

REPREZENTACIJA STANDARDNEGA IN NESTANDARDNEGA V VIRIH SSJ

Špela Arhar Holdt, Kaja Dobrovoljc

Trojina, zavod za uporabno slovenistiko, Ljubljana

Damjan Popič

Filozofska fakulteta, Ljubljana

UDK 811.163.6'271:81'322

Korpusni pristop omogoča neposredno primerjavo obstoječih jezikovnih pravil in njihovo udejanjanje v jezikovni rabi. Tovrsten združen pogled na normo in rabo lahko jezikovnemu uporabniku ponudi marsikaj novega, vendar samo v primeru, da je primerjava izvedena na pregleden, konsistenten in razumljiv način. Prispevek predstavlja možnosti sopostavitve korpusnih in normativnih podatkov, ki so preizkušene pri rezultatih projekta Sporazumevanje v slovenskem jeziku: leksikonu besednih oblik Sloleks, portalu Slogovni priročnik in Pedagoškem slovničnem portalu.

Sporazumevanje v slovenskem jeziku, Sloleks, Slogovni priročnik, Pedagoški slovnični portal, norma in raba

Corpus-driven approaches allow for immediate comparison of standardised language and language use. This compact and simultaneous overview of language norm and use provides users with a number of new possibilities, but only if the comparison is clear, consistent and understandable. The paper presents different means of contrasting normative and corpus-driven data. These resources have been implemented in the results of the project Communication in Slovene: the »Sloleks« lexicon of inflected forms, the »Style Manual«, and the »Pedagogical Grammar Portal«.

communication in Slovene, Sloleks, Manual of Style, Pedagogical Grammar Portal, language norm and use

1 Uvod

V sklopu projekta Sporazumevanje v slovenskem jeziku (SSJ) je med leti 2008 in 2013 potekala priprava raznovrstnih jezikoslovnih podatkovnih zbirk, virov in jezikovnih tehnologij.¹ Za vse projektne aktivnosti velja, da so izrazito korpusnojezikoslovne, usmerjene na eni strani v gradnjo in jezikoslovno označevanje novih korpusnih virov, na drugi pa v razvoj postopkov za uporabo

korpusnih podatkov pri pripravi sodobnih spletnih jezikovnih virov za slovenščino.

V tem prispevku se osredotočamo na tri projektne rezultate: leksikon besednih oblik Sloleks, portal Slogovni priročnik in Pedagoški slovnični portal. Vse tri združuje dejstvo, da se na tak ali drugačen način soočajo z vprašanjem sopostavitve podatkov o *standardni* in *nestandardni* jezikovni rabi v sodobni (pisni) slovenščini.² Korpusni pristop

¹ Več o rezultatih na www.slovenščina.eu. Projekt sta financirala Evropski socialni sklad in Ministrstvo za izobraževanje, znanost, kulturo in šport.

² Pri vseh projektnih rezultatih je »standardna« jezikovna raba, ki je skladna z obstoječim jezikovnim predpisom, »nestandardna« pa raba, ki od predpisa na tak ali drugačen način odstopa. Takšno poimenovanje, kljub temu da je v delnem neskladju z obstoječimi praksami v jezikoslovju, namreč najboljše ustreza novopripravljenim jezikovnim virom. O odločitvi v tem prispevku ne polemiziramo, predvidena pa je razlaga v drugi literaturi.

omogoča neposredno primerjavo obstoječih jezikovnih pravil in njihovo udejanjanje v jezikovni rabi. Tovrsten združen pogled na normo in rabo lahko jezikovnemu uporabniku ponudi marsikaj novega, vendar samo v primeru, da je primerjava izvedena na pregleden, konsistenten in razumljiv način.

V prispevku pojasnujemo, kako našeti trije viri odgovarjajo na omenjeni izziv sopolstitve informacij, pri čemer se sklicujemo na korpus, ki so izvor jezikovnih podatkov. Gre za pisni korpus Gigafida (Logar Berginc idr. 2012), govorni korpus GOS (Verdonik, Zwitter Vitez 2010) in Šolar, korpus besedil, ki so jih učenci napisali pri šolskem pouku (Rozman idr. 2012). Vsi trije korpusi so kot rezultati projekta SSJ prosto dostopni na spletu, kar seveda velja tudi za leksikon Sloleks, Slogovni priročnik in Pedagoški slovniki portal (za spletne naslove gl. seznam literature).

