

PRAVOPISENE NAPAKE GOVORCEV SLOVENŠČINE KOT NEPRVEGA JEZIKA: KORPUSNI POGLED

Mojca Stritar Kučuk

Filozofska fakulteta, Univerza v Ljubljani, Ljubljana
mojca.stritarkucuk@ff.uni-lj.si

DOI:10.4312/Obdobja.43.327-337

Čeprav je za govorce slovenščine kot neprvega jezika do neke mere pričakovano poznavanje pravopisnih pravil, se večina s tem srečuje posredno ali šele na višjih stopnjah učenja jezika. Prispevek predstavlja analizo treh kategorij pravopisnih napak v korpusu KOST: najpogostejše med njimi so napake črkovanja, sledi jim neustrezna raba velike oz. male začetnice, najmanj pa je neustreznega zapisa besed skupaj, narazen oz. z vezajem.

slovenščina kot drugi in tuji jezik, korpus usvajanja drugega in tujega jezika, napake črkovanja, neustrezna raba velike oz. male začetnice, neustrezen zapis skupaj oz. narazen

Although speakers of Slovenian as a FL/L2 are expected to know Slovenian orthographic rules to some extent, these are mostly presented indirectly to them or only at higher levels of language learning. This article analyzes three categories of orthographic errors in the KOST corpus. The most common of these are spelling errors, followed by inappropriate capitalization of words and inappropriately joined or divided words.

Slovenian as a second and foreign language, learner corpus, spelling errors, capitalization errors, inappropriately joined or divided words

1 Uvod

Z vprašanji slovenskega pravopisa se govorci¹ slovenščine kot neprvega jezika srečajo ob samem začetku usvajanja slovenščine, saj začetniki že v uvodnih urah učenja spoznajo osnovni izgovor slovenskih črk (prim. npr. 1. enoto v učbenikih Jerman idr. 2023; Markovič idr. 2023; Petric Lasnik idr. 2023). Po drugi strani pa jim je večina nadaljnjih pravopisnih vsebin predstavljena prek obravnave drugih jezikovnih tem (Pirih Svetina 2021: 158). Dejansko so sistematično vključene samo v delovna zvezka učnih kompletov za izpopolnjevalce *Slovenska beseda v živo 3a* (Markovič idr. 2011) in *3b* (Markovič idr. 2018). To še ne pomeni, da se od govorcev slovenščine kot neprvega jezika ne zahteva vsaj določena mera pravopisne ustreznosti. Pravopisna zmožnost, ki se nanaša na »sposobnost prepisovanja, pravnega zapisa, na pravilnost oblikovanja besedila ter pravilne rabe ločil« (Dodatek k SEJU 2023: 131), je v *Skupnem evropskem jezikovnem okviru* natančneje razdelana v lestvici za ocenjevanje pravopisnega nadzora (SEJO 2011: 143). Vključena je tudi v program

1 V prispevku moško obliko uporabljam za katerikoli spol.

Slovenščina kot drugi in tuji jezik (2020), pri čemer pa se od govorcev slovenščine kot neprvega jezika vendarle ne zahteva, da dosegajo standarde domačih govorcev slovenščine. Pri govoricah na višji ravni so denimo dopuščene »nekatero manjše (slovnične in pravopisne) pomanjkljivosti« (prav tam: 50), na ravni odličnosti pa so »težko opazne« (prav tam: 55).

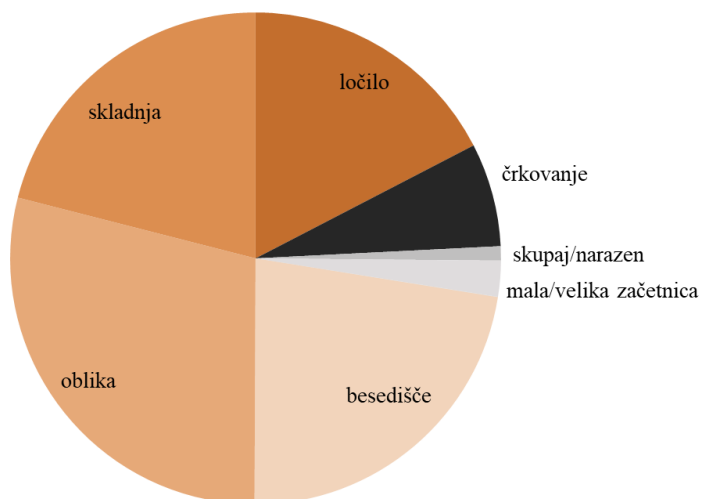
Namen pričujočega prispevka je na osnovi gradiva iz korpusa slovenščine kot tujega jezika KOST prepoznati in opisati pravopisna vprašanja, s katerimi imajo govorniki slovenščine kot neprvega jezika največ težav. S tem želim dopolniti starejše, pretežno na južnoslovanske govorce slovenščine osredotočene objave, ki so bile pripravljene na osnovi izkušenj avtorjev (prim. Šek 2021), anket (prim. Ulčnik, Nikolovski 2020) oz. priložnostno zbranih manjših korpusov (Požgaj Hadži, Ferbežar 2005; Balažič Bulc 2004). Prednost korpusnega gradiva je, da je do neke mere neizzvano – vsaj v tem smislu, da njegovi tvorci pri tvorjenju niso bili eksplicitno osredotočeni na pravopis. Po drugi strani je precej v korpus vključenih besedil nastalo brez zunanega nadzora nad okoliščinami tvorjenja (Stritar Kučuk 2022: 327), zato je možno, da so si tvorci pomagali z jezikovnimi orodji, ki lahko vplivajo na pravopisno ustreznost.

