



UDK 81'367.626

Prethodno priopćenje

Prihvaćen za tisak 18. 6. 2023.

<https://doi.org/10.29162/jez.2023.2>

Zrinka Kolaković¹
Edyta Jurkiewicz-Rohrbacher²
Jasmin Denise Gradischnig¹
¹ Alpen-Adria-Universität Klagenfurt
² Universität Regensburg

Akuzativne zamjeničke zanaglasnice *ju* i *je* u bosanskoj, hrvatskoj i srpskoj: kontrastivno korpusno istraživanje standardnih i razgovornih idioma*

Empirijska su istraživanja o varijaciji u uporabi zanaglasnica iznimno rijetka. Međutim, pažljiva i kontrastivna analiza bosanskih, hrvatskih i srpskih gramatika i ostalih jezičnih priručnika kao što su npr. jezični savjetnici, upućuje na postojanje određenih razlika u uporabi zanaglasnica ne samo među trima standardima nego i unutar njihovih razgovornih idioma. Ovaj se rad bavi razlikama u uporabi dviju varijanata akuzativne zamjeničke zanaglasnice trećega lica jednine ženskoga roda *ju* i *je*. Sudeći prema stanju opisanome u gramatikama, uporaba oblika *ju* znatno je ograničenija u bosanskoj i srpskoj nego u hrvatskoj standardu (cf. Silić i Pranjković 2007; Mrazović i Vukadinović 2009; Ridjanović 2012; Piper i Klajn 2014).

Kako bi se utvrdio stvarni raspon razlika u uporabi spomenutih zanaglasnica, primjeri za analizu crpljeni su ne samo iz tradicionalnih referentnih korpusa s jezičnom građom koja bi trebala odražavati stanje u trima standardima (Santos 1998; Utvić 2011; Čermák i Rosen 2012; Brozović Rončević i dr. 2018) nego i iz {bs, hr, sr}WaC-a, triju golemih mrežnih korpusa (Ljubešić i Klubička 2014) pretraživih s pomoću iznimno funkcionalnoga sučelja *NoSketchEngine* primjenom morfosintaktičkih oznaka, tj. tagova. Ručno je označeno 3916 primjera s akuzativnim zanaglasnim varijantama *ju* i *je*. Istraživanje potvrđuje kako jezični

* Rad je nastao u sklopu projekta *Baza hrvatskih morfoloških dubleta*. Autori zahvaljuju anonimnim recenzentima i urednicima te sudionicima 36. međunarodnoga znanstvenoga skupa HDPL-a *Standardni i nestandardni idiomi* na izrazito korisnim uputama za poboljšanje rada.

korpusi mogu poslužiti za dobivanje prijeko potrebnih uvida u raspon varijacije ne samo s obzirom na normu triju standarda nego i u stvarnoj svakodnevnoj uporabi, odnosno razgovornim idiomima. S pomoću frekvencije varijantnih oblika i konstrukcija u kojima se rabe utvrđuje se u kojoj se mjeri uzus razlikuje od propisane norme.

Uopćeni linearni regresijski model otkriva statistički značajne razlike u uporabi akuzativne zanaglasne varijante *ju* među standardnim idiomom hrvatskoga s jedne i standardnim idiomima bosanskoga i srpskoga jezika s druge strane. Osim toga, uočeno je da se razgovorni idiomi bosanskoga, hrvatskoga i srpskoga jezika značajno razlikuju u uporabi zanaglasne varijante *ju*. Nadalje, podatci pokazuju da je distribucija zanaglasnice *ju* značajno drukčija u razgovornim nego u standardnim idiomima analiziranih jezika. I na kraju, na temelju uopćenoga linearnoga regresijskoga modela, pokazano je da za razliku od *-(j)e* završetka riječi koja prethodi zanaglasnici trećega lica ženskoga roda u akuzativu, *-(j)u* završetak ima statistički značajan utjecaj na izbor među varijantama *ju* i *je*. Točnije govoreći, smanjuje vjerojatnost uporabe zanaglasnice *ju*.