2 Leksikon besednih oblik Sloleks

Leksikon besednih oblik Sloleks³ je obsežna podatkovna baza, ki ima znotraj projekta SSJ dve zelo različni, a tesno povezani vlogi. Po eni strani je leksikon besednih oblik namenjen uporabi v jezikovnotehnoloških aplikacijah, kot sta lematizator in oblikoslovnih označevalnik,⁴ ki za označevanje korpusnih besedil s podatki o lemi, besedni vrsti in drugih jezikoslovnih lastnostih poleg informacij iz ročno označenega učnega korpusa upoštevata tudi informacije iz obstoječega leksikona besednih oblik. Po drugi strani je leksikon namenjen neposredno človeškemu uporabniku, bodisi kot spletni »slovar« z informacijami glede pregibanja besed v slovenskem jeziku, bodisi kot osrednja podatkovna baza spletnega portala Slogovni priročnik.

Leksikon Sloleks trenutno vsebuje nekaj več kot 100.000 leksikonskih enot. Zapisan je v formatu Lexical Markup Framework

(LMF), ki omogoča, da vsaki leksikonski enoti oz. katerikoli izmed njenih oblik poleg oblikoslovnih lastnosti pripišemo poljubno število dodatnih informacij. Leksikon Sloleks tako poleg podatkov o oblikoslovnih lastnostih osnovne in pregibne oblike, pogostosti pojavljanja v korpusu Gigafida in morebitnih povezanih leksikonskih enotah vsebuje tudi podatke o normativni zaznamovanosti oblik glede na trenutno veljavni pravopisni standard.

Normativni podatki so v leksikonu besednih oblik zabeleženi v atributu *norma*, in sicer na ravni leme (če je nestandardna že osnovna oblika besede) ali posamezne oblike (če je nestandardna zgolj katera izmed oblik sicer standardne leme). Atribut *norma* lahko nato zavzame vrednost *nestandardno* (za oblike, ki niso v skladu s trenutnim pravopisnim standardom), *variantno* (za več oblik, ki so v skladu s trenutnim pravopisnim standardom) ali *nejasno* (za oblike, kjer norma ni jasno določljiva zaradi neskladij med pravopisnimi pravili in pravopisnim slovarjem). Odsotnost normativne oznake pomeni, da je oblika standardna, torej normativno nezaznamovana.

Čeprav je informacija o vseh variantnih zapisih oblik sicer dragocena tudi z vidika zagotavljanja čim večje natančnosti strojnega označevanja besednih oblik, kakršne se pojavljajo v realni sodobni slovenščini, imajo eksplicitni podatki o njihovi normativni zaznamovanosti zgolj jezikovnopriročniško namembnost, saj jih lematizator in označevalnik pri označevanju korpusnega besedišča ne pripisujeta. Slika 1 tako prikazuje primer leksikonske enote, kakršno najdemo v spletnem vmesniku leksikona Sloleks, ki prikazuje enake informacije kot izhodiščna baza v formatu LMF, le na preglednejši in uporabniku prijaznejši način. Vmesnik poleg tega vsebuje še hiperpovezave na konkordance vseh ali posameznih oblik v korpusu Gigafida ter povezave med posameznimi enotami zno-

³ Leksikon je v obliki spletnega iskalnika dostopen na <http://www.slovenscina.eu/sloleks>, več o njegovi zasnovi pa v Arhar: 2009.

⁴ Dostopno na <http://označevalnik.slovenščina.eu/>, več v Grčar, Krek in Dobrovoljc: 2012.

Kako pregibamo besede v slovenskem jeziku?

Išči

Legenda:
— standardna oblika
— nestandardna oblika

hči samostalnik, občno ime, ženski spol; **57.922** pojavitev

EDNINA			DVOJINA			MNOŽINA		
oblika	sklon	pogostost	oblika	sklon	pogostost	oblika	sklon	pogostost
hči	imenovalnik	29.717	hčeri	imenovalnik	163	hčere	imenovalnik	891
hčer	imenovalnik	736	hčer variantno	rodilnik	0	hčer variantno	rodilnik	22
hčera	imenovalnik	2.207	hčera variantno	rodilnik	72	hčera variantno	rodilnik	991
hčere	rodilnik	5.935	hčerama	dajalnik	102	hčeram	dajalnik	414
hčeri	dajalnik	533	hčeri	tožilnik	1.425	hčere	tožilnik	712
hčer	tožilnik	7.761	hčerah	mestnik	20	hčerah	mestnik	222
hčero	tožilnik	200	hčerama	orodnik	464	hčerami	orodnik	486
hči	tožilnik	500						
hčeri	mestnik	2.691						
hčerjo	orodnik	1.658						

