

## Prikazi, ocjene, osvrti

Ranko Matasović, *Gender in Indo-European*, Universitätsverlag Winter Heidelberg 2004

2004. je godine izišla još jedna indoeuropeistička knjiga praktički jedinoga aktivna hrvatskoga indoeuropeista – Ranka Matasovića. Ova knjiga govori o temi o kojoj se među indoeuropeistima već dugo i naširoko raspravlja – o rodu. Najopćenitije bi se moglo reći da je svrha knjige bila prikazati rod u indoeuropskom prajeziku i pojedinim indoeuropskim jezicima, i to ne samo povijesnolingvistički nego i tipološki i arealno. Kako sâm autor veli u uvodu, ova bi knjiga trebala poslužiti tipolozima i općim lingvistima koji se bave općenito problemom roda u jeziku/jezicima da se što bolje okoriste podacima koje im o tome mogu dati indoeuropski jezici kao najbolje proučena jezična porodica na svijetu. S obzirom na dugu pisanu povijest velika broja indoeuropskih jezika, oni nam pružaju dragocjene informacije i o općim načelima razvoja kategorije roda što se onda dalje može primijeniti na druge, slabije istražene, jezične porodice.

Kao što je i valjano, autor ponajprije polazi od pitanja *Što je to rod?* i pritom se ponajprije oslanja na paradigmatsku monografiju Grevillea Corbetta *Gender* iz 1991. koja predstavlja vrhunac tipološkoga bavljenja pitanjem roda. Nadalje se govori o rodu iz arealne perspektive što će reći da se ukratko daje pregled jezika svijeta s obzirom na prisutnost odnosno odsutnost kategorije roda u njima. Što se tiče sjevernoga euroazijskoga područja, kategoriju roda tu imaju gotovo svi indoeuropski jezici (osim armenskoga, perzijskoga i još nekih zapadnoiranskih jezika). Osim indoeuropskih jezika, rod postoji još i u jenisejskim jezicima, dočim ostali jezici ovoga područja (baskijski, uralski, altajski, čukotsko-kamčatski, korejski i japanski, ainu, sinotibetski, jukagirski i eskimsko-aleutski) ne poznaju kategorije roda. Za to usp. u hrvatskom primjerice *ón, òna, òno*, ali u mađarskom samo *ő*, u finskom samo *hän*, u turskom samo *o*, u kineskom samo *tā*, u baskijskom samo *hura* itd. Što se tiče kavkaskih jezika, rod ima većina naško-dagestanskih jezika te dva od pet abhasko-adigejska jezika. Nasuprot tomu, kategorije roda nema ni u jednom od četiriju kartvelskih jezika. Gotovo svi jezici indijskoga potkontinenta imaju rod (i ie. i dravidski), roda nema samo izdvojeni dravidski jezik brahui. Većina jezika jugoistočne Azije (austroazijski, austronezijski, dajski itd.) nema roda. Oko dvije trećine afričkih jezika imaju rod (od dvaju rodova u afrazijskim jezicima do 24 roda/imenske klase u fulbeu) – afrazijski svi imaju rod, osim omotskih koji ga izgubiše skupa s morfologijom; neki nilosaharski jezici imaju rod (što je čini se inovacija), a neki ne; većina nigersko-kordofanskih jezika ima kompleksne sustave roda/imenskih klasa (iznimku predstavljaju npr. mandski jezici); kojsanski jezici

u načelu imaju rod. Velik broj australskih jezika ima rod (uglavnom tri ili četiri roda), no ne i svi. Mnogi jezici Papue Nove Gvineje također imaju rod (od 2 do 12 njih), no ne i svi. Austronezijski jezici Oceanije, kao ni austronezijski jezici drugdje, nemaju roda. Većina jezika pacifičke obale Sjeverne Amerike ima rod (primjerice sališki), a isto tako i dosta jezika ostatka Sjeverne Amerike (primjerice algonkijski, irokijski). Većina pak jezika pacifičke obale Južne Amerike nema roda, dočim su sustavi s rodom česti među amazonijskim jezicima (primjerice u aravakanskoj porodici).

Nadalje autor kreće s pregledom sustava roda u ie. jezicima. Anatolijski jezici imaju dva roda – živi i neživi. Kod indoiranskih jezika staroindijski i srednjoindijski jezici čuvaju ie. 3 roda, dočim većina novoindijskih jezika ima samo 2 roda, izgubivši stari srednji rod (npr. hindski, pandžapski); rjedi su jezici koji su očuvali trorodni sustav (npr. gudžaratski, maratski) ili *ḥnī* koji su prvo bili izgubili rod pa su ga nanovo stvorili (npr. bengalski). Staroiranski jezici (aveistički i staroperzijski) čuvaju 3 roda. U srednjoiranskom razdoblju dosta zapadnih iranskih jezika (poput partskog i srednjoperzijskoga) gubi rod s gubitkom dočetnih slogova. Toga nema kod istočnih srednjoiranskih jezika (npr. sogdij-skoga ili hotanskoga), gdje nije bilo redukcije dočetnih slogova. Od novoiranskih su jezika rod očuvali neki istočni i izolirani zapadni iranski jezici. Rod nalazimo primjerice u sjevernokurdskim dijalektima, u paštu (afganskom) i jazgulamskom, dočim je on izgubljen recimo u modernom perzijskom, balučiju, osetskom i jagnopskom. I starogrčki su i novogrčki očuvali 3 ie. roda, kao i latinski, no svi su romanski jezici, osim rumunjskoga, izgubili srednji rod. Što se keltskih jezika tiče, staroirski je imao 3 roda, a srednji je rod nestao u srednjoirskom razdoblju. Britska pak grana keltskoga od samih početaka ima samo dva roda (kao suvremeni velški). Svi su rano posvjedočeni germanski jezici očuvali 3 ie. roda (npr. gotski i starovisokonjemački). Engleski je izgubio rod u potpunosti do 14. st., a suvremeni eng. ima zamjenički sustav roda (razlikuje rod samo kod zamjenica – *he, she, it*). Suvremeni njemački, kao i starovisokonjemački, čuva 3 ie. roda. U nizozemskom su se muški i ženski rod stopili u jedan pa on sada ima samo 2 roda. Od skandinavskih jezika, trorodni sustav čuvaju islandski, *nynorsk* i farski, dočim su muški i ženski rod, poput nizozemskoga, stopili švedski i danski. Armenski je izgubio rod već u pretpovijesno vrijeme, najvjerojatnije pod utjecajem substrata. Toharski, ovisno o analizi, ima 2 ili 3 roda i situacija je ondje donekle slična rumunjskoj (gdje je jednina srednjega roda jednaka muškomu rodu, a množina ženskomu). Toharski je A, zanimljivo, jedini ie. jezik koji razlikuje rod u 1. l. jd. osobnih zamjenica – *nāš* »ja (m. r.)«, ali *ñuk* »ja (ž. r.)«. Svi su slavenski jezici očuvali 3 ie. roda, kao i zapadnobaltijski pruski, dočim su istočnobaltijski jezici (litavski i latvijski) izgubili srednji rod. Staroalbanski je još imao 3 roda, no u suvremenom albanskom više nema srednjega roda. Mogu se ustanoviti 4 obrasca razvoja sustava roda u ie. jezicima – a) čuvanje triju rodova (m., ž., s.) kao u hrv., b) stapanje m. i s. roda kao u hindskom i litavskom, c) stapanje m. i ž. roda kao u švedskom i nizozemskom te d) potpun gubitak roda kao u armenskom ili engleskom (uz eventualno čuvanje razlike kod zamjenica).

