

Vojko Gorjanc: *Uvod v korpusno jezikoslovje*. Domžale: Izolit, 2005. 163 str.

Darinka Verdonik

Z Uvodom v korpusno jezikoslovje smo Slovenci in predvsem slovenisti dobili prvo monografijo v slovenskem jeziku, ki za razliko od tujih podobnih del prikazuje korpusno jezikoslovje predvsem glede na stanje v slovenskem (korpusnem) jezikoslovju, čeprav seveda ne zanemarija tujih tokov in dogajanj na tem zadnje desetletje zelo produktivnem področju raziskovanja jezika. Delo tako že po sami idejni zasnovi zapolnjuje pomembno vrzel in se uvršča med osnovne pripomočke za vsakogar, ki v gradnji in/ali uporabi korpusov vidi potencial za svoje raziskovalno delo, pa seveda tudi za vsakogar, ki se želi s korpusnim jezikoslovjem samo seznaniti.

Monografija *Uvod v korpusno jezikoslovje* avtorja Vojka Gorjanca je razdeljena v štiri večje razdelke, vsak od teh pa na več poglavij in podpoglavij. V uvodnem razdelku avtor predstavi tipologijo in karakteristike korpusov, na kratko oriše razvoj korpusov ter razpravlja o korpusih v razmerju do jezikoslovja. Drugi razdelek je namenjen gradnji korpusov. Po izhodiščnih premislekih avtor natančno razpravlja o načelih gradnje referenčnih korpusov ob primerih za večje angleške in ameriške korpuse ter srednjeevropske korpuse, seveda tudi ob obeh večjih slovenskih korpusih. Manj razpravljanja posveti označevanju korpusov, kjer predvsem predstavi vrste označevanja. Tretji razdelek je namenjen postopkom korpusne analize. Avtor ob primerih analize korpusa FIDA predstavi najpogostejše postopke analize, nazadnje pa še na kratko orodja za analizo. V zadnjem, četrtem razdelku nas avtor pripelje do srži korpusnega jezikoslovja: v njem navaja konkretne primere korpusne analize, rezultate analize pa nenehno sooča z obstoječimi jezikovnimi opisi v SSKJ in SP 2001. V prvem poglavju tako primerja listo besed v SSKJ in SP 2001 s pogostostno listo besed iz korpusa FIDA, drugo poglavje nameni vrednotenju nekaterih vidikov pravopisnega slovarja iz leta 2001 s stališča aktualnega zajetja slovenske družbene stvarnosti, v tretjem poglavju skuša s pomočjo korpusne analize slediti leksikalnim spremembam za primer poimenovanj s področja interneta, nazadnje pa prikaže primere, kako lahko zamejitev predmetnostnega področja (podkorpus) vpliva na rezultate korpusne analize. Na koncu monografije sledijo povzetek v slovenskem in angleškem jeziku ter angleško-slovenski in slovensko-angleški glosar s področja korpusnega jezikoslovja. Opremljenost s kazali je več kot zgledna, saj poleg standardnih vključuje med drugim izbor spletnih strani in seznam kratic.

Delo *Uvod v korpusno jezikoslovje* se po svoji zasnovi dotika mnogih vprašanj, ki jih korpusno jezikoslovje odpira. V nadaljevanju osvetlimo le nekaj osrednjih vidikov korpusnega jezikoslovja in pogledimo, kako jih obravnava avtor monografije Vojko Gorjanc. Osrednje vprašanje je gotovo, kako se korpusno jezikoslovje umešča v študije o jeziku oz. jezikoslovje. Ali gre za novo metodologijo ali tudi ločeno raziskovalno izhodišče in kakšno? Tega vprašanja se avtor na kratko dotakne v uvodu, kjer ugotavlja: »Korpusni pristop danes v jezikoslovju pogosto velja za komplementarnega tradicionalnim, vendar vsaj v nekaterih segmentih jezikoslovja, predvsem v leksikologiji in leksikografiji, vse bolj pa tudi pri drugih jezikovnih opisih pomeni bistveno spremembo metodologije raziskovanja in tako oblikuje samostojno raziskovalno izhodišče.« (Str. 7) Širše razpravljanje na temo položaja korpusnega jezikoslovja znotraj raziskovanja jezika je v četrtem poglavju prvega razdelka, kjer avtor razpravlja o korpusih in jezikoslovju. Opozori na diskusije med predstavniki tvorbnopretvorbene slovnice in zagovorniki raziskovanja na podlagi obsežnega, načrtno zbranega gradiva. Ugotavlja, da korpusno jezikoslovje pomeni zbirni pojem za niz različnih aktivnosti v zvezi s korpusi, znotraj jezikoslovja pa se vse bolj uveljavlja poimenovanje korpusni pristop. Ta je lahko delni ali popolni. Samo popolni korpusni pristop pomeni uporabo korpusa neodvisno od uveljavljenih jezikoslovnih interpretacij, sicer so analize obremenjene z obstoječimi jezikoslovnimi koncepti. Korpus je tako pogosto le vir za preverjanje hipotez, samo pri popolnem korpusnem pristopu, kjer so predmet analize le »surova« besedila brez kakršnih koli jezikoslovnih oznak, je korpus tudi vir za gradnjo hipotez. Avtor ugotavlja, da se jedro korpusnega jezikoslovja oblikuje ob raziskavah, kjer je korpus vir za gradnjo hipotez.