POVEZANE OBLIKE:
hčerin pridevnik

Slika 1: Prikaz leksikonske enote za samostalnik *hči* v spletnem vmesniku leksikona besednih oblik Sloleks

traj samega leksikona. Normativni podatki so prikazani s kombinacijo barvnih označevalcev (modro-*standardno* proti sivemu-*nestandardnemu*) oz. opisnega kvalifikatorja *variantno* v primeru več standardnih (modrih) oblik z identičnimi lastnostmi. Barvno razlikovanje je razloženo tudi v legendi v desnem zgornjem kotu vsake leksikonske enote.

3 Slogovni priročnik

Leksikon besednih oblik je poleg korpusa Gigafida obenem tudi osrednja podatkovna zbirka spletnega portala Slogovni priročnik,⁵ ki uporabnikom slovenskega jezika prinaša objektivne informacije o konkretnih jezikovnih zadregah, ne da bi ti za uporabo ali razumevanje potrebovali specialistično jezikoslovno znanje. Portal jim omogoča, da z vnosom konkretnih iskalnih pogojev o določenem problemu pridobijo vse potrebne podatke, da lahko z njim sami in v relativno kratkem času sprejmejo objektivizirano jezikovno odločitev. Vsebina sloni na seznamu 700 jezikovnih zadreg z različnih jezikovnih

ravnin: pravopis, pravorečje, oblikoslovje, besedotvorje, besedišče, skladnja, besedilo in razno.

Kot alternativni način opisa jezika, vezan na jezikovnotehnološko ogrodje, je priročnik v veliki meri vezan na dihotomijo med standardnim in nestandardnim. Eksplicirane normativne informacije se (poleg inherentnih normativnih informacij v leksikonu besednih oblik) v Slogovnem priročniku pojavljajo v treh oblikah, in sicer v:

- normativnih informacijah v kratkem odgovoru,
- informacijah o jezikovnem standardu znotraj dolgega odgovora in
- povezavah na zunanje kodifikacijske vire.

Pri vseh treh sekcijah znotraj priročnika so podane informacije o jezikovnem standardu, vendar na različne načine. Medtem ko so informacije v t. i. kratkem odgovoru predvsem jedrnat in namenjene hitri samopomoči pri jezikovnih odločitvah, se t. i. dolgi odgovor osredotoča predvsem na dihotomijo

⁵ Dostopno na <http://slogovni.slovenscina.eu/>, več v Krek 2012.

med rabo in standardom. Zadnji del odgovora, sekcija Za navdušence, je namenjen vsem, ki želijo o določeni zadregi izvedeti več, predvsem iz veljavnega in zgodovinskega predpisa.

3.1 Kratki odgovor

Vsako geslo na portalu tako poleg statičnega dela v obliki dolgega odgovora in povezav na zunanje vire vsebuje dinamični, kratki odgovor, ki s kombinacijo grafa in njegovega kratkega opisa prikazuje statistične podatke o jezikovni rabi za konkretne oblike, ki jih je vstopni mehanizem prepoznal kot relevantne za govorčevu trenutno zadrego. Kratki odgovor je sestavljen iz besedila v formatu XML, ki deluje kot univerzalni mehanizem, s katerim lahko opišemo statistično stanje pri vseh možnih kombinacijah standardnih ali nestandardnih oblik pri določeni kategoriji (Krek 2012: 227). To omogoča, da kratki odgovor vedno odseva realno stanje v jeziku v razmerju rabe do standarda. Ob vsakršni posodobitvi besedilnega korpusa, ki služi kot podstat slogovnega priročnika, se morebitne spremembe v jeziku, ki jih korpus registrira, takoj prikažejo na grafu in v spremljajočem opisu.

Za pravilno delovanje dinamičnega dela je tako ključna povezava z leksikonom besednih

oblik, ki portalu poleg podatkov o pogostosti oblik v besedilnem korpusu Gigafida zagotavlja še podatke o njihovi normativni (ne)zaznamovanosti. Slika 2 prikazuje, kako je normativna zaznamovanost na grafu prikazanih podatkov označena barvno (analogno s spletnim vmesnikom za Sloleks so standardne oblike označene z modro in nestandardne s sivo barvo, oblike z oznako nejasno pa so obarvane rumeno), pomen obarvanih stolpcev pa je vedno natančno pojasnjen tudi v kratkem odgovoru nad njimi.