Nakon I. dijela (Uvod) i II. dijela (Rod u ie. jezicima), u III. se dijelu govori o rodu u ie. prajeziku odnosno autor pokušava doći do načelā kojima se pojedinim riječima dodjeljivao rod u ie. prajeziku. Autor je stoga podijelio ie. riječi u sljedeća značenjska polja: 1) nazive srodstva, 2) ljudska i nadljudska bića, 3) životinje, 4) biljke, 5) voće i žitarice, 6) fizički svijet, 7) dijelovi tijela, 8) oruđe i oružje, 9) apstrakcije, 10) tekućine i 11) tvari. Kod nazivā srodstva, sve su imenice, očekivano, m. (\*ph<sub>2</sub>tēr »otac«) ili ž. r. (\*méh<sub>2</sub>tēr »mati«). U drugoj skupini, kod (nad)ljudskih bića, nailazimo na sva tri roda – m. (\*deywós »bog«), ž. (\*tewteh<sub>2</sub> »narod«) i sr. r. (\*génh<sub>1</sub>os »rod, porodica«). Zanimljivo je da je ie. riječ za »ženu«, \*g<sup>v</sup>enh<sub>2</sub>, ako je suditi po staroirskom *bé* bila srednjega roda. Kod životinja pak ima svega – m. (\*b<sup>h</sup>éb<sup>h</sup>rus »dabar«, ž. (\*h<sub>2</sub>enh<sub>2</sub>tis »patka, ūtva«, s. r. (\*péku »stoka«) ili m./ž. (\*muHs »miš«). Zanimljivo je da su u ie., za razliku od hrvatskoga, imena mladih životinja m. r. (\*pórkos »prase«, a manjih životinja, pogotovu kukaca, ž. r. (\*wób<sup>h</sup>seh<sub>2</sub> »osa«). Što se 4. skupine, biljaka, tiče, tu pak nema m. r., ali se za većinu naziva drveća rod ne može pouzdano rekonstruirati. Za ž. r. usp. \*b<sup>h</sup>eh<sub>2</sub>gós »bukva«, a za sr. r. \*b<sup>h</sup>ólyom »list« (čini se da su svi dijelovi biljaka bili sr. r.). Voće i žitarice su pak isključivo sr. r. u ie. (\*meh<sub>2</sub>lom »jabuka«). U 6. su skupini, kod riječi za fizički svijet, posvjedočena sva tri roda – m. (\*h<sub>2</sub>stēr »zvijezda«, \*h<sub>2</sub>ékmōn »kamen«, ž. (\*h<sub>2</sub>ēp- »voda«, \*nok<sup>w</sup>ts »noć«, sr. (\*mori »more«). Kod dijelova tijela opet nailazimo na sva tri roda, s tim da izgleda da postoji mala tendencija da unutarnji organi budu sr. r. (\*kērd »srce«. Usp. i m. (\*pōds »noga«) i ž. r. (\*b<sup>h</sup>ord<sup>h</sup>eh<sub>2</sub> »brada«. Kod oružja i oruđa i ostalih od čovjeka napravljenih objekata opet ima svega, iako se čini da su tu imenice m. r. (\*k<sup>w</sup>ek<sup>w</sup>los »kotač«) mlade, usp. i ž. (\*neh<sub>2</sub>us »brod«) i sr. r. (\*Hyugóm »jaram«). Izgleda da su apstrakcije u ie. u načelu bile ž. (\*mrtís »smrt«) ili sr. r. (\*kléwos »slava«, dočim su izgleda dvije imenice m. r. tu sekundarne (\*swépnos »san«). Imenice za tekućine su, osim jedne iznimke m. r. (\*sok<sup>w</sup>ós »mezgra, smola«, sr. r. (\*wódr »voda«. Što se tiče tvārī, òne su, osim nekoliko iznimaka poput riječi \*snōy<sup>wh</sup>s »snijeg« (ž.), sr. r. (\*dóru »drvo«. Zatim slijedi rasprava u kojoj se govori o više problema vezanih s rodom u ie. i dodatno analizira prethodno podastrt materijal. Očito je da što se ide od 1. prema 11. skupini – od imenica za živo prema imenicama za neživo, ima sve više sr. r. pa su tako imenice za životinje u načelu m. ili ž. roda – veće su obično m. r. (\*włk<sup>w</sup>os »vuk«, a manje ž. (\*g<sup>h</sup>eluh<sub>2</sub>s »kornjača« – u knjizi greškom stoji \*g<sup>h</sup>eluh<sub>2</sub>s). Kada dodemo do »najmanje živih« stvari, tipa iménā tekūcīnā, òna su isključivo sr. r. (\*krewh<sub>2</sub> »krv«). Između krājnostī, kod biljaka, dijelova tijela, apstrakcija itd. nalazimo i m./ž. i sr. r. U nastavku autor raspravlja o o-osnovama ž. r., tj. o imenicama koje su oblikom muške, ali su slaganjem (i značenjem) ženskoga roda kao \*snusós »sna-ha«. Postoji također i obrnut slučaj, kada su imenice oblikom ženske, a slaganjem muške (usp. hrv. *môj slúga*). No dočim su u prvom slučaju posrijedi prežitci iz ie., u drugom je slučaju riječ o inovaciji u pojedinim jezicima. Autor nadalje utvrđuje da tvorba ženskih imenica od muških (kao \*h<sub>1</sub>ékweh<sub>2</sub> »kobila« od \*h<sub>1</sub>ékwos »konj«) nije ie. porijekla nego je riječ o mladoj pojavi (u pridjeva je pak to već bilo moguće u ie.). Nadalje autor govori o rodu, naglasku i prijete-

voju u ie. i tu, zanimljivo, izražava sumnju u postojanje ie. amfidinamičkoga sklonidbenoga obrasca, no ne obrazlaže to. Autor zatim govori o slaganju po rodu u ie. – s imenicama su se u ie. slagale zamjenice, brojevi i pridjevi. Usp. \*h<sub>1</sub>ék<sup>wos</sup> wesus h<sub>1</sub>esti »konj je dobar«, ali \*h<sub>1</sub>wid<sup>h</sup>eweh<sub>2</sub> weswih<sub>2</sub> h<sub>1</sub>esti »udovica je dobra«. Slaganje je s pridjevima ž. r. jako mlado u ie., a dobrim dijelom čak i posljeie. Autor zatim govori o pojavi poznatoj iz grčkoga, avestičkoga i hetitskoga da se imenice srednjega roda u množini slažu s glagolom u jednini. To je, kako je već odavna utvrđeno, posljedica toga što je N. mn. sr. r. (formalno podudaran N. jd. ž. r.!) zapravo kolektiv, zbirna imenica, a ne prava množina (usp. hrv. *lišće pàdā*). Zanimljiva je i spekulacija da je to pravilo možda uzrok nestanku razlike između 3. l. jd. i mn. u baltijskim jezicima. I u ie. je izgleda, kao i u hrv., »neutralni«, »neodređeni« rod bio m. r. (usp. u hrv. *mùževi i žene su išli* gdje se glagol slaže s m. r.), no to je, kako ispravno zaključuje autor, prije posljedica sintaktičke arhaičnosti (iz doba prije nego što se razvio poseban ž. r. u pridjeva), negoli ie. patrijarhalnosti ili, sačuvaj bože, šovinizma.

IV. je dio knjige posvećen rodu u ranom indoeuropskom. Da je ie. sustav triju rodova nastao od starijega dvorodnoga sustava, znaju već i vrapci na grani. Ne samo da su nam iz lat. i grč. poznati tzv. pridjevi s dvama nastavcima (jedan za m./ž., a drugi za sr. r.), nego je tu i činjenica da u anatolijskom postoje samo dva roda (živi i neživi). Autor pokušava odgovoriti na dva pitanja – kako je ž. r. pridružen određenim imenicama u kasnoie. i kako je nastalo novo slaganje po rodu, odnosno kako nastadoše posebni oblici ž. r. u zamjenicama i pridjevima? Imenice su ž. r. nastale od sr. r. što je jasno po tome što im. ž. i sr. r. dijele isti dvojinjski nastavak (\*-ih<sub>1</sub>), po tome što su imenice ž. r. zapravo često samo kolektivi prema im. sr. r., po tome što imenicama ž. r. u kasnom, posljeanatolijskom ie. u načelu odgovaraju imenice neživoga roda u hetitskom itd. Slaganje je kolektiva sr. r. (na \*-eh<sub>2</sub>) s glagolom u jd. pogodovalo semantičkoj reinterpretaciji starih kolektiva kao jd. (npr. \*h<sub>1</sub>wid<sup>h</sup>eweh<sub>2</sub> »obitelj umrloga« > »udovica«). Ženskomu su rodu pridružene logično i im. koje su označavale ženska bića (tipa \*swésōr »sestra«), a stara je pridjevska tvorba na \*-ih<sub>2</sub> (od koje potječe latinsko-keltski nastavak G. jd. o-osnova -ī) također reinterpretirana kao ž. r. – \*włk<sup>wos</sup> »vuk« – \*włkwih<sub>2</sub> »vukov« > \*włk<sup>w</sup>ih<sub>2</sub> (poimeničeno) »ona koja je vukova« > »vučica« (e ovo sada jest pomalo seksistički!). No zašto bi baš stari kolektivi na \*-eh<sub>2</sub> i stari poimeničeni pridjevi na \*-ih<sub>2</sub> postali novim ženskim rodom? Čini se zbog slučajnosti! Naime imenica je za ženu izvršavala upravo na \*-h<sub>2</sub> (\*g<sup>w</sup>enh<sub>2</sub>) i to je shvaćeno kao nastavak ž. r. Dakle, ženski je rod u kasnom ie. zapravo nastao sasvim slučajno. Tada je staro slaganje \*so wesus g<sup>w</sup>enh<sub>2</sub> zamijenjeno novim \*seh<sub>2</sub> weswih<sub>2</sub> g<sup>w</sup>enh<sub>2</sub> i eto nam ženskoga roda!

