajo ne posluha, ne jezikovnega znanja. Ponovno pridemo do sklepa, da je najbolje vsaki besedi pripisati vse njene pregibne možnosti eksplicitno.

Hudo težavo povzroča dejstvo, da pri pisanju slovenščine ne uporabljamo naglasov. S tem so le-ti kot informacija za računalniško določitev pravih pregibnih oblik izgubljeni. Lep primer za dvoumnost in izgubo informacije pri izpuščanju naglasov je računalniška težava razlikovanja samostalnika od pridevnika v tako enostavnem primeru kot je naslov Levstikove povesti *Martin Krpan*. Zaradi opustitve naglasa računalnik ne zmore ločiti moškega imena *Martin* od pridevnika *Martin* ((od) Marte). Po drugi strani pa prav zaradi raznolikosti končnic slovenščina (in z njo večina slovanskih jezikov) nima toliko težav pri razlikovanju posameznih sklonov kot npr. angleščina, ki pozna le saški genitiv in nekaj množinskih oblik (Zupan 1999).

Seveda se tudi pri slovenščini ne smemo zanašati le na končnice. Poznamo celo vrsto dvoumnosti: nekateri skloni samostalnikov so oblikovno identični sklonom drugih besednih vrst. Primer za tako ekvivalenco so vsi samostalniki moškega spola, pri katerih roditeljske množine ni moč razlikovati od pridevniške oblike iste osnove: *gospod-ov*, *stol-ov*, *stroj-ev*. Takih primerov je še več, vendar niso toliko splošni, da bi jih lahko uokvirili v splošna pravila. Primer za netipično ekvivalenco (homografijo) besednih oblik (Popovič, Willet 1990) je samostalnik *puščava*, ki se v povsem enaki obliki pojavi tudi v drugi osebi dvojine sedanjika glagola *puščati*.

Vsi navedeni primeri potrjujejo, da je človekov razum bistveno boljje usposobljen za razpoznavanje in analizo besed, kot so ali morda bodo v bodočnosti usposobljeni računalniki. Ker so slovnice napisane za ljudi in ne za računalnike, je

za učinkovito besedno analizo na podlagi lingvističnih pravil iz učbenikov zelo težko sestaviti računalniške algoritme, tj. do vseh podrobnosti pojasnjene postopke.

Osnovna zamisel predstavljene rešitve problema avtomatske besedne analize temelji na ideji tako imenovane *računalniške osnove*. Računalniška osnova besede je definirana kot tisti najdaljši začetni del besede, ki se ohrani pri vseh pregibnih oblikah. Tako so računalniške osnove besed *miza*, *meter*, *lep*, *majhen*, *igrati* in *peti*: *miz-* (*miz-a*, *miz-e*,...), *met-* (*met-er*, *met-ra*,...), *lep-* (*lep-0*, *lep-a*,...), *majh-* (*majh-en*, *majh-na*,...), *igra-* (*igra-ti*, *igra-m*,...) in *p-* (*p-eti*, *p-ojem*,...). Računalniška osnova je lahko enaka lingvistični, kot v primerih *lep* in *igra-ti*, vendar marsikdaj ni, kar vidimo iz primera računalniške osnove *p-* pri glagolu *p-eti*: *p-ojem*, *p-oješ*, ... *bom p-el*, *sva p-ela* itd.

Največkrat so računalniške osnove krajše od oblikoslovnih osnov. To je lahko posledica izpuščanja polglasnikov, kot v primeru *majh-en*: *majh-na*, *majh-ne* itd., zaradi premen pri glagolu *i-skati*, *i-ščem*, *i-ščeš* itd. ali zaradi drugih posebnosti slovenščine.

Veliko računalniških osnov je povezanih z različnimi nizi računalniških končnic. Različni nizi računalniških končnic pripeljejo eno računalniško osnovo do različnih slovarskih gesel. Primer za to sta računalniški osnovi *d-* in *pet-*. Osnova *d-* ima tri nize računalniških končnic za geslo: *d-an* (*d-ne*, *d-neva*,...), za geslo *d-reti* (*d-erem*, *d-ereš*,...) in geslo *d-uti* (*d-mem*, *d-meš*,... (SP 1962)). Osnova *pet-* pa ima celo šest različnih nizov računalniških končnic, in sicer za gesla: *pet-a*, *pet-ek*, *Peter*, *pet-lja*, *pet-i* (števniki) in *pet-en/ni* (pridevnik).

Ker so v slovarju besedna gesla shranjena kot računalniške osnove in pripadajoči niz računalniških končnic, lahko vsaki besedi avtomatično poiščemo vse pregibne oblike. Velja tudi obratno: vsaki pregibni obliki lahko poiščemo pripadajoče slovarsko geslo (Popovič, Willet 1990; Zupan 1991). Edina stvar, ki jo moramo ob osnovi še poznati, so vsi možni nizi računalniških končnic. Zato je bil velik del raziskave osredotočen prav na določitev vseh možnih nizov slovenskih računalniških končnic v slovenskem jeziku.

Pregled nizov računalniških končnic

Nizi računalniških končnic so razdeljeni na samostalniške, pridevniške in glagolske. Samostalniški nizi se dele na računalniške končnice sklanjatev samostalnikov moškega, ženskega in srednjega spola ter na računalniške končnice množinskih samostalnikov. Ti zadnji so prav tako razdeljeni na nize za samostalnike moškega, ženskega in srednjega spola. Nizi pridevniških končnic so razdeljeni na take, ki vsebujejo pridevniške, števnike in zaimenske računalniške končnice. Samostalniški zaimki so zaradi specifičnosti obravnavani posebej: ne z nizi računalniških končnic, ampak kot samostojne nepregibne oblike. Najštevilčnejša skupina nizov so računalniške končnice glagolskih spregatev (89 različnih nizov). Pregled vseh nizov računalniških končnic kaže tabela I. Iz nje je razvidno, da je različnih nizov računalniških končnic 240.

Število računalniških končnic v nizih je pri posameznih besednih vrstah različno. Samostalniki imajo osemnajst, pridevniki dvanajst in glagoli enainvajset računalniških končnic. Končnice so v nizih navedene vedno v istem vrstnem redu. Poleg pravih računalniških končnic nastopata v nizih še dve oznaki: *-0* in *-#*: prva oznaka (*-0*) pomeni, da na tem mestu računalniška osnova nima končnice in lahko stoji samostojno, npr.: prvi sklon in četrti sklon prve moške sklanjatve ednine (tabela II, nizi 2, 4, 6 in 8); druga oznaka (*-#*) pa pomeni, da sklanjatev ali spregatev zadevnih končnic sploh ne pozna (npr.: množinske oblike besede *človek*, tabela II, niz 10).

Pri samostalnikih je vedno navedenih osemnajst računalniških končnic (za vse sklone v vseh številih) čeprav niso vse med seboj različne. Druga ženska sklanjatev ima sedem (npr.: *mis-el*, *-li*, *-lijo*, *-lima*, *-lih*, *-lim*, *-limi*), druga moška pa osem različnih računalniških končnic (*pob-a*, *-u*, *-om*, *-ov*, *-oma*, *-ih*, *-i*, *-e*). Tretje sklanjatve vseh spolov pa sploh nimajo računalniških končnic (primeri: *oča*, *mami* in *doma*). V tabelah II do IV samostalniških sklanjatev so računalniške končnice za ednino, dvojino in množino ločene s črto.

Tabela I: Pregled števila različnih nizov računalniških končnic za vse sklanjatve in spregatve

| Računalniške končnice | | Dolžina niza | Število nizov |
|------------------------------------|--|--------------|---------------|
| Navadne samostalniške sklanjatve | moški | 18 | 48 |
| | ženski | 18 | 29 |
| | srednji | 18 | 23 |
| Množinske samostalniške sklanjatve | moški | 6 | 5 |
| | ženski | 6 | 10 |
| | srednji | 6 | 7 |
| Pridevniške sklanjatve | pridevniki | 12 | 11 |
| | števniki | 12 | 9 |
| | pridevniški zaimki | 12 | 9 |
| | samostalniški zaimki imajo individualne oznake | – | – |
| Glagolske spregatve | | 21 | 89 |
| Skupaj | | | 240 |

Zaradi ločevanja živosti v tožilniku ednine se število nizov različnih računalniških končnic pri moških samostalnikih skoraj podvoji. Samostalnikom za živa bitja so prirejeni drugačni nizi (z izpremenjeno končnico v tožilniku ednine) kot tistim za nežive stvari (primerjaj lihe in sode nize v tabeli II – razlikujejo se le po četrti končnici!). Edninske računalniške končnice za sklanjanje besede *žerjav-0* (ptica) so: *žerjav-0*, *-a*, *-u*, *-a*, *pri -u*, *z -om*, za besedo *žerjav-0* (dvigalna naprava) pa: *žerjav-0*, *-a*, *-u*, *-0*, *pri -u*, *z -om*. Podobno je tudi pri vseh ostalih moških samostalnikih, npr.: sklanjatvi samostalnikov na *-em*: *komuniz-em* in *Eraz-em*.