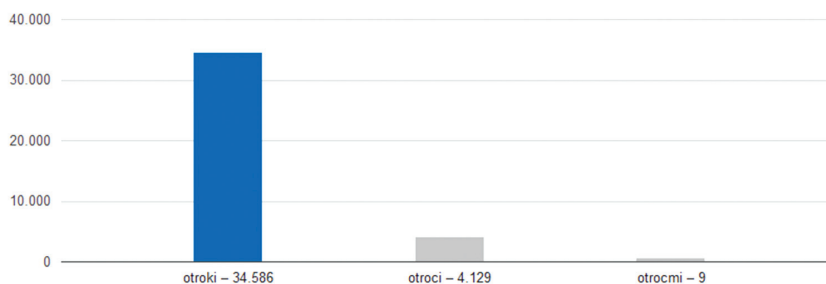
3.2 Dolgi odgovor

Za razliko od kratkega je dolgi odgovor, ki je zapisan v standardnem formatu HTML, statičen, kar pomeni, da ga lahko osvežujemo le z ročnim spreminjanjem HTML-kode in vsebine. Dolgi odgovor podaja naslednje informacije:

- podatke o jezikovnem standardu,
- podatke o rabi,
- sistemske opredelitve razmerja standard – raba,
- povezave: povezave na zunanje vire; poljudne razlage strokovnih terminov, rabljenih v dolgem odgovoru; povezave do spiskov besed, ki spadajo v isto problem-

KRATKO IN JEDRNATO

Na grafu si lahko ogledate podatke o rabi oblik za orodnik množine besede **otrok** v korpusu Gigafida. Oblika "[z] otroki" je ustrezna, obliki "[z] otroci" in "[z] otrocmi" pa nista skladni s trenutnim pravopisnim standardom.



Slika 2: Primer kratkega odgovora na spletnem portalu Slogovni priročnik za iskalni pogoj *Z otroki ali z otroci?*

NA DOLGO IN ŠIROKO

Na splošno **svojljne pridevnike** iz samostalnikov moškega spola naredimo tako, da jim dodamo **-ov** ali **-ev**, pri čemer je izbira **odvisna od glasu**, s katerim se samostalnik konča. Izjema so svojljni pridevniki iz samostalnikov, ki se končajo na izgovorjeni r. Pri teh imamo dve enakovredni možnosti.

1. Lahko jih podaljšamo z -j, kar pomeni, da bomo zaradi **preglasa** uporabili **-ev**, npr. **novinar – novinarjev, Gregor – Gregorjev**.
2. Lahko pa jim dodamo **-ov**, npr. **satir – satirov, Bor – Borov**.

Raba ene od obeh možnih oblik se je pri večini pridevnikov ustalila in na splošno prevladuje oblika z **-jev**. Oblika z **-ov** pa skoraj vedno uporabimo v dveh primerih:

1. kadar pri samostalniku pred končnim izgovorjenim r stoji **polglasnik**, zapisan s črko, ki pri sklanjanju izgine (Peter → Petra, Petru ... → Petrov oče).
2. pri **enozložnih** samostalnikih (Bor → Borov).

Obstajajo tudi **posamezna imena**, pri katerih izbira precej niha. Če nas zanima, katero obliko pisci raje uporabljajo, se o tem lahko prepričamo v **korpusu**.

Poseben problem so angleška, francoska in nekatera druga **lastna imena**, ki se končajo z izgovorjenim r, vendar črki r sledi še **nemi -e**, kot na primer v priimkih **Molière** [moljêr], **Baudelaire** [bodlêr], **Saussure** [sosîr], **Gilmore** [gîlmor], **Shakespeare** [šékspîr] in **podobnih**. Pri tvorjenju svojljnih pridevnikov iz teh lastnih imen **nemi -e** praviloma izpustimo in dodamo **-jev** ali **-ov**, kar pomeni, da sta ustrezni obe obliki, npr. **Molièrov/Molièrjev, Saussurov/Saussurjev, Baudelairov/Baudelairjev** itd. Pri nekaterih imenih se je raba ustalila pri eni od možnosti, pri drugih pa se uporabljata obe obliki, vendar ena navadno prevladuje. Podatek o tem je mogoče dobiti v **korpusu**.

