

Darinka Verdonik

Fakulteta za elektrotehniko, računalništvo in informatiko Univerze v Mariboru

Mesto govora v slovarju slovenskega jezika

O tem, kje je mesto govora v splošnem slovarju in na kak način opisovati značilnosti govornega jezika, lahko diskutiramo z veliko vidikov. Da bi konstruktivno prispevali k razmišljanju o novem slovarju slovenskega jezika, predstavimo v prispevku predlog načrta, kako opisovati govorni jezik v njem. Predlog izhaja iz prepričanja, da je treba posebnosti govorne rabe opisovati na vseh jezikoslovnih ravneh, ne samo na ravni izgovora in regijsko značilnih izgovornih različic.

Spoken language in the dictionary of Slovene

What is the place of speech in the general dictionary and how to describe the properties of spoken language? We could discuss about this from many viewpoint. With the aim to make a constructive contribution to discussions on new Slovene dictionary, this paper introduces a possible work plan for description of spoken language in it. The work plan is based on the belief that the characteristics of spoken language should be described on all linguistic levels, not merely on phonetic level and regionally specific pronunciations.

Ključne besede: govorni jezik, govor, izgovor, koncept, izvedba

Keywords: spoken language, speech, pronunciation, concept, realization

1 Uvod

Mesto govora v slovarju je do neke mere povezano z bolj temeljnim vprašanjem – mestom govora v jezikoslovju. Čeprav danes marsikje, tudi pri nas, še vedno prevladujoče velja iz zapisov po de Saussurju izpeljano načelo, da v »govoru« (tukaj v pomenu kot pri de Saussurju 1997) ni ničesar jezikoslovno relevantnega, kar ne bi bilo opazno že v »jeziku«, Škiljan (2008) ob pregledu zgodovine govora v jezikoslovju ugotavlja, »da so obdobja, v katerih je bil interes znanstvenikov usmerjen predvsem v govor, daljša (za 'prazgodovino' lingvistike pa morda tudi pomembnejša) od obdobj, ko je bil v ospredju jezikovni sistem« (Škiljan 2008:

30). Škiljan pri tem sicer misli »govor« v pomenu, kot ga uporablja de Saussure (1997), vendarle pa ta v veliki meri vključuje tudi govorno rabo jezika. V zvezi s tem navede Škiljan kot primer med drugim obsežno zanimanje za retoriko, ki je od antike naprej močno vplivala na izobraževalne sisteme vse do 18. stoletja. Vendar utemeljuje, da tudi strukturalno jezikoslovje, čeprav se skuša izogibati vprašanju »govora«, le-tem v večini primerov ne uide povsem – jezikoslovje je vendarle humanistična znanost in ne more (in tudi ne sme!) odmisлити Človeka. V sodobnem jezikoslovju zadnjih nekaj desetletij lahko opazujemo nov pohod jezikoslovja »govora« (sociolingvistika, psiholingvistika, pragmatično jezikoslovje, konverzacijska analiza, etnografija govora, diskurzne študije ipd.), ki obravnava najrazličnejše vidike jezikovne rabe (seveda ne samo govorne, ampak tudi te), vendar te smeri niso uspele in v resnici niti niso zares poskušale prenavljati temeljev jezikoslovja, ampak so ostale komplementarne jezikoslovju »jezika« kot sistema znakov. Izjema so morda funkcijske slovnice (Halliday 1994; Dik 1997), ki pa tudi niso povsem prepričljive zaradi pretirano poudarjene družbene razsežnosti jezika, medtem ko zanemarjajo njegov kognitivni vidik.

Ob tem se moramo ustaviti ob terminološki zadregi v zvezi z govorom: termin se lahko nanaša na jezikovno rabo v kontekstu (ki je pisna ali govorna) ali pa samo na govorno rabo, ki pa se po drugi strani odraža tudi v jezikovnem sistemu na ravni glasoslovja (in delno oblikoslovja – npr. naglasni vzorci pri pregibanju). Prvi pomen je vezan predvsem na strukturalistično jezikoslovje in zato precej specifičen; v tem prispevku govorimo o govoru v drugem pomenu, kot o govorni rabi jezika.