So še druge izjeme. Pri samostalnikih se za množice več od štiri v prvem sklonu uporablja drugačna končnica kot za tri ali štiri: *štirje slon-i*, *pet slon-ov* ali *štiri kače-e*, *pet kač-0* itd. Ta končnica je identična končnicama roditelja dvojine in

množine, ki sta v vseh samostalniških nizih na osmem oziroma štirinajstem mestu. Zato posebna oznaka ni potrebna.

Posebnosti so tudi dodatne računalniške končnice pri nekaterih množinskih oblikah: npr.: *sin-i* in *sin-ovi*, *grob-i* in *grob-ovi* itd. Posebnosti, netipičnosti in alternativne končnice moških sklanjatev so navedene v tabeli II v nizih od 37 do 43. Pri sklanjatih vseh spolov naletimo tudi na take nize računalniških kočnic, po katerih se sklanja le ena sama beseda, kar po definiciji predstavlja izjemo, npr.: *blag-or* (*blag-ra*, ...), *pes-em* (*pes-mi*, ...), *i-go* (*i-žesa*, ...), *u-ho* (*u-šesa*, ...) in *o-ko* (*o-česa*, ...) itd.

Navedene izjeme seveda še zdaleč niso vse. Vsi nizi različnih računalniških končnic, ki sem jih zasledil pri pregledu več kot 60.000 samostalnikov, so zbrani v tabelah II (moške), III (ženske) in IV (srednje).

Tabela II: Nizi samostalniških računalniških končnic moških sklanjatev

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------|-------|-------|-----|--------|-------|--------|--------|-------|--------|
| -0 | -0 | -0 | -0 | -0 | -0 | -0 | -0 | -0 | -0 |
| a | a | a | a | ja | ja | na | na | ta | a |
| u | u | u | u | ju | ju | nu | nu | tu | u |
| a | -0 | a | -0 | ja | -0 | na | -0 | ta | a |
| u | u | u | u | ju | ju | nu | nu | tu | u |
| om | om | em | em | jem | jem | nom | nom | tom | om |
| a | a | a | a | ja | ja | na | na | ta | a |
| ov | ov | ev | ev | jev | jev | nov | nov | tov | ov |
| oma | oma | ema | ema | jema | jema | noma | noma | toma | oma |
| a | a | a | a | ja | ja | na | na | ta | a |
| ih | ih | ih | ih | jih | jih | nih | nih | tih | ih |
| oma | oma | ema | ema | jema | jema | noma | noma | toma | oma |
| i | i | i | i | ji | ji | ni | ni | ti | -# |
| ov | ov | ev | ev | jev | jev | nov | nov | tov | -# |
| om | m | em | em | jem | jem | nom | nom | tom | -# |
| e | e | e | e | je | je | ne | ne | te | -# |
| ih | h | ih | ih | jih | jih | nih | nih | tih | -# |
| i | I | i | i | ji | ji | ni | ni | ti | -# |
| mulat | polip | garač | vrč | vratar | cekar | rabelj | nagelj | Pavle | človek |

| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|-------|-------|-------|--------|-------|--------|------|-------|--------|--------|
| ec | ec | eg | ej | ek | ek | el | el | elj | elj |
| ca | ca | ga | ja | ka | ka | la | la | lja | lja |
| cu | cu | gu | ju | ku | ku | lu | lu | lju | lju |
| ec | ca | eg | ja | ka | ek | la | el | lja | elj |
| cu | cu | gu | ju | ku | ku | lu | lu | lju | lju |
| cem | cem | gom | jem | kom | kom | lom | lom | ljem | ljem |
| ca | ca | ga | ja | ka | ka | la | la | lja | lja |
| cev | cev | gov | jev | kov | kov | lov | lov | ljev | ljev |
| cem | cema | goma | jema | koma | koma | loma | loma | ljema | ljema |
| ca | ca | ga | ja | ka | ka | la | la | lja | lja |
| cih | cih | gih | jih | kih | kih | lih | lih | ljih | ljih |
| cema | cema | goma | jema | koma | koma | loma | loma | ljema | ljema |
| ci | ci | gi | ji | ki | ki | li | li | lji | lji |
| cev | cev | gov | ij | kov | kov | lov | lov | ljev | ljev |
| cem | cem | gom | jem | kom | kom | lom | lom | ljem | ljem |
| ce | ce | ge | e | ke | ke | le | le | lje | lje |
| cih | cih | gih | jih | kih | kih | lih | lih | ljih | ljih |
| ci | ci | gi | ji | ki | ki | li | li | lji | lji |
| kršec | tepec | bezeg | zlodej | maček | sodček | osel | vatel | rabelj | nagelj |

| | | | | | | | | | |
|--------|-------|------|-------|--------|--------|-------|--------|------|------|
| 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| em | em | en | en | enj | enj | er | er | es | es |
| ma | ma | na | na | nja | nja | ra | ra | sa | sa |
| mu | mu | nu | nu | nju | nju | ru | ru | su | su |
| ma | em | na | en | nja | enj | ra | er | sa | es |
| mu | mu | nu | nu | nju | nju | ru | ru | su | su |
| mom | mom | nom | nom | njem | njem | rom | rom | som | som |
| ma | ma | na | na | nja | nja | ra | ra | sa | sa |
| mov | mov | nov | nov | njev | njev | rov | rov | sov | sov |
| moma | moma | noma | noma | njema | njema | roma | roma | soma | soma |
| ma | ma | na | na | nja | nja | ra | ra | sa | sa |
| mih | mih | nih | nih | njih | njih | rih | rih | sih | sih |
| moma | moma | noma | noma | njema | njema | roma | roma | soma | soma |
| mi | mi | ni | ni | nji | nji | ri | ri | si | si |
| mov | mov | nov | nov | njev | njev | rov | rov | sov | sov |
| mom | mom | nom | nom | njem | njem | rom | rom | som | som |
| me | me | ne | ne | nje | nje | re | re | se | se |
| mih | mih | nih | nih | njih | njih | rih | rih | sih | sih |
| mi | mi | ni | ni | nji | nji | ri | ri | si | si |
| Erazem | kosem | oven | boben | suženj | reženj | pater | popper | pes | oves |

| | | | | | | | | | |
|--------|-------|--------|------|-------|------|-------|-----|--------|------|
| 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| et | et | or | e | o | o | k | -0 | an | (an) |
| ta | ta | ra | a | a | a | ka | a | neva | -# |
| tu | tu | ru | u | u | u | ku | u | nevu | -# |
| ta | et | or | e | a | o | ka | a | an | -# |
| tu | tu | ru | u | u | u | ku | u | nevu | -# |
| tom | tom | rom | om | om | om | kom | em | nevom | -# |
| ta | ta | ra | a | a | a | ka | a | neva | ni |
| tov | tov | rov | -0 | ov | ov | k | -0 | nevov | ni |
| toma | toma | roma | oma | oma | oma | koma | ema | nevoma | nema |
| ta | ta | ra | a | a | a | ka | a | neva | ne |
| tih | tih | rih | ih | ih | ih | cih | eh | nevih | neh |
| toma | toma | roma | oma | oma | oma | koma | ema | nevoma | nema |
| ti | ti | ri | i | i | i | ci | je | nevi | ni |
| tov | tov | rov | -0 | ov | ov | k | -0 | nevov | ni |
| tom | tom | rom | om | om | om | kom | em | nevom | nem |
| te | te | re | a | e | e | ke | e | neve | ni |
| tih | tih | rih | ih | ih | ih | cih | eh | nevih | neh |
| ti | ti | ri | i | i | i | ki | mi | nevi | ni |
| valpet | hrbet | blagor | Čile | Marko | kino | otrok | mož | dan | dan |

