

Oblikoslovni vzorci za strojno procesiranje slovenščine

Špela ARHAR HOLDT

Fakulteta za računalništvo in informatiko Univerze v Ljubljani,
Filozofska fakulteta Univerze v Ljubljani,
spela.arharholdt@fri.uni-lj.si

Abstract

The paper presents a morphological database drawn on 96,290 entries (nouns, adjectives, verbs and adverbs) from the Sloleks Morphological Lexicon of Slovene. Each lemma in the database has been ascribed a code for a morphological pattern, marking different features of its inflection. The candidates for the morphological patterns were firstly automatically extracted from Sloleks, taking into account morphosyntactic information and mutable word parts as defined in the lexicon. The 1,043 candidates were then manually revised and hierarchically organised into 494 distinct patterns, which are illustrated in this paper with a unique code, a prototypical example, a short description and a frequency of lemmata corresponding to the pattern. In the process, a set of inconsistencies and lacunae in Sloleks has been unveiled, however, they are seen as a valuable starting point for further development of this invaluable language resource for modern Slovene.

Ključne besede: Sloleks, leksikon besednih oblik, oblikoslovje, oblikoslovni vzorci, slovenščina

Keywords: Sloleks, word form lexicon, morphology, morphological patterns, Slovene

1 Uvod

Nacionalni raziskovalni projekt Nova slovnica sodobne standardne slovenščine: viri in metode (NSSSS),¹ ki je potekal med leti 2017 in 2020, je bil usmerjen v razvoj metodologije za strojno podprto analizo sodobne slovenščine in izdelavo prosto dostopnih baz z jezikovnimi podatki, ki jih skupnost potrebuje za pripravo jezikovnotehnoloških orodij, empiričnih jezikoslovnih analiz in sodobnega slovničnega opisa. Projektne aktivnosti so bile organizirane po jezikovnih ravlinah in razdeljene na oblikoslovje in besedotvorje, kolokacije, stalne besedne zveze, vezljivost in besedne nize. V prispevku predstavljamo projektni rezultat s prvega od naštetih področij: bazo oblikoslovnih podatkov, v kateri je 96.290 enotam leksikona besednih oblik Sloleks (Dobrovoljc et al. 2019) pripisana koda oblikoslovnega vzorca, po katerem se pregibajo. Baza je dostopna na repozitoriju CLARIN.SI (Arhar Holdt et al. 2020) pod licenco CC-BY-SA-4.0.

Metodologija priprave oblikoslovnih vzorcev je interdisciplinarno združila strojno luščenje jezikovnih podatkov in njihovo jezikoslovno pregledovanje in urejanje. Namen dela in načela metodologije smo že predstavili v članku (Arhar Holdt in Čibej 2018), ki se je osredotočal na samostalniške vzorce. Vzorci za ostale tri besedne vrste, ki so bile vključene v projektno delo (pridevnik, glagol, prislov), v literaturi še niso bili popisani, prav tako še ni bil predstavljen končni rezultat, podatkovna baza. V prispevku želimo pregledno in celovito predstaviti sistem vzorcev in bazo, deloma pa tudi aktivnosti, ki so stekle po koncu projekta NSSSS in zagotavljajo, da bodo pripravljene podatki uporabljeni za svoj osnovni namen: razvoj strojno podprtega generiranja novih leksikonskih enot na podlagi korpusnih podatkov.

Čeprav je nova metodologija primarno usmerjena v strojno procesiranje slovenščine, ponuja svež pogled tudi na področju jezikoslovja. Zaradi potrebe po strojni obvladljivosti je izrazito formalistična in površinska, na drugi strani pa ponuja podatke o zastopanosti jezikovnih pojavov, ki lahko pomagajo pregledneje strukturirati jezikovni opis (ibid.: 57–59). V nadaljevanju bo zanimivo identificirati značilnosti, ki

1 Spletna stran, ki predstavlja projektne cilje, rezultate in sodelujoče partnerje: <https://slovnica.ijs.si/>.

jih prinese nova organizacija podatkov, pri čemer se bo mogoče osredotočiti na posamezne skupine vzorcev in izbrane jezikovne pojave. Poleg primerjave s trenutnim referenčnim slovničnim opisom v Slovenski slovnici (Toporišič 2004) je med prioritetaми primerjava s pregibnostno-naglasnimi vzorci za slovenščino, ki jih predstavlja Mirtič (2015). Primerjave načrtno in v celoti puščamo za kasnejše delo.

V prispevku najprej povzamemo konceptualni in metodološki okvir priprave oblikoslovnih vzorcev, nato opišemo podatkovno bazo na repozitoriju CLARIN.SI. Večina prispevka je namenjenega opisu sistema vzorcev za samostalnik, pridevnik, glagol in prislov. Vzorci, ki so urejeni hierarhično, beležijo pa tudi informacijo o oblikoslovnih variantah, so predstavljeni v tabelarični obliki, ki naniza kode, opiše značilnosti vzorca in poda tipski primer leme, ki se pregiba po obravnavanem vzorcu. Prispevek sklenemo z napovedjo bodočega dela.

2 Konceptualni in metodološki okvir

Ker metodologijo in načela priprave oblikoslovnih vzorcev natančneje predstavlja že prispevek Arhar Holdt in Čibej (2018), na tem mestu zgolj povzemamo glavne značilnosti, ki lahko služijo za lažje razumevanje rezultatov v nadaljevanju.

2.1 Sloleks kot jezikovni vir za pripravo vzorcev

Jezikovni vir za pripravo baze z oblikoslovnimi vzorci je leksikon Sloleks, odprto dostopna zbirka besednih oblik, ki v trenutni različici (Dobrovoljc et al. 2019) prinaša oblikoslovne informacije za 100.805 slovenskih besed različnih besednih vrst. Leksikon, ki je bil razvit v projektu Sporazumevanje v slovenskem jeziku (SSJ)², vsebuje nabor pregibnih oblik, podatke o pogostosti leme in pregibnih oblik iz referenčnega pisnega korpusa Gigafida (Logar et al. 2012), standardne in nestandardne oblikoslovne variante ter povezave na besedotvorno sorodne besede (kot je predstavljeno v Arhar 2009).

2 Spletna stran projekta z opisom aktivnosti in povezavami na rezultate: <http://ssj.slovenscina.eu/>. Projektne specifikacije za razvoj leksikona (Erjavec et al. 2008) so na voljo na strani: <http://projekt.slovenscina.eu/Vsebine/Sl/Kazalniki/K3.aspx>.

Pomemben del podatkov, ki so bili vključeni v Sloleks, izvira iz ročno pripravljene leksikalne zbirke Ases (Amebisov skupni elektronski slovar, predstavljeno v Arhar in Holozan 2009), ki je osnova za jezikovne tehnologije, kot sta slovnični pregledovalnik Besana in strojni prevajalnik Presis.³ Paradigme leksikona Sloleks so bile v projektu SSJ urejene skladno s sistemom za oblikoskladenjsko označevanje besedilnih korpusov JOS,⁴ ki je del mednarodne iniciative MULTEXT-East in prinaša nabor oznak in ustrežajočih jezikovnih značilnosti za različne besedne vrste (Erjavec in Krek 2008).

Po končanem projektu SSJ so bile kot del načrtov za korpusno osnovani slovarski opis sodobne slovenščine (Gorjanc et al. 2015) opredeljene tudi smernice nadaljnjega razvoja leksikona Sloleks (Dobrovoljc et al. 2015), v katerih so avtorji izpostavili potrebo po dopolnitvi leksikona s strojno berljivimi oblikoslovnimi vzorci, ki bi omogočili »validacijo pregibnih paradigem iztočnic v obstoječih priročnikih, pripisovanje paradigem novim leмам ter razvoj metod za njihovo samodejno prepoznavanje v besedilnih korpusih« (ibid.: 95).

Strojno generiranje leksikonskih enot za slovenščino je preizkusil Rejc (2017). Avtor samostalnike iz leksikona Sloleks najprej gruči v skupine s podobnimi oblikoskladenjskimi lastnostmi in na tej podlagi z naivnim Bayesovim klasifikatorjem zgradi model za napovedovanje paradigem za nove besede. Izkazana povprečna uspešnost klasifikatorja samostalnikov je 88,99 %, natančnost napovedi novih paradigem pa 87,69 %. Leksikonske enote je torej mogoče strojno generirati brez predhodno pripravljenih vzorcev, vendar je pričakovano, da bodo vzorci izboljšali natančnost in omogočili preizkus novih postopkov. Strojno berljivi oblikoslovnimi vzorci imajo vrednost tudi za druge naloge s področja strojnega procesiranja slovenščine in za dopolnitev jezikovnih virov za človeškega uporabnika, zato je bila njihova priprava vključena v načrt projekta NSSSS.⁵

3 Več o teh izdelkih na spletni strani podjetja Amebis: <https://www.amebis.si/>

4 Projektna spletna stran: <http://nl.ijs.si/jos/>. V času pisanja prispevka je v rabi posodobljena različica označevalnega sistema, MULTEXT-East 6: <http://nl.ijs.si/ME/V6/msd/html/msd-sl.html>.

5 Delo se je pričelo leta 2017, leta 2018 pa je bil Sloleks nadgrajen v novo različico, Sloleks 2.0, ki vsebuje z metodami nevronske mreže pripisane naglase in zapise v fonetični pisavi IPA (Krsnik 2018) in posodobljene frekvenčne podatke iz korpusa Gigafida 2.0 (Krek et al.

2.2 Metodološke značilnosti

Ideja metodologije je relativno preprosta (za podrobnejši opis glej Arhar Holdt in Čibej 2018). V prvem koraku programsko pregledamo vse paradigme leksikona Sloleks in izluščimo kandidate za oblikoslovne vzorce, pri čemer se za razlikovanje vzorcev upoštevajo oblikoskladenjske oznake in spremenljivi deli besednih oblik. Za razliko od običajnih jezikoslovnih pristopov vzorci torej ne temeljijo na ločevanju besednih osnov in končnic, temveč nas zanima (za strojne pristope enostavno določljiv) lom med nespremenljivim in spremenljivim delom besedne oblike (npr. *odstot-ek*; *odstot-ka*; *odstot-kom*). V drugem koraku strojno pridobljene kandidate za vzorce ročno pregledamo in hierarhično uredimo glede na oblikovno sorodnost, nato pa jih opremimo še s podatkom o številu pripadajočih lem in izberemo tipski primer.⁶

Izvedba, ki smo jo najprej preizkusili na naboru vzorcev za samostalnik, je razkrila vrsto nedoslednosti, pomanjkljivosti in tudi vsebinskih napak v leksikonu (ibid.: 55–56). Manjkajoče ali zatipkane besedne oblike, nedosledno beležene in naključno razvrščene oblikoslovne variante so povzročile, da je bilo med kandidati za vzorce veliko šuma. Za lažjo sledljivost postopka teh pomanjkljivosti sprti nismo odpravljali, smo pa težave beležili, s čimer bo mogoče v nadaljevanju zagotoviti vsebinske in strukturne popravke leksikona. Upoštevati je torej treba, da priprava vzorcev poteka po korakih in da bo šele popravljanje leksikona, strojno pridobivanje novega leksikonskega gradiva in jezikoslovno urejanje strojno pripravljenih leksikonskih enot omogočilo končne rešitve.

Za kasnejšo obravnavo smo načrtno pustili dve problematični skupini lem. (1) Vzorci, ki izkazujejo nestandardne variante, so bili zaenkrat uvrščeni k najbližjim standardnim vzorcem, saj so

2020). Nabor leksikonskih enot je ostal enak, zato nadgradnja vira na pripravo vzorcev ni vplivala.

6 V večini primerov je bila za tipski primer izbrana v korpusu najpogostejša lema, vendar je v določenih primerih slednja zavajajoča (npr. zaradi napak v leksikonu, glej Arhar Holdt in Čibej 2018: 39). Takšni primeri so bili ročno zamenjani. Trenutno so tipski primeri predvsem pomoč za hitrejše razumevanje vzorcev. Za potrebe jezikovnega opisa in za didaktične namene je mogoče tipske primere izbrati drugače oz. z upoštevanjem dodatnih značilnosti.

nestandardne variante v trenutnem leksikonu pripisane sporadično (ibid.: 56).⁷ V prihodnosti bo treba nestandardne variante celostno pregledati in urediti, nato pa vzorce posodobiti. (2) Druga pričakovana posodobitev je nadgradnja formata oz. načina zapisovanja leksikonskih enot, ki se lahko pregibajo po več vzorcih, npr. *lesketati* – *lesketam* / *leskečem*. Tovrstni primeri so v trenutnem leksikonu beleženi, kot da gre za eno samo paradigmo, ki ima varianto pri vseh oblikah, kar ni ustrezno.⁸

S strojnim luščenjem smo pridobili 1.043 kandidatov za oblikoslovne vzorce, ki smo jih ročno pregledali, selekcionirali in uredili v 494 vzorcev. Te smo pripisali nazaj v podatke leksikona Sloleks in dopolnjene podatke objavili kot odprto dostopno bazo na repozitoriju CLARIN.SI (Arhar Holdt et al. 2020).

3 Podatkovna baza z oblikoslovnimi vzorci

Podatkovna baza z oblikoslovnimi vzorci je pripravljena v formatu XML, ki je zaradi obvladljivosti razdeljen v več datotek glede na prvo črko leksikonske enote (npr. ločeno vse enote na *a*-, nato na *b*- in tako naprej). Oblikoslovni vzorec (angl. *lexeme_pattern*) je vpisan v glavo (*head*) leksikonske enote kot eden od *grammarFeature*, kot prikazuje primer v nadaljevanju. V nadaljevanju zapisa (*body*), ki ga v spodnji primer zaradi prostorske potratnosti ne vključujemo, sledijo posamezne besedne oblike z oblikoskladenjsko oznako in korpusno frekvenco.

7 Precej nestandardnih oblik je bilo v leksikon dodatnih že med ročno gradnjo leksikalne zbirke Ases, in sicer kot pomoč pri ciljnem razvoju jezikovnih tehnologij (npr. za lažjo prepoznavo nestandardnih oblik, ki so jih v strojni pregledovalnik ali prevajalnik vpisovali uporabniki in uporabnice orodij). Nadgradnja, ki je bila korpusno osnovana, vendar je pokrila samo izbrane primere tipičnih oblikoslovnih zadreg, je bila dodana pri razvoju portala Slogovni priročnik (Dobrovoljc in Krek 2013) v sklopu projekta Sporazumevanje v slovenskem jeziku.

8 Združeni prikaz je mogoče videti v vmesniku leksikona Sloleks 2.0 na primeru *lesketati*: https://viri.cjvt.si/sloleks/slv/headword/2339/lesketati?tab=-_oblike.

```

<entry>
  <head>
    <headword>
      <lemma>čas</lemma>
    </headword>
    <lexicalUnit id="51045" sloleksId="LE_f23a00b66caaf81d53515fe6ffd532ca"
      slolekskey="s_čas" type="single">
      <lexeme>čas</lexeme>
    </lexicalUnit>
    <grammar>
      <category>samostalnik</category>
      <grammarFeature name="gender">masculine</grammarFeature>
      <grammarFeature name="lexeme_pattern">Sm1.1.o</grammarFeature>
      <grammarFeature name="type">common</grammarFeature>
    </grammar>
    <measureList>
      <measure type="frequency" source="Gigafida 2.0">1869664</measure>
    </measureList>
    <relatedEntryList/>
  </head>

```

Primer 1: Zapis oblikoslovnega vzorca v glavi leksikonske enote čas.