Čeprav so korpus in potrebna orodja za analizo samo priprave za jezikoslovno (ali jezikovnotehnoško) raziskovanje, so zahtevni tako časovno in finančno kot tudi glede znanj, potrebnih za to delo. Korpusno jezikoslovje tako nikakor ni samo ali predvsem korpusna analiza, pač pa je veliko pozornosti treba posvetiti tudi gradnji korpusov in razvoju orodij za delo s korpusom. Monografija *Uvod v korpusno jezikoslovje* zajame vsa osrednja področja korpusnega jezikoslovja: predstavi obstoječe korpuse, opozori na dilemo vključevanja različnih vrst besedil v referenčne korpuse, navaja vrste jezikoslovnega označevanja korpusov, nas seznanja s postopki korpusne analize, na kratko opiše najbolj znana orodja za analizo in navaja primere korpusne analize. Pri tem ostaja predvsem na »jezikoslovnem« bregu korpusnega jezikoslovja: možnost uporabe korpusov za razvoj jezikovnotehnoških aplikacij je zgolj omenjena (str. 23). Nekoliko morda pogrešamo tudi kakšno napotilo bralcu,

da korpusna analiza ponuja še številne druge možnosti raziskovanja, če se naučimo sami programirati preproste iskalne algoritme (nekaj primerov takšnih analiz s kratkimi opisi najdemo npr. v delu Biberja idr. (1998)).

Pri navajanju postopkov korpusne analize in primerov je monografija *Uvod v korpusno jezikoslovje* osredotočena predvsem na uporabo korpusne analize v leksematiki in lekikalnem pomenoslovju. To je nedvomno utemeljeno s tem, da se osrednji prispevek korpusnega jezikoslovja kaže prav v slovaropisju in razvoj referenčnega korpusa slovenskega jezika FIDA je tesno povezan s slovaropisjem – kot vir služi pri gradnji Velikega angleško-slovenskega slovarja (Krek idr., 2005–), pa tudi v bolj individualnih projektih, pomembnejši med temi je poskusni zvezek besednodružinskega slovarja slovenskega jezika (Stramljič Breznik, 2004). Že sam avtor pa ob tem opozarja, da se korpusi v jezikoslovju uporabljajo za vrsto raziskav, od fonetično-fonoloških do sociolingvističnih. Na ta vidik nikakor ne smemo pozabiti. Tuji avtorji, npr. Kennedy idr. (1998), tako med jezikoslovnimi opisi angleščine, ki temeljijo na korpusu, poleg besedoslovnih opisov navaja oblikoslovne, skladenjske, pragmatične in besedilnovrstne študije, Biber idr. (1998) pa poleg tega opozarjajo še na uporabnost korpusov pri učenju jezikov in razvoju jezikovnih zmožnosti, jezikovnozgodovinskih in stilističnih študijah ipd. Tudi v slovenskem prostoru najdemo nekaj na korpusu temelječih študij, npr. Modrijan (2004) uporabi korpusni pristop za analizo naslavljanj, Gantar (2005) s korpusnim pristopom analizira mesto stalnih besednih zvez v slovenskem leksikalnem fondu, vendar lahko te štejejo bolj za začetke. Delo *Uvod v korpusno jezikoslovje* je za nadaljnji razvoj takšnih študij nedvomno vzpodbudno in aktualno za kar najširši krog raziskovalcev jezika. Pa ne samo za raziskovalce: vsak zaveden pišoči posameznik, ki pozna pomen slovarja, bo podobno uporabnost našel tudi v korpusu, *Uvod v korpusno jezikoslovje* pa mu bo dragocen vodnik pri uporabi korpusov.

Avtor v uvodu poudarja, da korpusni pristop v analizo jezika vnaša večjo verodostojnost, saj omogoča izpostavitve v jeziku tipičnega in zmanjšuje možnost obrobnega kot temeljnega. Seveda pa tudi za korpusni pristop velja, da je treba glede na postopke analize rezultate kritično vrednotiti, kot opozarja tudi avtor (npr. str. 72). Še bolj pa je treba kritično vrednotenje ohraniti do gradiva oz. do tako imenovane reprezentativnosti korpusa. Avtor se tega problema natančno zaveda in mu posveti veliko pozornosti: »Eno temeljnih vprašanj korpusnega jezikoslovja je zagotavljanje **reprezentativnosti korpusa**: kako določiti in uravnotežiti količino raznoterih besedil v korpusu.« (str. 30)

Naj sklenemo: delo *Uvod v korpusno jezikoslovje* prinaša dober pregled področja. Po zasnovi je namenjeno predvsem jezikoslovni publiki, po načinu obravnave problema pa lahko dostopno bralcem, ki se s korpusnim jezikoslovjem šele želijo seznaniti, kot tudi zahtevnejšim bralcem, ki najdejo v njem na enem mestu zbrane informacije in podatke o tem področju.

Literatura

Biber, D., S. Conrad, R. Reppen, 1998: *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.

Gantar, P., 2005: Mesto stalnih besednih zvez v (sodobnem) slovenskem leksikalnem fondu: korpusni pristop. V: Kržišnik, E. (ur.). *Frazeologija v jezikoslovju in drugih vedah*. Ljubljana: Filozofska fakulteta.

Kennedy, Graeme, 1998: *An Introduction to Corpus Linguistics*. London, New York: Logman.

Krek, S., Gabrovšek, D., Stabej, M., Prosenc Šegula, I., Smolej, M., Komar, S., Vesel, N. (ur.), 2005–: *Veliki angleško-slovenski slovar Oxford*. Ljubljana, DZS.

Modrijan, N., 2004: Korpusni pristop pri analizi naslavljanja. V: T. Erjavec, J. Žganec Gros (ur.). *Jezikovne tehnologije*. Ljubljana: Institut Jožef Stefan.

Stramljič Breznik, Irena, 2004: *Besednodružinski slovar slovenskega jezika: Poskusni zvezek za iztočnice na B*. Maribor: Slavistično društvo Maribor.