| | | | | | | | |
|-----|-------|-----|-------|------|-------|-----|-------|
| 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 |
| -0 | -0 | -0 | a | a | ja | -0 | i |
| u | -# | -# | e | a | je | -# | ega |
| -# | -# | -# | i | u | ji | -# | emu |
| -# | -# | -# | o | a | jo | -# | ega |
| -# | -# | -# | i | u | ji | -# | em |
| -# | -# | -# | o | om | jo | -# | im |
| -# | ova | -# | i | a | ja | -# | a |
| -# | ovov | -# | -0 | ov | ij | -# | ih |
| -# | ovoma | -# | ama | oma | jema | -# | ima |
| -# | ova | -# | i | a | ji | -# | a |
| -# | ovih | eh | ah | ih | jih | -# | ih |
| -# | ovoma | -# | ama | oma | jema | -# | ima |
| -# | ovi | -# | e | i | ji | -# | i |
| -# | ovov | -# | -0 | ov | ij | -# | ih |
| -# | ovom | -# | am | om | jem | -# | im |
| -# | ove | eh | e | e | je | -# | e |
| -# | ovih | -# | ah | ih | jih | -# | ih |
| -# | ovi | -# | ami | i | ji | -# | imi |
| med | vrt | dol | sluga | poba | vodja | oča | stari |

Tabela III: Nizi samostalniških računalniških končnic ženskih sklanjatev

| | | | | | | | | | |
|--------|------|------|--------|-------|--------|--------|--------|---------|--------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| a | a | a | a | (ca) | ja | (ka) | la | lja | ma |
| e | e | e | e | -# | je | -# | le | lje | me |
| i | i | i | ej | -# | ji | -# | li | lji | mi |
| o | o | o | o | -# | jo | -# | lo | ljo | mo |
| i | i | i | e | -# | ji | -# | li | lji | mi |
| o | o | o | o | -# | jo | -# | lo | ljo | mo |
| i | i | i | e | -# | ji | -# | li | lji | mi |
| -0 | j | i | e | ac | ij | ak | el | elj | em |
| ama | ama | ema | ema | -# | jama | -# | lama | ljama | mama |
| i | i | i | e | -# | ji | -# | li | lji | mami |
| ah | ah | eh | eh | -# | jah | -# | lah | ljah | mah |
| ama | ama | ema | ema | -# | jama | -# | lama | ljama | mama |
| e | e | i | e | -# | je | -# | le | lje | me |
| -0 | j | i | a | ac | ij | ek | el | elj | em |
| am | am | em | em | -# | jam | -# | lam | ljam | mam |
| e | e | i | e | -# | je | -# | le | lje | me |
| ah | ah | eh | eh | -# | jah | -# | lah | ljah | mah |
| ami | ami | emi | emi | -# | jami | -# | lami | ljami | mami |
| pišk-a | bo-a | sl-a | gosp-a | ov-ca | zar-ja | peč-ka | nič-la | kok-lja | tek-ma |

| | | | | | | | | | |
|--------|---------|--------|--------|---------|------|--------|------|------|-------|
| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| na | nja | ra | va | ev | o | -0 | -0 | i | i |
| ne | nje | re | ve | ve | e | i | i | vi | ere |
| ni | nji | ri | vi | vi | i | i | i | vi | eri |
| no | njo | ro | vo | ev | o | -0 | -0 | i | er |
| ni | nji | ri | vi | vi | i | i | i | vi | eri |
| no | njo | ro | vo | vijo | o | jo | jo | vjo | ero |
| ni | nji | ri | vi | vi | i | i | i | vi | eri |
| en | enj | er | ev | ev | -0 | i | i | vi | era |
| nama | njama | rama | vama | vima | ama | ima | ema | vema | erama |
| ni | nji | ri | vi | vi | i | i | i | vi | eri |
| nah | njah | rah | vah | vah | ah | ih | eh | veh | erah |
| nama | njama | rama | vama | vama | ama | ima | ema | vema | erama |
| ne | nje | re | ve | ve | e | i | i | vi | ere |
| en | enj | er | ev | ev | -0 | i | i | vi | era |
| nam | njam | ram | vam | vam | am | im | em | vem | eram |
| ne | nje | re | ve | ve | e | i | i | vi | ere |
| nah | njah | rah | vah | vah | ah | ih | eh | veh | erah |
| nami | njami | rami | vami | vami | ami | mi | mi | vmi | erami |
| him-na | luk-nja | šav-ra | smokev | belitev | Argo | plavut | klet | kr-i | hč-i |

| | | | | | | | | |
|--------|--------|--------|--------|-----------|--------|--------|------|--------|
| 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| el | em | an | en | enj | er | et | -0 | a |
| li | mi | ni | ni | nji | ri | ti | -0 | e |
| li | mi | ni | ni | nji | ri | ti | -0 | i |
| el | em | an | en | enj | er | et | -0 | o |
| li | mi | ni | ni | nji | ri | ti | -0 | i |
| lijo | mijo | nijo | nijo | njjo | rijo | tjo | -0 | o |
| li | mi | ni | ni | nji | ri | ti | -0 | i |
| lima | mim | ni | ni | njima | er | et | -0 | ih |
| li | mima | nema | nima | njima | rima | tema | -0 | ima |
| li | mi | ni | ni | nji | ri | tima | -0 | i |
| lih | mih | neh | nih | njima | rih | teh | -0 | ih |
| lima | mima | nema | nima | njima | rima | tima | -0 | ima |
| li | mi | ni | ni | nji | ri | ti | -0 | e |
| li | em | ni | ni | enj | er | ti | -0 | ih |
| lim | mim | nem | nim | njim | rim | tem | -0 | im |
| li | mi | ni | ni | nji | ri | ti | -0 | e |
| lih | mih | neh | nih | njih | rih | tmi | -0 | ih |
| limi | mimi | nmi | nimi | njimi | rimi | tmi | -0 | imi |
| mis-el | pes-em | rav-an | kaz-en | povod-enj | reb-er | lak-et | Ines | star-a |

Tabela IV: Nizi samostalniških računalniških končnic srednjih sklanjatev

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---------------------------------|---------------------------------|--|--|--|--|--|---------------------------------------|---------------------------------------|---------------------------------------|
| o a u o u om | o a u o u om | o a u o u om | O esa esu o esu esom | go žesa žesu go žesu žesom | ho šesa šesu ho šesu šesom | ko česa česu ko česu česom | lo la lu lo lu lom | mo ma mu mo mu mom | no na nu no nu nom |
| i - oma i ih oma | i i ima i ih ima | esi es esoma esi esih esoma | esi es esoma esi esih esoma | žesi žes žesoma žesi žesih žesoma | šesi šes šesoma šesi šesih šesoma | česi čes česoma česi česih česoma | li el loma li lih loma | mi em mima mi mih mima | ni en noma ni nih noma |
| a - om a ih i | a i im a ih i | esa es esom esa esih esi | esa es esom esa esih esi | žesa žes žesom žesa žesih žesi | šesa šes šesom šesa šesih šesi | či čes čem či čeh čmi | la el lom la lih li | ma em mom ma mih mi | na en nom na nih ni |
| blag-o | zl-o | neb-o | tel-o | i-go | u-ho | o-ko | ges-lo | pis-mo | ok-no |

| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---------------------------------------|---------------------------------------|----------------------------------|---------------------------------------|---|---|--|--------------------------------------|--------------------------------------|--------------------------------------|
| ro ra ru ro ru rom | vo va vu vo vu vom | e a u e u em | je ja ju je ju jem | lje lja lju lje lju ljem | nje nja nju nje nju njem | rje rja rju rje rju rjem | -0 na nu -0 nu nom | -0 sa su -0 su som | -0 ta tu -0 tu tom |
| ri er roma ri rih roma | vi ev voma vi vih voma | i -0 ema i ih ema | ji ij jema ji jih jema | lji elj ljema lji ljih ljema | nji enj njema nji njih njema | rji erij rjema rji rjih rjema | ni n noma ni nih noma | si s soma si sih soma | ti t toma ti tih toma |
| ra er rom ra rih ri | va ev vom va vom vi | a -0 em a ih i | ja ij jem ja jih ji | lja elj ljem lja ljih lji | nja enj njem nja njih nji | rja erij rjem rja rjih rji | na n nom na nih ni | sa s som sa sih si | ta t tom ta tih ti |
| jed-ro | ubošt-vo | kesan-je | žret-je | dup-lje | tr-nje | ned-rje | sle-me | oje-0 | revše-0 |

| 21 | 22 | 23 |
|--|----------------------------------|--|
| -0 va vu -0 vu vom | o ega emu o em im | je jega jemu je jem jim |
| vi vov voma vi vih voma | a ih ima i ima im | ji ij jema ji jih jema |
| vi ev vom ve vih vi | a ih im a ih imi | ja ij jem ja jih jimi |
| poldne | Laško | Kozje |