Zatim slijedi i jedan od možda najzanimljivijih dijelova knjige – poglavlje o rodu i rečeničnom ustroju ie. Gotovo je općeprihvaćeno među indoeuropeistima da je ie. u starijem stadiju bio ergativan jezik ili barem jezik aktivnoga tipa. No Matasović nije zadovoljan argumentima koji se navode u prilog toj tezi – smatra da to ponajviše ide na sljedeći način: »ergativni su jezici zanimljivi, ie. je jamačno bio zanimljiv > jamačno je i ie. bio nekoć ergativan!«. S jezične su

strane, misli Matasović, takove tvrdnje pak potpuno neosnovane. Obično se tvrdi da je ie. bio ergativan na osnovi sljedećega: ie. je nominativni nastavak \*-s rezerviran sàmo za agense. Kako imenice srednjega roda ne mogu biti agensi, one nemaju toga nastavka. Nastavak je \*-s isprva označavao sàmo subjekat prijelaznoga glagola – ergativ, dočim je subjekat neprijelaznoga glagola i objekat prijelaznoga – apsolutiv – bio nemarkiran. Akuzativni bi nastavak \*-m bio mladega, čestičnoga ili posljeložnoga postanja. To što imenice srednjega roda nemaju nastavak \*-s je posljedica toga što òne nikada nisu subjekat prijelaznoga glagola (jer znače neživo). S obzirom da imenice sr. r., koje su nežive, nisu mogle biti subjekti radnje, logično je da su mogle imati sàmo apsolutiv, odnosno da je \*-m nastavak i N. i A. jd. sr. r. u kasnoie. No, kako smatra autor, ovakove tvrdnje nisu pretjerano uvjerljive. Kao prvo, to što neživo ne može biti subjektom ne znači da je jezik ergativan (tako je primjerice u japanskom, a japanski nije ergativan). Kao drugo, nijedan ie. jezik nema zabrane po kojoj imenice srednjega roda ne bi mogle biti subjektima. Treće, nema očita razloga zašto bi se moralo smatrati da je nastavak \*-s stariji od nastavka \*-m. Dapače, vjerojatno je obrnuto, tvrdi uvjerljivo Matasović jer čak ni zamjenice koje nemaju u N. nastavka \*-s imaju u A. nastavak \*-m, usp. ie. \*so »taj«, A. \*tom (skr. *sa, tam*). Dapače, i usporedbe ie. s drugim prajezicima prije ukazuju na to da je \*-m stariji nastavak. Osim ergativne teorije, popularna je i teorija o ie. kao o jeziku aktivnoga tipa u starijoj fazi (Gamkrelidze & Ivanov). Pa ipak, kako dokazuje autor, pozivajući se pritom na tipologinju Johannu Nichols, osobine koje navodno ima ie. i koje ga navodno povezuju s aktivnim jezicima uopće nisu statistički relevantne (npr. nepostojanje glagola »imati« ili opozicija živo : neživo itd.). Autor također vješto banalizira i još jedan »dokaz« da je ie. bio jezik aktivnoga tipa – veli da to što je ie. imao dvije riječi za »vodu« i »vatra« nije pretjerano znakovito. Ne treba u tome odmah tražiti skrivenu aktivnost u predie. – jezici katkada jednostavno imaju po dvije riječi za određene pojmove (usp. hrv. *vàtra* i *òganj*), što ne mora odmah nužno značiti da je jedna riječ bila smatrana živim, a druga neživim. Konačno, ie. je imao i dvije riječi za »ruku« (\**meh<sub>2</sub>r* i \**g<sup>h</sup>es<sub>r</sub>*), ističe, ali svejedno nikomu ne pada na pamet tvrditi da je »ruka« mogla biti i agens u predie. Postavlja se pitanje zašto je u ie. u sr. r. N. = A. Kako je N. u jezicima svijeta često neobilježen, bez nastavka, može se pretpostaviti da je i rani ie. imao nulti nastavak u N. To je očuvano i u zamjenici \*so, osim u atematskim im. sr. r. Nastavak je A. pak bio \*-m. Također se može recimo pretpostaviti da je \*-m bio nastavak za A. sàmo određenih objekata. Kako su imenice sr. r. obično označavale kakove tvari ili apstrakcije, logično je da nisu u načelu bile određene nikada pa nisu ni imale nastavka \*-m. Poznata je stvar da je nastavak N. jd. *o*-osnova, \*-os, jednak nastavku G. jd. atematskih osnova (\*-os). Na osnovi toga Matasović iznosi novu hipotezu o nastanku tematskih osnova u ie. Krećemo primjerice od imenice \*Hyewgs »uprezanje« (lat. *con-iūx* »supružnik«) kojoj je G. bio \*Hyugós. G. se je te imenice rabio često atributno, kao pridjev, kao u hetitskom. Neki su se od tih G. posve okamenili, postali pravim pridjevima, a zatim su se i poimeničili. Zatim su takove nove imenice krivo morfemski analizirane – \*-o- je protumačeno kao vezni samoglas, a \*-s kao nominativni nastavak. Odatle u het. na-

stavak *-aš* i u N. i u G. *o*-osnova. Ovako je lako razumjeti i zašto nom.-ak. nastavak im. sr. r. nije gola osnova na *\*-o*. Jednostavno nikada nije bilo takova nastavka u paradigmi starih tematskih pridjeva. N/A. na *\*-om* je rezultat analogije: kako su atematske im. sr. r. imale isti N. i A. (čista osnova), tako je i A. na *\*-om* kod tematskih pridjeva uzet i za N., a pridjev se poslije može lijepo poimeničiti. Sve je ovo moguće stoga što u ie. nije bilo velike razlike između pridjeva i imenica. Najbolji je kandidat za srodstvo s ie. uralski prajezik. U njem je, zanimljivo, N. uvijek neobilježen, dočim je akuzativni nastavak ni više ni manje nego upravo naš *\*-m*! Osim toga, uralski je nominativno-akuzativni jezik što se sve sjajno slaže s ovdje predloženom unutaršnjom rekonstrukcijom ie. Na kraju poglavlja Matasović zaključuje da svrha dotičnoga poglavlja nije bila nepobitno dokazati da je *\*-m* doista najstariji ie. nastavak i da se je razvoj ie. doista mogao samo ovako dogoditi, nego samo da druge hipoteze (o ergativnom i aktivnom postanju ie.) nisu dovoljno osnovane.

Djelo završava kratkim poglavljem o arealnim i tipološkim aspektima roda u ie. Autor prvo promatra razmještaj jezika s kategorijom roda u Euraziji. Neie. ju jezici Europe, kao etrurski i baskijski nisu imali (u bilješci je omaškom citirano Pfister 1998 umjesto Pfiffig 1998). U ostatku sjeverne Eurazije nijedan jezik (osim ketskoga) nema roda. Rod imaju jezici grubo rečeno južno od ie. (afrazijski, dravidski, naško-dagestanski, abhasko-adigejski itd.), dočim ga oni istočno (uralski, altajski, sinotibetski itd.) nemaju. Ie. je sustav roda, kako zaključuje autor, tipološki najslbližiji ketskomu. Od ostalih se eurazijskih sustava rodova (kao od onih u sumerskom, elamskom, dravidskom itd.) on poprilično razlikuje. Što nam arealno-tipološko razmatranje sustava roda govori o mogućim dubinskim genetsko-arealnim vezama ie. i drugih jezika? Jedan je od većih problema nostratičke teorije to što pola navodnih nostratičkih jezika ima rod (ie., afrazijski, dravidski), a druga polovica ne (uralski, altajski, kartvelski). Autor smatra da bi to zastupnici nostratičke teorije svakako morali objasniti, uzevši u obzir da je rod dijakronijski prilično stabilna kategorija (teško se stječe i teško gubi). Pa ipak, valja napomenuti da ovaj argumenat možda i nije toliko jak koliko to može na prvi pogled izgledati. Ako je ie. srodan i s jednim jezikom, onda će to najprije biti uralski iako je jasno da uralski nema roda. Pitanje je dakle koliko nam rod može pomoći u dubinskom uspoređivanju jezika imamo li na umu da se po kategoriji roda ie. razlikuju čak i od svojega, po svem sudeći, najbližega srodnika.

I što na kraju zaključiti? Knjigu valja pohvaliti, zacijelo će biti zanimljivim i korisnim štivom indoeuropeistima i tipolozima, a u budućnosti će sasvim sigurno biti i nezaobilaznom referencom pri svakoj raspravi o temi roda ne samo u indoeuropeistici, gdje je nezaobilazno djelo, kao što prve reakcije već i pokazuju, nego i u široj sferi tipologije, kao što je sâm autor i planirao.

*Mate Kapović*

**Lelija Sočanac: *Hrvatsko–talijanski jezični dodiri s rječnikom talijanizama u standardnome hrvatskom jeziku i dubrovačkoj dramskoj književnosti*.  
Nakladni zavod Globus, Zagreb 2004.**

U izdanju izdavačke kuće Globus izašla je 2004. godine knjiga Lelije Sočanac pod naslovom *Hrvatsko–talijanski jezični dodiri s rječnikom talijanizama u standardnome hrvatskom jeziku i dubrovačkoj dramskoj književnosti*. Knjiga obaseže 404 stranice, od toga je 258 stranica teksta, dok ostatak čine dodaci: u podnaslovu navedeni rječnici talijanizama, abecedni popis svih talijanizama, varijanata i izvedenica u rječniku talijanizama u hrvatskome standardnom jeziku, zatim isti takav popis za rječnik talijanizama u dubrovačkoj dramskoj književnosti, veoma opsežna bibliografija, zaključak te napokon sažetak na engleskom jeziku.