2 Analiza

Analiza je potekala na gradivu iz pisnega korpusa slovenščine kot tujega jezika KOST,² ki trenutno obsega več kot 1,5 milijona pojavnih in vključuje besedila vsaj srednješolsko izobraženih tujih govorcev. Večinoma prihajajo z govornega območja južnoslovanskih jezikov, večji pa so še deleži tistih, za katere je prvi jezik angleščina, italijanščina, nemščina, nizozemščina, poljščina, španščina, ruščina ali ukrajinščina (prim. Stritar Kučuk 2022). Na približno četrtini besedil v korpusu KOST so označene jezikovne napake. Ena od štirih osnovnih kategorij napak so napake zapisa (Stritar Kučuk 2023: 7–8), pri katerih gre za pisno podobo besedila oz. pravopisna vprašanja in predstavljajo več kot četrtino vseh označenih napak (Graf 1). Napake zapisa so razdeljene še na nekaj podtipov. Med njimi so najpogostejše napake ločil, ki jim je namenjen poseben prispevek v tem zborniku (Stritar Kučuk, Pisek, Šter), tu pa se bom osredotočila na tri podtipove napak zapisa: napake črkovanja, neustrezno rabo velike oz. male začetnice ter neustrezen zapis skupaj oz. narazen. Napake krajšav kot peti podtip napak zapisa se v korpusu KOST pojavijo samo desetkrat, zato jih v analizo nisem vključila.

Za vsakega od treh obravnavanih podtipov napak sem iz korpusa KOST izvozila vse primere napak. V programu Microsoft Excel sem jih nato ročno pregledala in jim pripisala dodatne oznake glede na formalno podobo napake. Pri vseh treh podtipih napak zapisa sem nato bolj kvalitativno označila specifični jezikovni problem. V zadnji fazi analize pa sem primere spet združila v širše kategorije, ki omogočajo delno kvantifikacijo in s tem lažje predstavljanje rezultatov.

2 Prim. <https://kost.cjvt.si/>.



Graf 1: Osnovni tipi napak po pogostnosti v korpusu KOST; napake zapisa so razdelane na podtype.