Ključne riječi: bosanski, hrvatski, srpski, standardni i razgovorni idiom(i), varijacija, korpusi, zanaglasnica trećega lica jednine ženskoga roda u akuzativu

1. Uvod

Mnogobrojni su domaći i strani znanstveni radovi posvećeni zanaglasnicama u bosanskoj, hrvatskoj i srpskoj jeziku (Kolaković i dr. 2022: 4). Pomnim se iščitavanjem vrlo brzo uviđa kako su strani redom formalni, a domaći uglavnom nisu – među potonjima ističu se preskriptivni i analitički¹ (usp. Jonke 1953; Gortan 1956; Pešikan 1959; Babić 1963, 1964a, 1964b; Alerić 1964; Brabec 1964; Mulić 1964; Browne 1968, 1974, 2003, 2004, 2014; Popović 1979, 2004; Čavar i Wilder 1994; Wilder i Čavar 1994; Franks i Progovac 1994; Franks i King 2000; Franks i Peti-Stantić 2006; Franks 2010; Schütze 1994; Mamić 1995; Mišeska-Tomić 1996; Progovac 1996; Radanović-Kocić 1996; Peti-Stantić 2002, 2007; Ham 2006; Milićević 2007, 2009, 2023; Bošković 2004; Stjepanović 2004; Aljović 2005; Alexander 2008; Diesing i dr. 2009; Gračanin-Yukse 2019). Deskriptivni i inferencijski empirijski radovi temeljeni na velikim podacima, a posvećeni zanaglasnicama koji uzimaju u obzir varijaciju iznimno su rijetki (usp. Reinkowski 2001; Jurkiewicz-Rohrbacher i dr. 2017b; Hansen i dr. 2018; Kolaković i dr. 2019; Kolaković i dr. 2022). Upravo zbog toga cilj je ovoga rada na temelju korpusnih podataka i primjenom inferencijskih statističkih metoda pridonijeti boljemu opisu i razumijevanju čimbenika koji utječu

¹ Dok se preskriptivni radovi odlikuju svojom ateoretičnošću, analitički su radovi uglavnom teorijski utemeljeni.

na uporabu zanaglasnica u bosanskoj, hrvatskoj i srpskoj uzimajući u obzir varijaciju i to ne samo onu koja postoji na razini norme nego i uzusa.

Ovaj je rad posvećen varijaciji unutar zamjeničkih zanaglasnica. Naime, u bosanskoj, hrvatskoj i srpskoj postoje dvije varijante akuzativne zamjeničke zanaglasnice trećega lica jednine ženskoga roda: *ju* i *je*. Sudeći prema stanju u suvremenim gramatikama i jezičnim savjetnicima, uporaba se tih zanaglasnica u standardnim idiomima spomenutih jezika razlikuje, usp. primjere (1a–d).

- | | | |
|------|----------------------|--------------------------|
| (1a) | <i>Vidim ju.</i> | (Ham i dr. 2014: 74) |
| (1b) | * <i>Vidim je.</i> | (Ham i dr. 2014: 74) |
| (1c) | <i>On je vidi.</i> | (Piper i Klajn 2014: 97) |
| (1d) | * <i>On ju vidi.</i> | (Piper i Klajn 2014: 97) |

Mišljenja bosanskih, hrvatskih i srpskih gramatičara o tome u kojim se kontekstima moraju, mogu ili smiju rabiti akuzativne zamjeničke zanaglasnice *ju* i *je* značajno su se mijenjala tijekom prošlih stoljeća. Štoviše, spomenute su varijante dijelom bile i predmetom žučnih sporenja hrvatskih i srpskih jezikoslovaca tijekom razdoblja tzv. srpsko-hrvatskoga, odnosno hrvatsko-srpskoga jezika. Isti primjer koji u *Hrvatskim jezičnim savjetima* (Ham i dr. 2014: 74) služi ilustraciji poželjne, normativne uporabe u suvremenome standardnome idiomu hrvatskoga jezika u *Hrvatskome ili srpskome jezičnome savjetniku* (Maretić 1924: 39) smatrao se provincijalizmom, v. (1a).