Knjiga je podijeljena na šest poglavlja: 1. Jezici u kontaktu, 2. Hrvatsko–romanski dodiri: povijesni pregled, 3. Adaptacija modela na fonološkoj razini, 4. Adaptacija modela na morfološkoj razini, 5. Adaptacija modela na semantičkoj razini te 6. Prebacivanje kodova. Slijedi zaključak te dodaci koje smo spomenuli ranije.

U prvome poglavlju autorica donosi povijesni pregled istraživanja jezičnih dodira koji postaju predmetom intenzivnijeg interesa pojedinih jezikoslovaca u 18. stoljeću kada se tadašnji leksikografi pri skupljanju građe za svoje rječnike susreću s problemom posuđenica. Govori nadalje o miješanim jezicima koji su relativno dugo bili važnom temom jezikoslovnih rasprava bez obzira na stavove onih koji su taj termin zagovarali ili onih koji su ga osporavali. U ovom poglavlju iscrpno su nabrojani svi značajni predstavnici pojedinih struja i škola iz toga područja, počevši od Raska, Schuchardta, Windischa, preko de Saussurea i Bloomfielda, pa do Haugena i Weinreicha čije se teorije, bez obzira na neke revizije i kritička promatranja pojedinih dijelova, još i danas smatraju temeljima kontaktne lingvistike. Autorica nadalje spominje Rudolfa Filipovića koji je u Hrvatskoj nedvojbeno bio vodeći stručnjak iz te grane jezikoslovlja i koji, uz daljnji razvoj teorije, uvodi i domaće nazivlje. Na kraju se ukratko osvrće na najnovija strujanja u kontaktnoj lingvistici posljednjih godina, izdvajajući pri tom vrlo snažno izraženu interdisciplinarnost kao jednu od najistaknutijih osobina. Nakon povijesnoga pregleda detaljno se obrađuju najvažniji pojmovi vezani uz istraživanje jezičnih dodira, a to su dvojezičnost (u pojedinca i u zajednici), interferencija i transferencija na svim jezičnim razinama, integracija, te prebacivanje kodova koje se može promatrati s dva različita aspekta, sintaktičkoga i sociolingvističkoga.

Drugo je poglavlje posvećeno hrvatsko–romanskim jezičnim dodirima od samih početaka još u rimsko doba, zatim za vladavine Bizanta, za vrijeme sukoba hrvatsko–ugarske države i Mlečana, za francuske te naposljetku austrijske uprave. U svim se tim razdobljima vrlo detaljno zasebno obrađuju ona geografska područja jadranske obale u kojima se ti jezični dodiri ostvaruju, a to su

Istra, Dalmacija i Dubrovnik. Ta je činjenica veoma važna, jer su, bez obzira na relativnu blizinu, te tri regije imale različitu povijest pa kao posljedicu i različitu jezičnu situaciju. Posebni se odjeljak bavi dalmatskim jezikom koji je, prema Skoku, nastao iz balkanskog latiniteta koji se u doba naseljavanja Slavena govorio u jednom dijelu Istočnorimskoga Carstva i koji je u doba bizantske vladavine i hrvatske narodne dinastije do pada Dalmacije pod mletačku vlast bio jezik javne, ponajprije usmene, komunikacije. Autorica se također bavi različitim ulogama što su ih na ovim područjima imali mletački i toskanski: prvi je bio govorni jezik vladajućega sloja kao i dominantan jezik trgovine i plovidbe na Mediteranu; drugi u Dalmaciju ulazi kao jezik kulture koji su dalmatinski studenti učili na talijanskim sveučilištima u Bologni i Padovi. Kao rezultat svih tih složenih, dugotrajnih i intenzivnih hrvatsko–romanskih, a kasnije hrvatsko–talijanskih kulturnih i jezičnih dodira dolazi do veoma značajnoga intimnoga i kulturnoga posuđivanja što kao posljedicu ima mnogobrojne posuđenice talijanskoga podrijetla u hrvatskome jeziku. Dio tih posuđenica pripada danas mjesnim primorskim govorima kao izraz tradicije i znak lokalnoga identiteta, dok dio posuđenica koje pripadaju sferi kulturnoga posuđivanja ulazi u standardni jezik.

U sljedeća tri poglavlja slijedi analiza prilagodbe talijanskih posuđenica na tri jezične razine: fonološkoj, morfološkoj i semantičkoj. Pri tome se autorica potpuno oslanja na teoriju Rudolfa Filipovića u kojoj su razrađeni različiti stupnjevi prilagodbe koje tijekom integracije u jezik primatelj može ostvariti pojedini strani model. Budući da je Filipovićeva teorija izrasla ponajprije na temelju istraživanja posuđenica engleskoga podrijetla, Lelija Sočanac nužno je neke od postavki iz te teorije morala prilagoditi građi koju je obrađivala, dakle talijanizmima. U svim je tim poglavljima analiza provedena na dva bitno različita korpusa. Prvi korpus čine talijanizmi koji pripadaju suvremenomu hrvatskomu standardnom jeziku ili, u pojedinim slučajevima, njegovoj supstandardnoj razini; drugi je korpus načinjen od talijanizama skupljenih na građi starije dubrovačke književnosti. Tako različiti korpusi uzeti kao polazište za analizu, omogućili su autorici da opiše i dvije posve različite vrste jezičnoga posuđivanja: kulturno i intimno.

S obzirom na relativno visok stupanj sličnosti dvaju uspoređivanih jezičnih sustava na fonološkoj razini, autorica je zapazila umanjenu potrebu za fonološkom prilagodbom talijanskih posuđenica. Nadalje, kod talijanizama u hrvatskome dolazi i do različitih glasovnih promjena koje nisu uvijek nužno vezane uz prilagodbu sustavu jezika primatelja, nego su, posebice u primorskim govorima, najčešće posljedica dijalekatskih utjecaja.

Četvrto je poglavlje posvećeno prilagodbi talijanskih modela na morfološkoj razini. Ovdje je također primijenjena Filipovićeva metodologija prema kojoj treba razlikovati primarne i sekundarne pojave u tome procesu: u fazi primarne adaptacije strana se riječ prilagodava morfološkom sustavu jezika primatelja, dok sekundarna faza obuhvaća sve promjene do kojih dolazi nakon potpune integracije modela u domaći sustav. Morfološki sustavi talijanskoga i hrvatskoga znatno se razlikuju: talijanski je djelomično analitički, dok je hrvatski potpuno sintetički, s vrlo složenim sklonidbenim sustavom i slobodnim redom ri-



ječi. Autorica je u ovome dijelu analize uočila neke veoma zanimljive pojave, ponajprije vezane uz tzv. dubrovački korpus koji pripada kategoriji intimnoga posuđivanja. Tako se, primjerice, za mjesne i vremenske priloge dugo smatralo kako nisu podložni aloglotskim utjecajima. Zahvaljujući, međutim, činjenici da je ovdje riječ o bitno drukčijem dodiru dvaju jezika, dakle neposrednome, u dubrovačkom je korpusu naden veći broj primjera upravo te vrste. Nadalje, u tom je korpusu naden i znatan broj veznika koji pripadaju kategoriji funkcionalnih morfema koji se inače rijetko posuđuju. Autorica primjećuje kako je uporaba veznika u dubrovačkome korpusu olakšana visokim stupnjem sintaktičke konvergencije između hrvatskoga i talijanskoga te čestim prebacivanjem i miješanjem kodova.

U poglavlju koje se bavi semantičkom razinom, zapažene su najveće razlike između dva istraživana korpusa, i to ponajprije u odnosu na domene iz kojih potječu posuđenice talijanskoga podrijetla. Dok u standardnome hrvatskom jeziku najveći broj talijanizama pripada sferi kulture, glazbe, mode, kulinarstva, bankarstva i trgovine, u dubrovačkom je korpusu talijanizama uočen znatan broj posuđenica koje označuju posve apstraktne pojmove, dijelom vezane uz misaone procese, dijelom uz logičke kategorije. Velik se broj talijanizama iz dubrovačkoga korpusa odnosi na razne vidove međuljudskih odnosa i društvene interakcije. Sve je to, dakako, posljedica izravnoga dodira dvaju jezika, rezultat dvojezičnosti, što nužno dovodi i do posuđivanja leksičkih elemenata iz sfera koje u posrednome posuđivanju gotovo nikad, ili veoma rijetko, dolaze u obzir.

Šesto je poglavlje autorica posvetila prebacivanju kodova kao važnome aspektu dvojezičnosti. Dakako, ovdje je analiza provedena samo na dubrovačkome korpusu jer se jedino u uvjetima dvojezičnosti kakva je oslikana u dramskoj književnosti Dubrovnika može govoriti o toj kategoriji. Lelija Sočanac bavi se nekoliko različitih vidova prebacivanja kodova: jedno je sintaktički vid unutar kojega navodi brojne, nadasve zanimljive, primjere međurečeničnoga i unutarrečeničnoga prebacivanja kodova, a drugo sociolingvistički i pragmatički aspekt. Ovaj potonji u Dubrovniku je često bio vezan uz društveni status, obrazovanje, spol te opreku grad – selo. Autorica dolazi do zaključka kako se, gledajući u dijakronijskoj perspektivi, pragmatički aspekti prebacivanja kodova mijenjaju usporedno s promjenom sociolingvističke situacije.