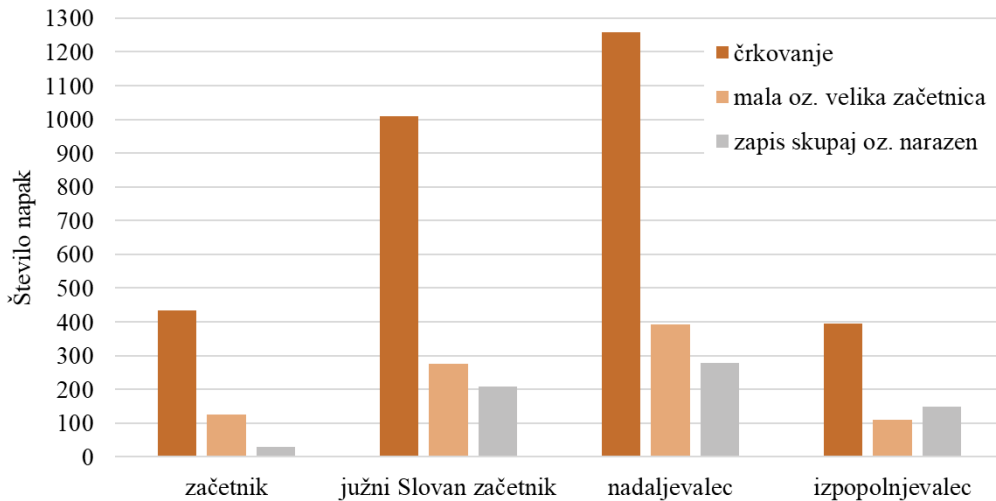
3 Rezultati

Po odstranitvi nerelevantnih oz. neustrezno označenih primerov so končne številke napak nekoliko nižje kot v različici korpusa KOST v spletnem konkordančniku. Za analizo je bilo relevantnih 2962 napak črkovanja, 902 primera neustrezne rabe velike oz. male začetnice in 414 primerov neustreznega zapisa skupaj oz. narazen. Podrobneje so posamezni tipi teh napak predstavljeni v nadaljevanju, osredotočimo pa se še na stopnje jezikovne zmožnosti tvorcev. V korpusu KOST je največ besedil južnoslovanskih začetnikov, sledijo deleži nadaljevalcev, izpopolnjevalcev in začetnikov (prim. Stritar Kučuk 2022). Pričakovali bi, da bo to razmerje ohranjeno tudi pri pravopisnih napakah (prim. Ulčnik, Nikolovski 2020: 287), vendar so v vseh treh kategorijah največ napak naredili nadaljevalci (Graf 2), samo pri izpopolnjevalcih pa je več napak zapisa skupaj oz. narazen kot napak velike oz. male začetnice, predvidoma zaradi rabe kompleksnejših, sestavljenih besed.

3.1 Napake črkovanja

Napake črkovanja, povezane z realizacijo zapisa slovenskih fonemov oz. z ustrezno črkovno podobo besed,³ so bile pri raziskovanju slovenščine kot neprvega jezika deležne največ pozornosti (prim. Ulčnik, Nikolovski 2022) in so bile v korpusu KOST najpogostejše med pregledanimi podtipi pravopisnih napak. Delajo jih govorniki vseh jezikov. Izrazito največ, več kot 470, so jih naredili govorniki srbsčine. Sledijo jim govorniki makedonščine (ok. 300), bosanščine in ruščine (več kot 250).

3 Ta tip napak se precej prekriva z drugimi kategorijami (prim. Može 2012: 51), zato so klasifikacije problemov v drugih analizah lahko drugačne (prim. rabo -o namesto -l v preteklem deležniku moškega spola v Požgaj Hadži, Ferbežar 2005 in Balažic Bulc 2004).



Graf 2: Razmerje med tipi pravopisnih napak in jezikovno zmožnostjo tvorcev.

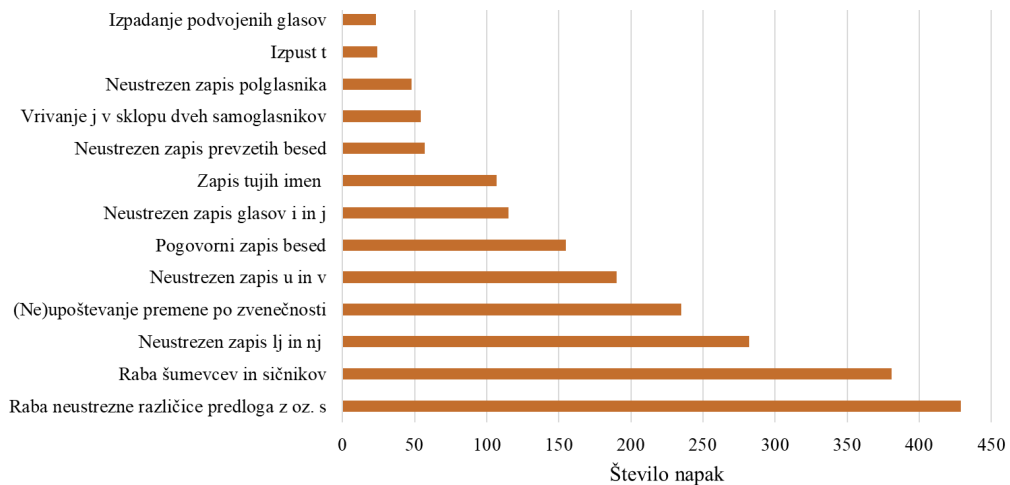
Največji del, torej slaba petina napak črkovanja (559), je v jezikoslovnem smislu manj zanimiv, saj zanj lahko s precejšnjo gotovostjo sklepam, da so posledica nepozornosti pri pisanju, bodisi na roko bodisi na računalnik: črke manjkajo (v *Soveniji*),⁴ so dodane (*rekreativnimi šport*) ali pa zamešane (*sva se odpeljlaa*), včasih je dodan oz. ponovljen celoten zlog (*razumumemo*). Tudi neustrezna naglasna znamenja (11 primerov, npr. nepotrebni ostrivec v primeru *človeške célice*) so verjetno posledica napake pri tipkanju.

Kot bolj sistemska pa lahko izpostavim vprašanja, predstavljena na Grafu 3. Najpogostejša je neustrezna raba predloga *z oz. s* (*s drugimi prijatelji*). Sledijo ji težave z rabo šumevcev, ki je verjetno po eni strani velikokrat posledica nepozornosti pri digitalnem načinu pisanja (Kranjc 2023: 17–18), po drugi strani pa je v mnogih primerih izpuščena npr. zgolj strešica na eni besedi v stavku (*vse vrste doživetij, ob cesti do Portoroza*). Tvorci uporabljajo šumevec namesto sičnika (*žvečer*) ali obratno (*božic*), težave imajo s sklopom *šč* (*slovenšine, praščiče*), pa tudi z zamenjevanjem *č* in *š* (*všasih*) ter *č* in *ž* (*zažel*).