Valja dodati kako je s dijakronijske točke gledišta ishodišni akuzativni oblik spomenute zanaglasnice bio *ju* (usp. Antić 1998: 9; Matasović 2008: 239), a s vremenom se na pojedinim jezičnim područjima, odnosno u tekstovima pojedinih autora umjesto njega počinje rabiti *je* – izvorno genitivni oblik zamjeničke zanaglasnice trećega lica jednine ženskoga roda na što se u svojim opisima osvrću i neki gramatičari (npr. Veber 1873: 41). Pod Maretićevim (1899: 186; 1924: 39) utjecajem krajem 19., odnosno početkom 20. stoljeća, uporaba se izvorno genitivnoga oblika zamjeničke zanaglasnice *je* počinje pojačano preporučivati i u većini akuzativnih konteksta, dok se uporaba izvorno akuzativne zanaglasnice *ju* značajno ograničava.² Zbog snage njegova autoriteta takvo je stanje, uz sitnije promjene, ostalo u mnogim jezičnim priručnicima sve do danas (Zoričić 1998: 98). Naime, iako je akuzativna zanaglasnica *ju* još uvijek vrlo živa ne samo u nekim hrvatskim nego i bosanskim i srpskim dijalektima (usp. Sekereš 1977: 331; Lisac 2003: 130; Hoyt 2012: 64; Peco 2007: 200; Stevanović 1950: 111; Bukumirić 2003: 223; Okuka 2008: 272–273, 254;

² Cilj ovoga rada nije istražiti razloge koji su doveli do spomenute odluke, odnosno normativne preporuke. Dijakronijska varijacija u uporabi akuzativnih zanaglasnica *ju* i *je* u okviru hrvatskih idioma od 16. do 21. stoljeća, kao i (jezično)politički konteksti i ostali čimbenici koji su utjecali na promjene u normativnim preporukama tema je posebnih radova, v. Kolaković i Blagus Bartolec (2022) te Kolaković i dr. (2022).

Mladenović 2010: 51; Wonisch 2012: 223; Kolaković i dr. 2022: 216), njezina je uporaba u standardnome idiomu bosanskoga i srpskoga jezika ograničena na određene kontekste. Za razliku od toga, mnogi hrvatski gramatičari (Barić i dr. 1997: 210; Raguž 1997: 78; Raguž 2010: 95; Silić i Pranjković 2007: 121) i autori jezičnih savjetnika (Matković 2006: 76) dopuštaju znatno širu uporabu akuzativne zanaglasnice *ju* pri čemu *ju* neki od njih (Mamić 1997: 62; Zoričić 1998: 99; Težak i Babić 2000: 127; Ham 2007: 60; Frančić i Petrović 2013: 143; Ham i dr. 2014: 74) gotovo izjednačuju s njezinom varijantom *je* ili joj čak daju prednost.

Mamić (1997: 61) tvrdi kako se ishodišno Maretijevo kruto pravilo o raspodjeli zanaglasnica *ju* i *je* tijekom 20. stoljeća postupno proširivalo u korist oblika *ju* i čini se da je to doista tako ne samo u hrvatskim nego i u srpskim jezičnim priručnicima. Međutim, kontrastivnim iščitavanjem preskriptivne i deskriptivne jezikoslovne literature uočavaju se određene razlike koje pak do sada nisu bile empirijski proučavane ni u standardnim ni u ostalim idiomima spomenutih jezika. U skladu s tim cilj je ovoga rada dvojak. S jedne strane empirijskim se podacima istražuje opseg varijacije u uporabi akuzativnih zanaglasnica trećega lica jednine ženskoga roda *ju* i *je* u okviru bosanskoga, hrvatskoga i srpskoga jezika. S druge strane nastoji se pridonijeti razumijevanju mehanizama koji stoje iza varijacije u uporabi spomenutih akuzativnih zanaglasnica u standardnim i razgovornim idiomima bosanskoga, hrvatskoga i srpskoga jezika.

Razdjel 2 donosi prikaz literature o uporabi akuzativnih zanaglasnica *ju* i *je* na temelju kojega se oblikuju istraživačka pitanja predstavljena u razdjelu 3. Istraživački nacrt, izvor jezične građe i prikupljeni uzorci primjera s akuzativnim zanaglasnicama *ju* i *je* opisani su u razdjelu 4. Nakon toga, u razdjelima 5, 6 i 7 slijede rezultati, rasprava i zaključci te smjernice za daljnja istraživanja.

2. Pregled literature

2.1. *Općenito o zanaglasnicama u gramatikama*

Nenaglasnice ili klitike riječi su bez naglasaka koje se u bosanskoj, hrvatskoj i srpskoj jeziku vežu ili uz prethodnu ili uz susljednu riječ u jednu izgovornu, odnosno naglasnu cjelinu (usp. Ham 2007: 24; Silić i Pranjković 2007: 21; Stanojčić i Popović 1992: 41; Mrazović i Vukadinović 2009: 55). Među tim se atoničnim riječima razlikuju prednaglasnice, tj. proklitike koje se naslanjaju na susljednu riječ i zanaglasnice, odnosno enklitike koje se naslanjaju na prethodnu riječ. Za razliku od enklitika, proklitike u standardnim idiomima mogu nositi naglasak (Težak i Babić 1994: 72; Barić i dr. 1997: 71; Ivić i dr. 2007: 72; Piper i Klajn 2014: 27; Jahić i dr. 2000: 125–126, 249).³

30 ³ V. Ćurković (2014: 186) za dijalektne primjere sa zanaglasnicama koje nose zanaglasnu duljinu.