Na repozitoriju CLARIN.SI sta poleg datotek XML na voljo tudi ustrezajoča XML shema in pa knjižnica vzorcev, ki prinaša razlikovalne značilnosti za vsakega od vključenih vzorcev. Zapis vzorca, po katerem se pregiba primer čas, prikazujemo spodaj.⁹

AUTOMATIC PARADIGM 1: Somei: -Ø, Somer: -a, Somed: -u, Sometn: -Ø, Somem: -u, Someo: -om, Somdi: -a, Somdr: -ov, Somdd: -oma, Somdt: -a, Somdm: -ih, Somdo: -oma, Sommi: -i, Sommr: -ov, Sommd: -om, Sommt: -e, Sommm: -ih, Sommo: -i

Primer 2: Zapis vzorca z oblikoskladenjskimi oznakami in spremenljivimi deli oblik.

Tabela 1 za vsako besedno vrsto prikaže, s koliko lemami je zastopana v leksikonu Sloleks 2.0 in kolikšnemu številu teh lem je bil pripisan oblikoslovni vzorec. Pokritost z vzorci je opredeljena tudi z odstotki. Podatke prikazujemo za besedne vrste, ki smo jih obravnavali na projektu NSSSS (samostalnik, pridevnik, glagol, prislov), kot tudi druge (ne)pregibne besedne vrste v leksikonu.

⁹ Spremenljivi del oblik pri tem primeru sovpadе z jezikoslovnim razumevanjem samostalniških končnic: čas-Ø, čas-a, kar pri številnih vzorcih ne velja, npr. *otro-k*, *otro-ka* ali *minist-er*, *minist-ra* in podobno.

Tabela 1: Podatki o pregledanih in manjkajočih vzorcih glede na izhodiščni Sloleks 2.0.

Besedna vrsta	Št. lem v leksikonu Sloleks 2.0	Št. lem z vzorcem v oblikoslovni bazi	Odstotek pokritosti
Samostalnik	54.260	53.662	98,8 %
Pridevnik	26.612	26.422	99,3 %
Glagol	10.242	9.996	97,6 %
Prislov	6.906	6.172	89,4 %
Ostale pregibne besedne vrste	2.409	/	0 %
Ostale besedne vrste	373	/	0 %
Skupaj pregibne besedne vrste	100.429	96.252 ¹⁰	95,8 %
Skupaj vse besedne vrste	100.802	96.252	95,5 %

Kot kaže Tabela 1, so vzorci na voljo za 95,5 % enot leksikona Sloleks. Za obravnavo so ostale enote drugih besednih vrst (pregibni so še števniki in zaimki), od obravnavanih pa leksikonske enote, ki vsebujejo že opredeljene pomanjkljivosti (razdelek 2.2). Ker so bili najbolj problematični primeri izpuščeni iz obravnave, se napačno uvrščene leme v trenutnih rezultatih pojavljajo redko. Predvidevati pa je mogoče, da so podatki o številu vzorcev in lem, ki se po njih pregibajo, nižji, kot bodo po urejanju (in v nadaljnjih korakih seveda dopolnjevanju) leksikona. Frekvenčne podatke v tem prispevku je treba interpretirati skladno s temi opozorili.

4 Pregled urejenih vzorcev

V tem razdelku predstavljamo oblikoslovne vzorce po vrsti in ločeno za vse štiri pregledane besedne vrste: samostalnik, pridevnik, glagol in prislov. Vzorci za samostalnik so že bili predhodno predstavljeni, vendar v preliminarni različici, ki se je po nadgradnji metodologije še nekoliko spremenila.

¹⁰ Pričakovano število lem s pripisanim vzorcem je 96.290, vendar je preverba med pripravo prispevka razkrila, da nekaj vzorcev (večinoma samostalniških z varianto) pri avtomatskem vpisu podatkov v bazo ni bilo upoštevanih, zato se podatki razlikujejo za 38 lem. Identificirane nedoslednosti med sistemom vzorcev in bazo bodo v nadaljnjem delu odpravljene.

Podatkovne tabele so precej obširne, zato smo jih skušali urediti v strnjene in za branje pregledne oblike. V tabelah so vzorci urejeni hierarhično, kar prikazuje koda posameznega vzorca. Na prvem nivoju so vzorci, ki jih družijo podobne osnovne značilnosti, zbrani v skupine. Na drugem nivoju je izražen osnovni vzorec, ki je na tretjem nivoju nadalje členjen glede na izbrane oblikoslovne značilnosti: samostalniški vzorci so deljeni glede na to, ali je lema občno ali lastno ime; pridevniški glede na vrsto pridevnika; in glagolski glede na dovršnost glagola. Tem značilnostim se natančneje posvetimo pri vsaki posamezni tabeli v nadaljevanju.

Omeniti je treba še četrti nivo členjenja vzorcev, ki podaja informacijo o tem, ali vzorec vsebuje oblikoslovne variante oz. ali je v paradigmi omejen na posamezno slovnično kategorijo, kot npr. velja za določene samostalniške leme, ki izkazujejo oblike samo v ednini ali samo v množini. Te vzorce zaradi večje preglednosti prikazujemo v ločenih tabelah, kjer so variante in omejitve tudi natančneje opisane.

Koda vzorca je strukturirana stopenjsko po nivojih. Koda *Sm1.1.o*, pripisana samostalniku *čas*, ki smo ga uporabili kot zgled v Primeru 1 in 2, denimo pomeni:

- S: samostalnik,
- m: moški spol,
- 1.1: osnovni nepreglašeni vzorec za neživo: *čas-Ø*, *čas-a*, tož. *čas-Ø*; je del skupine *Sm1*, ki združuje vzorce te vrste za živo in neživo,
- o: občni samostalnik.

Primer 3: Stopenjski zapis vzorca, ki odraža hierarhično urejenost na nivoje.

V tabelah uporabljamo simbola * in **, ki pomenita:

* Vzorec v trenutnih podatkih ni izpričan. Kot bo razvidno iz tabel, je pri določenih primerih moč pričakovati, da bo vzorec s pojmom novega besedišča aktualiziran, ne pa vedno.

** Vzorec se v trenutnih podatkih pojavlja samo z oblikoslovnimi variantami ali z omejitvami. Vsi tovrstni vzorci so naštetni ločeno v tabelah, ki opisujejo variante in omejitve.

Kot je že bilo opozorjeno pri opisu metodologije, rezultati v tabelah odslkavajo stanje leksikona Sloleks, kakršen je v času priprave prispevka. To vključuje določene nedoslednosti pri beleženju variantnih oblik in paradigem, vključenost oblik ali paradigem, ki so glede na korpusne podatke redke ali se v referenčnem korpusu sploh ne pojavljajo,¹¹ potencialno vprašljivo beleženje lastnosti, kot sta živost ter dovršnost in podobno. Ko bo Sloleks nadgrajen in popravljen, bodo odpravljene tudi naštetе zadrege.

4.1 Samostalnik

Največji delež leksikona Sloleks 2.0 zajemajo samostalniki, ki jih je skupno 54.260, od tega 43.908 občnoimenskih in 10.352 lastnoimenskih. Samostalniške paradigme so ločene glede na slovnični spol in vsebujejo visoko število možnih variant in omejitev, zato je bil izhodiščni nabor kandidatov za samostalniške vzorce precej obsežen. Tudi urejene vzorce v nadaljevanju razdelka prikazujemo ločeno glede na spol.

4.1.1 Moški spol

Oblikoslovni vzorci za samostalnike moškega spola so razdeljeni v 10 skupin, ki prinašajo 61 vzorcev drugega nivoja. Ti so nadalje deljeni glede na to, ali je lema označena kot občnoimenski ali lastnoimenski samostalnik. Na tretji ravni je izpričanih 73 vzorcev. Razlikovalne značilnosti za umestitev vzorcev v skupine so (ne)preglašenost,¹² potencialno izpuščanje polglasnika ali podaljševanje osnove. Ločeni so tudi vzorci, ki se pregibajo z vezajem, ničto končnico¹³ ali so

11 V projektu Sporazumevanje v slovenskem jeziku je bil leksikon Sloleks opremljen s podatki o zastopanosti besednih oblik v referenčnem korpusu, vendar pogostnost oblik (še) ni bila upoštevana za naknadni pregled vsebine leksikona. Tipičen primer stanja, ki ga odražajo tudi vzorci v tabelah, lahko prikažemo s pomočjo glagola *zatreti*, ki se glede na podatke leksikona Sloleks (https://viri.cjvt.si/sloleks/slv/headword/10084/zatreti?tab=-_oblike) pojavlja z dvema variantnima paradigmama: *zatreti* – *zatrem* ter *zatreti* – *zatarēm*, pri čemer pa podatki iz korpusa nakazujejo, da je v rabi le prva varianta.

12 Premena vokala *o* v *e* za *c*, *č*, *ž*, *š* in *j*, ki v vzorcih npr. loči primere tipa *čas*, *čas-om* od *razvoj*, *razvoj-em*.

13 Simbol \emptyset uporabljamo le na mestih, kjer v jezikoslovnem smislu ustreza ničti končnici, v okviru zadane metodologije pa mu ustreza primerljivi pomen praznega mesta za spre-

podobni pridevniškimi ali samostalniškimi ženskim vzorcem. Urejene vzorce prikazuje Tabela 2.

Tabela 2: Oblikoslovni vzorci za samostalnike moškega spola.

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Sm1		Osnovni nepreglašeni vzorci za neživo in živo. Vključuje leme na -Ø, -o in -e.	
Sm1.1	čas	Osnovni nepreglašeni vzorec, neživo: <i>čas-Ø</i> , <i>čas-a</i> , tož. <i>čas-Ø</i> .	Sm1.1.o (čas, 4.886) Sm1.1.l (Windows, 2)
Sm1.2	predsednik	Osnovni nepreglašeni vzorec, živo: <i>predsednik-Ø</i> , <i>predsednik-a</i> , tož. <i>predsednik-a</i> .	Sm1.2.o (predsednik, 2.529) Sm1.2.l (Janez, 1.556)
Sm1.3	Helsinki	Osnovni nepreglašeni vzorec, kjer so samo množinske oblike, ki jih zato ne moremo strojno umeščati glede na živost: <i>Helsinki</i> , <i>Helsinkov</i> .	**Sm1.3.o **Sm1.3.l
Sm1.4	evro	Nepreglašeni vzorec, neživo, lema na -o: <i>evr-o</i> , <i>evr-a</i> , tož. <i>evr-o</i> .	Sm1.4.o (evro, 99) Sm1.4.l (Yugo, 2)
Sm1.5	dodo	Nepreglašeni vzorec, živo, lema na -o: <i>dod-o</i> , <i>dod-a</i> , tož. <i>dod-a</i> .	Sm1.5.o (dodo, 25) Sm1.5.l (Branko, 201)
Sm1.6	polfinale	Nepreglašeni vzorec, neživo, lema na -e: <i>polfinal-e</i> , <i>polfinal-a</i> , tož. <i>polfinal-e</i> .	Sm1.6.o (polfinale, 10) **Sm1.6.l
Sm1.7	kamikaze	Nepreglašeni vzorec, živo, lema na -e: <i>kamikaz-e</i> , <i>kamikaz-a</i> , tož. <i>kamikaz-a</i> .	Sm1.7.o (kamikaze, 1) Sm1.7.l (Mike, 32)
Sm1.8	las	Nepreglašeni vzorec, neživo, posebnosti: <i>las-Ø</i> , <i>las-a</i> , mn. <i>las-je</i> .	Sm1.8.o (las, 2) *Sm1.8.l
Sm1.9	otrok	Nepreglašeni vzorec, živo, posebnosti: <i>otro-k</i> , <i>otro-ka</i> , mn. <i>otro-ci</i> .	Sm1.9.o (otrok, 1) *Sm1.9.l
Sm1.10	človek	Nepreglašeni vzorec, živo, posebnosti: <i>človek</i> , mn. <i>ljudje</i> .	Sm1.10.o (človek, 7) *Sm1.10.l
Sm2		Osnovni preglašeni vzorci za neživo in živo. Vključuje leme na -Ø in -o. ¹⁴	
Sm2.1	razvoj	Preglašeni vzorec, neživo: <i>razvoj-Ø</i> , <i>razvoj-a</i> , tož. <i>razvoj-Ø</i> .	Sm2.1.o (razvoj, 640) Sm2.1.l (Andrej, 9)
Sm2.2	prijatelj	Preglašeni vzorec, živo: <i>prijatelj-Ø</i> , <i>prijatelj-a</i> , tož. <i>prijatelj-a</i> .	Sm2.2.o (prijatelj, 855) Sm2.2.l (Franc, 843)
Sm2.3	Radenci	Preglašeni vzorec, kjer imamo samo množinske oblike, ki jih zato ne moremo strojno umeščati glede na živost: <i>Radenci</i> , <i>Radencev</i> .	**Sm2.3.o **Sm2.3.l
Sm2.4	pončo	Preglašeni vzorec, neživo, lema na -o: <i>ponč-o</i> , <i>ponč-a</i> , tož. <i>ponč-o</i> .	Sm2.4.o (pončo, 7) *Sm2.4.l

menljivi del oblike. V prispevku sicer govorimo o ničti končnici, ko poimenujemo vrsto pregibanja, pri kateri oblike ostanejo enake lemi. Tako kot končnica v jezikoslovnem smislu tudi »spremenljivi del oblike« v teh primerih ni izražen.