Glavnina težav govorcev slovenščine kot neprvega jezika pri zapisu pa je verjetno posledica morfonološkega pravopisnega principa (Unuk 2009: 35). Težave s sklopoma *lj* in *nj* se lahko realizirajo v različnih oblikah. Beseda *Ljubljana* je v korpusu KOST zapisana tudi *Ljubjana, Ljublana, Ljubliana, Ljubljana, Ljubljana, v Ljublani*. Črka *j* se vriva v besede, kjer je v slovenščini niti ne izgovarjamo (*računalnjik*). Neustrezen zapis glasov zaradi neupoštevanja premene po zvonečnosti (*dogotki* nam. *dogodki, iskušnje* nam. *izkušnje, angleš* nam. *Anglež, jas* namesto *jaz*) je problem za govorce vseh jezikov. Pri četrtini primerov neustreznega zapisa *u* in *v* gre za predlog *v* (*u Ljublani*). Med govorce, ki slovenščino očitno usvajajo tudi iz okolja, je pogost

4 Vsi primeri so iz korpusa KOST. Prepisani so v izvorni obliki. Morebitne poudarke sem naredila avtorica tega prispevka.

fonetični zapis besed (*tut* ali *tud* nam. *tudi*, *blo* nam. *bilo*, *tuki* nam. *tukaj*, *drugač* nam. *drugače*). Razmeroma pogosta težava govorcev slovenščine kot neprvega jezika pa je še zapis *i* in *j*. Pri tem izpuščajo *i* (*Italja*) ali *j* (*tretina*), na mesta, kjer jih v standardni slovenščini ni, vstavljajo *i* (*piščančijo hrano*) ali *j* (*mladji*) ter pišejo *i* namesto *j* (*pogovarjamo*) ali obratno (*funkcionirati*). Pri zapisu tujih imen se pojavlja podomačeni zapis namesto tujega (*Ridli Skot* nam. *Ridley Scott*) ali obratno (*Paris* nam. *Pariz*). Zapisovalne težave so pogoste pri besedah, ki jih tvorci dojemajo kot prevzete oz. tujega izvora (*corona* nam. *korona*). Še dva pričakovana pravopisna problema v zvezi s slovenskim črkovanjem sta vrivanje oz. izpust črke *j* v zevu oz. sklopu dveh samoglasnikov (*karijera* nam. *kariera*, *violičen* nam. *vijoličen*) in vrivanje ali izpust črke *e* pri zapisu polglasnika (*magisterski* nam. *magistrski*, *semestr* nam. *semester*). V nekaterih besedah tvorci izpuščajo črko *t* (*korisne* nam. *koristne*). Nekajkrat se pojavijo tudi težave v besedah, ki imajo podvojene enake črke (*štiriindvajset* nam. *štiriindvajset*, *oodih* nam. *oddih*).



Graf 3: Najpogostejši jezikovni problemi, povezani z neustreznim črkovanjem v korpusu KOST.

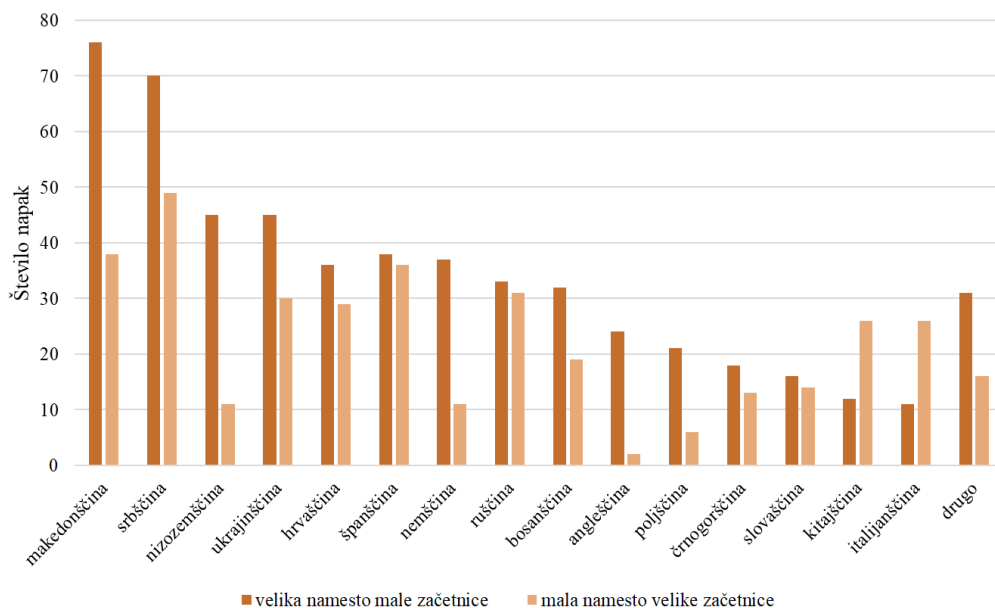
Med problemi, ki se v korpusu KOST pojavijo z 20 primeri ali manj, so še raba neslovenskih črk (*božič*, *myslim*), zamenjevanje *m* in *n* (*hrama*), *b* in *v* (*uporavljam*), vstavljanje odvečnih *m* (*oktobra*), *n* (*Istanbul*) ali *a* (*amapak*), izpust *m* (*pomebno*) ali *n* (*na konferenci*), zapis *ie* namesto *i* (*veliko treniera*).⁵