Za uspešno reševanje zadrege moramo poznati pravilen izgovor tujega imena!

Slika 3: Dolgi odgovor za problem tvorbe svojljnih pridevnikov iz moških imen, ki se končajo na govorjeni r

- sko kategorijo (in so v leksikonu besednih oblik označene z isto oznako),
- jezikovnosistemske podatke, ki pojasnjujejo ozadje jezikovne zadrege (pri jezikovnih zadregah, pri katerih je to mogoče),
- etimološke in zgodovinoslovnične osvetlitve,
- morebitna opozorila,
- problemskospecifične informacije.

Namen dolgega odgovora je torej ta, da poda podrobnejše, strokovne in argumentirane informacije o danem jezikovnem problemu, istočasno pa uporabnika opremi še z informacijami o drugih dostopnih virih, ki so mu lahko kakorkoli v pomoč, obenem pa

služi kot ogledalo jezikovnega predpisa in varovalka pri nereprezentativnih kodifikacijskih določilih v jezikovnem standardu.

3.3 Za navdušence

Zadnji del dolgega odgovora je sestavljen iz zbranih hiperpovezav na digitalizirane kodifikacijske priročnike, kjer lahko uporabniki pridobijo teoretično ozadje o določeni jezikovni zadregi. Zajeti so vsi kodifikacijski viri, ki so prosto dostopni, hiperpovezave pa kažejo neposredno na relevantna poglavja, ki so tudi poimenovana. Vključen je tudi Googleov iskalni modul, s katerim lahko uporabniki iščejo po starejših virih (starih pravopisih in slovnica), ki so bili digitalizirani. Nabor

ZA NAVDUŠENCE

Slovenski pravopis – pravila (2001):

- Stran 89 - Posebnosti 1. moške (o-jevske) sklanjatve – krajšanje osnove
- Stran 90 - Posebnosti 1. moške (o-jevske) sklanjatve – daljšanje z j
- Stran 114 - Težji primeri iz besedotvorja (pridevnik) – priponsko obrazilo -ov/-ev oz. -in

Preverite tudi, kaj o vašem iskalnem pogoju pravijo **digitalizirani slovenski pravopisi in starejše slovnice**, ki so izšli v obdobju od 1899 do 2001.

Slika 4: Rubrika Za navdušence s povezavami na zunanje kodifikacijske vire

virov je odvisen od tehnične obdelave in pa od avtorskih pravic – predvsem od slednjega vprašanja je odvisno, s kakšno hitrostjo se bo fond virov dopolnjeval.

4 Pedagoški spletni portal

Pedagoški spletni portal⁶ je za razliko od Slogovnega priročnika namenjen za uporabo pri poučevanju oz. učenju slovenščine. Ciljni uporabniki portala so učenci zadnje triade osnovnih šol in dijaki vseh vrst srednjih šol. Teme, ki jih portal predstavlja, so neposreden odgovor na jezikovne težave, identificirane v korpusu Šolar – torej jezikovne napake, ki se pojavljajo v šolskih besedilih, primarno spisih in esejih, napisanih v času šolskega pouka (Rozman idr. 2012).

Portal sestavljajo interaktivna slovnicična poglavja, pripravljena na osnovi podatkov iz besedilnih korpusov.⁷ Kot je že bilo izpostavljeno (Krek, Arhar 2010; Kosem idr. 2012), je združevanje pedagoškega in korpusnega pogleda na jezik svojevrsten izziv, ki pa prinaša s seboj številne nove možnosti. Med najpomembnejšimi so: (I) uporaba izključno avtentičnega jezikovnega gradiva, (II) upoštevanje realne distribucije jezikovnih pojavov, (III) združevanje dihotomij, ki povzročajo nerealno parcialen pogled na jezik (pisni vs.

govorjeni jezik, standardno vs. nestandardno, norma vs. raba, vseslovensko vs. regionalno).

Ker je portal namenjen uporabi v šolski situaciji, pri predstavitvi jezikovnih pravil vedno sledi obstoječi jezikovni normi. Mogoče pa je ob zglede, ki jezikovnim pravilom sledijo, postaviti tudi primere rabe, ki od pravil odstopajo. Takšna sopostavitve se uporabljajo v primerih, opisanih v nadaljevanju.