Nakon Zaključka u kojemu se rezimiraju najvažniji rezultati provedenoga istraživanja, slijede rječnici talijanizama u oba korpusa. Rječnici su izrađeni u skladu s koncepcijom koju je u okviru svojega projekta *Engleski element u europskim jezicima* osmislio Rudolf Filipović i u skladu s načelima postavljenim upravo za taj projekt. Tako je natuknica talijanski model, slijedi izvorni, dakle talijanski, izgovor, zatim talijanizam u hrvatskome s izgovorom, oznaka stupnja prilagodbe na morfološkoj razini, na semantičkoj razini te napokon značenje posuđenice. Takav tip rječnika nedvojbeno je veoma složen, ali zauzvrat donosi mnoštvo podataka za svaku pojedinu posuđenicu.

Na kraju su knjige abecedni popisi svih talijanizama iz jednoga i drugoga korpusa, uključivši tu i sve varijante i izvedenice.

Vrlo iscrpna bibliografija na više od dvanaest stranica bilježi zaista maksimalan broj naslova iz područja kontaktne lingvistike, etimoloških istraživanja, romanistike i drugih domena relevantnih za ovo veoma opsežno istraživanje.

Knjiga Lelije Sočanac ispunila je svojim pojavljivanjem najmanje dvije važne zadaće: s jedne je strane s najrazličitijih aspekata ocrtala utjecaje i dodire talijanskoga i hrvatskoga jezika u dijakronijskoj i sinkronijskoj perspektivi, s druge je strane uspješno pokazala kako se primjenom pojedinih teorija bez obzira na to koji se jezici istražuju može provesti valjana jezična analiza. *Hrvatsko-talijanski jezični dodiri* djelo su koje će nedvojbeno privući pažnju ne samo talijanista nego i svih onih koje zanimaju pojave vezane uz jezično posuđivanje u najširem smislu.

Vesna Muhvić-Dimanovski

**Marko Tadić, *Jezične tehnologije i hrvatski jezik*, Exlibris, Zagreb 2003.**

Knjiga je podijeljena u dvije osnovne cjeline: I. *Jezične tehnologije* u kojoj se nalaze poglavlja 1. Što su jezične tehnologije?, 2. Jezične tehnologije za hrvatski jezik, 3. Jezične tehnologije i pretraživanje informacija; II. *Primjena jezičnih tehnologija na hrvatski jezik* s poglavljima 4. Portal jezičnih tehnologija za hrvatski jezik, 5. Hrvatski nacionalni korpus, 6. Izgradnja Hrvatskoga morfološkoga leksikona, 7. Morfosintaktičko označavanje hrvatskih korpusa, 8. Primjena jezičnih tehnologija pri prevodenju pravne stečevine EU na hrvatski jezik i 9. Pogovor. Na kraju knjige nalazi se opsežan popis bibliografskih jedinica i pojmovnik jezičnih tehnologija.

Napokon se pojavila knjiga koja objedinjuje i usustavljuje znanja iz ovog, do sada neobrađenog područja za hrvatski jezik. Građa se zasniva na autorovu višegodišnjem radu u području korpusne i računalne lingvistike, brojnim objavljenim člancima, predavanjima, kao i saznanjima stečenim u radu na domaćim i međunarodnim znanstvenim projektima.

U prvom se poglavlju definira područje rada, opisuje razlika između *računalne lingvistike* i *računalne obradbe (prirodnoga) jezika*. Prva disciplina koristi računala kao pomoćna sredstva u proučavanju jezika i zasebna je lingvistička grana definirana svojom odjelitom metodologijom. Drugoj je pak u žarištu proučavanja sam računalni sustav s težnjom za što učinkovitijom i što bržom obradbom (jezičnih) podataka uz što manji utrošak računalnih resursa. Jezične se tehnologije komparativno svrstavaju uz bok ostalim suvremenim tehnologijama i analitički se uspoređuju dijelovi formalne aparature. Budući da se sve tehnologije zasnivaju na sirovini koju obrađuju, a rezultat je konačni proizvod, u ovom je slučaju sirovina sam prirodni jezik, a proizvod su sustavi koji omogućuju olakšanu uporabu jezika u računalnome okružju. Autor u ovom poglavlju razgraničuje područje jezičnih tehnologija od drugih, po predmetu proučava-

nja bliskih znanstvenih područja, što je originalna argumentacija u svjetlu međunarodnih publikacija. U poglavlju *Sastavnice jezičnih tehnologija* donosi podjelu jezičnih tehnologija na jezične resurse, jezične alate i komercijalne proizvode, gdje se po posebnim potpoglavljima daje usustavljeni prikaz svih resursa.

Hrvatski jezik u digitalnom obliku, kao i u digitalnim kanalima (ponajprije na internetu) danas nije dovoljno zastupljen. Ovo poglavlje prožeto mnogim kritičkim osvrtima daje uvid u današnje stanje hrvatskih jezičnih resursa i alata. Ističući strateški položaj jezičnih tehnologija u suvremenom informacijskom društvu, te usporedbom s nekim drugim jezicima, autor motivacijski nastoji mobilizirati širu lingvističku zajednicu u smjeru zajedničke suradnje i bržeg razvoja jezičnih tehnologija u nas.

Većina strukturiranih podataka/informacija danas kodirana je u obliku teksta (digitalnog), a većina pretraživača informacija usmjerena je na engleski jezik. Za hrvatski, koji je za razliku od engleskoga flektivno bogat jezik, morfološki »osjetljivi« upiti (pretraživanje po svim flektivnim oblicima riječi) znatno pospješuju pretraživanje informacija. Uključivanjem takvog morfološkoga modula u sustav za pretraživanje tekstova na hrvatskome jeziku omogućilo bi se puno pretraživanje teksta (*full-text search*). Razvoj sustava za automatsko prepoznavanje naziva (*named entities*) iz teksta dodatno bi i znatno pospješio prepoznavanje obavijesti. Nazivi izravno povezuju tekst s izvanjskim svijetom i pružaju odgovore na temeljna pitanja koja nose najveću količinu obavijesti u tekstu *tko? kada? što? gdje? koliko?* Sljedeća razina koja pospješuje kvalitetu pretraživanja jest semantička. Osnovna je nakana omogućiti pojmovno umjesto abecednog pretraživanja rječnika, gdje su u leksičkoj bazi kodirani značenjski odnosi među riječima. Takvi specijalizirani tezaursi koji se izrađuju za pojedine europske jezike izvode se u okviru projekta Euro-WordNet u koji bi svakako trebalo uključiti hrvatski jezik. Posljednja cjelina/razina koja bi poboljšala pretraživanje informacija odnosi se na višejezične upite. Sustav koji podržava višejezične upite omogućio bi pretraživanje repozitorija znanja pohranjenog u tekstovima na raznim jezicima. Takav bi složeni sustav obuhvaćao dvo- ili višejezične leksikone, razvijenu morfološku sastavnicu, ali i sintaktičku komponentu. Sintaktička bi komponenta sezala do razine imeničnih izraza (NP) ili koordiniranih imeničnih izraza (NP i NP). U razvijenoj inačici, takva bi tražilica obuhvaćala i sustave za sažimanje teksta (*summarizing systems*). Čitavo poglavlje intonirano je potrebom za nužnošću promptnoga razvijanja izloženih sustava za hrvatski jezik, jer bi se u budućnosti moglo dogoditi da se do određenih znanja može doći samo putem »odabranih« jezika, a ne i na ostalima. Takav se apel osobito očitava u rečenici posljednjeg, zaključnog odlomka poglavlja »Ta je razdjelnica između jezika s razvijenim jezičnim tehnologijama i onih bez njih, nalik razdjelnici između jezika s pismom i jezika bez pisma«, a rezultat nerazvijanja jezičnih tehnologija bila bi *funkcionalna* nepismenost za dotični jezik.

Drugi dio knjige iznosi i opisuje postojeće jezične tehnologije za hrvatski jezik. Polazište za proučavanje jezičnih tehnologija, ali i središte zainteresiranih strana

za razmjenu i preuzimanje relevantnih informacija, referencija i resursa je WWW–portal *Jezične tehnologije za hrvatski jezik* (<http://www.hnk.ffzg.hr/jthj/>). Portal je ponajprije namijenjen svima zainteresiranima za primjenu jezičnih tehnologija na hrvatski jezik, studentima i istraživačima, tvrtkama s područja informacijskih tehnologija i telematike, kao i uredima za prevodenje. Informativna sastavnica portala pruža iscrpne obavijesti relevantne za hrvatski jezik osobito sa stajališta jezičnih tehnologija. Interaktivna sastavnica pruža mogućnost izravne pretrage *Hrvatskoga nacionalnoga korpusa* i *Hrvatskoga morfološkoga leksikona*. Cilj je hrvatskoga portala za jezične tehnologije služiti kao servis hrvatskoj jezičnotehnološkoj zajednici i kao koncentracijska točka u kojoj se ocrtavaju interesi, aktivnosti i rasprave.