5 Med napake črkovanja sodi tudi neustrezen zapis števil in datumov (*01. 01. 2022* nam. *1. 1. 2022*) ter števnikov v odvisnih sklonih (*od 66h skladišč* nam. *iz 66 skladišč*). Ker pri tem ne gre zgolj za zapis glasov, sem jih izpustila iz tu predstavljene analize.

3.2 Neustrezna raba velike oz. male začetnice

Neustrezna raba velike oz. male začetnice je tipično pravopisni problem, ki v primerjavi z napakami črkovanja manj vpliva na razumljivost besedila in je pri učenju slovenščine zelo v ozadju – ne nazadnje ni vključen v nobeno obstoječe učno gradivo.

Po formalni plati je v skoraj dveh tretjinah primerov tvorec neustrezno uporabil veliko namesto male začetnice, v dobri tretjini primerov pa obratno. Pregled teh napak glede na prvi jezik tvorcev (Graf 4)⁶ pa pokaže, da južni Slovani in večina drugih govorcev večkrat zapisujejo veliko namesto male začetnice, govorniki španščine, ruščine in slovaščine precej izenačeno delajo obe vrsti napak, medtem ko govorniki kitajščine in italijanščine izrazito pogosteje uporabljajo malo namesto velike začetnice. Upoštevati moramo, da vzorci besedil za nekatere jezike niso dovolj veliki, da bi na njihovi osnovi lahko posploševali brez natančnejše analize.



Graf 4: Napake velike oz. male začetnice glede na prvi jezik tvorcev.

3.2.1 Neustrezna raba velike namesto male začetnice

Vsaj desetina primerov neustrezne rabe velike namesto male začetnice v korpusu KOST se pojavi le nekajkrat in pri posameznem govorniku, zato lahko predvidevam, da gre za lapsuse ali nesistemske težave, kar je v zanimivem nasprotju z domačimi govorniki slovenščine (Može 2012: 48). Tisti problemi, ki se v korpusu KOST večkrat ponovijo, so prikazani v Tabeli 1.

⁶ Med jeziki, zajetimi v kategoriji *drugo*, so še bolgarščina, češčina, francoščina, igbo, indonezijsščina, kirgiščina, korejščina, madžarščina, portugalsščina, romunščina, slovensščina in turščina.

Jezikovni problem: neustrezen zapis ...	Primer⁷	Št. pojavitev
Vrstni pridevnik na <i>-ski</i> oz. <i>-ški</i>	v <i>Dunajsko zimo</i>	63
Zemljepisna imena, deli naselij ipd.	v <i>Črni Gori, na Prelaz Vršič, na Štepanjskom Naselju</i>	53
Jeziki	<i>malo Slovensko</i>	44
Prazniki	<i>za Novo leto</i>	37
Imena fakultet ali šol	<i>na Fkulteti za Računalništvo in Informatiko</i>	34
Raba za vejico, ki zaključuje nagovor v pismu	<i>Spoštovani kolegi in profesorica, Pišem vam iz, ne tako sončne, Ljubljane.</i>	33
Prireditve, zgodovinski dogodki ali obdobja	<i>po Druge Svetovne Vojne</i>	23
Imena podjetij ali ustanov	<i>na Veleposlaništvu Slovenije, uslužbenci Evropske Komisije</i>	23
Meseci	<i>od Septembra do Novembra</i>	20
Študijske smeri ali šolski predmeti	<i>študirati Kemijo</i>	16
Raba za dvopičjem ali oklepajem	<i>imam štiri elektronske naslove: Enega za službo, drugega za univerzo</i>	12
Dnevi	<i>v Nedeljo</i>	10
Avtomobilske in druge blagovne znamke	<i>vozijo Vespo</i>	10

Tabela 1: Jezikovni problemi, povezani z neustrezno rabo velike namesto male začetnice v korpusu KOST.

3.2.2 Neustrezna raba male namesto velike začetnice

Obratni tip težav je precej redkejši (Tabela 2). Tudi tu gre v veliki meri za individualne probleme, pri katerih je mogoče predvidevati, da gre za lapsuse, ali pa je težko iskati skupne točke med njimi. To lahko ugotovim že za neustrezen zapis prve besede v povedi z malo začetnico, ki se v korpusu KOST pojavi 64-krat: ker ima večina sodelujočih tvorcev vsaj srednješolsko izobrazbo, je do tega verjetno prišlo zaradi nepozornosti pri pisanju in ne zaradi pomanjkljive jezikovne zmožnosti.