4.1 Tipične napake

Pri predstavitvi jezikovnih napak, značilnih za v poglavju obravnavano jezikovno težavo, in pri nadaljnjem opisu ter kategorizaciji teh napak se uporabljajo primeri rabe iz korpusa Šolar. Napake so v besedilu označene z rdečo barvo, popravljena različica prinaša zeleno barvo, kar prikazuje Slika 5.

Zgledi, ki vsebujejo tipične napake, se na spletnem pedagoškem portalu uporabljajo tudi pri nekaterih vrstah slovnicičnih vaj (npr. vajah, ki od učenca zahtevajo, naj v besedilnem odlomku identificira oz. označi jezikovne napake).

4.2 Nestandardna besedila

V določenih slovnicičnih poglavjih se uporabljajo tudi primeri rabe iz nestandardnih besedil korpusa Gigafida, kot so denimo

The screenshot shows a web interface for a pedagogical portal. At the top, there's a title bar with the text "Zapis opisnega deležnika na -l". Below it, a yellow box contains the text "Tukaj te bom počakaval." and "Tukaj te bom počakala." with a red 'X' next to the first sentence. To the right, a table compares incorrect and correct sentence constructions. The table has two columns: the left column shows incorrect sentences with a red 'X' and the right column shows the corrected sentences with a green checkmark. The examples are: 1. "Zelo veliko se je igrav, ko je imel prosti čas." vs. "Zelo veliko se je igral, ko je imel prosti čas." 2. "Takrat sem si misliv: 'Ko bom starejši, bom šel s kolesom na Lisco.'" vs. "Takrat sem si mislil: 'Ko bom starejši, bom šel s kolesom na Lisco.'" 3. "Kdo naj bi biu vitez?" vs. "Kdo naj bi bil vitez?". At the bottom, there are three navigation buttons: "RAZLAGA", "POVZETEK", and "VAJE".

To poglavje je pravo zate, če imaš tovrstne težave:	
✗ Zelo veliko se je igrav, ko je imel prosti čas.	✓ Zelo veliko se je igral, ko je imel prosti čas.
✗ Takrat sem si misliv: "Ko bom starejši, bom šel s kolesom na Lisco."	✓ Takrat sem si mislil: "Ko bom starejši, bom šel s kolesom na Lisco."
✗ Kdo naj bi biu vitez?	✓ Kdo naj bi bil vitez?

RAZLAGA > POVZETEK > VAJE >

Slika 5: Uvodni prikaz tipičnih napak pri zapisu deležnika na -l

⁶ Portal bo prosto dostopen s septembrom 2013 na <http://slovnica.slovenščina.eu/>.

⁷ Primeri tem: *Preglas*, *Sklanjanje samostalnika otrok*, *Nedoločnik in kratki nedoločnik*, *Namenilnik*, *Predlog z/s* in *Predlog k/h*, *Zapis zanikanih glagolov* itn.

zapisi na različnih spletnih forumih. S temi primeri rabe želimo učencu pokazati razliko med pravili, ki veljajo za pisanje v standardnem jeziku, in pa jezikom »elektronskih neformalnih žanrov«, pri katerih je najti namensko približevanje zapisa govorjenemu jeziku (kar ndr. pomeni opuščanje pravopisnih pravil in zapisovanje besed sledeč njihovi slušni podobi).

Kot kaže Slika 6, v poglavju *Zapis opisnega deležnika na -l* proti koncu razlage učencem pokažemo, da v nestandardni slovenščini pogosto vidimo deležnik zapisan tako, kot ga izgovorimo. Nato pa učenca opozorimo, da v standardni slovenščini tovrstni zapisi veljajo za napako, zaradi česar je potrebna pozornost.

Takšen potek razlage učencem omogoča, da pravilo standardne slovenščine spoznajo sočasno z informacijo, da ni nujno, da bodo vsa besedila, s katerimi se bodo srečevali, pravilu tudi v resnici sledila. Tako lahko učenec uzavesti razlike med različnimi komunikacijskimi situacijami, v katerih veljajo različne možnosti izbire – in seveda tudi iz tega izhajajoče različne komunikacijske posledice.

4.3 Razhajanja med jezikovno normo in rabo

Kadar identificiramo, da se odstopanja od določenega jezikovnega pravila pojavljajo razpršeno po različnih besedilnih vrstah (tudi v časopisih, revijah, leposlovju ...), učenca na razhajanja eksplicitno opozorimo.