Naslednji nizi računalniških končnic, ki vsebujejo le po šest končnic, pripadajo množinskim samostalnikom, ki jih je v slovenščini ni niti 500. Največ je ženskih (307). Ti imajo tudi največ različnih nizov računalniških končnic (tabela VI). Najmanj množinskih samostalnikov je srednjega spola – le 30, a imajo kljub temu več različnih nizov računalniških končnic kot moški (primerjaj tabeli V in VII). Pri moških množinskih samostalnikih je dodan še niz z eno samo računalniško končnico *-je* (tabela V, niz 5). Ta opiše posebno obliko imenovalnika množine samostalnikov kot so *očet-je*, *mož-je*, *škof-je* itd.

Tabela V: Nizi računalniških končnic moških množinskih samostalniških sklanjatev

| 1 | 2 | 3 | 4 | 5 |
|-------|----------|---------|---------|---------|
| i | i | i | je | (je) |
| ov | ev | ih | i | -# |
| om | em | im | em | -# |
| e | e | e | je | -# |
| ih | ih | ih | eh | -# |
| i | i | imi | mi | -# |
| Žir-i | Radenc-i | domač-i | ljud-je | očet-je |

Tabela VI: Nizi računalniških končnic ženskih množinskih samostalniških sklanjatev

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------|-------|-------|--------|--------|--------|---------|-------|--------|--------|
| i | i | i | e | je | le | lje | ne | re | ve |
| i | i | i | -0 | ij | el | elj | en | er | ev |
| im | im | em | am | jam | lam | ljam | nam | ram | vam |
| i | i | i | e | je | le | lje | ne | re | ve |
| ih | ih | eh | ah | jah | lah | ljah | nah | rah | vah |
| imi | mi | mi | ami | jami | lami | ljami | nami | rami | vami |
| gosl-i | prs-i | san-i | man-ce | gar-je | org-le | dim-lje | os-ne | cit-re | buk-ve |

Tabela VII: Nizi računalniških končnic srednjih množinskih samostalniških sklanjatev

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---------|----------|--------|-------|------|--------|------|
| a | a | a | a | la | ra | (a) |
| -0 | -0 | -0 | -0 | al | er | -# |
| am | em | om | om | lom | rom | -# |
| a | a | a | a | la | ra | -# |
| ah | ih | ih | eh | leh | rih | -# |
| i | i | i | mi | li | ri | emi |
| pljuč-a | železc-a | črev-a | drv-a | t-la | ned-ra | tl-a |

Samostalniškim nizom končnic slede pridevniški. Ker se morajo pridevniške sklanjatve poleg s sklonom in številom ujemati s samostalnikom tudi v spolu, bi lahko bilo v vsakem pridevniškem nizu trikrat toliko računalniških končnic kot pri samostalnikih, torej štiriinpetdeset. K sreči je v veliki večini nizov med vsemi teoretično možnimi računalniškimi končnicami le po dvanajst različnih. Zato tudi različnih nizov ni veliko: osem standardnih nizov računalniških končnic in trije nizi z eno samo, dodatno računalniško končnico (tabela VIII). Edina večja razlika med nizi končnic je pri pridevnikih na *-ski* in *ški*. Pri njih se določna in nedoločna oblika ne razlikujeta. Zato se pri teh nizih prva končnica *-i* ponovi na četrtem mestu (drugi nizi v tabelah VIII–X). Posebne končnice *-ak*, *-al* in *-an* nastopajo le pri nekaj pridevnikih, npr.: *meh-ek* (*meh-ak*), *svet-el* (*svet-al*) in *tem-en* (*tem-an*). Tem pridevnikom so poleg navadnih končnic pripisani še dodatni nizi, tako da ima npr.

pridevnik *moč-en* v računalniškem slovarju poleg niza končnic na *-en* (tabela VIII, niz 5) naveden še niz z eno samo končnico *-an* (tabela VIII, niz 11). Podobno je v vseh ostalih primerih. Če se beseda sklanja (ali sprega) na dva različna načina, ima dva niza računalniških končnic.

Pri pridevniških sklanjatvah števnikov in zaimkov je število različnih nizov še manjše. V bistvu gre pri obeh tabelah le za pet oziroma tri standardne (tabela IX, nizi 1–5 in tabela X, nizi 1–3) in za dodatne nize računalniških končnic, ki se pojavljajo samo pri izjemah (tabela IX, nizi 6–9 in tabela X, nizi 4–9). Standardni nizi števnških in zaimenskih standardnih končnicah so identični s standardnimi nizi pridevniških. Identične nize imamo v različnih tabelah samo zaradi tega, da računalniški slovar razlikuje med števnik, zaimki in pridevniki, ki se sicer sklanjajo z enakimi računalniškimi končnicami. S tem omogočimo, da računalnik pri analizi avtomatično sporoči, ali gre pri besedi za pridevnik, števnik ali zaimek. Deležniki v slovarju nimajo posebnih nizov računalniških končnic, ampak iste kot pridevniki. V izboljšavah so predvidene posebne tabele in oznake tudi za deležnike, ki se sklanjajo kot pridevniki.

Tabela VIII: Nizi pridevniških računalniških končnic za pridevnike

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|------|-----------|---------|---------|--------|---------|---------|--------|---------|---------|--------|
| -0 | i | ek | el | en | ev | eljk | er | (-ak) | (-al) | (-an) |
| a | a | ka | la | na | va | lja | ra | -# | -# | -# |
| e | e | ke | le | ne | ve | lje | re | -# | -# | -# |
| i | i | ki | li | ni | vi | ljki | ri | -# | -# | -# |
| o | o | ko | lo | no | vo | ljo | ro | -# | -# | -# |
| em | em | kem | lem | nem | vem | ljem | rem | -# | -# | -# |
| im | im | kim | lim | nim | vim | ljim | rim | -# | -# | -# |
| ih | ih | kih | lih | nih | vih | ljih | rih | -# | -# | -# |
| ega | ega | kega | lega | nega | vega | ljeга | rega | -# | -# | -# |
| emu | emu | kemu | lemu | nemu | vemu | ljemu | remu | -# | -# | -# |
| ima | ima | kima | lima | nima | vima | ljima | rima | -# | -# | -# |
| imi | imi | kimi | limi | nimi | vimi | ljimi | rimi | -# | -# | -# |
| blag | piransk-i | slad-ek | svet-el | moč-en | plit-ev | dup-elj | pik-er | slad-ak | svet-al | moč-an |

Tabela IX: Nizi pridevniških računalniških končnic za števnik

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-----|-------|---------|-------|------|-------|-------|--------|---------|
| -0 | i | en | em | -0 | den | a | lje | je |
| a | a | na | ma | ta | -# | e | i | i |
| e | e | ne | me | te | -# | eh | eh | ih |
| i | i | ni | mi | ti | -# | em | em | im |
| o | o | no | mo | to | -# | ema | ema | ima |
| em | em | nem | mem | tem | -# | emi | emi | imi |
| im | im | nim | mim | tīm | -# | emima | emima | -# |
| ih | ih | nih | mih | tih | -# | emimi | emimi | -# |
| ega | ega | nega | mega | tega | -# | -# | -# | -# |
| emu | emu | nemu | memu | temu | -# | -# | -# | -# |
| ima | ima | nima | mima | tīma | -# | -# | -# | -# |
| imi | imi | nimi | mimi | tīmi | -# | -# | -# | -# |
| en | prv-i | enoj-en | os-em | sto | e-den | dv-a | tr-ije | štūr-je |