Hrvatski nacionalni korpus (HNK) temeljni je sinkronijski i dijakronijski referentni korpus hrvatskoga jezika. Izgrađuje se u Zavodu za lingvistiku Filozofskoga fakulteta Sveučilišta u Zagrebu od 1998. godine u okviru znanstvenih projekata MZT–a. Autor u knjizi iznosi osnovne korpusne parametre i teorijske postavke, a za dublji uvid nudi dodatne izvore literature. Interesantan je podatak da je korpus u potpunosti dostupan, odnosno pretraživ putem Interneta.

Pored korpusa, osnovnim jezičnim resursima svakako pripada leksikon. Hrvatski morfološki leksikon (HML) je leksikon u kojem su pohranjeni svi generirani potencijalni oblici općih riječi hrvatskoga jezika. HML sadrži oznake što interpretiraju svaki leksički unosak, a osim oznaka koje opisuju gramatičke kategorije riječi (POS) unesene su i oznake koje nose podatak o pripadajućim morfosintaktičkim opisima (MSD) kodirane prema MULTEXT–East v2 preporuci. Leksikon je također slobodno pretraživ putem interneta, a moguće je i direktno hipertekstovno odašiljanje svih flektivnih oblika riječi WWW–tražilici Google, tj. uporaba morfološki »osjetljivog« upita.

Morfosintaktičko označavanje hrvatskih korpusa sljedeći je veliki korak koji će se izvoditi u okviru računalnojezikoslovnih i korpusnih projekata unutar spomenutog Zavoda. Iako sustavi za označavanje (*taggers*) korpusa za velik broj zapadnoeuropskih jezika već postoje, pa i za neke slavenske, za hrvatski još ne postoji cjelovit i pouzdan sustav. Osnovna je zadaća sustava za označavanje razriješiti inherentnu višeznačnost pojava iz korpusa na POS i MSD razini. Jedna je od prvih prepreka kod označavanja riječi nailaženje na mnoštvo istopisnih riječi. Autor istopisnost dijeli u dvije osnovne skupine. *Unutarnja istopisnost* je slučaj kad ista pojava (ili niz pismena) predstavlja različite oblike iste leme, tj. oblike s različitim morfosintaktičkim opisima, no koji još uvijek pripadaju istoj lemi (npr. dativ, lokativ i instrumental množine u hrvatskoj često imaju identičan lik, ali predstavljaju zapravo tri različita oblika iste leme s tri različita morfosintaktička opisa). Kod *vanjske istopisnosti* ista pojava (ili niz pismena) predstavlja oblike dviju ili više različitih lema. Tako je na primjer na hrvatskom jeziku oblik *cijene* istopisan unutar dviju različitih gramatičkih kategorija, imenice i glagola. No to je samo jedan od uvodnih problema, a u knjizi se analiziraju znatno složenije zagonetke koje nameće obrada prirodna jezika na ovoj razini.

Usustavljen prikaz područja neupućenom čitatelju daje odličan uvid u jezične tehnologije i njihovu primjenu na hrvatski jezik, ali ne libeći se zakoračiti u neka otvorena pitanja računalne lingvistike, ova knjiga vrijedan je izvor informacija i stručnjacima iz samog područja. Osim brojnih izvornih rasprava proizašlih iz višegodišnjeg bavljenja autora ovim područjem, dodatnu kvalitetu knjizi daje opisivanje problematike obrade jezika na mnogim razinama jezika/obrade lingvistike na primjerima hrvatskog jezika. Pročitavši knjigu, čitalac će lako razumjeti potrebu za proučavanjem hrvatskoga jezika iz nove perspektive, jer iznesena empirijska metodologija koju pruža računalna lingvistika baca novo svjetlo na dosad nedostupne ili teško dostupne jezične činjenice.

*Božo Bekavac*

**John R. Taylor, *Cognitive Grammar*, Oxford University Press, 2002.**

Knjiga J. Taylora *Cognitive Grammar* predstavlja uvod u kognitivnu gramatiku, koja je svoj razvitak započela sedamdesetih godina dvadesetoga stoljeća pod vodstvom Ronalda Langackera, koji je do danas ostao njezinim vodećim predstavnikom. Knjiga se sastoji od dvadeset i osam poglavlja podijeljenih u sedam dijelova. U skladu s uvodnim karakterom knjige, na kraju svakoga poglavlja nalaze se prijedlozi literature za produbljivanje znanja iz obrađene teme s kratkim objašnjenjima, te pitanja čija je namjera ponukati čitatelja na daljnja razmišljanja o izloženim problemima. Na samome kraju nalazi se praktični glosar s definicijama osnovnih pojmova kognitivne gramatike koji se u prethodnome tekstu opširnije objašnjavaju.

U prvome dijelu (*Background*, poglavlja 1 do 6), J. Taylor određuje mjesto kognitivne gramatike kao specifične teorije jezika u širem okviru kognitivne lingvistike s kojom dijeli osnovne teorijske postavke. Pritom se posebno osvrće na pojam 'kognitivnoga' koji se odnosi na tvrdnju da je jezik integralni dio ljudskih kognitivnih procesa, te se u tom svjetlu u lingvistici valja i proučavati. Drugo poglavlje donosi temeljne postavke kognitivne gramatike i uvod u njezin teorijski aparat. Priroda jezika jest simbolička, što znači da između semantičkih struktura kao značenja jezičnih izraza i fonoloških struktura kao njihovih vanjskih manifestacija djeluju simbolički odnosi. Sintaksa ne postoji kao zasebna razina jezične organizacije već se i ona promatra u okviru simboličkih odnosa između semantičkih i fonoloških struktura. Simboličke jedinice svih razina postaju, dakle, osnovnim predmetom bavljenja kognitivne gramatike. Posljednja dva poglavlja prvoga dijela govore o fonološkoj, odnosno semantičkoj strukturi. S obzirom na to da fonologija ima vlastita organizacijska načela koja mogu djelovati neovisno o semantičkim strukturama koje jezični izraz simbolizira, Taylor objašnjava zašto se fonološkim strukturama pridaje takva autonomija. Na sličan se način i semantičkoj strukturi pripisuje određen stupanj autonomije. Postoji, naime, razina konceptualne organizacije koja utječe na značenjski sadržaj simboličkih jedinica koje su sastavnice nekog jezičnog izraza,

što za posljedicu ima da je značenje izraza bogatije od značenja njegovih sastavnica. Kao vrijednost uvodnoga dijela knjige valja spomenuti i vrlo kratke usporedno–kritičke osvrtne na ne–kognitivne lingvističke pristupe, posebice na čomskijevsku lingvistiku te na Saussureovo poimanje jezičnoga znaka kao poveznicu s videnjem simboličkih jedinica u okviru kognitivne gramatike.

U drugome dijelu (*Basic concepts*, poglavlja 7 do 13) Taylor objašnjava temeljne pojmove kognitivne gramatike. Jedan od osnovnih odnosa u jeziku kao strukturiranom skupu jezičnih jedinica, jest 'vertikalni' odnos između sheme, vrlo općenitog i podrobnostima siromašnog prikaza, te njezinih instanci, koje su specifičnije od nadređenog im shematskog prikaza. Taj odnos postoji na svim razinama jezičnih jedinica, a Taylor redom ističe problem taksonomija na razini semantičkih jedinica, zatim fonoloških segmenata kao elemenata u odnosima fonoloških shema i njihovih instanci, te na kraju shematskih odnosa među simboličkim jedinicama, pri čemu je, primjerice, riječ *tree* instanca više shematične simboličke jedinice [IMENICA]. Taylor nadalje čitatelja na vrlo jasan način upoznaje s pojmovima profila, baze i domene, ključnima za analizu značenja u okviru kognitivne gramatike. Profil izraza jest ono što izraz označava (*designation*), a za konceptualizaciju profiliranog entiteta nuždan je kontekst u vidu konceptualne strukture koju nazivamo bazom. Pojam domene odnosi se pak na širu konfiguraciju znanja potrebnu za ostvarivanje konceptualizacije. Ova tri Langackerova pojma izvrsno su prikazana, a navode se i usporedbe sa sličnim pojmovima u drugim lingvističkim pristupima.