14 Predvidoma obstaja tudi različica za -e, npr. *Djordj-e*, ki v podatkih trenutno še ni izpričana.

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Sm2.5	<i>Franjo</i>	Preglašeni vzorec, živo, lema na -o: <i>Franj-o</i> , <i>Franj-a</i> , tož. <i>Franj-a</i> .	*Sm2.5.o Sm2.5.l (Franjo, 9)
Sm2.6	<i>mož</i>	Preglašeni vzorec, živo, posebnosti: <i>mož-Ø</i> , <i>mož-a</i> , im. mn. <i>mož-je</i> .	Sm2.6.o (mož, 2) *Sm2.6.l
Sm2.7	<i>prakonj</i>	Preglašeni vzorec, živo, posebnosti: <i>prakonj-Ø</i> , <i>prakonj-a</i> , rod. mn. <i>prakonj-Ø</i> .	Sm2.7.o (prakonj, 1) *Sm2.7.l
Sm3		Nepreglašeni vzorci za neživo in živo, pri katerih se izpušča polglasnik.	
Sm3.1	<i>odstotek</i>	V spremenljivem delu je -k-, neživo: <i>odstot-ek</i> , <i>odstot-ka</i> , tož. <i>odstot-ek</i> .	Sm3.1.o (odstotek, 829) *Sm3.1.l
Sm3.2	<i>deček</i>	V spremenljivem delu je -k-, živo: <i>deč-ek</i> , <i>deč-ka</i> , tož. <i>deč-ka</i> .	Sm3.2.o (deček, 232) Sm3.2.l (Božiček, 128)
Sm3.3	<i>sejem</i>	V spremenljivem delu je -m-, neživo: <i>sej-em</i> , <i>sej-ma</i> , tož. <i>sej-em</i> .	Sm3.3.o (sejem, 264) **Sm3.3.l
Sm3.4	<i>Erazem</i>	V spremenljivem delu je -m-, živo: <i>Eraz-em</i> , <i>Eraz-ma</i> , tož. <i>Eraz-ma</i> .	Sm3.4.o (mikroorganizem, 1) Sm3.4.l (Erazem, 2)
Sm3.5	<i>meter</i>	V spremenljivem delu je -r-, neživo: <i>met-er</i> , <i>met-ra</i> , tož. <i>met-er</i> .	Sm3.5.o (meter, 152) Sm3.5.l (Peter, ¹⁵ 2)
Sm3.6	<i>minister</i>	V spremenljivem delu je -r-, živo: <i>minist-er</i> , <i>minist-ra</i> , tož. <i>minist-ra</i> .	Sm3.6.o (minister, 41) Sm3.6.l (Aleksander, 30)
Sm3.7	<i>Zadar</i>	V spremenljivem delu je -r-, v im. ed. -ar, neživo: <i>Zad-ar</i> , <i>Zad-ra</i> , tož. <i>Zad-ar</i> .	*Sm3.7.o **Sm3.7.l
Sm3.8	<i>Aleksandar</i>	V spremenljivem delu je -r-, v im. ed. -ar, živo: <i>Aleksand-ar</i> , <i>Aleksand-ra</i> , tož. <i>Aleksand-ra</i> .	*Sm3.8.o Sm3.8.l (Aleksandar, 2)
Sm3.9	<i>posel</i>	V spremenljivem delu je -l-, neživo: <i>pos-el</i> , <i>pos-la</i> , tož. <i>pos-el</i> .	Sm3.9.o (posel, 43) **Sm3.9.l
Sm3.10	<i>osel</i>	V spremenljivem delu je -l-, živo: <i>os-el</i> , <i>os-la</i> , tož. <i>os-la</i> .	Sm3.10.o (osel, 7) Sm3.10.l (Pavel, 9)
Sm3.11	<i>kamen</i>	V spremenljivem delu je -n-, neživo: <i>kam-en</i> , <i>kam-na</i> , tož. <i>kam-en</i> .	Sm3.11.o (kamen, 31) **Sm3.11.l
Sm3.12	<i>oven</i>	V spremenljivem delu je -n-, živo: <i>ov-en</i> , <i>ov-na</i> , tož. <i>ov-na</i> .	Sm3.12.o (oven, 3) Sm3.12.l (Domen, 25)
Sm3.13	<i>mozeg</i>	V spremenljivem delu je -g-, neživo: <i>moz-eg</i> , <i>moz-ga</i> , tož. <i>moz-eg</i> .	Sm3.13.o (mozeg, 1) *Sm3.13.l
Sm3.14	<i>mezeg</i>	V spremenljivem delu je -g-, živo: <i>mez-eg</i> , <i>mez-ga</i> , tož. <i>mez-ga</i> .	Sm3.14.o (mezeg, 1) *Sm3.14.l
Sm3.15	<i>hrbet</i>	V spremenljivem delu je -t-, neživo: <i>hrb-et</i> , <i>hrb-ta</i> , tož. <i>hrb-et</i> .	Sm3.15.o (hrbet, 1) *Sm3.15.l
Sm3.16	<i>valpet</i>	V spremenljivem delu je -t-, živo: <i>valp-et</i> , <i>valp-ta</i> , tož. <i>valp-ta</i> .	Sm3.16.o (valpet, 1) *Sm3.16.l

15 Nekateri primeri osebnih lastnih imen, ki v rabi nastopajo tudi kot del zemljepisnih lastnih imen (npr. *Sveti Peter*), so v leksikon vključeni tudi z različico, ki izraža neživost.

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Sm3.17	<i>pes</i>	V spremenljivem delu je <i>-s-</i> , živo: <i>p-es, p-sa</i> , tož. <i>p-sa</i> .	Sm3.17.o (<i>pes</i> , 1) *Sm3.17.l
Sm3.18	<i>oves</i>	V spremenljivem delu je <i>-s-</i> , živo: <i>ov-es, ov-sa</i> , tož. <i>ov-es</i> .	**Sm3.18.o *Sm3.18.l
Sm3.19	<i>veter</i>	V spremenljivem delu je <i>-r-</i> , neživo, podaljšava v množini: <i>vet-er, vet-ra</i> , tož. <i>vet-er</i> , mn. <i>vet-rovi</i> .	Sm3.19.o (<i>veter</i> , 1) *Sm3.19.l
Sm3.20	<i>blagor</i>	V spremenljivem delu je <i>-r-</i> , neživo: <i>blag-or, blag-ra</i> , tož. <i>blag-or</i> .	Sm3.20.o (<i>blagor</i> , 1) *Sm3.20.l
Sm3.21	<i>Russell</i>	V spremenljivem delu je <i>-ll-</i> , živo: <i>Russ-ell, Russ-lla</i> tož. <i>Russ-lla</i> .	*Sm3.21.o Sm3.21.l (<i>Russell</i> , 3)
Sm4		Preglašeni vzorci za neživo in živo, pri katerih se izpusti polglasnik.	
Sm4.1	<i>marec</i>	V spremenljivem delu je <i>-c-</i> , neživo: <i>mar-ec, mar-ca</i> , tož. <i>mar-ec</i> .	Sm4.1.o (<i>marec</i> , 406) Sm4.1.l (<i>Žalec</i> , 7)
Sm4.2	<i>igralec</i>	V spremenljivem delu je <i>-c-</i> , živo: <i>igral-ec, igral-ca</i> , tož. <i>igral-ca</i> .	Sm4.2.o (<i>igralec</i> , 1.906) Sm4.2.l (<i>Avstrijec</i> , 163)
Sm4.3	<i>čevelj</i>	V spremenljivem delu je <i>-lj-</i> , neživo: <i>čev-elj, čev-lja</i> , tož. <i>čev-elj</i> .	Sm4.3.o (<i>čevelj</i> , 52) **Sm4.3.l
Sm4.4	<i>rabelj</i>	V spremenljivem delu je <i>-lj-</i> , živo: <i>rab-elj, rab-lja</i> , tož. <i>rab-lja</i> .	Sm4.4.o (<i>rabelj</i> , 10) Sm4.4.l (<i>Avbelj</i> , 43)
Sm4.5	<i>ogenj</i>	V spremenljivem delu je <i>-nj-</i> , neživo: <i>og-enj, og-nja</i> , tož. <i>og-enj</i> .	Sm4.5.o (<i>ogenj</i> , 12) **Sm4.5.l
Sm4.6	<i>suženj</i>	V spremenljivem delu je <i>-nj-</i> , živo: <i>suž-enj, suž-nja</i> , tož. <i>suž-nja</i> .	Sm4.6.o (<i>suženj</i> , 4) *Sm4.6.l
Sm4.7	<i>Mengeš</i>	V spremenljivem delu je <i>-š-</i> , neživo: <i>Meng-eš, Meng-ša</i> , tož. <i>Meng-eš</i> .	*Sm4.7.o **Sm4.7.l
Sm4.8	<i>Badovinac</i>	V spremenljivem delu je <i>-ac-</i> , živo: <i>Badovin-ac, Badovin-ca</i> , tož. <i>Badovin-ca</i> .	*Sm4.8.o Sm4.8.l (<i>Badovinac</i> , 9)
Sm5		Vzorci za sklanjanje z uporabo vezaja, pri čemer se pojavljajo preglašene in nepreglašene končnice. ¹⁶	
Sm5.1	<i>CD</i>	Kratice, ki se sklanjajo preglašeno, neživo: <i>CD, CD-ja</i> , tož. <i>CD</i> .	Sm5.1.o (<i>CD</i> , 8) **Sm5.1.l
Sm5.2	<i>KUD</i>	Kratice, ki se sklanjajo nepreglašeno, neživo: <i>KUD, KUD-a</i> , tož. <i>KUD</i> .	**Sm5.2.o **Sm5.2.l
Sm6		Vzorci za sklanjanje z ničtimi končnicami.	
Sm6.1	<i>foto</i>	Vzorec za sklanjanje z ničtimi končnicami: <i>foto</i> .	Sm6.1.o (<i>mio</i> , 3) **Sm6.1.l

16 Izpričani so samo primeri s kategorijo neživo, vendar je mogoče predvideti, da obstajajo tudi paradigme za živo (npr. *DJ*, tož. *DJ-ja*).

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Sm7		Vzorci za leme na <i>-a</i> ali <i>-ja</i> , ki so enaki ženskim vzorcem in se alternativno pregibajo po preglašanih in nepreglašanih vzorcih za moški spol.	
Sm7.1	<i>panda</i>	Nepreglašeno, živo: <i>pand-a</i> , <i>pand-e</i> / <i>pand-a</i> .	Sm7.1.o (panda, 28) Sm7.1.l (Miha, 134)
Sm7.2	<i>zborovodja</i>	Preglašeno, živo, vriva se <i>-i-</i> v rodilniku dv. in mn.: <i>zborovod-ja</i> , <i>zborovod-je</i> / <i>zborovod-ja</i> ; rod. mn. <i>zborovod-ij</i> / <i>zborovod-jev</i> .	Sm7.2.o (zborovodja, 19) Sm7.2.l (Mitja, 11)
Sm7.3	<i>kuža</i>	Preglašeno, živo, ne vriva se <i>-i-</i> v rodilniku dv. in mn.: <i>kuž-a</i> , <i>kuž-e</i> / <i>kuž-a</i> , rod. mn. <i>kuž-Ø</i> / <i>kuž-ev</i> .	Sm7.3.o (kuža, 16) Sm7.3.l (Matija, 56)
Sm7.4	<i>tesla</i>	Preglašeno, živo, vriva se polglasnik, zato lom pri <i>-el</i> : <i>tes-la</i> , <i>tes-le</i> / <i>tes-la</i> , rod. mn. <i>tes-el</i> / <i>tes-lov</i> .	Sm7.4.o (tesla, 1) *Sm7.4.l
Sm8		Vzorci za živo in neživo, ki v dvojini in množini izražajo podaljšavo z <i>-ov-</i> .	
Sm8.1	<i>sok</i>	Vzorec za neživo, ki izraža podaljšavo z <i>-ov-</i> : <i>sok-Ø</i> , dvojina <i>sok-ova</i> .	Sm8.1.o (sok, 6) *Sm8.1.l
Sm8.2	<i>bog</i>	Vzorec za živo, ki izraža podaljšavo z <i>-ov-</i> : <i>duh-Ø</i> , dvojina <i>duh-ova</i> .	Sm8.2.o (bog, 2) *Sm8.2.l
Sm9		Vzorci za živo in neživo, ki se podaljšujejo z <i>-j-</i> , <i>-t-</i> ali <i>-n-</i> .	
Sm9.1	<i>denar</i>	Vzorec za neživo, ki izraža podaljšavo z <i>-j-</i> : <i>denar-Ø</i> , <i>denar-ja</i> , tož. <i>denar-Ø</i> .	Sm9.1.o (denar, 535) *Sm9.1.l
Sm9.2	<i>direktor</i>	Vzorec za živo, ki izraža podaljšavo z <i>-j-</i> : <i>direktor-Ø</i> , <i>direktor-ja</i> , tož. <i>direktor-ja</i> .	Sm9.2.o (direktor, 1.154) Sm9.2.l (Igor, 632)
Sm9.3	<i>kofe</i>	Vzorec za neživo, ki izraža podaljšavo s <i>-t-</i> : <i>kofe-Ø</i> , <i>kofe-ta</i> , tož. <i>kofe-Ø</i> .	Sm9.3.o (kofe, 1) *Sm9.3.l
Sm9.4	<i>pezde</i>	Vzorec za živo, ki izraža podaljšavo s <i>-t-</i> : <i>pezde-Ø</i> , <i>pezde-ta</i> , tož. <i>pezde-ta</i> .	Sm9.4.o (pezde, 7) Sm9.4.l (Jože, 109)
Sm9.5	<i>buhtelj</i>	Vzorec za neživo, ki izraža podaljšavo z <i>-n-</i> : <i>buhtelj-Ø</i> , <i>buhtelj-na</i> , tož. <i>buhtelj-Ø</i> .	Sm9.5.o (buhtelj, 3) *Sm9.5.l
Sm10		Vzorci, podobni pridevniškim.	
Sm10.1	<i>moški</i>	Vzorec, podoben pridevniškemu, živo: <i>moški</i> , <i>moškega</i> , tož. <i>moškega</i> .	Sm10.1.o (moški, 5) Sm10.1.l (Cetinski, 15)

Samostalniški vzorci moškega spola izkazujejo 4 variante, razen tega se v leksikonu Sloleks pojavljajo paradigme, ki imajo le oblike za ednino ali množino, kar je pri bodočem urejanju leksikona mogoče ohraniti ali (zlasti pri trenutno edninskih samostalnikih) spremeniti. Variante ali omejitve se pojavljajo pri 52 raznolikih vzorcih, kot prikazuje Tabela 3.

Tabela 3: Variante in omejitve v samostalniških vzorcih moškega spola.

Koda	Opis variante / omejitve	Vzorci in št. lem v bazi
V1	Imenovalnik množine: <i>-i/-je</i> (npr. <i>gospodi / gospodje</i>).	Sm1.1.o-V1 (ud, 1), Sm1.2.o-V1 (gospod, 32), Sm1.2.l-V1 (Hrvat, 5), Sm1.2.o-V1+V3 (gost, 1), Sm9.4.o-V1 (oče, 2)
V2	Rodilnik ednine, pri samostalnikih, ki izražajo živost, tudi tožilnik ednine: <i>-a/-u</i> (npr. <i>tata / tatu</i>)	Sm1.1.o-V2 (mir, 5), Sm1.1.o-ednina+V2 (sram, 7), Sm1.1.o-V2+V3 (nos, 2), Sm8.1.o-V2 (strah, 12), Sm8.1.o-V2+V3+V4 (most, 1), Sm8.2.o-V2 (tat, 1)
V3	Mestnik množine, v določenih primerih tudi dvojine: <i>-eh/- (ov)ih</i> (npr. <i>gosteh / gostih</i>).	Sm1.1.o-V2+V3 (nos, 2), Sm1.1.o-V3+V4 (kol, 1), Sm1.2.o-V1+V3 (gost, 1), Sm4.1.o-V3 (konec, 1), Sm8.1.o-V2+V3+V4 (most, 1)
V4	Orodnik množine: <i>-(ov)i/-mi</i> (npr. <i>koli / kolmi</i>).	Sm1.1.o-V3+V4 (kol, 1), Sm8.1.o-V2+V3+V4 (most, 1)
ednina	Oblike so (trenutno) samo za ednino.	Sm1.1.o-ednina (promet, 524), Sm1.1.l-ednina (Maribor, 445), Sm1.1.o-ednina+V2 (sram, 7), Sm1.4.o-ednina (vaterpolo, 18), Sm1.4.l-ednina (Nato, 43), Sm1.6.o-ednina (pasodoble, 1), Sm1.6.l-ednina (Google, 8), Sm2.1.o-ednina (hokej, 124), Sm2.1.l-ednina (Kranj, 69), Sm3.1.o-ednina (nameček, 6), Sm3.1.l-ednina (Podčetrtek, 5), Sm3.3.o-ednina (turizem, 207), Sm3.3.l-ednina (Videm, 1), Sm3.5.o-ednina (koper, 9), Sm3.5.l-ednina (Koper, 6), Sm3.7.l-ednina (Zadar, 2), Sm3.9.l-ednina (Basel, 7), Sm3.11.o-ednina (česen, 2), Sm3.11.l-ednina (München, 16), Sm3.13.o-ednina (bezeg, 1), Sm3.15.o-ednina (ocet, 1), Sm3.18.o-ednina (oves, 1), Sm4.1.o-ednina (svinec, 24), Sm4.1.l-ednina (Gradec, 18), Sm4.3.o-ednina (žajbelj, 3), Sm4.3.l-ednina (Bruselj, 7), Sm4.5.l-ednina (Sovodenj, 1), Sm4.7.l-ednina (Mengeš, 1), Sm5.1.o-ednina (DDV, 81), Sm5.1.l-ednina (BMW, 437), Sm6.1.o-ednina (foto, 16), Sm6.1.l-ednina (New, 132), Sm9.1.o-ednina (humor, 42), Sm9.1.l-ednina (Tivoli, 67)
množina ¹⁷	Oblike so (trenutno) samo za množino.	Sm1.3.o (otrobi, 12), Sm1.3.l (Helsinki, 25), Sm2.3.o (tisoči, 5), Sm2.3.l (Radenci, 34)
Somei	Oblike so (trenutno) samo za imenovalnik ednine.	Sm6.1.o-Somei (EUR, 53)
Sometn	Oblike so (trenutno) samo za tožilnik ednine.	Sm6.1.o-Sometn (poštev, 1)

¹⁷ Zaradi napake pri strojnem vpisu ti vzorci v oblikoslovni bazi trenutno nimajo potrebne opredelilne delnosti (*množina*), kar bo popravljeno pri naslednji nadgradnji baze.