7 Neustrezno zapisana črka je odebeljena. V primerih so ohranjene tudi morebitne druge napake.

Jezikovni problem: neustrezen zapis ...	Primer	Št. pojavitev
Pripadniki narodnosti	<i>veliko prijateljev makedoncev, bošnjakov, slovencev, črnogorcev in srbov</i>	79
Države	<i>v sovjetski zvezi</i>	38
Druga zemljepisna imena	<i>na jugu jadranskega morja</i>	34
Imena fakultet	<i>Študiram violončelo na akademiji za glasbo v Ljubljani.</i>	19
Imena drugih podjetij in ustanov	<i>Zdaj sem vsak dan v centralni tehniški knjižnici</i>	13
Naslovi filmov, knjig, glasbenih skupin in raznih umetniških stvaritev	<i>v seriji mama, poslušam depeche mode</i>	13

Tabela 2: Jezikovni problemi, povezani z neustrezno rabo male namesto velike začetnice v korpusu KOST.

3.3 Neustrezen zapis skupaj, narazen oz. z vezajem

Tudi napake zapisa besed skupaj oz. narazen so v veliki meri vezane na formalno plat jezika in nimajo večjega vpliva na razumljivost besedil. Po zunanji podobi tvorci v skoraj polovici primerov besedo, ki bi morala biti zapisana skupaj, pišejo narazen (196 napak), malo manjši delež napak tega tipa je obratnih (167 napak). Izmed teh sem iz nadaljnje analize izločila 31 primerov, pri katerih gre za očiten lapsus (*mi jepomagalo duševno*). Pri 48 napakah gre za neustrezno rabo vezaja, največkrat v povezavi s pregibanjem prevzetih besed (*v nevarni svet Facebook-a*). Nekajkrat se vezaj pojavi pri pisanju samostalniških besednih zvez (*v Sankt-Peterburgu*). Za razliko od napak velike oz. male začetnice, pri katerih gre v veliki meri za individualne težave, pa se zdi, da je večina neustreznih primerov zapisa skupaj oz. sistemske narave.

3.3.1 Neustrezen zapis narazen namesto skupaj

Iz Tabele 3 je razvidno, da imajo govorci slovenščine kot neprvega jezika pri besedah, ki jih pišejo narazen namesto skupaj, največ težav, ko prvi del besede razumejo kot predlog (*do poldne* nam. *dopoldne*, *k malu* nam. *kmalu*). Pri nekaterih od teh besed je v določenih kontekstih dejansko mogoč tudi zapis narazen (*potem*, *zato*). Takšno razumevanje besed kot predložnih zvez se prenaša tudi na nekatere nezložene veznike (*am pak* nam. *ampak*). Sorodna težava so sklopljeni členki (*za res* nam. *zares*, *vse eno* nam. *vseeno*), tuji govorci pa imajo pogosto težave še z zapisom naveznih oblik osebnih zaimkov (prim. Požgaj Hadži, Ferbežar 2005: 141).

Jezikovni problem	Primer	Št. pojavitev
Zapis nekaterih sestavljenih prislovov, predlogov in veznikov	<i>Za radi tega</i>	43
Zapis naveznih oblik osebnih zaimkov	<i>službo za me</i>	36
Zapis prislovov s <i>krat</i>	<i>en krat na teden</i>	25
Zapis zanikanega glagola <i>biti</i>	<i>ni so dolgočasne</i>	18
Zapis členkov	<i>prav za prav</i>	16

Tabela 3: Jezikovni problemi, povezani z neustreznim zapisom besed narazen namesto skupaj v korpusu KOST.

Redkeje se v KOST-u pojavijo še težave pri zapisu tvorjenih pridevnikov (*južno ameriškem* nam. *južnoameriškem*), zanikanih pridevnikov oz. prislovov, nekaterih sodobnih oz. tujih imen, ki ne sledijo ustaljenim pravilom slovenskega pravopisa (*Power Point* nam. *PowerPoint*), glavnih ali vrstilnih števnikov (*štiri deset* nam. *štirideset*), tvorjenih samostalnikov (*trideset letniki* nam. *tridesetletniki*), stopnjevanih oblik pridevnika ali prislova (*naj pogostejši* nam. *najpogostejši*). 14 primerov pa so tvorci pisali narazen namesto z vezajem: sestavljenke s številko ali krajšavo v prvem delu (*5 krat* nam. *5-krat*) in priredno zložene pridevnike (*črno bela* nam. *črno-bela*).