Treći dio (*Morphology*, poglavlja 14 do 17) posvećen je unutrašnjoj strukturi složenih jedinica, s posebnim osvrtom na strukturu riječi. Iako se sam pojam *morfologija* koristi u svrhu lingvističke analize, naglašava se kako ona nije poseban gramatički modul odvojen od sintakse. Morfemska analiza prikazana je, između ostalih, na primjeru engleskih riječi *book/books*, a pokazuje kako korijenski morfem *book* i množinski sufiks *-s* nose različite fonološke i konceptualne karakteristike. Posebno se naglašava važnost shema (npr. sheme za tvorbu množine, sheme za tvorbu prošlog vremena) koje zamjenjuju tradicionalna tvorbena pravila i dovode u pitanje stalno stvaranje oblika prema zadanom pravilu, te navode na shvaćanje da su složeni morfološki oblici pohranjeni kao gotovi i spremni za uporabu. Kao takvi se nalaze u stalnom međudjelovanju sa svojom shemom i svakom uporabom učvršćuju njezin položaj u sustavu simboličkih jedinica.

Za četvrti dio (*Nouns, Verbs, and Clauses*, poglavlja 18 do 21) može se reći da predstavlja težište cijele knjige jer govori o imenicama, glagolima i rečenicama (*clauses*) – temeljnim kategorijama svakoga jezika i glavnim predmetima analize u svakoj gramatičkoj teoriji. Kako se može i očekivati u kognitivnolingvističkom pristupu, oni su definirani u odnosu na svoj konceptualni sadržaj. Tako imenica profilira stvar (*thing*) kao relativno autonomnu konfiguraciju konceptualnoga prostora, dok glagol profilira proces (*process*), to jest relaciju koja se odvija u vremenu. Imenicama i imeničnim izrazima posvećena su dva poglavlja, u kojima Taylor objašnjava kognitivne procese nužne, prema Langackeru, za karakterizaciju imenica (*grounding, instantiation, specification*). Posebno se osvrće na razliku između dviju kategorija imenica u engleskome jeziku

(*count* i *mass*), koja je konceptualno utemeljena, a zasniva se na distinkciji između 'tvari' i 'individualiziranih predmeta'. Od glagolskih tema izdvajamo glagolsko vrijeme i vid koji zauzimaju posebno mjesto. Ove dvije kategorije u engleskome su jeziku usko povezane, a njihova analiza nužno uključuje i ispitivanje konceptualnih odnosa između svršenih (*perfective*) i nesvršenih (*imperfective*) procesa. Posljednje poglavlje ovoga dijela usredotočuje se na unutrašnji ustroj rečenice – sudionike, njihove semantičke uloge i sintaktički izraz, te na pitanje kako se otvoren skup semantičkih uloga preslikava na vrlo ograničen broj sintaktičkih kategorija (subjekt, izravni objekt itd.).

U petome dijelu (*More on Meaning*, poglavlja 22 i 23) Taylor razrađuje neka pitanja neophodna za proučavanje značenja. Ponovno se vraća na pojam domene, razrađujući ga uz raspravu o enciklopedijskoj naravi značenja. Naime, puna karakterizacija značenja neke riječi (Taylor se uz izravne Langackerove citate koristi pojmom *banane*) obuhvaća sve domene relevantne za potpuno razumijevanje i sposobnost uporabe zadanoga pojma – oblik, boju, okus, miris, činjenicu da se banane jedu, da rastu na stablima i tako dalje. Takav skup domena naziva se matricom domena, a njegovo postojanje potvrđuje enciklopedijsko poimanje značenja. Osim toga, značenje je i fleksibilno. Koristeći se engleskom riječju *photograph*, Taylor pokazuje kako pojam fotografije nije jasno definiran, statički entitet, već se njegovo značenje kristalizira u specifičnim kontekstima. Drugim riječima, pojam se 'konstruira' od elemenata koji počivaju u mreži enciklopedijskoga znanja. Posljedica je toga pogled na značenje kao dinamički proces, a ne fiksni entitet. Nadalje, Taylor razvija zanimljivu i dobro organiziranu raspravu o pitanjima polisemije i srodnih pojmova, koja su u kognitivnoj lingvistici popularna od samih njezinih početaka. Odgovor na pitanje o prirodi unutrašnjega značenjskog ustroja riječi nalazi se, smatra Taylor, u odnosu nadređene sheme i instanci značenja. Zadana riječ je polisemna ako su prototipno značenje i njegove ekstenzije kognitivno istaknutije (*salient*) od svoje nadređene sheme (primjerice riječ *tree* sa značenjima stablo i obiteljsko stablo) pa pojedina značenja postoje kao gotove simboličke jedinice. Ako je pak shema kognitivno istaknutija od svojih instanci, radi se o monosemiji (primjerice riječ *grandfather*).

Šesti dio (*Approaches to Metaphor*, poglavlja 24 do 26) posvećen je metafori, jednoj od najistraživanijih jezičnih pojava u okviru kognitivne lingvistike, a započinje kratkim pregledom teorije metafora G. Lakoffa koja se temelji na postojanju temeljnih konceptualnih metafora. One omogućavaju razumijevanje neke apstraktne domene (ciljna domena) pomoću konkretnoga iskustva (izvorna domena). Uz kritički osvrt na Lakoffov pristup metaforama, Taylor umjesto uvriježenoga preslikavanja između dviju domena predlaže gledište prema kojemu konceptualne metafore imaju status shema među kojima se, kao i u drugih jezičnih jedinica, stvaraju taksonomski odnosi. Jedan od glavnih nedostataka klasifikacije metafore u okviru Lakoffova pristupa upravo je formulacija konceptualnih metafora s različitih shematskih razina, jer nema organizacije prema zajedničkim shemama koje bi omogućile prikaz pomoću jasnijih taksonomskih odnosa. Kao dokaze u prilog ovoj tvrdnji, Taylor navodi radove R. Jackendoffa i R. Langackera o prostornim i neprostornim uporabama engleskoga gla-

gola *go*, gdje se naglasak umjesto na preslikavanje među domenama, stavlja na njihove sličnosti te na konceptualne sheme koje ocrtavaju zajedničke elemente različitih domena.

Posljednji, sedmi dio (*Idioms and Constructions*, poglavlja 27 i 28) govori o idiomima i konstrukcijama. Čitatelju se može učiniti da su oni ovdje neopravdano združeni, no Taylor ih smatra povezanimi budući da i jedni i drugi predstavljaju složene simboličke jedinice, ali različitoga stupnja shematičnosti. Pritom su konstrukcije obično određene na višoj razini shematičnosti i kao svoje instance imaju otvoren skup jezičnih izraza, dok idiome karakterizira nizak stupanj shematičnosti. Ipak, i idiomi se najčešće mogu raščlaniti (na planu značenja ili izraza), što znači da se mogu svesti pod apstraktnije sheme. Kao neobičnost ovoga dijela ističemo činjenicu da konstrukcijama Taylor posvećuje tek dvadesetak stranica, dok one u radovima nastalim u okviru kognitivne gramatike zauzimaju iznimno važno mjesto, a u prilog tomu govori i činjenica da se u kognitivnoj lingvistici kao općem teorijskom okviru razvio i poseban pristup pod nazivom konstrukcijska gramatika.

Knjiga *Cognitive Grammar* Johna Taylora koristan je uvod u problematiku kognitivne gramatike koja, prema autorovim riječima, predstavlja najobuhvatniju i u potpunosti zaokruženu teoriju u širem okviru kognitivne lingvistike. Iako na ozbiljnijoj razini bavljenja kognitivnom lingvistikom ona ne može biti zamjenom za izvorne radove R. Langackera, njezina je korist za lingvistu i studenta lingvistike dvojaka. Prvo, ona donosi jasan i sustavan pregled osnovnih pojmova kao preduvjeta za razumijevanje kognitivne gramatike kao lingvističke teorije, te prikaz njihove primjene na svim jezičnim razinama. Drugo, uz većinu obrađenih tema Taylor nudi kratki, ali kritički osvrt na radove s istom tematikom unutar kognitivne lingvistike, ali i iz teorija poput strukturalizma ili generativne gramatike, što omogućava usporedbu ili pak praćenje zajedničkih niti u različitim lingvističkim pristupima, koje se u lingvistici često zanemaruju.

Nina Tuđman Vuković

***The Oxford Handbook of Computational Linguistics*  
 Ruslan Mitkov, (ur.), Oxford University Press, 2002**

*The Oxford Handbook of Computational Linguistics* opsežan je uvod u široko područje strojne obrade jezika koje se naziva računalnom lingvistikom. Priručnik ima za cilj pružiti uvid u metodologiju koja se primjenjuje u toj interdisciplinarnoj grani te jezične razine i teoretske okvire na kojima se zasnivaju važniji suvremeni pristupi.

Priručnik je podijeljen u tri veće cjeline. U prvom dijelu, pod naslovom *Osnove (Fundamentals)* iznose se temeljne postavke jezikoslovnih disciplina poput fonologije, morfologije, sintakse, semantike, proučavanja diskursa, leksikografije itd. na kojima se temelji računalni pristup problemima unutar tih disciplina.