4.1.2 Ženski spol

Oblikoslovni vzorci za samostalnike ženskega spola so razdeljeni v 9 skupin, ki prinašajo 29 vzorcev drugega nivoja. Ti so nadalje deljeni glede na to, ali so občnoimenski ali lastnoimenski, na tretji ravni je izpričanih 27 vzorcev. Razlikovalne značilnosti za umestitev vzorcev v skupine so spremenljivi del leme (npr. na *-a*, *-ev* ali drugo) ter vri-vanje ali izpuščanje polglasnika oz. vokala v paradigmi. V posebni skupini so umeščeni samostalniki, ki se pregibajo z ničto končnico, ter samostalniki, ki se pregibajo podobno pridevnikom. Urejene vzorce prikazuje Tabela 4.

Tabela 4: Oblikoslovni vzorci za samostalnike ženskega spola.

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Sz1		Vzorci za samostalnike ženskega spola, ki se končajo na <i>-a</i> . Vključuje primere, kjer se lema konča na zaporedna vokala.	
Sz1.1	<i>država</i>	Leme na <i>-a</i> : <i>držav-a</i> , <i>držav-e</i> , rod. mn. <i>držav-Ø</i> .	Sz1.1.o (država, 12.102) Sz1.1.l (Amerika, 465)
Sz1.2	<i>alinea</i>	Leme na <i>-a</i> , ki se končajo na zaporedna vokala: <i>aline-a</i> , <i>aline-e</i> , rod. mn. <i>aline-j</i> .	Sz1.2.o (alinea, 15) Sz1.2.l (Maria, 23)
Sz1.3	<i>gospa</i>	Vzorec za pregibanje samostalnika <i>gospa</i> , ki vsebuje posebnosti pri več oblikah: <i>gosp-a</i> , <i>gosp-e</i> , rod. mn. <i>gosp-a</i> .	Sz1.3.o (gospa, 1) *Sz1.3.l
Sz1.4	<i>Golte</i>	Vzorec za pregibanje samostalnika <i>Golte</i> , ki vsebuje samo množino in posebnosti: <i>Golt-e</i> , <i>Golt-Ø</i> , z <i>Golt-emi</i> .	**Sz1.4.o *Sz1.4.l
Sz2		Osnovni vzorci za samostalnike ženskega spola, ki se ne končajo na <i>-a</i> (in ne na <i>-ev</i> , ker so slednji pri Sz3).	
Sz2.1	<i>možnost</i>	Leme, ki niso na <i>-a</i> ali <i>-ev</i> : <i>možnost-Ø</i> , <i>možnost-i</i> .	Sz2.1.o (možnost, 5.237) **Sz2.1.l
Sz3		Vzorci za samostalnike ženskega spola, ki se končajo na <i>-ev</i> .	
Sz3.1	<i>odločitev</i>	Leme na <i>-ev</i> : <i>odločit-ev</i> , <i>odločit-ve</i> .	Sz3.1.o (odločitev, 823) **Sz3.1.l
Sz4		Vzorci za samostalnike ženskega spola, ki se ne končajo na <i>-a</i> ali <i>-ev</i> in imajo v množini v določenih sklonih v spremenljivem delu <i>-e-</i> .	

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Sz4.1	<i>stran</i>	Leme, ki se ne končajo na -a ali -ev in imajo v množini v določenih sklonih v spremenljivem delu -e-: <i>stran-Ø, stran-i</i> , rod. mn. <i>stran-eh</i> .	Sz4.1.o (stran, 101) *Sz4.1.l
Sz4.2	<i>kri</i>	Vzorec za pregibanje samostalnika <i>kri</i> , ki vsebuje samo ednino in posebnosti pri več oblikah: <i>kr-i, kr-vi</i> .	**Sz4.2.o *Sz4.2.l
Sz4.3	<i>Žiri</i>	Vzorec za pregibanje samostalnika <i>Žiri</i> , ki vsebuje samo množino in posebnosti pri več oblikah: <i>Žir-i, Žir-ov</i> .	*Sz4.3.o **Sz4.3.l
Sz5		Vzorci za samostalnike ženskega spola na -a, kjer se v rodilniku dvojine in množine vriva -e- ali -i-, redko tudi -a-.	
Sz5.1	<i>igra</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -r-: <i>ig-ra, ig-re</i> , rod. mn. <i>ig-er</i> .	Sz5.1.o (igra, 64) Sz5.1.l (Petra, 6)
Sz5.2	<i>izkušnja</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -nj-: <i>izkuš-nja, izkuš-nje</i> , rod. mn. <i>izkuš-enj</i> .	Sz5.2.o (izkušnja, 54) **Sz5.2.l
Sz5.3	<i>kaplja</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -lj-: <i>kap-lja, kap-lje</i> , rod. mn. <i>kap-elj</i> .	Sz5.3.o (kaplja, 34) **Sz5.3.l
Sz5.4	<i>megla</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -l-: <i>meg-la, meg-le</i> , rod. mn. <i>meg-el</i> .	Sz5.4.o (megla, 33) **Sz5.4.l
Sz5.5	<i>tekma</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -m-: <i>tek-ma, tek-me</i> , rod. mn. <i>tek-em</i> .	Sz5.5.o (tekma, 28) *Sz5.5.l
Sz5.6	<i>opna</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -n-: <i>op-na, op-ne</i> , rod. mn. <i>op-en</i> .	Sz5.6.o (opna, 24) Sz5.6.l (Vesna, 2)
Sz5.7	<i>spužva</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -v-: <i>spuž-va, spuž-ve</i> , rod. mn. <i>spuž-ev</i> .	Sz5.7.o (spužva, 15) *Sz5.7.l
Sz5.8	<i>ladja</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -j-: <i>lad-ja, lad-je</i> , rod. mn. <i>lad-ij</i> .	Sz5.8.o (ladja, 6) Sz5.8.l (Katja, 6)
Sz5.9	<i>ovca</i>	Leme na -a, kjer se v rodilniku dvojine in množine vriva vokal, v spremenljivem delu je -c-: <i>ov-ca, ov-ce</i> , rod. mn. <i>ov-(a)c</i> .	**Sz5.9.o *Sz5.9.l
Sz5.10	<i>mati</i>	Vzorec za sklanjanje samostalnikov <i>mati</i> in <i>hči</i> , ki vsebuje posebnosti: <i>mat-i, mat-ere</i> .	Sz5.10.o (mati, 2) *Sz5.10.l
Sz6		Vzorci za sklanjanje z ničtimi končnicami.	

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Sz6.1	<i>lady</i>	Vzorec za sklanjanje z ničtimi končnicami: <i>lady</i> .	Sz6.1.o (<i>lady</i> , 7) Sz6.1.l (Jennifer, 3.326)
Sz7		Vzorci za samostalnike ženskega spola, ki se ne končajo na <i>-a</i> in vsebujejo izpuštljiv polglasnik. V dv. in mn. je <i>-i-</i> (<i>bolez-nima</i>).	
Sz7.1	<i>bolezen</i>	Leme na <i>-en</i> , vzorec vsebuje izpuštljiv polglasnik: <i>bolez-en</i> , <i>bolez-ni</i> , daj. dv. <i>bolez-nima</i> .	Sz7.1.o (<i>bolezen</i> , 12) *Sz7.1.l
Sz7.2	<i>misel</i>	Leme na <i>-el</i> , vzorec vsebuje izpuštljiv polglasnik: <i>mis-el</i> , <i>mis-li</i> , daj. dv. <i>mis-lima</i> .	Sz7.2.o (<i>misel</i> , 3) *Sz7.2.l
Sz7.3	<i>povodenj</i>	Leme na <i>-enj</i> , vzorec vsebuje izpuštljiv polglasnik: <i>povod-enj</i> , <i>povod-nji</i> , daj. dv. <i>povod-njima</i> .	Sz7.3.o (<i>povodenj</i> , 2) *Sz7.3.l
Sz7.4	<i>pesem</i>	Leme na <i>-em</i> , vzorec vsebuje izpuštljiv polglasnik: <i>pes-em</i> , <i>pes-mi</i> , daj. dv. <i>pes-mima</i> .	Sz7.4.o (<i>pesem</i> , 1) *Sz7.4.l
Sz8		Vzorci za samostalnike ženskega spola, ki se ne končajo na <i>-a</i> in vsebujejo izpuštljiv polglasnik. V dv. in mn. je <i>-e-</i> (<i>ravnema</i>).	
Sz8.1	<i>raven</i>	Vzorci, ki vsebujejo izpuštljiv polglasnik in v katere se vriva <i>-e-</i> . Leme na <i>-en</i> : <i>rav-en</i> , <i>rav-ni</i> , daj. dv. <i>rav-nema</i> .	**Sz8.1.o *Sz8.1.l
Sz8.2	<i>ravan</i>	Vzorci, ki vsebujejo izpuštljiv polglasnik in v katere se vriva <i>-e-</i> . Leme na <i>-an</i> : <i>rav-an</i> , <i>rav-ni</i> , daj. dv. <i>rav-nema</i> .	**Sz8.2.o *Sz8.2.l
Sz8.3	<i>reber</i>	Vzorci, ki vsebujejo izpuštljiv polglasnik in v katere se vriva <i>-e-</i> . Leme na <i>-er</i> : <i>reb-er</i> , <i>reb-ri</i> , daj. dv. <i>reb-rema</i> .	**Sz8.3.o *Sz8.3.l
Sz8.4	<i>lahet</i>	Vzorci, ki vsebujejo izpuštljiv polglasnik in v katere se vriva <i>-e-</i> . Leme na <i>-et</i> : <i>lah-et</i> , <i>lah-ti</i> , daj. dv. <i>lah-tema</i> .	**Sz8.4.o *Sz8.4.l
Sz9		Vzorec, podoben pridevniškemu.	
Sz9.1	<i>častita</i>	Vzorec, podoben pridevniškemu: <i>častit-a</i> , <i>častit-e</i> , mn. <i>častit-ih</i> .	Sz9.1.o (<i>častita</i> , 2) *Sz9.1.l

Samostalniški vzorci ženskega spola izkazujejo 3 variante. Razen tega se v leksikonu Sloleks pojavljajo paradigme, ki imajo le oblike za ednino ali množino, kar je pri bodočem urejanju leksikona mogoče ohraniti ali (zlasti pri trenutno edninskih samostalnikih) spremeniti. Variante ali omejitve se pojavljajo pri 37 raznolikih vzorcih, kot prikazuje Tabela 5.

Tabela 5: Variante in omejitve v samostalniških vzorcih ženskega spola.

Koda	Opis variante / omejitve	Vzorci in št. lem v bazi
V1	Rodilnik dvojine in množine: -Ø/-a (npr. <i>vod / voda</i>)	Sz1.1.o-V1 (voda, 9), Sz3.1.o-V1 (cerkev, 1), Sz5.1.o-V1 (sestra, 1), Sz5.3.o-V1 (zemlja, 1), Sz5.4.o-V1 (metla, 2)
V2	Rodilnik dvojine in množine: -ac/-c (npr. <i>ovc / ovac</i>)	Sz5.9.o-V2 (ovca, 1)
V3	Orodnik ednine:-ijo/-jo (npr. <i>rebrjo / rebrijo</i>)	Sz8.1.o-V3 (raven, 2), Sz8.2.o-V3 (ravan, 1), Sz8.3.o-V3 (reber, 2), Sz8.4.o-V3 (lahet, 3)
ednina	Oblike so (trenutno) samo za ednino.	Sz1.1.o-ednina (nafta, 865), Sz1.1.l-ednina (Slovenija, 499), Sz2.1.o-ednina (last, 23), Sz2.1.l-ednina (Podpeč, 5), Sz3.1.l-ednina (Lokev, 1), Sz4.2.o-ednina (kri, 1), Sz6.1.o-ednina (madame, 18), Sz6.1.l-ednina (Karmen, 85)
množina	Oblike so (trenutno) samo za množino.	Sz1.1.o-množina (finance, 58), Sz1.1.l-množina (Jesenice, 187), Sz1.4.l-množina (Golte, 1), Sz2.1.o-množina (obresti, 4), Sz3.1.l-množina (Ponikve, 2), Sz4.3.l-množina (Žiri, 1), Sz5.1.o-množina (citre, 2), Sz5.1.l-množina (Pekre, 1), Sz5.2.l-množina (Bitnje, 4), Sz5.3.o-množina (grablje, 2), Sz5.3.l-množina (Trbovlje, 14), Sz5.4.o-množina (orgle, 3), Sz5.4.l-množina (Murgle, 1), Sz5.6.l-množina (Ravne, 3), z5.8.o-množina (škarje, 2), Sz5.8.l-množina (Nazarje, 10), Sz6.1.o-množina (OI, 1), Sz6.1.l-množina (ZDA, 1), Sz7.2.o-množina (jasli, 4)

4.1.3 Srednji spol

Oblikoslovni vzorci za samostalnike srednjega spola so razdeljeni v 7 skupin, ki prinašajo 24 vzorcev drugega nivoja. Nadalje so deljeni glede na to, ali so občnoimenski ali lastnoimenski. Na tretji ravni je izpričanih 19 vzorcev. Razlikovalne značilnosti za umestitev vzorcev v skupine so v prvi vrsti preglašenost (leme na -e ali -o) ter vrivanje ali izpuščanje vokala v paradigmi. V posebni skupini so umeščeni samostalniki, ki se pregibajo z ničto končnico, ter samostalniki, ki se pregibajo podobno pridevnikom. Urejene vzorce prikazuje Tabela 6.

Tabela 6: Oblikoslovni vzorci za samostalnike srednjega spola.