3.3.2 Neustrezen zapis skupaj namesto narazen

Obratni tip težav, torej neustrezen zapis skupaj namesto narazen, je nekoliko redkejši (Tabela 4). Prednjači zapis predloga skupaj z besedo, ki mu sledi (*naprimer* nam. *na primer*). Bistveno manj je primerov, v katerih je nikalnica pisana skupaj z glagolom v sedanjiku, kar je sicer najpogostejši problem pri domačih govornicah slovenščine (Može 2012: 50).

Jezikovni problem	Primer	Št. pojavitev
Zapis predloga	<i>Bolj mi ustreza učenje v živo</i>	62
Zapis zanikanega glagola	<i>Nevem, zakaj me to toliko presenetilo.</i>	22
Zapis merske enote stično s številko	<i>Visok je 215cm in ima 115kg.</i>	18

Tabela 4: Jezikovni problemi, povezani z neustreznim zapisom besed skupaj namesto narazen v korpusu KOST.

V nekaj primerih so skupaj neustrezno pisane še zveze s *čim* (*čimprej* nam. *čim prej*), *vsak dan* in *novo leto*.

4 Zaključek

Moja analiza pisne produkcije govorcev slovenščine kot neprvega jezika iz korpusa KOST je pokazala, da se tudi oni soočajo s pravopisnimi težavami: najbolj pri zapisu slovenskih glasov, kjer je vrsta problemov v veliki meri odvisna od prvega jezika učečega se. Precej sistemskih problemov imajo pri zapisu besed skupaj oz. narazen, medtem ko je težav z rabo velike oz. male začetnice sicer več, a se zdi, da so bolj

nesistemske. Zanimivo splošno opažanje je, da pojavitev napake ni nujno povezana s primanjkljajem v tvorčevi jezikovni zmožnosti, pač pa lahko kaže samo na njegov slab nadzor nad pisanjem. Pri vseh pregledanih tipih napak se je zgodilo, da je tvorec v isti povedi napako naredil in hkrati ne. Tako je npr. z zapisom zanikanega glagola v naslednjem primeru: *Sem visok ni sem suh, ampak nisem ni debel*.

Omeniti moram omejitve same korpusne metode: nekateri tipični problemi, ki jih omenja strokovna literatura, so se v korpusu KOST pojavili le nekajkrat ali pa sploh ne. Zamenjevanje črk *h* in *j*, ki bi ga lahko pričakovali pri govorecih španščine, je denimo v korpusu zelo redko, saj se pojavi trikrat (*včasij* nam. *včasih*). To nikakor ne pomeni, da ti problemi za govorce slovenščine kot neprvega jezika niso relevantni. Na treh četrtinah gradiva iz korpusa KOST jezikovne napake niso označene in to gradivo ni bilo vključeno v analizo. Morda se tovrstne napake niso pojavile tudi zaradi specifičnih tem pisanja ali pa zaradi dejstva, da so tvorca v KOST vključenih besedil predvidoma sorazmerno dobro pisмени v svojem jeziku (prim. Ulčnik, Nikolovski 2022: 288). Analiza korpusnega gradiva drugačne populacije bi verjetno prinesla drugačne rezultate.

Kako te ugotovitve aplicirati v praksi? Ker se pravopisna ustreznost predvsem na višjih ravneh jezikovne zmožnosti do neke mere zahteva od učečih se, bi kazalo tovrstne vsebine nekoliko bolj vključiti v učna gradiva in pedagoški proces, seveda z zavedanjem potreb ciljne publike (prim. Piriš Svetina 2021: 160) in dejstva, da morajo prednost vendarle imeti vidiki jezikovne zmožnosti, ki so primarnjši za uspešno sporazumevanje. S korpusnim gradivom bi bilo smiselno dopolniti tudi črkovalnike in sorodne računalniške programe (Leacock idr. 2015), ki, kadar so zasnovani samo na gradivu domačih govorcev, na gradivu učečih se večinoma ne delujejo dobro (prim. Bexte idr. 2022).

Literatura

- BALAŽIČ BULC, Tatjana, 2004: Jezikovni prenos pri učenju sorodnih jezikov: na primeru slovenščine in srbohrvaščine. *Jezik in slovstvo* XLIX/3–4. 77–89.
- BEXTE, Marie, LAARMANN-QUANTE, Ronja, HORBACH, Andrea, ZESCH, Torsten, 2022: LESPELL – A Multi-Lingual Benchmark Corpus of Spelling Errors to Develop Spellchecking Methods for Learner Language. Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Jan Odič, Stelios Piperidis (ur.): *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*. Marseille: European Language Resources Association. 697–706.
- Dodatek k SEJU = Svet Evrope, 2023: *Skupni evropski jezikovni okvir: učenje, poučevanje, ocenjevanje – dodatek*. Ljubljana: Zavod RS za šolstvo. www.zrss.si/pdf/SEJO.pdf (dostop april 2024).
- JERMAN, Tanja, PISEK, Staša, STRAJNAR, Anja, 2023: *Slovenščina ekspres 1: učbenik za začetnike na tečajih slovenščine kot drugega in tujega jezika*. Ljubljana: Založba Univerze v Ljubljani.
- KRANJC, Simona, 2023: Od medmrežnih pogovorov do ChatGPT-ja. Jerca Vogel (ur.): *Slovenski jezik, literatura, kultura in digitalni svet(ovi)*. 59. seminar slovenskega jezika, literature in kulture. Ljubljana: Založba Univerze v Ljubljani. 11–20.
- LEACOCK, Claudia, CHODOROW, Martin, TETREAULT, Joel, 2015: Automatic Grammar- and Spell-Checking for Language Learners. Sylviane Granger, Gaëtanelle Gilquin, Fanny Meunier (ur.): *The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press. 567–586.