Za primer takšnega razhajanja lahko navedemo slovnično poglavje *Zapis predloga z/s*, v katerem učencu ndr. predstavimo jezikovno pravilo, da pri izbiri predložne variante sledimo zapisu besede in ne njenemu izgovoru (pravičen zapis je: *z Javierjem*, *s SMS-om*). Obenem opozorimo, da se pisci v praksi pogosto odločajo v nasprotju s pravilom (sledijo izgovoru besede: *s Javierjem*, *z SMS-om*). Na tem mestu razlage ponudimo tudi povezavo na Slogovni priročnik, kjer si lahko učenec prebere več o tej temi.

5 Zaključek

V prispevku smo pokazali tri načine sopostavljanja podatkov o jezikovni normi in jezikovni rabi. V oblikoslovnem leksikonu je posamezna besedna oblika opredeljena kot standardna ali nestandardna, uporabnik pa dobi tudi podatek o morebitni normativni

PEDAGOŠKI SLOVNIČNI PORTAL

Zapis opisnega deležnika na -l

Priše, vide, zmaga

RAZLAGA
POVZETEK
ZA NAVDUŠENE
KOROŠKA IN GORENJSKA, POZOR!
VAJE

Včasih opisne deležnike na -l tudi zapišemo tako, kot jih slišimo. To se dogaja pri pisanju v **nestandardni slovenščini** (ko klepetamo na spletu, pišemo SMS-e prijateljem in podobno). Spodaj lahko vidiš dva primera s spletnih forumov:

- men je šow tud ful na živce Paris in še na koncu Ahila ubije
- Ko je pa Messi rekeu da bi igrau za barco zastojn ste usi pljuvali po njemu.

V **standardni slovenščini** tak zapis **velja za napako**. Še zlasti pogosta je ta napaka v šolskih spisih in esejih.

- Upam, da sem vas skozi ta spis **popeljav** prek življenja Alfreda Pačnika.
- Duhovnik se je ukazu **upru** zato, ker je imel slovenščino zelo rad.

Slika 6: Prikaz primerov nestandardnega zapisa deležnika na -l

nejasnosti in pa variantnosti oblik. Obenem so na voljo podatki o pogostnosti posamezne oblike v korpusu Gigafida, z neposredno povezavo na relevantne korpusne podatke. Za Slogovni priročnik je značilna dodatna stopnja jezikoslovne obravnave korpusnih podatkov. Poleg kratke grafične predstavitev, kako pogosto se določen jezikovni pojav pojavlja v korpusu Gigafida, si lahko uporabniki preberejo tudi daljši sestavek, ki natančneje opisuje obravnavani jezikovni problem. Tretji način sopostavljanja pa je preizkušen v slovničnih poglavjih Pedagoškega spletnega portala, kjer se omenjenim pridružuje še jezikovnodidaktični vidik s stalno prisotnim vprašanjem, kako učencem sočasno s predstavitvijo jezikovnih pravil pokazati tudi mavrično in pogosto kaotično realnost jezikovne rabe.

S tovrstnim sopostavljanjem je omogočen združen pogled na jezikovna pravila in informacije o tem, ali se v rabi ta pravila upoštevajo ali ne, tj. na tematiko, ki je bila v preteklosti v jezikovnih priročnikih pogosto na silo (oz. po sili razmer) ločena. Novi pogled ni dragocen le za jezikovne uporabnike, ki imajo po novem možnost na enem mestu dobiti raznovrstne, a urejene podatke, ampak tudi za jezikoslovce. Delo, kakršnega opisujemo, namreč izredno dobro nakazuje težava na mesta oz. nerešene naloge tako za normativistiko kot za slovaropisce in slovničarje, pripravljalce učnih gradiv za pouk slovenščine, in najbrž še koga.

Viri SSJ so sicer v precejšnji meri pilotski: nikakor ne zaključeni, vendar pa so stabilno izhodišče za razmislek o vprašanju, kako na identificirana razhajanja med določenimi jezikovnimi pravili in jezikovno rabo reagirati. Avtorjem prispevka se kažeta dva odgovora: kadar se zdi, da je težava v uporabnikih, je treba jezikovno pravilo jasneje predstaviti in se z njim pri opismenjevanju šolajoče se populacije osredotočeno (in morda na drugačen način, kot do zdaj) ukvarjati; kadar se zdi, da je težava v pravilu, je treba pravilo spremeniti. Težave z zapisovanjem opisnega deležnika na -l, ki se jih dotika