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Ss1		Osnovni vzorci za preglašene samostalnike.	
Ss1.1	<i>življenje</i>	Osnovna preglašena paradigma: <i>življenj-e</i> , <i>življenj-a</i> .	Ss1.1.o (življenje, 5.742) **Ss1.1.l
Ss2		Osnovni vzorci za nepreglašene samostalnike.	
Ss2.1	<i>delo</i>	Osnovna preglašena paradigma: <i>del-o</i> , <i>del-a</i> .	Ss2.1.o (delo, 422) **Ss2.1.l
Ss2.2	<i>drva</i>	Vzorec za lemo <i>drva</i> , ki ima samo delno paradigmo in posebnosti: <i>drva</i> , <i>drveh</i> .	**Ss2.2.o *Ss2.2.l
Ss2.3	<i>Rova</i>	Vzorec za lemo <i>Rova</i> , ki ima samo delno paradigmo in posebnosti: <i>Rova</i> , <i>Rov</i> .	*Ss2.3.o **Ss2.3.l
Ss3		Vzorci za preglašene samostalnike, kjer se v roditeljski dv. in mn. vriva vokal (-i- ali -e-).	
Ss3.1	<i>podjetje</i>	Vzorci, pri katerih se vriva -i-: <i>podjet-je</i> , <i>podjet-ja</i> , rod. mn. <i>podjet-ij</i> .	Ss3.1.o (podjetje, 564) **Ss3.1.l
Ss3.2	<i>ozemlje</i>	Vzorci, pri katerih se vriva -e-: <i>ozem-lje</i> , <i>ozem-lja</i> , rod. mn. <i>ozem-elj</i> .	Ss3.2.o (ozemlje, 5) *Ss3.2.l
Ss4		Vzorci za nepreglašene samostalnike, kjer se v roditeljski dv. in mn. vriva vokal -e-.	
Ss4.1	<i>ministrstvo</i>	Vzorci, pri katerih je v spremenljivem delu -v-, vriva se -e-: <i>ministrst-vo</i> , <i>ministrst-va</i> , rod. mn. <i>ministrst-ev</i> .	Ss4.1.o (ministrstvo, 443) *Ss4.1.l
Ss4.2	<i>geslo</i>	Vzorci, pri katerih je v spremenljivem delu -l-, vriva se -e-: <i>ges-lo</i> , <i>ges-la</i> , rod. mn. <i>ges-el</i> .	Ss4.2.o (geslo, 22) *Ss4.2.l
Ss4.3	<i>okno</i>	Vzorci, pri katerih je v spremenljivem delu -n-, vriva se -e-: <i>ok-no</i> , <i>ok-na</i> , rod. mn. <i>ok-en</i> .	Ss4.3.o (okno, 10) *Ss4.3.l
Ss4.4	<i>jutro</i>	Vzorci, pri katerih je v spremenljivem delu -r-, vriva se -e-: <i>jut-ro</i> , <i>jut-ra</i> , rod. mn. <i>jut-er</i> .	Ss4.4.o (jutro, 7) *Ss4.4.l
Ss4.5	<i>pismo</i>	Vzorci, pri katerih je v spremenljivem delu -m-, vriva se -e-: <i>pis-mo</i> , <i>pis-ma</i> , rod. mn. <i>pis-em</i> .	Ss4.5.o (pismo, 2) *Ss4.5.l
Ss4.6	<i>tla</i>	Vzorec za lemo <i>tla</i> , ki obstaja samo v množini in izkazuje vrivanje -a: <i>t-la</i> , <i>t-al</i> .	**Ss4.6.o *Ss4.6.l
Ss4.7	<i>dno</i>	Vzorec za lemo <i>dno</i> , ki vsebuje številne posebnosti.	Ss4.7.o (dno, 1) *Ss4.7.l
Ss5		Vzorci za leme na -e, ki vsebujejo podaljševanje osnove s -t-, -n- ali -s-.	
Ss5.1	<i>dekle</i>	Vzorec za leme na -e, ki vsebuje podaljšave s -t-: <i>dekle-Ø</i> , <i>dekle-ta</i> .	Ss5.1.o (dekle, 17) *Ss5.1.o
Ss5.2	<i>ime</i>	Vzorec za leme na -e, ki vsebuje podaljšave z -n-: <i>ime-Ø</i> , <i>ime-na</i> .	Ss5.2.o (ime, 11) *Ss5.2.l
Ss5.3	<i>oje</i>	Vzorec za leme na -e, ki vsebuje podaljšave s -s-: <i>oje-Ø</i> , <i>oje-sa</i> .	Ss5.3.o (oje, 2) *Ss5.3.l

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
Ss6		Vzorci za sklanjanje z ničtimi končnicami.	
Ss6.1	SP	Vzorec za sklanjanje z ničtimi končnicami: SP.	**Ss6.1.o **Ss6.1.l
Ss7		Vzorci za leme na -o, ki imajo podaljšavo z -es-.	
Ss7.1	telo	Osnovni vzorec za leme na -o, ki imajo podaljšavo z -es-: <i>tel-o, tel-esa</i> .	Ss7.1.o (telo, 10) *Ss7.1.l
Ss7.2	uho	Vzorec za leme na -o, ki imajo podaljšavo z -es-, premena h-š: <i>u-ho, u-šesa</i> .	Ss7.2.o (uho, 1) *Ss7.2.l
Ss7.3	oko	Vzorec za leme na -o, ki imajo podaljšavo z -es-, premena k-č: <i>o-ka, o-česa</i> .	Ss7.3.o (oko, 1) *Ss7.3.l
Ss7.4	igo	Vzorec za leme na -o, ki imajo podaljšavo z -es-, premena g-ž: <i>i-go, i-žesa</i> .	Ss7.4.o (igo, 1) *Ss7.4.l
Ss7.5	črevo	Vzorec za leme na -o, ki imajo podaljšavo z -es-, posebnosti v množini: <i>črev-o, črev-esa, mn. črev-a</i> .	Ss7.5.o (črevo, 1) *Ss7.5.l
Ss8		Vzorci, podobni pridevniškim.	
Ss8.1	valentinovo	Vzorec, podoben pridevniškemu, lema na -o: <i>valentinov-o, valentinov-ega</i> .	Ss8.1.o (valentinovo, 9) **Ss8.1.l
Ss8.2	Trebnje	Vzorec, podoben pridevniškemu, lema na -e: <i>Trebnj-e, Trebnj-ega</i> .	*Ss8.2.o **Ss8.2.l

Samostalniški vzorci srednjega spola izkazujejo eno samo varianto. Razen tega se v leksikonu Sloleks pojavljajo paradigme, ki imajo le oblike za ednino ali množino, kar je pri bodočem urejanju leksikona mogoče ohraniti ali spremeniti. Variante ali omejitve se pojavljajo pri 17 raznolikih vzorcih, kot prikazuje Tabela 7.

Tabela 7: Variante in omejitve v samostalniških vzorcih srednjega spola.

Koda	Opis variante / omejitve	Vzorci in št. lem v bazi
V1	Mestnik dvojine in množine: <i>-ih/-eh</i> (npr. <i>sencih / senceh</i>)	Ss1.1.o-V1 (sence, 1)
ednina	Oblike so (trenutno) samo za ednino.	Ss1.1.o-ednina (zdravje, 197), Ss1.1.l-ednina (Celje, 18), Ss2.1.o-ednina (mleko, 469), Ss2.1.l-ednina (Kosovo, 17), Ss6.1.o-ednina (DP, 4), Ss6.1.l-ednina (MID, 3), Ss8.1.l-ednina (Laško, 73), Ss8.2.l-ednina (Trebnje, 5)
množina	Oblike so (trenutno) samo za množino.	Ss1.1.o-množina (vratca, 9), Ss2.1.o-množina (vrata, 17), Ss2.1.l-množina (Selca, 8), Ss2.2.o-množina (drva, 1), Ss2.3.l-množina (Rova, 1), Ss4.6.o-množina (tla, 1), Ss4.4.o-množina (jetra, 2)
Sosei	Oblike so (trenutno) samo za imenovalnik in tožilnik ednine.	Ss6.1.o-Sosei (popoldne, 3)

4.2 Pridevnik

Oblikoslovni vzorci za pridevnik so razdeljeni v 6 skupin, ki prinašajo 37 vzorcev drugega nivoja. Ti so nadalje deljeni glede na vrsto pridevnika, ki je pripisana v leksikonu,¹⁸ in sicer: splošni pridevnik (p), npr. *nov*, deležniški pridevnik (d), npr. *poslan*, ter svojilni pridevnik (s), npr. *človekov*. Ker je pregibanje pridevnikov povezano z njihovo vrsto, se deležniški in svojilni vzorci pojavljajo v omejenem naboru skupin, kot prikazuje Tabela 8. Na tretjem nivoju je realiziranih 38 vzorcev.

Razlikovalne značilnosti pri pridevniških vzorcih so preglašenost (ali se 1. oseba ednine srednjega spola pridevnika konča na *-o* ali *-e*), vsebovanost ne/določnih oblik v paradigmi moškega spola ter značilnosti stopnjevanja (obstoj obrazilnega stopnjevanja in morebitne premene). Posebej so vzorci, kjer v paradigmi prihaja do izpuščanja vokala (*-e-*, *-a-*), in vzorec za pregibanje z ničto končnico.

¹⁸ Teh kategorij med pripravo vzorcev nismo posebej preverjali in urejali, mogoče pa je to zagotoviti pri nadaljnjem razvoju leksikona. Za kategorizacijo je nekoliko zahtevnejša predvsem deležniška skupina.

Tabela 8: Oblikoslovnji vzorci za pridevnike.

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
P1		Osnovna nepreglašena paradigma (srednji spol na -o), ki lahko izkazuje določnost ter obrazilno stopnjevanje ali pa ne.	
P1.1	<i>človekov</i>	Nepreglašena paradigma, samo nedoločne oblike in nestopnjevano: <i>človekov-Ø</i> , <i>človekov-ega</i> .	P1.1.p (sam, 4) *P1.1.d P1.1.s (človekov, 8.185)
P1.2	<i>slovenski</i>	Nepreglašena paradigma, samo določne oblike in nestopnjevano: <i>slovensk-i</i> , <i>slovensk-ega</i> .	P1.2.p (slovenski, 5.152) *P1.2.d, *P1.2.s
P1.3	<i>znan</i>	Nepreglašena paradigma, nedoločne ter določne oblike in nestopnjevano: <i>znan-Ø/znan-i</i> , <i>znan-ega</i> .	P1.3.p (bel, 2.132) P1.3.d (znan, 3.358) *P1.3.s
P1.4	<i>nov</i>	Nepreglašena paradigma, nedoločne ter določne oblike in stopnjevano na -ejši: <i>nov-Ø/nov-i</i> , <i>nov-ega</i> ; <i>nov-ejši</i> .	P1.4.p (nov, 149) *P1.4.d, *P1.4.s
P1.5	<i>lep</i>	Nepreglašena paradigma, nedoločne ter določne oblike in stopnjevano na -ši: <i>lep-Ø/lep-i</i> , <i>lep-ega</i> ; <i>lep-ši</i> .	P1.5.p (lep, 5) *P1.5.d, *P1.5.s
P2		Osnovna preglašena paradigma (srednji spol na -e), ki lahko izkazuje določnost ter obrazilno stopnjevanje ali pa ne.	
P2.1	<i>zadnji</i>	Preglašena paradigma, samo določne oblike in nestopnjevano: <i>zadnj-i</i> , <i>zadnj-ega</i> .	P2.1.p (zadnji, 184) *P2.1.d, *P2.1.s
P2.2	<i>tekoč</i>	Preglašena paradigma, nedoločne ter določne oblike in nestopnjevano: <i>tekoč-Ø/tekoč-i</i> , <i>tekoč-ega</i> .	P2.2.p (mogoč, 54) P2.2.d (tekoč, 759) *P2.2.s
P3		Osnovna paradigma z izpuščenim polglasnikom (<i>držav-(e)n-ega</i>), ki lahko izkazuje določnost ter obrazilno stopnjevanje ali pa ne.	
P3.1	<i>državen</i>	Pridevniki na -en/-ni, nedoločne ter določne oblike in nestopnjevano: <i>držav-en/držav-ni</i> , <i>držav-nega</i> .	P3.1.p (državen, 5.570) *P3.1.d, *P3.1.s
P3.2	<i>pomemben</i>	Pridevniki na -en/-ni, nedoločne ter določne oblike in stopnjevano: <i>pomemb-en/pomemb-ni</i> , <i>pomemb-nega</i> ; <i>pomemb-nejši</i> .	P3.2.p (pomemben, 591) *P3.2.d, *P3.2.s
P3.3	<i>dobrodošel</i>	Pridevniki na -el/-li, nedoločne ter določne oblike in nestopnjevano: <i>dobrodoš-el/dobrodoš-li</i> , <i>dobrodoš-lega</i> .	P3.3.p (dobrodošel, 31) P3.3.d (pretek, 72) *P3.3.s
P3.4	<i>topel</i>	Pridevniki na -el/-li, nedoločne ter določne oblike in stopnjevano: <i>top-el/top-li</i> , <i>top-lega</i> ; <i>top-lejši</i> .	P3.4.p (topel, 5) *P3.4.d, *P3.4.s
P3.5	<i>plehek</i>	Pridevniki na -ek/-ki, nedoločne ter določne oblike in nestopnjevano: <i>pleh-ek/pleh-ki</i> , <i>pleh-kega</i> .	P3.5.p (grenek, 30) *P3.5.d, *P3.5.s
P3.6	<i>redek</i>	Pridevniki na -ek/-ki, nedoločne ter določne oblike in stopnjevano: <i>red-ek/red-ki</i> , <i>red-kega</i> ; <i>red-kejši</i> .	P3.6.p (red, 9) *P3.6.d, *P3.6.s

P3.7	<i>jeder</i>	Pridevniki na <i>-er/-ri</i> , nedoločne ter določne oblike in nestopnjevano: <i>jed-er/jed-ri, jed-rega</i> .	P3.7.p (jeder, 15) *P3.7.d, *P3.7.s
P3.8	<i>hiter</i>	Pridevniki na <i>-er/-ri</i> , nedoločne ter določne oblike in stopnjevano: <i>hit-er/hit-ri, hit-rega; hit-rejši</i>	P3.8.p (hiter, 8) *P3.8.d, *P3.8.s
P3.9	<i>mrtev</i>	Pridevniki na <i>-ev/-vi</i> , nedoločne ter določne oblike in nestopnjevano: <i>mrt-ev/mrt-vi, mrt-vega</i> .	P3.9.p (mrtev, 2) *P3.9.d, *P3.9.s
P3.10	<i>plitev</i>	Pridevniki na <i>-ev/-vi</i> , nedoločne ter določne oblike in stopnjevano: <i>plit-ev/plit-vi, plit-vega</i> .	P3.10.p (plitev, 1) *P3.10.d, *P3.10.s
P3.11	<i>bolan</i>	Pridevniki na <i>-an/-ni</i> , nedoločne ter določne oblike in nestopnjevano: <i>bol-an/bol-ni, bol-nega</i> .	P3.11.p (bolan, 1) *P3.11.d, *P3.11.s
P4		Pridevniki, ki se obrazilno stopnjujejo in pri tem pride do premene v <i>-ž(-)ji, -šji, -čji, -lji</i> . V tej skupini so predvsem posebnosti.	
P4.1	<i>drag</i>	Pridevniki na <i>-g/-gi</i> , nedoločne ter določne oblike in stopnjevano z <i>-žji</i> : <i>dra-g/dra-gi, dra-gega; dra-žji</i> .	P4.1.p (drag, 4) *P4.1.d, *P4.1.s
P4.2	<i>nizek</i>	Pridevniki na <i>-zek/-zki</i> , nedoločne ter določne oblike in stopnjevano z <i>-žji</i> : <i>ni-zek/ni-zki, niz-kega; ni-žji</i> .	P4.2.p (nizek, 3) *P4.2.d, *P4.2.s
P4.3	<i>lahek</i>	Pridevniki na <i>-hek/-hki</i> , nedoločne ter določne oblike in stopnjevano z <i>-žji</i> : <i>la-hek/lah-hki, la-hkega; la-žji</i> .	**P4.3.p *P4.3.d, *P4.3.s
P4.4	<i>težek</i>	Pridevniki na <i>-ek/-ki</i> , nedoločne ter določne oblike in stopnjevano z <i>-ji</i> : <i>tež-ek/tež-ki, tež-jega; tež-ji</i> .	**P4.4.p *P4.4.d, *P4.4.s
P4.5	<i>tih</i>	Pridevniki na <i>-h/-hi</i> , nedoločne ter določne oblike in stopnjevano s <i>-šji</i> : <i>ti-h/ti-hi, ti-hega; ti-šji</i> .	P4.5.p (tih, 1) *P4.5.d, *P4.5.s
P4.6	<i>visok</i>	Pridevniki na <i>-sok/-soki</i> , nedoločne ter določne oblike in stopnjevano s <i>-šji</i> : <i>vi-sok/vi-soki, vi-sokega; vi-šji</i> .	P4.6.p (visok, 1) *P4.6.d, *P4.6.s
P4.7	<i>velik</i>	Pridevniki na <i>-lik/-liki</i> , nedoločne ter določne oblike in stopnjevano s <i>-čji</i> : <i>ve-lik/ve-liki, vel-likega; ve-čji</i> .	P4.7.p (velik, 1) *P4.7.d, *P4.7.s
P4.8	<i>globok</i>	Pridevniki na <i>-ok/-oki</i> , nedoločne ter določne oblike in stopnjevano z <i>-lji</i> : <i>glob-ok/glob-oki, glob-okega; glob-lji</i> .	P4.8.p (globok, 1) *P4.8.d, *P4.8.s
P5		Pridevniki, ki se obrazilno stopnjujejo in pri tem pride do premene v <i>-(j)ši</i> . V določenih primerih je lom tudi prej, npr. pri <i>-njši</i> , ali pa je stopnjevana oblika povsem drugačna od osnove (<i>dober, boljši</i>). V tej skupini so predvsem posebnosti.	
P5.1	<i>širok</i>	Pridevniki na <i>-ok/-oki</i> , nedoločne ter določne oblike in stopnjevano s <i>-ši</i> : <i>šir-ok/šir-oki, šir-okega; šir-ši</i> .	P5.1.p (širok, 1) *P5.1.d, *P5.1.s
P5.2	<i>trd</i>	Pridevniki na <i>-d/-di</i> , nedoločne ter določne oblike in stopnjevano s <i>-ši</i> : <i>tr-d/tr-di, tr-dega; tr-ši</i> .	P5.2.p (trd, 2) *P5.2.d, *P5.2.s
P5.3	<i>mlad</i>	Pridevniki na <i>-d/-di</i> , nedoločne ter določne oblike in stopnjevano z <i>-jši</i> : <i>mła-d/mła-di, mla-dega; mla-jši</i> .	P5.3.p (mlad, 2) *P5.3.d, *P5.3.s
P5.4	<i>sladek</i>	Pridevniki na <i>-dek/-dki</i> , nedoločne ter določne oblike in stopnjevano z <i>-jši</i> : <i>sla-dek/sla-dki, sla-dkega; sla-jši</i> .	**P5.4.p *P5.4.d, *P5.4.s