Drugi dio, pod naslovom *Obrade, metode i resursi (Processes, Methods and Resources)* posvećen je zadacima, metodama i resursima, bilo već postojećim ili onima koje je tek potrebno razviti za strojnu obradu određenog prirodnog jezika. Istražuju se teme poput strojne segmentacije teksta, obilježavanja vrsta riječi, sintaktičke analize, razlučivanja značenja riječi, razrješavanja problema anafore, generiranja teksta, prepoznavanja govora, crpljenja podataka, pretvaranja teksta u govor, strojnog učenja i usvajanja riječi, korpusa, ontologija. U drugom se dijelu također razmatraju teoretski pristupi koji se temelje na automatima s konačnim brojem stanja, statističkim izračunima te raznim formalizmima (npr. *tree-adjointing grammars*).

U trećem dijelu knjige, nazvanom *Primjene (Applications)* opisuju se sustavi razvijeni metodologijom računalne lingvistike i njihova uporaba u područjima poput strojnog prevodenja, crpljenja podataka, strojnog odgovaranja na upite, izrade tekstualnih sažetaka te prepoznavanja i obilježavanja termina. Nadalje, opisuju se multimedijalni/multimodalni sustavi, sustavi za govorne dijaloge, računalno potpomognuto učenje i obradu višejezičnih podataka putem interneta.

Priručnik se sastoji od ukupno 38 članaka različitih autora. Autor svakog članka priznati je stručnjak za određeno područje.

Autor predgovora je urednik Ruslan Mitkov, a uvoda, u kojem se ukratko ocrta razvoj strojnog prevodenja, računalne lingvistike i obrade prirodnog jezika i u kojem se također navode relevantni časopisi iz navedenih područja je Martin Kay.

U prvom dijelu knjige (*Osnove*) Steven Bird autor je članka o primjeni računala u fonološkoj obradi u kojem se uz osnovne crte opisne fonologije daje uvod u računalnu fonologiju kao i u primjenu formalnih i računalnih tehnika za obradu i prikaz fonoloških podataka. Autor članka o morfologiji je Harald Trost koji iznosi osnovne teoretske postavke tradicionalnih i suvremenih pristupa i govori o širokoj primjeni računala u područjima poput sustava za pravopisno ispravljanje, automatsku razdiobu riječi prilikom pisanja itd. Patrick Hanks napisao je članak o leksikografiji u kojem iznosi pregled računalne leksikografije sa dva aspekta: 1) funkcije leksikona u računalnim programima i 2) uporabe računalnih metoda prilikom sastavljanja novih rječnika. Hanks također daje povijesni pregled leksikografskih postupaka i metoda. Članak o sintaksi u kojem se daje uvod u problematiku kojom se bave računalne sintaktičke analize napisao je Ronald M. Kaplan. U njemu ukratko ocrta formalne sintaktičke opise koji imaju istaknutu ulogu u računalnom pristupu. Poglavlje o semantici napisao je Shalom Lappin u kojem ističe da je jedan od glavnih ciljeva računalne semantike strojno modeliranje značenja fraza i rečenica na temelju značenja njihovih sastavnih elemenata. Istraživanja na polju računalne semantike dijeli u dvije veće skupine. Prva se odnosi na prikaz značenja, a druga na analizu sintaktičkih fraza te njihova međusobnog kombiniranja prilikom tvorbe značenjskih jedinica. Autor također istražuje odnos između sintakse i semantike te uloge elemenata konteksta i diskursa prilikom strojnog pristupa definiranju značenja. Allan Ramsay autor je članka o diskursu, a Geoffrey Leech i Martin Weisser o mogućnostima uporabe računala u istraživanju pragmatičkih i dijaloških elemenata u pisanim i govorenim tekstovima. Carlos

Martin–Vide autor je članka u kojem daje uvod u gramatike formalnih jezika kao sredstva za jezično generiranje, te u automate s konačnim brojem stanja kao sredstva za prepoznavanje jezičnih elemenata. Bob Carpenter, autor posljednjeg članka prvog dijela knjige pod naslovom *Complexity*, govori o asimptotičkoj kompleksnosti algoritama.

Drugi dio knjige započinje člankom pod naslovom *Segmentacija teksta* autora Andreia Mikheeva. Autor govori o prvom koraku strojne segmentacije teksta – tokenizaciji, odnosno o raščlanjivanju teksta na riječi, brojke, interpunkcije itd. Također govori o segmentaciji teksta na rečenice te o problemima na koje se nailazi unutar pojedinih jezika. Atro Voutilainen u svome članku o strojnom određivanju vrsta riječ opisuje postupke prilikom izrade automatskih označivača vrsta riječi, daje njihovu podjelu s obzirom na teoretske postavke na kojima počivaju te govori o nužnim koracima koje prethodno treba izvršiti u pogledu jezičnih resursa. Treći članak drugog dijela knjige posvećen je strojnoj sintaktičkoj analizi. Autor John Carroll daje pregled i opis vrsta takve analize te navodi formalne pristupe za razvoj nužnih alata. Mark Stevenson i Yorick Wilks autori su poglavlja posvećenog automatskom prepoznavanju smisla odnosno određenog značenja riječi. Autori navode da je *word–sense disambiguation* postupak identificiranja značenja riječi u kontekstu. Daju uvid u razvoj problematike od samih početaka razvoja strojnog prevodenja te obrazlažu osnovne pristupe koji se temelje na rječnicima, statističkim metodama ili su konekcionalističkog ishodišta. Članak zaključuju pregledom postupaka za evaluaciju rezultata i moguću primjenu opisanih sustava. Urednik cjelokupnog priručnika Ruslan Mitkov također je autor članka o razrješavanju anafore. U prvom dijelu članka Mitkov daje teoretski prikaz i probleme povezane s anaforizacijom te važnost njezina razrješavanja u strojnoj obradi teksta. U drugom dijelu članka autor opisuje povijesni razvitak na tom polju, kao i suvremena dostignuća. John Bateman i Michael Zock autori su članka o strojnom generiranju prirodnog jezika, Lori Lamel i Jean–Luc Gauvain članka o strojnom prepoznavanju govora, a Thierry Dutoit i Yannis Stylianou članka o sintezi teksta u govor u kojima se daje pregled dosadašnjih postignuća u tim područjima. Lauri Karttunen autor je članka o tehnologiji temeljenoj na automatima s konačnim brojem stanja, a Ch. Samuelson o statističkim metodama u računalnoj lingvistici. U članku pod naslovom *Machine Learning* Raymond J. Mooney piše o postupcima nužnim za strojno prepoznavanje jezičnih jedinica na različitim razinama, dok u članku o strojnom usvajanju leksika Yuji Matsumoto opisuje neophodne jezične resurse za takav postupak te navodi dosadašnja postignuća. Lynette Hirshman i Inderjeet Mani autorice su članka o postupcima evaluacije rezultata alata i sustava za računalnu obradu jezika, a Richard I. Kitteridge članka o jeziku određenih domena te karakteristikama sustava za strojnu obradu jezika koji se primjenjuju na jezik graničenih i kontroliranih područja. Tony McEnery autor je članka pod naslovom *Korpusna lingvistika* u kojem daje pregled uporabe i razvitka korpusa u strojnoj obradi jezika. Autor također govori o vrsti podataka koji se unose u korpuse, o mogućnostima obilježavanja korpusa te o obradi korpusa s unesenim jezikoslovnim podacima ili bez njih. Piek Vossen autor je članka o ontologijama u kojem raspravlja o njihovim karakteristikama

i uporabi u strojnoj obradi jezika, dok u posljednjem članku drugog dijela knjige Aravind Joshi iznosi teoretske postavke tzv. *tree-adjointing grammars*.

Treći dio knjige započinje člankom Johna Hutchinsa u kojem autor iznosi povijesni razvitak strojnog prevođenja, dok u sljedećem članku, također posvećenom istom području, Harold Somers daje uvid u sadašnje stanje te postojeće sustave za strojno i strojno potpomognuto prevođenje. Članak E. Tzoukermann, J. L. Klavans i T. Strazalkowskog pod naslovom *Information retrieval* posvećen je strojnom crpljenju podataka, a članak Ralpa Grishama bliskom području – *information extraction*. S. Harabagiu i D. Moldovan opisuju razvoj i metodologiju strojnog odgovaranja na upite, a E. Hovy postupke prilikom automatske izrade sažetaka dokumenata. Ch. Jacquemin i D. Bourigalt autori su članka o automatskom obilježavanju i crpljenju termina iz tekstova, M. A. Hearst članka pod naslovom *Text Data Mining*, a I. Androutsopolos i M. Aretoulaki članka pod naslovom *Natural Language Interaction*. E. Andre govori o primjeni računalne obrade prirodnog jezika u multimedijalnim i multimodalnim sustavima, a J. Nerbonne o uporabi računala i strojne obrade jezika u postupku strojno potpomognutog učenja kod ljudi. Posljednji članak knjige, kojeg su autori G. Grefenstette i F. Segond, posvećen je području strojne obrade višejezičnih podataka putem interneta.

Na kraju svakog članka nalazi se iscrpan popis literature, kao i savjeti za daljnje čitanje unutar određenog područja, čime ovaj opsežni priručnik svoju publiku pronalazi kod početnika, kao i kod stručnjaka koji se поближе žele uputiti u određene probleme širokog i interdisciplinarnog područja računalne lingvistike.

*Krešimir Šojat*