- MARKOVIČ, Andreja, KNEZ, Mihaela, ŠOBA, Nina, STRITAR, Mojca, 2011: *Slovenska beseda v živo 3a: delovni zvezek*. Ljubljana: Znanstvena založba Filozofske fakultete.
- MARKOVIČ, Andreja, KNEZ, Mihaela, ŠOBA, Nina, STRITAR, Mojca, 2018: *Slovenska beseda v živo 3b: delovni zvezek*. Ljubljana: Znanstvena založba Filozofske fakultete.
- MARKOVIČ, Andreja, STRITAR, Mojca, JERMAN, Tanja, PISEK, Staša, 2023: *Slovenska beseda v živo 1a: učbenik za začetni tečaj slovenščine kot drugega/tujega jezika*. Ljubljana: Založba Univerze v Ljubljani.
- MOŽE, Sara, 2012: Napake zapisa. Iztok Kosem, Mojca Stritar, Sara Može, Ana Zwitter Vitez, Špela Arhar Holdt, Tadeja Rozman: *Analiza jezikovnih težav učencev: korpusni pristop*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- PETRIC LASNIK, Ivana, PIRIH SVETINA, Nataša, PONIKVAR, Andreja, 2023: *A, B, C ... gremo: učbenik za začetnike na kratkih tečajih slovenščine kot drugega ali tujega jezika*. Ljubljana: Založba Univerze v Ljubljani.
- PIRIH SVETINA, Nataša, 2021: Od kod do kod slovenščina?: pravorečna in pravopisna vprašanja pri poučevanju slovenščine kot tujega jezika. *Jezik in slovstvo* LXVI/2–3. 157–165.
- POŽGAJ HADŽI, Vesna, FERBEŽAR, Ina, 2005: Tekst između slovenskog i srpskog jezika: tipične pisane pogreške na ispitima iz znanja slovenskog kao drugog jezika. *Prilozi proučavanju jezika* XXXVI. 137–146.
- SEJO = *Skupni evropski jezikovni okvir: učenje, poučevanje, ocenjevanje*, 2011. Ljubljana: Ministrstvo RS za šolstvo in šport, Urad za razvoj šolstva.
- Slovenščina kot drugi in tuji jezik: izobraževalni program za odrasle*, 2020. Ljubljana: Center za slovenščino kot drugi in tuji jezik, Filozofska fakulteta Univerze v Ljubljani. <https://centerslo.si/knjige/ucbeniki-in-prirocniki/prirocniki-in-ucno-gradivo/slovenscina-kot-drugi-in-tuji-jezik/> (dostop april 2024).
- STRITAR KUČUK, Mojca, 2022: KOST med korpusi usvajanja tujega jezika. Nataša Pirih Svetina, Ina Ferbežar (ur.): *Na stičišču svetov: slovenščina kot drugi in tuji jezik. Obdobja 41*. Ljubljana: Založba Univerze v Ljubljani. 323–334.
- STRITAR KUČUK, Mojca, 2023: *Priročnik za označevanje napak*. <https://www.cjvt.si/korpus-kost/wp-content/uploads/sites/24/2023/10/Prirocnik-za-oznacevanje-napak-v-KOST-u-2023-10-17.pdf> (dostop 26. 10. 2023).
- ŠEK, Polonca, 2021: Razvijanje pravopisne zmožnosti pri učencih in učenkah, ki jim je slovenščina drugi ali tuji jezik. Dragica Haramija, Simona Pulko (ur.): *Pogledi na slovenščino kot neprvi jezik v zgodnjem izobraževanju*. Maribor: Univerza v Mariboru, Univerzitetna založba. 171–193.
- ULČNIK, Natalija, NIKOLOVSKI, Gjoko, 2022: Zapisovalne težave govork in govorcev južnoslovanskih jezikov pri učenju slovenščine. Simona Pulko, Melita Zemljak Jontes (ur.): *Slovenščina kot drugi in tuji jezik v izobraževanju*. Maribor: Univerza v Mariboru, Univerzitetna založba. 273–290.
- UNUK, Drago, 2009: Pravopisna načela v slovenskem pravopisu. *Revija za elementarno izobraževanje* II/4. 27–36.