P5.5	<i>tanek</i>	Pridevniki na <i>-ek/-ki</i> , nedoločne ter določne oblike in stopnjevano z <i>-jši</i> : <i>tan-ek/tan-ki, tan-kega; tan-jši</i> .	P5.5.p (tanek, 1) *P5.5.d, *P5.5.s
P5.6	<i>kratek</i>	Pridevniki na <i>-tek/-tki</i> , nedoločne ter določne oblike in stopnjevano z <i>-jši</i> : <i>kra-tek/kra-tki, kra-tkega; kra-jši</i> .	P5.6.p (kratek, 1) *P5.6.d, *P5.6.s
P5.7	<i>majhen</i>	Pridevniki na <i>-jhen/-jhni</i> , nedoločne ter določne oblike in stopnjevano z <i>-njši</i> : <i>ma-jhen/ma-jhni, ma-jhnega; ma-njši</i> .	P5.7.p (majhen, 1) *P5.7.d, *P5.7.s
P5.8	<i>dolg</i>	Pridevniki na <i>-olg/-olgi</i> , nedoločne ter določne oblike in stopnjevano z <i>-aljši</i> : <i>d-olg/d-olgi, d-olgega; d-aljši</i> .	P5.8.p (dolg, 2) *P5.8.d, *P5.8.s
P5.9	<i>dober</i>	Popolna razlika med osnovno in stopnjevano obliko: <i>dober, boljši</i> .	P5.9.p (dober, 1) *P5.9.d, *P5.9.s
P5.10	<i>poceni</i>	Popolna razlika med osnovno in stopnjevano obliko: <i>poceni, cenejši</i> .	P5.10.p (poceni, 1) *P5.10.d, *P5.10.s
P6		Pridevnik se ne pregiba oz. se pregiba z ničto končnico.	
P6.1	<i>super</i>	Pridevnik se ne pregiba oz. se pregiba z ničto končnico. Določne in nedoločne oblike, brez stopnjevanja: <i>super-Ø</i> .	P6.1.p (super, 42) *P6.1.d, *P6.1.s

Pridevniški vzorci izkazujejo eno samo varianto, ki se pojavlja pri 9 vzorcih, kot kaže Tabela 9.

Tabela 9: Variante in omejitve v pridevniških vzorcih.

Koda	Opis variante / omejitve	Vzorci in št. lem v bazi
v1	Imenovalnik in tožilnik ednine moškega spola (nedoločna oblika): <i>-en/-an</i> (npr. <i>močen / močan</i>).	P3.1.p-V1 (cveten, 12), P3.2.p-V1 (hladen, 11), P3.3.p-V1 (presvetel, 1), P3.4.p-V1 (svetel, 1), P3.5.p-V1 (plehek, 11), P3.6.p-V1 (šibek, 5), P4.3.p-V1 (lahek, 2), P4.4.p-V1 (težek, 1), P5.4.p-V1 (sladek, 1)

4.3 Glagol

Oblikoslovni vzorci za glagol so razdeljeni v 10 skupin, ki prinašajo 72 vzorcev drugega nivoja. Ti so nadalje deljeni glede na dovršnost glagola, ki je pripisana v leksikonu,¹⁹ in sicer: dovršni glagol (d), npr. *odkriti*, nedovršni glagol (n), npr. *govoriti*, ter dvovidski glagol (v), npr. *pomagati*. Na tretjem nivoju je realiziranih 153 vzorcev, kot prikazuje Tabela 10.

¹⁹ Teh kategorij med pripravo vzorcev nismo posebej preverjali in urejali, mogoče pa je to zagotoviti pri nadaljnjem razvoju leksikona. Za kategorizacijo je nekoliko zahtevnejša predsem skupina dvovidskih glagolov.

Razlikovalne značilnosti za umestitev glagolskih vzorcev v skupine so v prvi vrsti mesto, kjer se lomi nedoločnik (npr. pri *-ti*, *-iti*, *-eti*, *-ati*), vrivanje vokalov in premene, skupaj z drugimi kazalci, med katerimi je običajno 1. oseba sedanjika ednine, včasih tudi velelniške ali deležniške oblike.

Tabela 10: Oblikoslovni vzorci za glagole.

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
G1		Lom spremenljivega dela oblik je v nedoločniku pri <i>-ti</i> . V nekaterih primerih je lahko lom tudi pri <i>-dati</i> in <i>-jti</i> .	
G1.1	<i>igrati</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-m</i> in velelniku <i>-j</i> : <i>igra-ti</i> , <i>igra-m</i> , <i>igra-j</i> .	G1.1.d (končati, 1.135) G1.1.n (igrati, 1.935) G1.1.v (pomagati, 1.345)
G1.2	<i>govoriti</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-m</i> in velelniku $-\emptyset$: <i>govori-ti</i> , <i>govori-m</i> , <i>govori-\emptyset</i> .	G1.2.d (odločiti, 1.832) G1.2.n (govoriti, 362) G1.2.v (praviti, 358)
G1.3	<i>šteti</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-jem</i> : <i>šte-ti</i> , <i>šte-jem</i> .	G1.3.d (odkriti, 129) G1.3.n (šteti, 11) G1.3.v (vpiti, 12)
G1.4	<i>tresti</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-em</i> : <i>tres-ti</i> , <i>tres-em</i> .	G1.4.d (prenesti, 28) G1.4.n (tresti, 2) G1.4.v (doprinesiti, 1)
G1.5	<i>ostati</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-nem</i> : <i>osta-ti</i> , <i>osta-nem</i> .	G1.5.d (ostati, 21) G1.5.n (stati, 1) G1.5.v (vzdeti, 1)
G1.6	<i>pleti</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-vem</i> : <i>ple-ti</i> , <i>ple-vem</i> .	G1.6.d (opleti, 3) G1.6.n (pleti, 1) *G1.6.v
G1.7	<i>gledati</i>	V nedoločniku je <i>-dati</i> , v sedanjiku <i>-dam</i> : <i>gle-dati</i> , <i>gle-dam</i> .	G1.7.d (pogledati, 9) G1.7.n (gledati, 1) G1.7.v (nagledati, 1)
G1.8	<i>napovedati</i>	V nedoločniku je <i>-dati</i> , v sedanjiku <i>-m</i> : <i>napove-dati</i> , <i>napove-m</i> .	G1.8.d (napovedati, 6) *G1.8.n, *G1.8.v
G1.9	<i>najti</i>	V nedoločniku je <i>-jti</i> , v sedanjiku <i>-jdem</i> : <i>na-jti</i> , <i>na-jdem</i> (lom zaradi deležnikov: <i>na-šel</i>).	G1.9.d (najti, 3) *G1.9.n, *G1.9.v
G1.10	<i>priiti</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-dem</i> : <i>pri-ti</i> , <i>pri-dem</i> (lom zaradi deležnikov: <i>pri-šel</i>).	G1.10.d (priiti, 1) *G1.10.n, *G1.10.v
G1.11	<i>dati</i>	V nedoločniku je <i>-ti</i> , v sedanjiku <i>-m</i> ; posebnost zaradi oblik <i>-ste</i> , <i>-sta</i> : <i>da-m</i> , <i>da-ste</i> .	**G1.11.d *G1.11.n, *G1.11.v
G2		Lom spremenljivega dela oblik je v nedoločniku pri <i>-ovati</i> in <i>-evati</i> . Sedanjik 1. os. ednine na <i>-ujem</i> .	

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
G2.1	<i>pričakovati</i>	V nedoločniku je <i>-ovati</i> , v sedanjiku <i>-ujem: del-ovati, del-ujem</i> .	G2.1.d (poškodovati, 42) G2.1.n (pričakovati, 304) G2.1.v (imenovati, 81)
G2.2	<i>nadaljevati</i>	V nedoločniku je <i>-evati</i> , v sedanjiku <i>-ujem: boj-evati, boj-ujem</i> .	G2.2.d (privarčevati, 5) G2.2.n (nadaljevati, 362) G2.2.v (oglaševati, 45)
G3		Lom spremenljivega dela oblik je v nedoločniku pri <i>-iti</i> . V nekaterih primerih je lom tudi pri <i>-niti</i> .	
G3.1	<i>riniti</i>	V nedoločniku je <i>-iti</i> , v sedanjiku <i>-em: rin-iti, rin-em</i> .	G3.1.d (vriniti, 376) G3.1.n (riniti, 6) G3.1.v (utegniti, 88)
G3.2	<i>dobiti</i>	V nedoločniku je <i>-iti</i> , v sedanjiku <i>-im: dob-iti, dob-im</i> .	G3.2.d (spomniti, 5) **G3.2.n, **G3.2.v
G3.3	<i>oditi</i>	V nedoločniku je <i>-iti</i> , v sedanjiku <i>-idem: od-iti, od-idem</i> (lom zaradi deležnikov: <i>od-šel</i>).	G3.3.d (oditi, 9) *G3.3.n, *G3.3.v
G3.4	<i>iti</i>	V nedoločniku je <i>iti</i> , v sedanjiku <i>grem</i> .	*G3.4.d, *G3.4.n G3.4.v (iti, 1)
G3.5	<i>sniti</i>	V nedoločniku je <i>-niti</i> , v sedanjiku <i>-nidem: s-niti, s-nidem</i> (lom zaradi deležnikov: <i>s-ešel</i>).	*G3.5.d, *G3.5.n G3.5.v (sniti, 1)
G4		Lom spremenljivega dela oblik je v nedoločniku pri <i>-eti</i> . V nekaterih primerih je lahko lom tudi pri <i>-deti</i> ali <i>-jeti</i> .	
G4.1	<i>sedeti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-im: sed-eti, sed-im</i> .	G4.1.d (naleteti, 193) G4.1.n (sedeti, 97) G4.1.v (videti, 68)
G4.2	<i>zreti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-em</i> , v deležniških oblikah <i>-e-</i> izpade: <i>zr-eti, zr-em, zr-l</i> .	G4.2.d (upreti, 27) G4.2.n (zreti, 3) G4.2.v (ucvreti, 5)
G4.3	<i>uspeti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-em</i> , v deležniških oblikah <i>-e-</i> ostane: <i>usp-eti, usp-em, usp-el</i> .	G4.3.d (uspeti, 8) G4.3.n (vreti, 2) G4.3.v (razumeti, 3)
G4.4	<i>verjeti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-amem: verj-eti, verj-amem</i> .	G4.4.d (prevzeti, 20) *G4.4.n G4.4.v (verjeti, 1)
G4.5	<i>pometi</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-anem: pom-eti, pom-anem</i> .	G4.5.d (pometi, 3) G4.5.n (meti, 1) *G4.5.v
G4.6	<i>požeti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-anjem: pož-eti, pož-anjem</i> .	G4.6.d (nažeti, 2) *G4.6.n G4.6.v (obžeti, 1)
G4.7	<i>začeti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-nem: zač-eti, zač-nem</i> .	G4.7.d (povzpeti, 14) *G4.7.n, **G4.7.v
G4.8	<i>sprejeti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-mem: sprej-eti, sprej-mem</i> .	G4.8.d (sprejeti, 3) *G4.8.n, *G4.8.v

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
G4.9	<i>prijeti</i>	V nedoločniku je <i>-jeti</i> , v sedanjiku <i>-mem: pri-jeti, pri-mem</i> .	G4.9.d (prijeti, 6) *G4.9.n, *G4.9.v
G4.10	<i>vedeti</i>	V nedoločniku je <i>-deti</i> , v sedanjiku <i>-m: ve-deti, ve-m</i> .	**G4.10.d, **G4.10.n G4.10.v (ovedeti, 1)
G4.11	<i>peti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-ojem: p-eti, p-ojem</i> .	G4.11.d (prepeti, 2) **G4.11.n, *G4.11.v
G4.12	<i>umreti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-em: umr-eti, umr-jem</i> . Pojavlja se kot variantna paradigma k <i>umreti, umrem</i> .	G4.12.d (zamreti, 5) G4.12.n (mreti, 1) *G4.12.v
G4.13	<i>doumeti</i>	V nedoločniku je <i>-eti</i> , v sedanjiku <i>-ejem: doum-eti, doum-ejem</i> . Pojavlja se kot variantna paradigma k <i>doumeti, doumem</i> .	G4.13.d (doumeti, 1) *G4.13.n, *G4.13.v
G5		Lom spremenljivega dela oblik je v nedoločniku pri <i>-ati</i> . V nekaterih primerih je lom pri <i>-jati</i> .	
G5.1	<i>spati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-im, velebnik -i: sp-ati, sp-im, sp-i</i> .	G5.1.d (zaslišati, 31) G5.1.n (spati, 23) G5.1.v (slišati, 28)
G5.2	<i>dremati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-am: drem-ati, drem-am</i> . Pojavlja se kot variantna paradigma k <i>dremati, dremljem</i> .	G5.2.d (dokopati, 22) G5.2.n (dremati, 36) G5.2.v (pokopati, 17)
G5.3	<i>peljati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-em, velebnik -i: pelj-ati, pelj-em, pelj-i</i> .	G5.3.d (pripeljati, 23) G5.3.n (peljati, 3) G5.3.v (užgati, 3)
G5.4	<i>sejati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-em, velebnik -∅: sej-ati, sej-em, sej-∅</i> .	G5.4.d (posejati, 10) G5.4.n (sejati, 2) G5.4.v (odsmejati, 3)
G5.5	<i>majati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-em, velebnik -aj: maj-ati, maj-em, maj-aj</i> . Pojavlja se kot variantna paradigma k <i>majati, majam</i> .	G5.5.d (zamajati, 6) G5.5.n (majati, 5) G5.5.v (primajati, 1)
G5.6	<i>orati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-jem: or-ati, or-jem</i> .	G5.6.d (zaorati, 7) G5.6.n (orati, 1) G5.6.v (poorati, 1)
G5.7	<i>jemati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-ljem: jem-ati, jem-ljem</i> .	G5.7.d (pozobati, 3) G5.7.n (jemati, 2) G5.7.v (preklepati, 1)
G5.8	<i>smejati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-im, velebnik -∅: smej-ati, smej-im, smej-∅</i> . Pojavlja se kot variantna paradigma k <i>smejati, smejem</i> .	G5.8.d (nasmejati, 2) G5.8.n (smejati, 1) G5.8.v (posmejati, 2)
G5.9	<i>bati</i>	V nedoločniku je <i>-ati</i> , v sedanjiku <i>-ojim: b-ati, b-ojim</i> .	G5.9.d (zbati, 4) **G5.9.n, *G5.9.v
G5.10	<i>prizadejati</i>	V nedoločniku je <i>-jati</i> , v sedanjiku <i>-m: prizade-jati, prizade-m</i> . Pojavlja se kot variantna paradigma k <i>prizadejati, prizadejam</i> .	G5.10.d (prizadejati, 3) *G5.10.n, *G5.10.v