Tabela X: Nizi pridevniških računalniških končnic za pridevniške zaimke

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-----|--------|---------|--------|-----|--------|-------|----------|------|
| -0 | i | en | do | a | le | ole | enle | es |
| a | a | na | a | a | ale | ale | nale | sa |
| e | e | ne | e | e | ele | ele | nele | se |
| i | i | ni | i | i | ile | ile | nile | si |
| o | o | no | o | o | ole | ole | nole | so |
| em | em | nem | om | em | emle | emle | nemle | sem |
| im | im | nim | im | ej | imle | ejle | nimle | sej |
| ih | ih | nih | ih | eh | ihle | ehle | nihle | seh |
| ega | ega | nega | oga | ega | egale | egale | negale | sega |
| emu | emu | nemu | omu | emu | emule | emule | nemule | semu |
| ima | ima | nima | ima | ema | imale | emale | nimale | sema |
| imi | imi | nimi | imi | emi | imile | emile | nimile | semi |
| moj | tist-i | takb-en | nek-do | t-a | tak-le | t-ole | takš-ele | v-es |

Pri zaimkih je treba omeniti še računalniški zapis sklanjatev samostalniških zaimkov. Ker večina samostalniških zaimkov sploh nima računalniške osnove (pri sklanjanju se spreminja celotna beseda, npr.: *jaz, mene; on, njega; ona, nje; mi, nas* itd.), so v računalniškem slovarju oblike vseh samostalniških zaimkov navedene eksplicitno. Vsaka oblika je opisana s štirimestno kodo: s številom med 5000 in 5999. Prva cifra kode: »5« signalizira, da je beseda samostalniški zaimek, druge tri pa povedo sklon s katerim se le-ta veže. To pomeni, da so v računalniškem slovarju kot samostojne besede posamično navedeni vsi samostalniški zaimki v vseh sklonih. Za bodočo avtomatsko besedno in stavčno analizo, čemur je opisani računalniški slovar namenjen, je zanesljivo prepoznavanje vseh oblik samostalniških zaimkov ena pomembnejših nalog.

Daleč največ različnih nizov računalniških končnic je pri glagolskih spregatvah. Med vsemi devetinosemdesetimi nizi (tabela XI) je enainosemdeset popolnih in sedem delnih. Večino popolnih nizov (75) sestavljajo končnice spregatev glagolov na *-ti*, ostanek pa šest nizov računalniških končnic spregatev glagolov na *-či*. Ostalih osem delnih nizov vsebuje računalniške končnice izjem in nepravilnosti. V popolnih nizih glagolskih spregatev je 21 različnih računalniških končnic. Poleg nedoločniške končnice (*-ti* ali *-či*) so vse končnice za tvorjenje sedanjikovih oblik (*sem, boš, bo, bova, bosta, bomo, boste, bodo/bojo*), vse oblike deležnikov prihodnika in preteklika (*-l, -la, -le, -li in -lo*), vse računalniške končnice velelnega naklona (5 računalniških končnic) in ena trpniška končnica (*-t*). Zanimivo je, da se velikanska večina glagolov sprega z natanko 21 različnimi računalniškimi končnicami in da je izjem pri tem sorazmeroma malo ne glede na to, ali je glagol pravilni ali nepravilni.

Nizi računalniških končnic so pri glagolih urejeni po nekoliko prirejenem abecednem redu in dolžini nedoločniške končnice (tabela XI). Najprej so nizi za spreganje glagolov na *-ti*. To so: *-ti, -ati, -eti, -iti, -sti, uti*, nato pa še glagoli na *-či* in *-eči*.

Zadnjih osem nizov so končnice pri spregatvah *računalniško* nepravilnih glagolov in izjem. Računalniško nepravilni glagoli so dvojni. Prvi so tisti, ki se pri spreganju spremenijo v **celoti** in sploh nimajo računalniške osnove, npr.: *biti (sem, si, je, ...)*, *ne imeti (nimam, nimaš, ...)*, *gnati (ženem, ženeš, ...)* in *scati (ščijem,*

ščiješ, ...). Drugi pa imajo alternativne spregatve brez nekaterih oblik, npr.: *iti* (*idem, ideš, ...*; z varianto: *grem, greš, ...*; in še z drugačnimi deležniki: *šel, šla, ...*),.

Posebnost slovenske glagolske spregatve je alternativna možnost računalniške končnice pri tretji osebi sedanjika v množini, npr. glagol *rasti*: (z oblikama *rast-ejo* in *rast-o*). Obe ti dve računalniški končnici (običajna in alternativna) nastopata v tabelah glagolskih nizov vedno na devetem in desetem mestu. Poleg glavne ima alternativno računalniško končnico 30 nizov, ostali pa ne.

Spregatvene oblike pomožnih glagolov *biti* (*sem, si, je, sva, smo, ... bom, boš, ... bodo ...*) in *ne biti* (*nisem, nišva, ni, ... niste, ... niso...*) so v računalniškem slovarju podane kot nepregibne besede, podobno kot samostalniški zaimki, torej vsaka oblika posebej (zadnja vrsta tabele XII).

Tabela XI: Nizi glagolskih računalniških končnic

| | | | | | | | | | |
|--------|--------|--------|--------|-------|--------|-------|---------|----------|---------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ti | ti | ti | ti | ti | ti | ti | ti | ti | ti |
| m | m | m | em | dem | dem | jem | nem | tem | vem |
| š | š | š | eš | deš | deš | ješ | neš | teš | veš |
| -0 | -0 | -0 | e | de | de | je | ne | te | ve |
| va | va | va | eva | deva | deva | jeva | neva | teva | veva |
| ta | ta | sta | eta | deta | deta | jeta | neta | teta | veta |
| mo | mo | mo | emo | demo | demo | jejo | nemo | temo | vemo |
| te | te | ste | ete | dete | dete | jemo | nete | tete | vete |
| jo | jo | jo | ejo | dejo | dejo | jete | nejo | tejo | vejo |
| -# | -# | -# | o | do | do | jo | -# | to | -# |
| l | l | l | el | del | šel | l | l | tel | l |
| la | la | la | la | dla | šla | la | la | tla | la |
| le | le | le | le | dle | šle | le | le | tle | le |
| li | li | li | li | dli | šli | li | li | tli | li |
| lo | lo | lo | lo | dlo | šlo | lo | lo | tlo | lo |
| -0 | j | j | i | di | di | j | ni | ti | vi |
| va | jva | jva | iva | diva | diva | jva | niva | tiva | viva |
| ta | jta | jta | ita | dita | dita | jta | nita | tita | vita |
| mo | jmo | jmo | imo | dimo | dimo | jmo | nimo | timo | vimo |
| te | jte | jte | ite | dite | dite | jte | nite | tite | vite |
| t | t | t | t | t | t | t | t | t | t |
| cepiti | tacati | dati | nesti | najti | najti | biti | vstati | rasti | pleti |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| ati | ati | ati | ati | ati | ati | ati | cati | dati | dati |
| em | em | em | im | jem | ljem | ojim | čem | m | dam |
| eš | eš | eš | iš | ješ | lješ | ojiš | češ | š | dam |
| e | e | e | i | je | lje | oji | če | - | daš |
| eva | eva | eva | iva | jeva | ljeva | ojiva | čeva | va | dava |
| eta | eta | eta | ita | jeta | ljeta | ojita | četa | sta | data |
| emo | emo | emo | imo | jemo | ljemo | ojimo | čemo | mo | damo |
| ete | ete | ete | ite | jete | ljete | ojite | čete | ste | date |
| ejo | ejo | ejo | ijo | jejo | ljejo | ojijo | čejo | jo | dajo |
| o | o | -# | e | -# | -# | oje | -# | do | -# |
| al | al | al | al | al | al | al | cal | dal | dal |
| ala | ala | ala | ala | ala | ala | ala | cala | dala | dala |
| ale | ale | ale | ale | ale | ale | ale | cale | dale | dale |
| ali | ali | ali | ali | ali | ali | ali | cali | dali | dali |
| alo | alo | alo | alo | alo | alo | alo | calo | dalo | dalo |
| -0 | i | aj | i | ji | lji | oj | či | j | j |
| va | iva | ajva | iva | jiva | ljiva | ojva | čiva | jva | jva |
| ta | ita | ajta | ita | jita | ljita | ojta | čita | jta | jta |
| mo | imo | ajmo | imo | jimo | ljimo | ojmo | čimo | jmo | jmo |
| te | ite | ajte | ite | jite | ljite | ojte | čite | jte | jte |
| at | at | at | at | at | at | at | cat | dat | dat |
| sijati | žgati | majati | tičati | orati | kopati | stati | klicati | povedati | gledati |