poglavje 4, bi lahko uvrstili v prvo skupino težav, torej med težave, ki so predvsem izziv opismenjevanja v slovenskem jeziku (kamor sodi tudi razumevanje razlike med standardno in nestandardno pisno slovenščino). Na drugi strani pa se pri izbiri pravega variantnega predloga (pogl. 4.3) v določeni meri kažejo težave na ravni pravila. Pravilo za izbiro ustrezne variante predloga z/s zahteva upoštevanje zapisa sledeče besede (npr. z *Javierjem* in ne s *Javierjem*), medtem ko naj bi pri izbiri variante predloga k/h sledili izgovoru sledeče besede (npr. h *Carmen* in ne k *Carmen*). Slednje pravilo obenem v pravopisnih pravilih ni eksplicitno navedeno, zato lahko uporabnik o njem sklepa zgolj iz podatkov, navedenih v pravopisnem slovarju (pri geslu »h«).

Vprašanje, kaj je vzrok za neupoštevanje jezikovnih pravil, je seveda veliko bolj zapleteno, kot ga predstavljamo v tem prispevku, pa tudi zgoraj navedeni možnosti reševanja težav nista preprosto izvedljivi in ne metodološko enoznačni. V vsakem primeru pa je mogoče zaključiti, da gre za probleme, ki jih mora (sproti) reševati uporabnojezikoslovni del stroke – in vrzeli med normo in rabo ne prepuščati v premoščanje lektorjem, učiteljem ali jezikovnim uporabnikom na splošno.

Literatura

- ARHAR, Špela, 2009: Učni korpus SSJ in leksikon besednih oblik za slovenščino. *Jezik in slovstvo* 54/3–4. 43–56.
- Gigafida, Korpus pisne slovenščine: <http://www.gigafida.net>.
- GOS, Korpus govornjene slovenščine: <http://www.korpus-gos.net>.
- GRČAR, Miha, KREK, Simon, DOBROVOLJC, Kaja, 2012: Obeliks: statistični oblikoskladenjski označevalnik in lematizator za slovenski jezik. Tomaž Erjavec, Jerneja Žganec Gros (ur.): *Zbornik Osme konference Jezikovne tehnologije*. Ljubljana: Institut Jožef Stefan. 89–94.
- KOSEM, Iztok, STRITAR, Mojca, MOŽE, Sara, ZWITTER VITEZ, Ana, ARHAR HOLDT, Špela, ROZMAN, Tadeja, 2012: *Analiza jezikovnih težav učencev: korpusni pristop*.

- Ljubljana: Trojina, zavod za uporabno slovenistiko.
- KREK, Simon, 2012: Spletni portal Slogovni priročnik. Boža Krakar Vogel (ur.): *Slavistika v regijah – Koper: Zbornik Slavističnega društva Slovenije 23*. Ljubljana: Zveza društev Slavistično društvo Slovenije. 225–231.
- KREK, Simon, ARHAR HOLDT, Špela, 2010: Slovenski besedilni korpusi: kako v razred? *Sodobna pedagogika* 61/1. 224–241.
- Leksikon besednih oblik Sloleks: <http://www.slovenscina.eu/sloleks>.
- LOGAR BERGINČ, Nataša, GRČAR, Miha, BRAKUS, Marko, ERJAVEC, Tomaž, ARHAR HOLDT, Špela, KREK, Simon, 2012: *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: gradnja, vsebina, uporaba*. Ljubljana: Trojina, zavod za uporabno slovenistiko in Fakulteta za družbene vede.
- Oblikoslovni označevalnik Obeliks: <http://oznacevalnik.slovenscina.eu>.
- Pedagoški slovnici portal: <http://slovnica.slovenscina.eu>.
- Projekt Sporazumevanje v slovenskem jeziku: www.slovenscina.eu.
- ROZMAN, Tadeja, KRAPŠ VODOPIVEC, Irena, STRITAR, Mojca, KOSEM, Iztok, 2012: *Empirični pogled na pouk slovenskega jezika*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Spletni portal Slogovni priročnik: <http://slogovni.slovenscina.eu>.
- Šolar, Korpus besedil učencev in dijakov: <http://www.korpus-solar.net>.
- VERDONIK, Darinka, ZWITTER VITEZ, Ana, 2011: *Slovenski govorni korpus Gos*. Ljubljana: Trojina, zavod za uporabno slovenistiko.