Zadrega s terminom govor sicer ni povsem slučajna: v jezikoslovju 20. stoletja lahko ločimo smeri, ki raziskujejo jezik kot sistem, ki je zunaj- ali vsaj nadkontekstni, in smeri, ki zagovarjajo potrebo po preučevanju jezikovne rabe v kontekstu (»govor«, kot ga opredeli de Saussure), vendar pregled slednjih pokaže, da pogosto izvirajo prav iz opazovanja govorne rabe jezika (npr. Austin 1990; Grice 1975; Firth 1964; Hymes 1974; konverzacijska analiza). Zdi se, da je zadrega povezana z razlikami med govorno in pisno rabo, ki se jih zavedamo globoko v sebi. Tako že Malinowski (1946) na primer pripomni, da je govorni prenosnik vpet v vsakdanje aktivnosti in zato bistveno bolj odvisen od konteksta kot pisni. Različne situacije in nameni, v katerih se rabi en ali drugi kanal, zagotovo veliko prispevajo k razlikam v jeziku govorne in pisne rabe. S tem povezani sta interaktivnost na eni in neinteraktivnost na drugi strani. Tudi delna zunanja reguliranost enega in nereguliranost drugega (vsaj v zasebni sferi) ima pomemben vpliv. Medtem ko se govorna raba vsaj v zasebni sferi, delno pa tudi javni

razvija prek nezavednih vzorcev, je pisna raba delno umetno ustvarjena kot plod človekove zavedne dejavnosti in je tudi na institucionalni ravni regulirana/normirana. Čeprav torej globoko v sebi vemo za kar nekaj pomembnih vzvodov, ki signalizirajo, da se govorjena in pisna raba zelo razlikujeta, pa prav veliko sistematičnih študij o tem še vedno ni narejenih. Ena redkih obsežnejših (Henrichsen, Allwood 2005) pa dokazuje, da so razlike med obema veliko večje, kot je predpostavljala strukturalizem. V kontrastivni korpusni analizi leksike v govorjeni in pisni rabi namreč ugotavlja več sorodnosti med govorjeno rabo v dveh različnih sorodnih jezikih (švedskim in danskim) kot pa med govorjeno in pisno rabo v istem jeziku. Tak rezultat je kljub vsemu zgoraj zapisanemu skoraj šokanten.

Če se torej vprašamo, ali je jezikoslovje pisnega jezika in »jezika« kot sistema zadostno jezikoslovje, je odgovor na dlani: ni. In če to širše vprašanje zožimo na tematiko, ki nas zanima tukaj – ali je splošni slovar, ki opisuje samo pisni jezik, in ne tudi govorjenega, zadosten – je odgovor enak: ni. Pritrjuje mu domača slovaropisna tradicija: v SSKJ je jasna tendenca, da se vključujejo gesla, očitno izhajajoča iz govorjene rabe – narečni polnopomenski izrazi *šenkati*, *šnofati*, *gate*, *kripa*, *malora*, medmetni izrazi tipa *ma*, *ej*, prislovni *ja*, *dobro*, *veste/veš* (vse v pogovorni rabi), če naštejemo samo nekaj naključnih, so vsi zajeti vanj. Na govor pa se seveda nanaša tudi izgovor besed, prav tako del informacij v slovarjih. Svoje mesto v slovarju govor tako brez dvoma ima in mora imeti.

Toda kje je to mesto? Na kak način opisovati značilnosti govorjenega jezika v sodobnem slovarju? O tem bi lahko diskutirali z veliko vidikov in prostora za vse tukaj ni. Da bi konstruktivno prispevala k razmišljanju o novem slovarju slovenskega jezika, predstavim v nadaljevanju predlog načrta, kako opisovati govorjeni jezik v njem.¹

2 Koncept opisa govorjenega jezika v slovarju

V prispevku izhajam iz predloga treh avtorjev o Slovarju sodobnega slovenskega jezika (SSSJ) (Krek idr. 2013), ki je ta hip edini javno dostopen in podrobno izdelan predlog zasnove novega slovarja.