| | | | | | | | | | |
|--|--|---|--|--|--|---|--|--|---|
| 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| dati jem ješ je jeva jeta jemo jete jejo -# | gati žem žeš že ževa žeta žemo žete žejo -# | gnati ženem ženeš žene ženeva ženeta ženemo ženete ženejo ženo | hati šem šeš še ševa šeta šemo šete šejo -# | jati nem neš ne neva neta nemo nete nejo -0 | kati čem češ če čeva četa čemo čete čejo -# | skati ščem ščeš šče ščeva ščeta ščemo ščete ščejo -# | lati eljem elješ elje eljeva eljeta eljemo eljete eljejo -# | lati oljem olješ olje oljeva oljeta oljemo oljete oljejo -# | rati erem ereš ere ereva ereta eremo erete erejo ero |
| jil jila jile jili jilo | gal gala gale gali galo | gnal gnala gnale gnali gnalo | hal hala hale hali halo | jal jala jale jali jalo | kal kala kale kali kalo | skal skala skale skali skalo | lal lala lale lali lalo | lal lala lale lali lalo | ral rala rale rali ralo |
| ji jiva jita jimo jite dat | ži živa žita žimo žite gat | ženi ženiva ženita ženimo ženite gnat | ši šiva šita šimo šite hat | ni niva nita nimo nite jat | či čiva čita čimo čite kat | šči ščiva ščita ščimo ščite skat | elji eljiva eljita eljimo eljite lat | olji oljiva oljita oljimo oljite lat | eri eriva erita erimo erite rat |
| glodati | trgati | gnati | pihati | dejati | jokati | iskati | stlati | klati | prati |

| | | | | | | | | | |
|--|--|--|--|---|--|--|--|---|---|
| 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| rati erjem erješ erje erjeva erjeta erjemo erjete erjejo -# | rati orjem orješ orje orjeva orjeta orjemo orjete orjejo -# | sati šem šeš še ševa šeta šemo šete šejo -# | scati ščijem ščiješ ščije ščijev ščijet ščijem ščijet ščijet ščijet ščijet -# | slati šljem šlješ šlje šljeva šljeta šljemo šljite šljejo -# | vati jem ješ je jeva jeta jemo jete jejo -# | vati ujem uješ uje ujeva ujeta ujemo ujete ujeto -# | vati ovem oveš ove oveva oveta ovemo ovete ovejo -# | evati ujem uješ uje ujeva ujeta ujemo ujete ujejo -# | ovati ujem uješ uje ujeva ujeta ujemo ujete ujejo -# |
| ral rala rale rali ralo eri | ral rala rale rali ralo orji | sal sala sale sali salo ši | scal scala scale scali scalo ščij | slal slala slale slali slalo šlji | val vala vale vali valo j | uval uvala uvale uvali uvalo uj | val vala vale vali valo ovi | eval evala evale evali evalo uj | oval ovala ovale ovali ovalo uj |
| eriva erita erimo erite rat | orjiva orjita orjimo orjite rat | šiva šita šimo šite sat | ščijva ščijta ščijmo ščijte scat | šljiva šljita šljimo šljite slat | jva jta jmo jte vat | ujva ujta ujmo ujte uvat | oviva ovita ovimo ovite vat | ujva ujta ujmo ujte evat | ujva ujta ujmo ujte ovat |
| srati | prati | risati | scati | poslati | ruvati | rvati | pozvati | ločevati | Strahovati |

| | | | | | | | | | |
|--|--|---|--|--|--|---|---|--|---|
| 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 |
| tati čem češ če čeva četa čemo čete čejo -# | zati žem žeš že ževa žeta žemo žete žejo -# | eti am aš a ava ata amo ate ajo -# | eti em eš e eva eta emo ete ejo o | eti em eš e eva eta emo ete ejo o | eti im iš i iva ita imo ite ijo e | eti jem ješ je jeva jeta jemo jete jo jo | eti mem meš me meva meta memo mete mejo mo | eti amem ameš ame ameva ameta amemo amete amejo emo | eti nem neš ne neva neta nemo nete nejo no |
| tal tala tale tali talo | zal zala zale zali zalo | el ela ele eli elo | l la le li lo | el ela ele eli elo | el eli ele ela elo | l la le li lo | el ela ele eli elo | el ela ele eli elo | el ela ele eli elo |
| či čiva čita čimo čite tat | ži živa žita žimo žite zat | ej ejva ejta ejmo ejte et | i iva ita imo ite et | i iva ita imo ite et | i iva ita imo ite et | ji jiva jita jimo jite t | mi miva mita mimo mite et | emi emiva emita emimo emite et | ni niva nita nimo nite et |
| metati | rezati | imeti | zreti | uspjeti | pršeti | mreti | oteti | najeti | načeti |

| | | | | | | | | | |
|--|---|--|--|--|--|--|---|--|--|
| 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
| eti anem aneš ane aneva aneta anemo anete anejo eno | eti anjem anješ anje anjeva anjeta anjemo anjete anjeto -# | eti ojem oješ oje ojeva ojeta ojemo ojete ojejo ojo | deti m š -0 va sta mo ste jo do | jeti mem meš me meva meta memo mete mejo -# | leti eljem elješ elje eljeva eljeta eljemo eljete eljeto -# | leti olnem olneš olne olneva olneta olnemo olnete olnejo -# | reti erem ereš ere ereva ereta eremo erete erejo ero | reti arem areš are areva areta aremo arete arejo -# | teti čem češ če čeva četa čemo čete čejo -# |
| el ela ele eli elo | el ela ele eli elo | el ela ele eli elo | del dela dele deli delo | jel jela jele jeli jelo | lel lela lele leli lelo | lel lela lele leli lelo | rl rla rle rli rlo | rl rla rle rli rlo | tel tela tele teli telo |
| ani aniva anita animo anite et | anji anjiva anjita anjimo anjite et | oj ojva ojta ojmo ojte et | di diva dita dimmo dite det | mi miva mita mimo mite jet | elji eljiva eljita eljimo eljite let | olni olniva olnita olnimo olnite let | ri driva rita rimo rite ret | ri rita riva rimo rite ret | -# -# -# -# -# tet |
| meti | žeti | peti | vedeti | prijeti | mleti | klati | dreti | treti | hoteti |

| | | | | | | | | | |
|---|--|--|---|---|---|---|---|---|---|
| 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 |
| iti em eš e eva eta emo ete ejo -# | iti im iš i iva ita imo ite ijo e | iti idem ideš ide ideva ideta idemmo idete idejo -# | iti jdem jdeš jde jdeva jdeta jdemo jdete jdejo -# | iti jdem jdeš jde jdeva jdeta jdemo jdete jdejo -# | sti m š -0 va sta mo ste do jo | sti em eš e eva eta emo ete ejo -# | sti dem deš de deva deta demo dete dejo -# | sti dem deš de deva deta demo dete dejo -# | sti tem teš te teva teta temo tete tejo to |
| il ila ile ili ilo | il ila ile ili ilo | šel šla šle šli šlo | šel šla šle šli šlo | šel šla šle šli šlo | del dla dle dli dlo | el la le li lo | l la le li lo | del dla dle dli dlo | tel tla tle tli tlo |
| i iva ita imo ite it | i iva ita imo ite it | idi idiva idita idimo idite it | jdi jdiva jdita jdimo jdite it | jdi jdiva jdita jdimo jdite it | j jva jta jmo jte st | i iva ita imo ite st | di diva dita dimmo dite st | di diva dita dimmo dite st | ti tiva tita timo tite st |
| crkniti | gubiti | zaiti | zaiti | najti | jesti | tepsti | krasti | mesti | mesti |