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
G6		Lom spremenljivega dela oblik je v nedoločniku pri <i>-uti</i> . Sedanjik 1. os. ednine na <i>-ovem</i> .	
G6.1	<i>rjuti</i>	V nedoločniku je <i>-uti</i> , v sedanjiku <i>-ovem</i> : <i>rj-uti</i> , <i>rj-ovem</i> . Pojavlja se kot variantna paradigma k <i>rjuti</i> , <i>rjujem</i> .	G6.1.d (zarjuti, 1) G6.1.n (rjuti, 1) *G6.1.v
G7		Lom spremenljivega dela oblik je v nedoločniku pri <i>-sti</i> .	
G7.1	<i>krasti</i>	V nedoločniku je <i>-sti</i> , v sedanjiku <i>-dem</i> : <i>kra-sti</i> , <i>kra-dem</i> .	G7.1.d (navesti, 60) G7.1.n (krasti, 8) G7.1.v (prisesti, 9)
G7.2	<i>zlesti</i>	V nedoločniku je <i>-sti</i> , v sedanjiku <i>-zem</i> : <i>zle-sti</i> , <i>zle-zem</i> .	G7.2.d (zlesti, 21) G7.2.n (lesti, 5) G7.2.v (odlesti, 2)
G7.3	<i>tepsti</i>	V nedoločniku je <i>-sti</i> , v sedanjiku <i>-em</i> : <i>tep-sti</i> , <i>tep-em</i> .	G7.3.d (pretepsti, 20) G7.3.n (tepsti, 6) G7.3.v (natepsti, 8)
G7.4	<i>plesti</i>	V nedoločniku je <i>-sti</i> , v sedanjiku <i>-tem</i> : <i>ple-sti</i> , <i>ple-tem</i> .	G7.4.d (zapesti, 18) G7.4.n (plesti, 3) G7.4.v (izplesti, 2)
G7.5	<i>jesti</i>	V nedoločniku je <i>-sti</i> , v sedanjiku <i>-m</i> : <i>je-sti</i> , <i>je-m</i> .	G7.5.d (objesti, 2) **G7.5.n G7.5.v (odcvesti, 1)
G8		Lom spremenljivega dela oblik je v nedoločniku pri <i>-či</i> . V nekaterih primerih je lahko lom tudi pri <i>-eči</i> .	
G8.1	<i>vleči</i>	V nedoločniku je <i>-či</i> , v sedanjiku <i>-čem</i> : <i>vle-či</i> , <i>vle-čem</i> .	G8.1.d (izreči, 46) G8.1.n (vleči, 3) G8.1.v (odtolči, 2)
G8.2	<i>doseči</i>	V nedoločniku je <i>-či</i> , v sedanjiku <i>-žem</i> : <i>dose-či</i> , <i>dose-žem</i> .	G8.2.d (doseči, 33) G8.2.n (streči, 2) G8.1.v (zaleči, 7)
G8.3	<i>premoči</i>	V nedoločniku je <i>-či</i> , v sedanjiku <i>-rem</i> : <i>premo-či</i> , <i>premo-rem</i> .	**G8.3.d *G8.3.n, **G8.3.v
G8.4	<i>odvreči</i>	V nedoločniku je <i>-eči</i> , v sedanjiku <i>-žem</i> : <i>odvr-eči</i> , <i>odvr-žem</i> .	G8.4.d (odvreči, 10) *G8.4.n, *G8.4.v
G9		Spremenljivi del oblik po premenah (palatalizacija, jotacija) vključuje konzonante: nedoločnik na <i>-zati</i> , <i>-sati</i> , <i>-cati</i> , <i>-kati</i> , <i>-hati</i> , <i>-gati</i> , <i>-tati</i> , <i>-vati</i> .	
G9.1	<i>pisati</i>	V nedoločniku je <i>-sati</i> , v sedanjiku <i>-šem</i> : <i>pi-sati</i> , <i>pi-šem</i> .	G9.1.d (napisati, 56) G9.1.n (pisati, 8) G9.1.v (pasati, 6)
G9.2	<i>kazati</i>	V nedoločniku je <i>-zati</i> , v sedanjiku <i>-žem</i> : <i>ka-zati</i> , <i>ka-žem</i> .	G9.2.d (pokazati, 49) G9.2.n (kazati, 5) G9.2.v (nalizati, 1)

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
G9.3	<i>klicati</i>	V nedoločniku je <i>-cati</i> , v sedanjiku <i>-čem</i> : <i>kli-cati</i> , <i>kli-čem</i> .	G9.3.d (poklicati, 10) G9.3.n (klicati, 1) G9.3.v (doklicati, 1)
G9.4	<i>priskakati</i>	V nedoločniku je <i>-kati</i> , v sedanjiku <i>-čem</i> : <i>priska-kati</i> , <i>priska-čem</i> .	G9.4.d (priskakati, 1) G9.4.n (pritikati, 1) G9.4.v (preskakati, 1)
G9.5	<i>lagati</i>	V nedoločniku je <i>-gati</i> , v sedanjiku <i>-žem</i> : <i>la-gati</i> , <i>la-žem</i> .	G9.5.d (nalagati, 2) G9.5.n (lagati, 1) *G9.5.v
G9.6	<i>jahati</i>	V nedoločniku je <i>-hati</i> , v sedanjiku <i>-šem</i> : <i>ja-hati</i> , <i>ja-šem</i> . Pojavlja se kot variantna paradigma k <i>jahati</i> , <i>jaham</i> .	G9.6.d (zavijati, 12) G9.6.n (jahati, 2) G9.6.v (prijahati, 2)
G9.7	<i>metati</i>	V nedoločniku je <i>-tati</i> , v sedanjiku <i>-čem</i> : <i>me-tati</i> , <i>me-čem</i> .	G9.7.d (zmetati, 9) G9.7.n (metati, 1) *G9.7.v
G9.8	<i>zahoteti</i>	V nedoločniku je <i>-teti</i> , v sedanjiku <i>-čem</i> : <i>zaho-teti</i> , <i>zaho-čem</i> .	G9.8.d (zahoteti, 1) *G9.8.n, *G9.8.v
G9.9	<i>ugnati</i>	V nedoločniku je <i>-gnati</i> , v sedanjiku <i>-ženem</i> : <i>u-gnati</i> , <i>u-ženem</i> .	G9.9.d (ugnati, 12) G9.9.n (gnati, 1) *G9.9.v
G9.10	<i>iskati</i>	V nedoločniku je <i>-skati</i> , v sedanjiku <i>-ščem</i> : <i>i-skati</i> , <i>i-ščem</i> .	G9.10.d (obiskati, 4) G9.10.n (iskati, 1) *G9.10.v
G9.11	<i>razposlati</i>	V nedoločniku je <i>-slati</i> , v sedanjiku <i>-šljem</i> : <i>razpo-slati</i> , <i>razpo-šljem</i> .	G9.11.d (razposlati, 3) *G9.11.n, *G9.11.v
G9.12	<i>razruvati</i>	V nedoločniku je <i>-vati</i> , v sedanjiku <i>-jem</i> : <i>ru-vati</i> , <i>ru-jem</i> . Pojavlja se kot variantna paradigma k <i>ruvati</i> , <i>ruvam</i> .	G9.12.d (razruvati, 2) *G9.12.n, *G9.12.v
G10		Spremenljivi del oblik v nedoločniku vključuje zvočnike <i>r</i> , <i>l</i> in <i>v</i> , pred katere se v 1. os. ednine vriva vokal, npr. <i>-rati/-erem</i> .	
G10.1	<i>prati</i>	V nedoločniku je <i>-rati</i> , v sedanjiku <i>-erem</i> : <i>p-rati</i> , <i>p-erem</i> .	G10.1.d (izbrati, 14) G10.1.n (prati, 1) G10.1.v (odprati, 2)
G10.2	<i>sрати</i>	V nedoločniku je <i>-rati</i> , v sedanjiku <i>-erjem</i> : <i>s-rati</i> , <i>s-erjem</i> .	G10.2.d (posrati, 6) G10.2.n (srati, 1) *G10.2.v
G10.3	<i>predreti</i>	V nedoločniku je <i>-reti</i> , v sedanjiku <i>-erem</i> : <i>pred-reti</i> , <i>pred-erem</i> . Pojavlja se kot variantna paradigma k <i>predreti</i> , <i>predrem</i> .	G10.3.d (predreti, 6) G10.3.n (dreti, 1) *G10.3.v
G10.4	<i>zatreти</i>	V nedoločniku je <i>-reti</i> , v sedanjiku <i>-arem</i> : <i>zat-reti</i> , <i>zat-arem</i> . Pojavlja se kot variantna paradigma k <i>zatreти</i> , <i>zatrem</i> .	G10.4.d (zatreти, 6) G10.4.n (treti, 1) *G10.4.v

Koda	Tipski primer	Opis	Vzorci in št. lem v bazi
G10.5	<i>klati</i>	V nedoločniku je <i>-lati</i> , v sedanjiku <i>-oljem</i> : <i>k-lati</i> , <i>k-oljem</i> .	G10.5.d (zaklati, 8) G10.5.n (klati, 1) G10.5.v (preplati, 2)
G10.6	<i>postlati</i>	V nedoločniku je <i>-lati</i> , v sedanjiku <i>-eljem</i> : <i>post-lati</i> , <i>post-eljem</i> .	G10.6.d (postlati, 3) *G10.6.n, *G10.6.v
G10.7	<i>mleti</i>	V nedoločniku je <i>-leti</i> , v sedanjiku <i>-eljem</i> : <i>m-leti</i> , <i>m-eljem</i> .	G10.7.d (zmleti, 3) G10.7.n (mleti, 1) G10.7.v (pomleti, 1)
G10.8	<i>kleti</i>	V nedoločniku je <i>-leti</i> , v sedanjiku <i>-olnem</i> : <i>k-leti</i> , <i>k-olnem</i> .	G10.8.d (prekleti, 4) G10.8.n (kleti, 1) *G10.8.v
G10.9	<i>odzvati</i>	V nedoločniku je <i>-vati</i> , v sedanjiku <i>-ovem</i> : <i>odz-vati</i> , <i>odz-ovem</i> .	G10.9.d (odzvati, 4) *G10.9.n, *G10.9.v

Glagolski vzorci izkazujejo 5 variant, ki se pojavljajo (tudi v kombinacijah) pri 38 raznolikih vzorcih. Razen tega se v leksikonu Sloleks pojavljajo paradigme, ki ne izkazujejo veledniških oblik, kar je pri bodočem urejanju leksikona mogoče ohraniti ali spremeniti. Nabor vzorcev z varianto ali omejitvijo prikazuje Tabela 11.

Tabela 11: Variante in omejitve v glagolskih vzorcih.

Koda	Opis variante / omejitve	Vzorci in št. lem v bazi
V1	Atematske končnice v 2. in 3. osebi dvojine ter 3. osebi množine (npr. <i>dodate / dodaste</i>).	G1.1.d-V1 (dodati, 11), G1.1.v-V1 (podati, 3), G1.1.d-V1+V2 (pridati, 1), G1.1.n-V1+V2 (zadati, 1)
V2	Tretja oseba množine: <i>-do/-jo</i> (npr. <i>gredo / grejo</i>).	G1.1.d-V1+V2 (pridati, 1), G1.1.n-V1+V2 (zadati, 1), G4.10.d-V2 (zvedeti, 6), G4.10.n-V2 (vedeti, 1), G7.5.d-V2 (pojesti, 14), G7.5.n-V2 (jesti, 1), G7.5.v-V2 (ovesti, 1)
V3	Tretja oseba množine: <i>-ijo/-e</i> (npr. <i>dobijo / dobe</i>).	G3.2.d-V3 (dobiti, 52), G3.2.n-V3 (slediti, 31), G3.2.v-V3 (dolžiti, 3), G4.1.d-V3 (poskrbeti, 21), G4.1.n-V3 (živeti, 30), G4.1.v-V3 (zagomazeti, 3), G4.1.n-velelnik+V3 (zdeti, 1), G5.1.d-V3 (obdržati, 6), G5.1.n-V3 (držati, 7), G5.1.v-V3 (pribežati, 1)

Koda	Opis variante / omejitve	Vzorci in št. lem v bazi
V4	Tretja oseba množine: -ejo/-o (npr. <i>nesejo / neso</i> , tudi z drugimi konzonanti v končnici).	G1.3.n-V4 (gniti, 1), G1.4.n-V4 (nesti, 1), G1.4.d-V4 (prinesti, 1), G4.2.d-V4 (odpreti, 6), G4.7.d-V4 (začeti, 4), G4.7.v-V4 (početi, 1), G4.7.n-V4 (peti, 1), G4.11.n-V4 (peti, 1), G8.1.d-V4 (poreči, 2), G8.1.v-V4 (obteči, 2), G8.1.n-V4 (teči, 2), G10.1.n-V4 (brati, 1)
V5	Tretja oseba množine: -ijo/-o (npr. <i>odcvetijo / odcveto</i>).	G4.1.v-V5 (odcveteti, 1)
velelnik	Paradigma je (trenutno) brez velelniških oblik.	G4.1.d-velelnik (zazdeti, 7), G4.1.n-velelnik (srbeti, 2), G4.1.n-velelnik+V3 (zdeti, 1), G7.3.d-velelnik (zazebsti, 3), G7.3.n-velelnik (zebsti, 1), G8.3.d-velelnik (pripomoči, 4), G8.3.v-velelnik (premoči, 2)

4.4 Prislov

Za razliko od ostalih besednih vrst prinašajo prislovi manj vzorcev, ki so tudi hierarhično manj kompleksni in zato urejeni samo v dve ravnini. Na drugi strani pa se pri prislovih pojavlja večje število redko zastopanih vzorcev, ki izkazujejo tudi redke in specifične variante (R5 in R6 v Tabeli 12). Zaradi specifik smo se odločili, da pri slednjih dveh skupinah dovolimo vzorce, ki imajo inherentne variantne možnosti, npr. *dolgo – dlje / dalj*.²⁰ Na ureditev ostalih skupin vplivajo odločitve, ki so bile sprejete pri razvoju označevalnega sistema za slovensko oblikoskladnjo JOS. Skupina R2 tako prinaša (v leksikonu avtomatsko identificirana) deležja in skupina R4 prislove, ki zaradi specifik označevanja v leksikonu obstajajo kot samostojne leme v primerniku ali presežniku. R1 prinaša splošne prislove, ki v paradigmi ne prinašajo obrazilnega stopnjevanja, R3 pa prislove na -o, ki imajo obrazilno stopnjevanje z -eje, -ejše ali -še. Urejene vzorce prikazuje Tabela 12.

²⁰ Variante so sicer razumljene kot dopolnilo k osnovnemu vzorcu, ker se običajno pojavljajo pri različnih besedah, pogosti pa so tudi osnovni vzorci brez variant. Pri primerih R5 in R6 se zdi taka obravnava kontraproduktivna, tudi tipskih primerov ni dovolj za ločevanje (ime-na bi morala biti npr. *dolgo_1*, *dolgo_2*, kar je neskladno s sistemom). Trenutno odločitev bomo evalvirali in po potrebi nadgradili.