¹ Predlog je izdelan samo vsebinsko, ni pa tudi finančno ocenjen in ni zajet v izvirnem predlogu avtorjev SSSJ (Krek idr. 2013).

Tukaj in zdaj lahko razmišljamo samo o enotnem opisu govornega in pisnega jezika v enem slovarju. Pri tem pa iz namena in tradicije slovarja jasno izhaja prvenstvena vloga pisnega jezika, govorni jezik tako v slovarju opisujemo v razmerju do pisnega. To pomeni, da se vsi elementi standardnega pisnega jezika predpostavijo kot nevtralni za govorno rabo nasploh, razen kjer so dodani opisi, ki podrobneje specificirajo govorno rabo. Ti opisi bodo zaradi majhnega obsega govornega gradiva spočetka zelo omejeni, vendar upamo, da se bodo sčasoma dopolnjevali in postajali vse bolj izpopolnjeni (in seveda tudi ažurirani).

Vpetost opisa govornega jezika v slovar mora biti čim bolj sistematična na prav vseh ravneh obdelave podatkov ter priprave, zasnove in redakcije iztočnic. Opis se mora torej začeti že na ravni sestavljanja geslovnika: neverbalni in polverbalni izrazi tipa *ija, jo*, ekspresivni izrazi tipa *alora, porkaduš, majkemi*, pozdravni izrazi tipa *ajde*, široko razširjeni izrazi tipa *pol* (kot prislov v pomenu *potem*), regionalno značilni in narečni izrazi tipa *gljeh, hcoj* in funkcijske besede (*ta rdeč, kaj pa te ti tu delaš, čakaj ka vidim*) – vsem takim in drugim posebnostim besedišča govornega jezika je treba slediti in jih sistematično, po izdelanih kriterijih, vključevati.

Naslednji nivo opisa govornega jezika je pomenski. Ne samo besede, značilne za govorni jezik, ampak tudi nekatere splošne besede lahko dobijo nove pomene in vloge, nepoznane v pisni rabi, ki jih je treba ustrezno opisati pod zavihkom o pomenu. Morda najbolj značilne take besede so razni pragmatični izrazi v interakciji, npr. diskurzni označevalci tipa *dobro (samo pač dobro rabiš več zemle ne), lej (no skratka lej), veš (ja veš je kar ogromno), zdaj (zdaj če pogledate še bolj natančno)*, ekspresivni izrazi tipa *madona (madona to je pa dobr), pismo (pismo ta ta kikla mi more tako prstajat)*, po inventivni rabi jezika je znan mladostniški jezik, npr. *carsko (o fak hudo carsko)*, itd. Morda še več posebnosti lahko pričakujemo na ravni stalnih besednih zvez, kjer kot značilne za govorno rabo najdemo razne pragmatične izraze, kot so *pa to, ali pa tako, v bistvu, pa daj*, ekspresivne izraze tipa *ti boga, jezus marija, daj ga svirat*, slengovske zveze tipa *ful fajn, ful dobro, ni panike* itd.

Opis posebnosti govora se nadaljuje na ravni opisa oblikoskladenjskih vzorcev, torej v zavihku o oblikah, npr. za besedo *dež* najdemo v petem sklonu obliko *po deži* namesto standardne *po dežju*, polglasnik marsikje ne izpada itd. Vendar pa oblike v slovarjih podajamo praviloma skupaj z informacijo o izgovoru (tako v tradiciji slovaropisja kot tudi v najnovejšem predlogu SSSJ), s tem pa se posebnosti oblikoskladenjskih vzorcev hitro

pomnožijo na precejšen obseg. Poznamo npr. široko razširjene glasovne preme pri glagolskem deležniku na -l v tretji osebi ednine (poleg normativnega *r/ekeU* še *r/eku*, *r*eko* in številne druge, niha pa lahko tudi samó naglasno mesto, npr. glagolske dvojnice tipa *pl/esala* proti *ples/ala*).