| | | | | | | | | | |
|---|---|--|---|---|--|--|--|--|--|
| 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 |
| sti zem zeš ze zeva zeta zemo zete zejo -# | lesti olzem olzeš olze olzeva olzeta olzemo olzete olzejo -# | uti mem meš me meva meta memo met mejo me | uti ovem oveš ove oveva oveta ovemo ovete ovejo -# | uti pem peš pe peva peta pemo pete pejo po | či čem češ če čeva četa čemo čete čejo ko | či rem reš re reva reta remo rete reje -# | či rem reš re reva reta remo rete rejo -# | či žem žeš že ževa žeta žemo žete žejo -# | či žem žeš že ževa žeta žemo žete žejo -# |
| zel zla zle zli zlo | olzel olzla olzle olzli olzlo | mel mela mele meli melo | ul ula ule uli ulo | spel pela pele peli pelo | kel kla kle kli klo | gel gla gle gli glo | gel gla gle gli glo | gel gla gle gli glo | gel gla gle gli glo |
| zi ziva zita zimo zite st | olzi olziva olzita olzimo olzite lest | mi miva mita mimo mite mit | uj ujva ujta ujmo ujte ut | pi piva pita pimo pite ut | ci civa cita cimo cite č | zi ziva zita zimo zite č | -# -# -# -# č | zi ziva zita zimo zite č | ži živa žita žimo žite č |
| gristi | mlesti | duti | pluti | vsuti | teči | pomoči | zmoči | vpreči | leči |

| | | | | | | | | |
|-------|-------|------|-------|-----------|--------|----------|-------|-----|
| 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 |
| eči | (eči) | (či) | iti | (iti)(em) | (ti) | (ti) | (ti) | |
| žem | -# | # | om | m | em | m | -# | -# |
| žeš | -# | # | oš | š | eš | š | -# | -# |
| že | -# | # | o | -0 | e | - | -# | -# |
| ževa | -# | # | ova | va | eva | va | -# | -# |
| žeta | -# | # | osta | sta | eta | ta | -# | -# |
| žemo | -# | # | omo | mo | emo | mo | -# | -# |
| žete | -# | # | oste | ste | ete | te | -# | -# |
| žejo | -# | # | odo | do | ejo | jo | -# | -# |
| -# | -# | -# | il | jo | o | -# | -# | -# |
| gel | -# | # | ila | -# | -# | -# | -l | el |
| gla | -# | # | ile | -# | -# | -# | -la | la |
| gle | -# | # | ili | -# | -# | -# | -le | le |
| gli | -# | # | ilo | -# | -# | -# | -li | li |
| glo | -# | # | odem | -# | -# | -# | -lo | lo |
| zi | ži | či | odeš | -# | i | -# | -# | # |
| ziva | Živa | čiva | odiva | -# | iva | -# | -# | -# |
| zimo | Žita | čita | odita | -# | ita | -# | -# | -# |
| zita | Žimo | čimo | odimo | -# | imo | -# | -# | -# |
| zite | Žite | čite | odite | -# | ite | -# | -# | -# |
| eč | eč | č | it | -# | it | -# | t | # |
| vreči | vreči | teči | biti | iti | iti | ne imeti | gnati | iti |
| | | | | grem | pojdem | | | šel |

Nepregibne besedne vrste

Vse ostale slovarske besede ali gesla so nepregibna in zato v računalniškem slovarju nimajo niza končnic. Nepregibne besede so označene s števili, ki povedo besedno vrsto. Oznake nepregibnih vrst so podane v tabeli XII.

Tabela XII: Oznake nepregibnih besed v računalniškem slovarju. Oznaka sklona je označena z drugo do četrto cifro kode. Za kombinacije več kot treh sklonov se uporablja več oznak. Pri predlogih pomeni oznaka sklona s katerim sklonom se vežejo.

| | Dodatni kriterij | Oznaka | Oznaka | Oznaka |
|------------------|------------------|----------|-------------|------------|
| Prislovi | krajevni | 2001 | | |
| | časovni | 2002 | | |
| | vzročni | 2003 | | |
| | načinovni | 2004 | | |
| | | Predlogi | Sam. zaimki | Pom. glag. |
| | 1. sklon | 3001 | 5001 | 9001 |
| | 2. sklon | 3002 | 5002 | 9002 |
| | 3. sklon | 3003 | 5003 | 9003 |
| | 4. sklon | 3004 | 5004 | 9004 |
| | 5. sklon | 3005 | 5005 | 9005 |
| | 6. sklon | 3006 | 5006 | 9006 |
| | 2 in 5. sklon | 3025 | 5025 | 9025 |
| | 2., 4., 6. sklon | 3246 | 5246 | 9246 |
| Vezniki | | 4000 | | |
| Kratice in tujke | | 6000 | | |
| Medmeti | | 7000 | | |
| Členki | | 8000 | | |

V računalniškem slovarju je zbranih nekaj več kot 156.000 računalniških osnov in nepregibnih gesel. Vsa gesla omogočajo računalniku, da s pomočjo računalniških

končnic prepozna približno 2 milijona različnih besednih oblik. Tabela XIII kaže podroben pregled vsebine računalniškega slovarja.

Na tem mestu je možna krajša primerjava z gesli SSKJ, ki jih opisuje Jakopin (1995) in pregled, kje bi se dalo z združitvijo obeh korpusov pridobiti. Oba slovarja se dokaj dobro pokrivata, do večjih razlik pride le pri pridevniki, ki jih je v računalniškem slovarju precej več kot v SSKJ (70.000 proti 21.000). Vzrok je v tem, da ima v računalniškem slovarju skoraj vsak samostalnik tudi pridevniško obliko. Pri samostalniki ima računalniški slovar približno 60.000 besed, SSKJ pa 51.500. Razlika v prid računalniškega slovarja gre na račun pomanjševalnic in ženskih oblik, ki jih SSKJ nima toliko. Glagolov pa ima SSKJ skoraj 2000 več kot računalniški slovar (16.479 proti 14.760). Menim, da gre razlika največ na račun različnih predpon. Ostale besedne vrste, ki so v glavnem nepregibne, so bistveno manj številčne in so zaradi tega tudi razlike precej manjše.

Glavna razlika med SSKJ v računalniški obliki in med računalniškim slovarjem, opisanim v tem delu, ni v številu obdelanih gesel, ampak v tem, da ima pričujoči slovar pri vsakem geslu implicitno (preko enega od ustreznih 240 nizov končnic) podane vse možne sklanjatvene ali spregatvene oblike gesla.

Tabela XIII: Število posameznih gesel in približno število pregibnih oblik, ki jih prepozna računalnik.

| Besedna vrsta | Število besed v slovarju | Povprečno št. različnih računalniških končnic v enem nizu | Približno št. besednih oblik, ki jih prepozna računalniški slovar |
|-----------------------|--------------------------|---|---|
| samostalniki: | | | |
| moški | 23.148 | 9 | 364.000 |
| ženski | 26.686 | 9 | 240.000 |
| srednji | 12.246 | 8 | 98.000 |
| množinski | 475 | 5 | 2.400 |
| pridevniške besede | 72.872 | 12 | 875.000 |
| glagoli | 14.760 | 21 | 310.000 |
| prislovi | 5.199 | | |
| predlogi | 76 | | |
| vezniki | 52 | | |
| samostalniški zaimki | 207 | | |
| kratice in tujke | 79 | | Vseh |
| medmeti | 255 | | nepregibnih |
| členki | 25 | | besed |
| pom. glagolske oblike | 58 | | cca 6.000 |
| Skupaj | 156.138 | | 1.895.400 |

Dva primera uporabe slovarja z računalniškimi osnovami in nizi končnic

Opisani računalniški slovar se da uspešno uporabiti pri raziskavah različnih slovenskih besedil. Žal ročnega dela tak slovar zaenkrat še ne zmore nadomestiti. Da opravi svojo nalogo, mora računalnik pri vsaki besedi preiskovanega besedila:

1. poiskati vse možne računalniške osnove in končnice te besede,
2. poiskati vse možne nize končnic, ki pripadajo vsaki možni osnovi in

3. kombinirati vse končnice preiskovane besede z vsemi končnicami možnih nizov.

Program mora upoštevati več različnih situacij. V splošnem velja zahteva, da mora biti kombinacija računalniške osnove in računalniške končnice iz ustreznega niza enaka preiskovani besedni obliki. Če tako kombinacijo najde, je to veljavno geslo. V primeru, ko nobena kombinacija ne ustreza preiskovani besedi, jo računalnik proglasi za »neznano«. V primeru, da je ustreznih kombinacij več, je dana beseda lahko nastala iz različnih slovarskih gesel. Ta primer, ki še ni rešen, je za nadaljnji razvoj avtomatične besedne in kasneje tudi stavčne analize najbolj zanimiv, predstavlja glavno smer nadaljnjega razvoja.