Tabela 12: Oblikoslovni vzorci za prislov.

Koda	Tipski primer	Opis	Št. lem v bazi
R1		Splošni prislov brez obrazilnega stopnjevanja.	
R1.1	<i>tako</i>	Splošni prislov brez obrazilnega stopnjevanja: <i>tako</i> .	5.509
R2		(Avtomatsko identificirano) deležje, brez obrazilnega stopnjevanja.	
R2.1	<i>rekoč</i>	(Avtomatsko identificirano) deležje, brez obrazilnega stopnjevanja: <i>rekoč</i> .	618
R3		Splošni prislov na -o, ki ima obrazilno stopnjevanje z -eje, -ejše ali -še.	
R3.1	<i>pozno</i>	Splošni prislov na -o, ki ima obrazilno stopnjevanje z -eje: <i>pozn-o, pozn-eje</i> .	2
R3.2	<i>novo</i>	Splošni prislov na -o, ki ima obrazilno stopnjevanje z -ejše: <i>nov-o, nov-ejše</i> .	2
R3.3	<i>lepo</i>	Splošni prislov na -o, ki ima obrazilno stopnjevanje z -še: <i>lep-o, lep-še</i> .	5
R4		Splošni prislovi, ki nimajo osnovne oblike.	
R4.1	<i>bolj</i>	Splošni prislovi v primerniku, kjer ne obstaja osnovna oblika: <i>bolj</i> .	5
R4.2	<i>najbolj</i>	Splošni prislovi v presežniku, kjer ne obstaja osnovna oblika: <i>najbolj</i> .	5
R5		Prislovi, ki imajo zaradi premen v stopnjevanih oblikah -ž(j)e oz. -ž-(j)e, -š(j)e in -lje. V tej skupini so predvsem posebnosti, ki se pogosto stopnjujejo na dva načina.	
R5.1	<i>drago</i>	Osnovna oblika se lomi na -go, stopnjevane oblike na -žje: <i>dra-go, dra-žje</i> .	3
R5.2	<i>strogo</i>	Osnovna oblika se lomi na -go, stopnjevane oblike na -ž(j)e: <i>stro-go, stro-ž(j)e</i> .	1
R5.3	<i>lahko</i>	Osnovna oblika se lomi na -hko, stopnjevane oblike na -ž(j)e: <i>la-hko, la-ž(j)e</i> .	2
R5.4	<i>nizko</i>	Osnovna oblika se lomi na -zko, stopnjevane oblike na -ž(j)e: <i>ni-zko, ni-ž(j)e</i> .	2
R5.5	<i>blizu</i>	Osnovna oblika se lomi na -zu, stopnjevane oblike na -ž(j)e: <i>bli-zu, bli-ž(j)e</i> .	1
R5.6	<i>težko</i>	Osnovna oblika se lomi na -ko, stopnjevane oblike na -(j)e: <i>tež-ko, tež-(j)e</i> .	1
R5.7	<i>trdo</i>	Osnovna oblika se lomi na -do, stopnjevane oblike na -še: <i>tr-do, tr-še</i> .	2
R5.8	<i>tiho</i>	Osnovna oblika se lomi na -ho, stopnjevane oblike na -š(j)e: <i>ti-ho, ti-š(j)e</i> .	1
R5.9	<i>široko</i>	Osnovna oblika se lomi na -oko, stopnjevane oblike na -še: <i>šir-oko, šir-še</i> .	1
R5.10	<i>visoko</i>	Osnovna oblika se lomi na -soko, stopnjevane oblike na -š(j)e: <i>vi-soko, vi-š(j)e</i> .	1
R5.11	<i>globoko</i>	Osnovna oblika se lomi na -oko, stopnjevane oblike na -lje: <i>glob-oko, glob-lje</i> .	1
R5.12	<i>daleč</i>	Osnovna oblika se lomi na -aleč, stopnjevane oblike na -lje: <i>d-aleč, d-lje</i> .	1

Koda	Tipski primer	Opis	Št. lem v bazi
R5.13	<i>dolgo</i>	Osnovna oblika se lomi na <i>-olgo</i> , stopnjevane oblike na <i>-lje/-alj</i> : <i>d-olgo, d-lje, d-alj</i> .	1
R6		Prislovi, ki imajo zaradi premen v stopnjevanih oblikah <i>-j(š)e</i> . V tej skupini so predvsem posebnosti. Pogosto se stopnjujejo na dva načina.	
R6.1	<i>mlado</i>	Osnovna oblika se lomi na <i>-do</i> , stopnjevane oblike na <i>-jše</i> : <i>m-la-do, m-la-jše</i> .	1
R6.2	<i>hudo</i>	Osnovna oblika se lomi na <i>-do</i> , stopnjevane oblike na <i>-je</i> : <i>hu-do, hu-je</i> .	1
R6.3	<i>sladko</i>	Osnovna oblika se lomi na <i>-dka</i> , stopnjevane oblike na <i>-j(š)e</i> : <i>sla-dka, sla-j(š)e</i> .	1
R6.4	<i>kratko</i>	Osnovna oblika se lomi na <i>-tko</i> , stopnjevane oblike na <i>-jše</i> : <i>kra-tko, kra-jše</i> .	1
R6.5	<i>tanko</i>	Osnovna oblika se lomi na <i>-ko</i> , stopnjevane oblike na <i>-j(š)e</i> : <i>tan-ko, tan-j(š)e</i> .	1
R6.6	<i>dobro</i>	Pri stopnjevanju se zamenja tudi osnova: <i>dobro, bolj(š)e</i> .	1
R6.7	<i>poceni</i>	Pri stopnjevanju se zamenja tudi osnova: <i>poceni, cenej(š)e</i> .	1
R6.8	<i>počasi</i>	Podaljšava pri stopnjevanju: <i>počas-i, počas-neje</i> .	1

5 Zaključek in nadaljnje delo

V prispevku smo predstavili hierarhično urejene oblikoslovne vzorce za slovenske samostalnike, pridevnike, glagole in prislove, ki so bili v projektu NSSSS zabeleženi kot del enot oblikoslovnega leksikona Sloleks. Odprto dostopna baza s pripisanimi vzorci bo skupaj s sezname identifikiranih pomanjkljivosti trenutne različice leksikona omogočila nadgradnjo leksikona in s tem večjo natančnost strojnega označevanja in pridobivanja jezikovnih podatkov iz besedilnih korpusov.

Priložnost za uporabo oblikoslovne baze za razvoj strojno podprtega generiranja novih leksikonskih enot na podlagi korpusnih podatkov se je ponudila s projektom Razvoj slovenščine v digitalnem okolju (RSDO),²¹ ki ga med leti 2020 in 2022 financirata Ministrstvo za kulturo in Evropski sklad za regionalni razvoj. Projekt razvija jezikovne vire, govorne in semantične tehnologije, strojno prevajanje in pripravo terminološkega portala. V projekt je vključena obsežna

21 Spletna stran projekta, ki predstavlja cilje, rezultate in sodelujoče partnerje: <https://www.slovenscina.eu/>

povečava leksikona Sloleks in s tem povezana priprava cevovoda za strojno podprto širjenje leksikona. Kandidati za nove leksikonske enote bodo pridobljeni iz besedilnega korpusa Gigafida 2.0 in primerljivih jezikovnih virov oz. baz. Lemam bodo strojno pripisani oblikoslovni vzorci in ustrezajoče oblike, pa tudi naglasne informacije, ki se jim projekt RSDO posveča v večji meri. Generirane enote bomo uvozili v program za urejanje leksikona, kjer bodo (deloma že med projektom, deloma po zaključku projekta) jezikoslovno pregledane in urejene.

Nadgradnja leksikona se umešča v širšo aktivnost razvoja Digitalne slovarske baze (Gantar 2020). V bazi bodo leksikonske informacije povezane z drugimi vrstami jezikovnih podatkov, kar bo (poleg vseh ostalih do sedaj naštetih aktivnosti) razkrilo potrebe po morebitni nadgradnji sistema vzorcev in nakazalo optimalne dolgoročne rešitve.

Na drugi strani želimo oblikoslovne vzorce – v ustrezno prilagojeni, uporabniško prijazni obliki – vključiti v leksikonski vmesnik, ki je namenjen uporabniški skupnosti.²² Jezikovni uporabniki in uporabnice prek vmesnika iščejo odgovore na jezikovne zadrege, povezane z oblikoslovjem (Dobrovoljc 2015). S tega vidika je bil Sloleks prepoznani kot dragocen pripomoček za uporabo pri pouku slovenščine (Stritar in Dobrovoljc 2013). Vključitev vzorcev v vmesnik bo omogočila pregled nad besediščem, ki se oblikoslovno obnaša primerljivo, kar je izrednega pomena za jezikovno didaktiko, tako za usvajanje slovenščine kot prvega kot drugega oz. tujega jezika.

Zahvala

Prispevek je nastal s financiranjem Agencije za raziskovalno dejavnost Republike Slovenije, in sicer raziskovalnega projekta Nova slovnica sodobne standardne slovenščine: viri in metode (J6-8256) ter programske skupine Jezikovni viri in tehnologije za slovenski jezik (P6-0411).

22 Vmesnik je dostopen na: <https://viri.cjvt.si/sloleks/slv/>.

Reference

- Arhar Holdt, Š., Čibej, J., Laskowski, C. in Krek, S. (2020). Morphological patterns from the Sloleks 2.0 lexicon 1.0, Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1411>.
- Arhar Holdt, Š. in Čibej, J. (2018). Oblikoslovni vzorci v leksikonu Sloleks: izhodiščni nabor za samostalnike. *Slovenščina 2.0: empirične, aplikativne in interdisciplinarne raziskave*, 6 (2), 33–66. <http://dx.doi.org/10.4312/slo2.0.2018.2.33-66>.
- Arhar, Š. (2009). Učni korpus SSJ in leksikon besednih oblik za slovenščino. *Jezik in slovstvo* 54 (3–4), 43–56. Dostopno prek: <https://jezikinslovstvo.com/pdf/2009-03-04-Razprave-Spela-Arhar.pdf>.
- Arhar, Š. in Holozan, P. (2009). Leksikalna podatkovna zbirka ASES (Amebisov skupni elektronski slovar). V V. Mikolič (ur.), *Jezikovni korpusi v medkulturni komunikaciji* (str. 30–51). Koper: Univerza na Primorskem, Znanstveno-raziskovalno središče, Založba Annales: Zgodovinsko društvo za južno Primorsko.
- Dobrovoljc, K. in Krek, S. (2013). Spletni portal Slogovni priročnik: luščenje in prikaz podatkov o jezikovni rabi. V A. Žele (ur.), *Družbena funkcijnost jezika: vidiki, merila, opredelitve, Obdobja* 32 (str. 101–107), 1. natis. Ljubljana: Znanstvena založba Filozofske fakultete. Dostopno prek: <https://centerslo.si/wp-content/uploads/2015/10/32-DobrovoljcK.pdf>.
- Dobrovoljc, K., Krek, S., Holozan, P., Erjavec, T., Romih, M., Arhar Holdt, Š., Čibej, J., Krsnik, L. in Robnik-Šikonja, M. (2019). Morphological lexicon Sloleks 2.0, Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1230>.
- Dobrovoljc, K. (2015). Oblikoslovne informacije v sodobnih slovarskih priročnikih. V V. Gorjanc et al. (ur.), *Slovar sodobne slovenščine: problemi in rešitve* (str. 64–79). Ljubljana: Znanstvena založba Filozofske fakultete. Dostopno prek: <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/15/47/486-1>.
- Dobrovoljc, K., Krek, S. in Erjavec, T. (2015). Leksikon besednih oblik Sloleks in smernice njegovega razvoja. V V. Gorjanc, Gantar, P., Kosem, I. in Krek, S. (ur.), *Slovar sodobne slovenščine: problemi in rešitve* (str. 80–105). Ljubljana: Znanstvena založba Filozofske fakultete. Dostopno prek: <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/15/47/489-1>.

- Erjavec, T. in Krek, S. (2008). Oblikoskladenjske specifikacije in označeni korpusi JOS. V T. Erjavec in J. Žganec Gros (ur.), *Zbornik Šeste konference Jezikovne tehnologije: zbornik 11. mednarodne multikonference Informacijska družba* (str. 49–53). Ljubljana: Institut Jožef Stefan. Dostopno prek: http://nl.ijs.si/jos/bib/jos_islct08.pdf.
- Erjavec, T., Holozan, P., Krek, S., Pivec, M., Rigač, S., Rozman, S. in Velušček, A. (2008). Specifikacije za leksikon besednih oblik – projekt Sporzumevanje v slovenskem jeziku, kazalnik 3. Kamnik. Dostopno prek: <http://projekt.slovenscina.eu/Vsebine/Sl/Kazalniki/K3.aspx>.
- Gantar, P. (2020). Dictionary of modern Slovene: from Slovene lexical database to digital dictionary database. *Rasprave Instituta za hrvatski jezik i jezikoslovlje*, 46 (2): 589–602. <https://doi.org/10.31724/rihjj.46.2.7>.
- Gorjanc, V., Gantar, P., Kosem, I. in Krek, S. (ur.) (2015). *Slovar sodobne slovenščine: problemi in rešitve*. Ljubljana: Znanstvena založba Filozofske fakultete. E-izdaja (2017). Ljubljana: Znanstvena založba Filozofske fakultete. <https://doi.org/10.4312/9789612379759>.
- Krek, S., Arhar Holdt, Š., Erjavec, T., Čibej, J., Repar, A., Gantar, P., Ljubešić, N., Kosem, I. in Dobrovoljc, K. (2020). Gigafida 2.0: the reference corpus of written standard Slovene. V N. Calzolari (ur.), *LREC 2020: Twelfth International Conference on Language Resources and Evaluation: conference proceedings* (str. 3340–3345). Pariz: European Language Resources Association. Dostopno prek: <http://www.lrec-conf.org/proceedings/lrec2020/LREC-2020.pdf>.
- Krsnik, L. (2018). *Napovedovanje naglasa slovenskih besed z metodami strojnega učenja*. Magistrsko delo. Univerza v Ljubljani, Fakulteta za računalništvo in informatiko. Dostopno prek: <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=98276&lang=slv>.
- Logar, N., Grčar, M., Brakus, M., Erjavec, T., Arhar Holdt, Š. in Krek, S. (2012). *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in cckRES: gradnja, vsebina, uporaba*. Ljubljana: Trojina, zavod za uporabno slovenistiko: Fakulteta za družbene vede. E-izdaja (2020). Ljubljana: Znanstvena založba Filozofske fakultete. Dostopno prek: <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/233/333/5394-1>.
- Mirtič, Tanja (2015). *Pregibnostno-naglasni vzorci knjižne slovenščine*. Doktorska disertacija. Univerza v Ljubljani, Filozofska fakulteta.
- Rejc, R. (2017). *Generiranje slovenskih besednih oblik s pomočjo strojnega učenja*. Diplomsko delo. Univerza v Ljubljani, Fakulteta za računalništvo

in informatiko. Dostopno prek: <https://repositorij.uni-lj.si/IzpisGradiva.php?lang=slv&id=91151>.

Stritar, M. in Dobrovoljc, K. (2013). Korpusi na poti v šole: jezikovnotehnološko izpopolnjevanje učiteljev. *Slovenščina 2.0: empirične, aplikativne in interdisciplinarne raziskave*, 1 (1), 181–194. <https://revije.ff.uni-lj.si/slovenscina2/article/view/6922>.

Toporišič, J. (2004). *Slovenska slovnica*. Maribor: Založba Obzorja.