Z upoštevanjem naglasnih premen pri pregibanju smo že globoko v vprašanih izgovora. Govorjeni jezik je poseben na vseh ravneh, vendar pa je samo zunanja zvočna podoba – s stališča slovarja predvsem izgovor besed tisti, ki ga ne moremo opisati na podlagi opusa pisnega jezika, ampak je v celoti odvisen od govornega gradiva. V slovarju, ki mora biti, če naj bo v koraku s sodobnimi jezikoslovnimi tokovi, gradivno utemeljen, mora tudi opis izgovora kolikor mogoče temeljiti na gradivu. Vendar mora slovar vsaj t. i. standardni izgovor predvideti za prav vsako iztočnico in vsako posamezno besedno obliko te iztočnice, tudi če govornega gradiva zanjo ni – in marsikje ga ni. Opisovanje standardnega izgovora predvsem na podlagi jezikoslovne teorije bo zato v slovarju še vedno nujno, vseeno pa je treba teorijo na vsakem koraku preverjati, ali so na razpolago novejši gradivni podatki, ki jo ali potrjujejo ali zahtevajo spremembo.

Ob tem bo potreben ponoven razmislek, kaj nam pomeni standardni izgovor. Stroka se mora pri tem zavedati vseh vidikov procesa standardizacije in normativnosti in v zvezi z govorom ob tem napotujem bralca vsaj na zapis Stabeja (2008), čeprav seveda ni edini, ki se je spraševal o tem (gl. npr. še Pogorelec 2008). Standardizacija je družbena, jezikovnopolitična »analoga« jezikoslovja in zdi se, da jo bo iz takih ali drugačnih razlogov še nekaj časa opravljalo. Splošni slovar ima pri tem, če hoče ali ne, zelo pomembno, morda celo ključno vlogo. V zvezi s temi vprašanji, ki so preobsežna, da bi jih lahko obravnavala tukaj, predlagam med drugim vsaj ponoven razmislek, katere situacije so v današnjem času tiste, v katerih se rabi pretežno standardni govorni jezik. Na podlagi tega se lahko poskuša oblikovati uravnotežen podkorpus korpusa GOS in ta je lahko v pomoč pri odločanju o standardnih izgovorih besed.

Čeprav smo načeloma predpostavili, da je standardni jezik, opisan v slovarju, nevtralen za govorno rabo nasploh, je dejstvo, da to (bistveno bolj kot v pisni rabi) ne velja za prav vse govorne situacije: v zasebnem pogovoru raba (vsaj določenih) standardnih oblik (nekaterih tudi sistemsko gledano, npr. dolgega nedoločnika) ni primerna (pa spet to ne velja nujno za cel spekter rab: dolgi nedoločnik na primer je ustrezna oblika za rabo v zasebnem govoru v

SV-koncu Slovenije). Kadar standardne oblike niso primerne, bi si torej želeli informacije o tem, katere oblike pa so primerne ter kdaj in kje. O kako obsežni nalogi ob tem govorimo, si pravzaprav težko predstavljamo, pa tudi ni potrebno, da si: omejitev obsega informacij namreč predstavlja govorjeno gradivo. Informacije o nestandardnih govorjenih rabah vsaj danes nikakor ne morejo biti celostne, ampak se podajo samo tiste, ki so gradivno ustrezno dokumentirane.² V tem segmentu bo v slovarju ostala velika vrzel in upamo, da se bo v prihodnosti izpopolnjevala.

Razmisliti moramo tudi o opremljanju podatkov s kvalifikatorji. Za ločevanje podatkov, ki temeljijo na govornem gradivu, od podatkov, ki temeljijo na pisnem gradivu, predlagam splošni kvalifikator »v govoru« ali »govor« ali podobno. Ta kvalifikator torej pomeni, da je raba tipično govorjena, medtem ko se v pisnem jeziku ne pojavlja ali se pojavlja takrat, kadar ta imitira govorjeni jezik. Lahko je dodan na ravni iztočnice, posameznega pomena, stalne besedne zveze ali oblike. Naslednja raven informacije, ki se zdi pomembna za uporabnika, je, ali je nekaj v govoru primerno samo za določene situacije. Te lahko ločimo na mnogo načinov, vendar je najbolj relevantno za uporabnika predvsem dvoje: ali je nekaj primerno samo v zasebni rabi, ne pa tudi v javni; ter ali je nekaj značilno samo za rabo na določenih regijskih območjih in katerih. Možni so še kakšni dodatni, bolj podrobni kvalifikatorji, npr. za označevanje ekspresivnosti, mladostniškega govora ali vulgarnosti, vendar je to že zelo povezano s sistemom kvalifikatorskega označevanja na ravni celotnega slovarja, ne samo segmenta govora v njem.