Kot prvi primer bom opisal, kako program, ko sreča besedo *bral*, ugotovi, da gre za glagolsko besedo *brati* in hkrati še, da je to njegov deležnik. Program najprej razstavi besedo *bral* na vse možne računalniške osnove in končnice. Te možnosti so štiri:

bral + -0 *bra* + -l *br* + -al in *b* + -ral

Nato pogleda, če so te štiri osnove v slovarju in kateri nizi računalniških končnic so pripisani k vsaki osnovi. Ob osnovah najde program v slovarju naslednje nize končnic:

- ob osnovi *bral*- dva niza:

| | |
|--|---------------------------------|
| samostalniški, tabela II, 12. niz – končnice | -ec, -ca, -cu, -ec itd., |
| pridevniški, tabela VIII, 5. niz – končnice | -en, -na, -ne, -ni itd., |
- ob osnovah *bra*- *br*- v slovarju ni nobenega niza,
- ob osnovi *b*- so trije nizi:

| | |
|---|-----------------------------------|
| glagolski, tabela XI, 17. niz: končnice | -ati, -ojim, -ojiš itd., |
| glagolski, tabela XI, 30. niz: končnice | -rati, -erem- ereš itd. in |
| glagolski, tabela XI, 84. niz: končnice | -iti, -om, -oš itd. |

V sklepnih fazi program kombinira vse osnove in pripadajoče končnice. Pri prvi osnovi *bral* računalnik ugotovi, da niti prvi niti drugi niz ne vsebujeta končnice (-0), kar pomeni, da gesli *bral-ec* in *bral-en* ne dasta kombinacije *bral-0*. Zato računalnik pravilno odloči, da beseda *bral* ni samostojno geslo. Ker druga in tretja osnova (*bra*- in *br*-) nimata pridruženega nobenega niza, program ti dve osnovi izpusti. Med preverjanjem zadnjih treh nizov glagolskih končnic (pri osnovi *b*-) ugotovi, da ustreza le drugi niz, ki ima v 11 vrstici končnico *-ral*. Ta končnica da skupaj z osnovo *b*- pravilno vhodno besedo *b-ral*.

Zato da dobi slovarsko geslo, mora program povezati računalniško osnovo *b*- s končnico na prvem mestu 30. glagolskega niza, kar da iskano geslo: *b-rati*. Ker so končnice istih sklonov ali istih spregatvenih oblik v nizih vendno na istih mestih, je enajsto mesto, na kateri se nahaja računalniška končnica *-ral*, hkrati tudi podatek, da gre za deležnik preteklega (ali prihodnjega časa) v moški edninski obliki (torej: *sem, si, je, bom, boš, bo* ali *bi* bral). Točna določitev časa je možna le v kombinaciji z opredelitvijo pomožnega glagola, ki ga računalniški slovar določi pred tem.

Drugi primer uporabe računalniškega slovarja, sestavljenega iz računalniških osnov in pripadajočih nizov računalniških končnic, je iskanje konkordanc.

Če želimo na primer izvedeti, kam in kako je pesnik Janez Menart vtikal v svojo poezijo (Menart 1960; 1963; 1990) besedo *sreča*, moramo med njegovimi verzi poiskati vse t. i. konkordance gesla *sreča*. To storimo tako, da vpišemo zahtevano geslo v imenovalniku ednine. Program iz vpisanega geslo avtomatično sestavi s pomočjo računalniških končnic seznam devetih možnih besed z osnovo *sreč-* (v prvem nizu tabele III je devet različnih končnc) in z njimi primerja vsako besedo pesnikovih poezij. Predlog, da bi se dalo isto delo opraviti z navadnim iskanjem v urejevalniku besedil, je dober le navidezno. Če bi namreč iskali po besedilu samo z računalniško osnovo *sreč-*, bi dobili tudi veliko zvez (konkordanc) s povsem drugimi besedami, kot so: *srečanje*, *srečnež*, *srečnik*, *srečnica*, *srečka*, *srečolov*, *srečen*, *srečati*, *srečanje*, *srečevati*, *srečavati*, *srečevanje*, *srečavanje* in vse njihove izpeljanke.

Kot je pozoren bralec lahko ugotovil iz vseh konkordanc v tabeli XIV, se je nekaj podobnega primerilo tudi našemu računalniškemu slovarju. V Menartovi pesmi *Tristan in Izolda* (šesti primer) je računalnik našel konkordanco besede *srečati* in ne besede *sreča*. Takih napak je pri pravilni uporabi računalniškega slovarja malo in jih zato z lahkoto poiščemo, odstranimo ali premestimo na ustrezno mesto.

Tabela XV: Vse konkordance besede *sreča*, ki jih je računalniški program našel v 32.000 besed obsegajočem korpusu delčka Menartovih poezij (1960; 1963; 1990)

| Naslov pesmi | | | | |
|-------------------|---------------------------------|----------------|--------|-------------------------------|
| Parizina | | Poznal ni | sreče | ne gorja: ostale so noči brez |
| Anno Domini | Tam zunaj je življeneje, je | | sreča | in je beda, so rojstva |
| Topoli | | Rjoveli smo od | sreče, | streljali, tolkli, klali |
| Pogreb | ji vzela je življenje, odnesla | | srečo | let... Ne misli nanjo |
| Jesenski dan | poslušal o junaku, ki je | | srečo | šel iskat. |
| Tristan in Izolda | boš oči zaprla in | | sreča | me prijatelj |
| Poetska | sem ga potresal, žal za | | srečo, | da sem v kaj verjel |
| Nežnost v mraku | blodi mehka bolečina, ki ji | | sreča | je ime, ki zaljubljenca |
| Pričakovanje | in vse se staplja v | | srečo: | to je ona. |
| Pričakovanje | bohota se in peni in ko že | | sreča | skoraj zakipi |
| Otrok | srčeca drobni trepet kmalu te v | | sreči | začudi, potlej se odebelim |
| Pesnik se popravi | napis vklesan: umrl je od | | sreče. | Star sto let. |
| Pesnik se popravi | v zgodovino kot vzorec | | sreče | za vso domovino. |

Zaključek

Možnosti uporabe računalniškega slovarja so velike. Dobro lahko služi pri večini slovničnih obdelav slovenskih besedil. Posebej je treba poudariti, da mora kljub avtomatizaciji precejšen odstotek dela še vedno opraviti jezikoslovec sam.

Opisani računalniški slovar je seveda potreben še izpopolnjevanja, dopolnjevanja morda boljše sistematike in prevsem preverjanja in popravljanja napak. V delu, ki obsega preko 150,000 ročno vnešenih opredelitev gesel (besedne analize in določitve pravičnega niza računalniških končnic), seveda ne moremo pričakovati, da bi bilo povsem brez napak.

LITERATURA

- P. JAKOPIN, 1995: Nekaj števil iz Slovarja slovenskega knjižnega jezika. *Slavistična revija* 43/3. 341–375.
- P. JAKOPIN, A. BIZIAK, 1997: O strojno podprtem oblikoslovnem označevanju slovenskega besedila. *Slavistična revija* 47/3–4. 451–580.
- R. L. LENČEK, 1982: *The Structure and History of the Slovene Language*. New York: Columbia.
- J. MENART, 1960: *Parizina*. Ljubljana: DZS.
– – 1963: *Semafori mladosti*. Ljubljana: DZS.
– – 1990: *Srednjeveške pridige in balade*. Ljubljana: CZ.
- M. POPOVIČ, P. WILLET, 1990: Processing of Documents and Queries in a Slovene Language Free Text Retrieval System. *Literary and Linguistic Comp.* 5/2. 182–190.
- SS – *Slovenski pravopis*, 1962 Ljubljana: DZS.
- J. TOPORIŠIČ, 1975: Main Characteristics of the Slovene Language. D. Komac, R. Škerlj: *Angleško-slovenski in slovensko-angleški slovar*. Ljubljana: CZ.
– – 1984: *Slovenska slovnica*. Maribor: Obzorja.
– – 1994: *Slovenski jezik in sporočanje*. Maribor: Obzorja.
- J. ZUPAN, 1991: Slonček, program za geslenje slovenskega teksta. *Informatica* 3. 15–21.
- J. ZUPAN, 1999: The Application of Neural Networks in Linguistics. *Lecture Note on Physics*. Vol. 522. Eds. J. W. Clark, T. Lindenau, M. L. Ristig. Berlin: Springer. 1999, 224–241.
- F. ŽAGAR, 1991: *Slovenska slovnica in jezikovna vadnica*. Maribor: Obzorja.