3 Praktična implementacija predstavljenega koncepta

Kako se tako zasnovan opis govorjenega jezika izvede v slovarju, bom predstavila, izhajajoč iz predloga SSSJ. Osnovno gradivo za podatke o govorjenem jeziku je korpus GOS (Verdonik idr. 2013), ki pa ga je treba nujno obravnavati ločeno od pisnih korpusov, ne pa vključevati ga vanje, saj se sicer podatki v pisnih korpusih zaradi majhnega obsega GOS-a statistično gledano povsem izgubijo.

² Z ustrezno dokumentiranostjo mislim, da je treba za vsako obliko in izgovor dokazovati ustrezno število rab pri več različnih govorcih (predlagam vsaj tri) v (vsaj dveh) različnih situacijah za vsako posamezno kvalifikatorsko enoto, da se jim lahko pripiše ustrezna verodostojnost, ne pa samo naključnost.

Predlog SSSJ predvideva za prioritetno obravnavo besedišča frekvenčno listo korpusa KRES (Logar Berginc idr. 2011). Za ustrezno opredelitev morebitnih dodatnih gesel govornega jezika predlagam dodatne analize na korpusu GOS, ki bodo privedle do seznama gesel, tipičnih za govorni jezik. Pri tem se za kontrastivno analizo lahko uporabi korpus KRES, prilagojen tako, da se iz njega izločijo besedilni tipi, v katerih je imitiran govorni jezik, za definiranje tipičnega besedja pa metoda izločanja ključnih besed (npr. Scott 1997). Končni seznam je seveda treba ročno pregledati in analizirati. Vse besede, ki bodo po tem postopku pridobljene iz korpusa GOS in vključene v slovar, dobijo kvalifikator »v govoru/govor«. Analizirajo se na vseh ravneh enako kot vse druge iztočnice, temelječ na korpusu GOS.

Analiza pomenov se mora tudi za vse ostale iztočnice, ne samo tiste, ki izhajajo iz govora, preveriti tudi v korpusu GOS. Predvidevamo, da velikega dela iztočnic v korpusu ali sploh ne bo ali pa vsaj ne bo posebnih novih pomenov; za manjši del iztočnic, nekaj primerov je navedenih v sekciji 2, pa lahko pričakujemo, da bo treba k splošnemu pomenskemu opisu s kvalifikatorjem »v govoru« (ali pa še dodatnim kvalifikatorjem o zasebni rabi in morda tudi o regiji/-ah rabe) dodati nov pomen. Čeprav se sliši na videz zahtevno dodatno delo, ga v praksi zaradi majhnosti GOS-a ni pričakovati prav veliko.

Pri vseh iztočnicah, pri katerih se raba v zadosti frekventnem obsegu pojavlja tudi v korpusu GOS, pa čeprav ni izkazanih posebnih novih pomenov, je priporočljivo navesti vsaj en zglede iz tega korpusa; dobrodošla bi bila tudi hiperaktivna povezava, ki vodi uporabnika do zvočnega posnetka tega zglede.

Posebna analiza kolokacij na govornem korpusu, razen za tiste iztočnice, ki so označene kot značilne za govor, ni smiselna, saj je korpus premajhen, da bi lahko pričakovali zanesljive dodatne informacije v večjem obsegu. Je pa smiselna posebna analiza stalnih besednih zvez s poudarkom na frazeoloških enotah, na podlagi katere se lahko dodajo za govor specifične dodatne informacije v zavihek o frazeologiji, kot tudi določijo morebitne dodatne večbesedne iztočnice.

Največ obravnave v zvezi z govornim jezikom bo pri zavihku o oblikah. Informacije za ta zavihek v predlogu SSSJ izhajajo iz leksikona Sloleks (Arhar 2009). Ta je za zdaj samo oblikoslovni in ne vključuje fonetičnega zapisa besed in besednih oblik, vendar je predvidena razširitev v to smer. Pri razširjanju in pregledovanju leksikona so pomembne vse tiste

smernice, ki jih v sekciji 2 že podrobno opisujem pri vprašanju o oblikah in izgovoru. Prva faza izdelave fonetičnega slovarja – tj. izdelava t. i. standardnega izgovora – lahko temelji na polavtomatskih postopkih, po možnosti tudi na izkoriščanju informacij v digitalnih bazah SSKJ ali pravopisnega slovarja. Vse zapise je treba zatem seveda ročno pregledati in po potrebi popraviti ter označiti s kvalifikatorji. Najprimernejši zapis v samem leksikonu je računalniško berljiva fonetična abeceda MRPA, ki se lahko avtomatsko pretvarja tako v IPO kot v za uporabnika bolj berljiv in prijazen, a poenostavljen prikaz. Druga faza dopolnjevanja fonetičnega leksikona – informacije o nestandardnih oblikah in izgovorih – mora temeljiti na podobnih analizah na korpusu GOS, kot jih avtorji SSSJ sicer predvidijo pod zavihkom o govoru. S predpostavko, da je korpus GOS predhodno popravljen na nivoju lematizacije in oblikoslovnega označevanja, lahko z avtomatskimi postopki za vsako posamezno besedno obliko izluščimo dovolj frekventne različne pogovorne zapise (kot so imenovani v korpusu GOS). Ti morajo biti nato ročno pregledani in dopolnjeni, in sicer je treba: (a) izločiti korpusne šume in napake, (b) izločiti tiste zapise, ki ustrezajo standardnim in so že bili izdelani v prvi fazi, (c) pretvoriti sedanji pogovorni zapis iz korpusa GOS v fonetični zapis, po potrebi pa hkrati dopolniti nove oblike na oblikoskladenjski ravni; pri tem je treba opozoriti, da lahko pričakujemo potrebo po razširitvi abecede MRPA in vzporedno tudi IPE (glej Zemljak idr. 2002); (č) dodati potrebne kvalifikatorje.

Kot rečeno, so podobni podatki v sedanjem predlogu SSSJ predvideni pod zavihkom Govor. Če je upoštevan predlog, predstavljen zgoraj, potrebe po posebnem zavihku o govoru ni, različne nestandardne izgovorjave smo zajeli že v zavihku o oblikah. Prikazovanje nestandardnih pogovornih oblik in izgovorov je lahko v vzporednem stolpcu (ki je lahko po potrebi tudi prikrit in prikazan na klik). Če bi se zaradi opremljenosti s številnimi regijskimi kvalifikatorji pokazalo, da lahko informacije postanejo težko berljive, je še vedno možna rešitev, da se s klikanjem izberejo samo posamezne regije.

V slovarju SSSJ je podatek o osnovnem izgovoru besede predviden tudi pri iztočnici. Tudi tja se črpa iz leksikona Sloleks, vendar samo za standardni izgovor. Če tudi za osnovno obliko obstajajo še nestandardne izgovorne različice (primer, bogat s številnimi različicami, je prislov *lahko*), pa se te opisujejo samo pod zavihkom Oblike.

4 Zaključek

V članku sem predstavila predlog, na kak način konceptualno vpeti in nato praktično izvesti prikaz informacij o govornem jeziku v novem slovarju slovenskega jezika. Predlog izhaja iz prepričanja, da je treba posebnosti govornje rabe opisovati na vseh jezikoslovnih ravneh, ne samo na ravni izgovora in regijsko pogojenih izgovornih različic.

Vseh vprašanj, ki bi jih bilo treba prediskutirati v zvezi z govorom v slovarju, tukaj zaradi omejitve prostora žal ni bilo mogoče niti načeti, kaj šele ustrezno predstaviti. Če jih nekaj samo nakažem: vprašanja norme in normativnosti, strokovna vprašanja v glasoslovju (gl. npr. Jurgec 2011), vprašanja opisa pragmatičnih prvin v govoru, pričakovanja uporabnikov o podatkih o govoru in mnoga druga. Upajmo, da bodo doživela ustrezno obravnavo ob izdelavi slovarja.

Literatura

- ARHAR, Špela, 2009: Učni korpus SSJ in leksikon besednih oblik za slovenščino. *Jezik in slovstvo* 54/3–4. 43–56.
- AUSTIN, John L., 1990: *Kako napravimo kaj z besedami*. Ljubljana: ŠKUC: Znanstveni inštitut Filozofske fakultete.
- DIK, Simon C., 1997: *The Theory of Functional Grammar*. Berlin, New York: Mouton de Gruyter.
- FIRTH, John R., 1964: *The Tongues of Men & Speech*. London: Oxford University Press.
- GRICE, Herbert Paul, 1975: Logic and conversation. P. Cole, J. L. Morgan (ur.): *Speech Acts*. New York: Academic Press. 41–58.
- HALLIDAY, M. A. K., 1994: *An Introduction to Functional Grammar*. Arnold.
- HENRICHSEN, P. J., ALLWOOD, J., 2005: Swedish and Danish, spoken and written language: A statistical comparison. *International Journal of Corpus Linguistics* 10/3. 367–399.
- HYMES, Dell, 1974: *Foundations in Sociolinguistics: An Ethnographic Approach*. London: Tavistock Publications.
- JURGEC, Peter, 2011: Slovenščina ima 9 samoglasnikov. *Slavistična revija* 59/3. 243–268.
- KREK, Simon, KOSEM, Iztok, GANTAR, Polona, 2013: Predlog za izdelavo Slovarja sodobnega slovenskega jezika. Ljubljana: s. n. http://trojina.org/slovar-predlog/datoteke/Predlog_SSSJ_v1.1.pdf.

- LOGAR BERGINC, Nataša, GRČAR, Miha, ERJAVEC, Tomaž, ARHAR HOLDT, Špela, KREK, Simon, 2011: *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: gradnja, vsebina, priprava*. Ljubljana: Založba Trojina.
- MAŁINOWSKI, Bronislaw, 1946: The problem of meaning in primitive languages. C. K. Ogden, I. A. Richards (ur.): *The Meaning of Meaning: A Study of the Influence of language upon Thought*. New York: Harcourt, Brace & World, Inc. 296–336.
- POGORELEC, Breda, 2008: Jezikovno načrtovanje govornega jezika pri Slovencih. Primož Vitez (ur.): *Spisi o govoru*. Ljubljana: Znanstvenoraziskovalni inštitut Filozofske fakultete. 75–86.
- de SAUSSURE, Ferdinand, 1997: *Predavanja iz splošnega jezikoslovja*. Ljubljana: ISH – Fakulteta za podiplomski humanistični študij.
- SCOTT, M., 1997: PC analysis of key words – and key key words. *System* 25/1. 1–13.
- STABEJ, Marko, 2008: Kako pa govoriš!? (Spis o normi in normalnem). Primož Vitez (ur.): *Spisi o govoru*. Ljubljana: Znanstvenoraziskovalni inštitut Filozofske fakultete. 87–100.
- ŠKILJAN, Dubravko, 2008: Dolga zgodovina jezikoslovja govora. Primož Vitez (ur.): *Spisi o govoru*. Ljubljana: Znanstvenoraziskovalni inštitut Filozofske fakultete. 29–38.
- VERDONIK, Darinka, KOSEM, Iztok, ZWITTER VITEZ, Ana, KREK, Simon, STABEJ, Marko, 2013: Compilation, transcription and usage of a reference speech corpus: The case of the Slovene corpus GOS. *Language Resources and Evaluation* 47/4. 1031–1048.
- ZEMLJAK JONTES, Melita, KAČIČ, Zdravko, DOBRIŠEK, Simon, ŽGANEC GROS, Jerneja, WEISS, Peter, 2002: Računalniški simbolni fonetični zapis slovenskega govora. *Slavistična revija* 50/2. 159–169.