Jezične jedinice koje ulaze u potonju grupu raznolike su i pripadaju različitim vrstama riječi. Sažeto govoreći, među njima se nalaze oblici zamjenica (ličnih i povratne) u genitivu, dativu i akuzativu,⁴ nesvršeni nenaglašeni prezentski oblici glagola *biti* te nenaglašeni pomoćni oblici istoga glagola koji služe tvorbi kondicionala, zatim nenaglašeni prezentski oblici pomoćnoga glagola *htjeti* i upitna čestica *li* (Hudeček, Matković i Čutuk 2012: 105; Fekete i dr. 2005: 310; Ivić i dr. 2007: 161; Piper i Klajn 2014: 28; Jahić i dr. 2000: 126).

2.2. Zamjeničke zanaglasnice u gramatikama

Budući da se u središtu zanimanja ovoga rada nalazi samo varijacija u uporabi ličnih zanaglasnica, u Tablici 1 donosi se samo njihov usporedan prikaz u standardnim idiomima bosanskoga, hrvatskoga i srpskoga jezika (usp. Babić i dr. 1991: 647; Silić i Pranjković 2007: 118–121; Mrazović i Vukadinović 2009: 363–364; Piper i Klajn 2014: 28, 95, 97; Jahić i dr. 2000: 249; Čedić 2008: 111–112).⁵

⁴ Sudeći prema dostupnoj literaturi inventar zanaglasnica u bosanskoj, hrvatskoj i srpskoj standardu nije identičan. Naime, autori bosanskih i srpskih gramatika (npr. Jahić i dr. 2000: 249; Stanojčić i Popović 1992: 90) ili uopće ne navode *si* kao dativni zanaglasnički oblik povratne zamjenice ili pak izričito ističu da u bosanskoj, odnosno srpskoj standardu ne postoji takav oblik (npr. Mrazović i Vukadinović 2009: 367; Ridjanović 2012: 440). Nadalje, pitanje je postoji li uopće *se* kao genitivni zanaglasnički oblik povratne zamjenice (usp. Brabec i dr. 1952: 73; Stevanović 1986: 304; Težak i Babić 1994: 108; Težak i Babić 2000: 127; Raguž 1997: 77; Barić i dr. 1997: 72, 208; Ham 2007: 60; Stanojčić i Popović 1992: 90; Mrazović i Vukadinović 2009: 367). Jahić i dr. (2000: 249) ne navode čak ni *se* kao akuzativni zanaglasnički oblik.

⁵ Valja napomenuti da autori hrvatskih i bosanskih gramatika (Babić i dr. 1991: 647; Barić i dr. 1997: 208–209; Raguž 1997: 77; Silić i Pranjković 2007: 118–121; Ham 2007: 59; Jahić i dr. 2000: 249–250; Čedić 2008: 111–112) prikazujući paradigmu ličnih zamjenica naglašene oblike GDA prvoga i drugoga lica te trećega lica muškoga i srednjega roda jednine označavaju kratkosilaznim naglaskom, a srpskih kratkouslaznim (Stevanović 1986: 303; Stanojčić i Popović 1992: 90; Mrazović i Vukadinović 2009: 363–364; Piper i Klajn 2014: 95–97). Mrazović i Vukadinović (2009: 364) u tekstu napominju kako je i kratkosilazni naglasak moguć, a Babić i dr. (1991: 647) i Jahić i dr. (2000: 250) dodaju kako je moguć i kratkouslazni naglasak, pri čemu se u prvospomenutoj gramatici precizira kako se kratkouslazni naglasak ostvaruje u pjesničkome jeziku. Ivić i dr. (2007: 72) pak za genitiv, odnosno akuzativ prvoga lica jedine navode oba naglasaka, a za zamjenicu trećega lica muškoga roda jednine u istome padežu samo kratkouslazni. Osim toga treba dodati kako se naglasna norma u standardnome idiomu hrvatskoga jezika tijekom dvadesetoga stoljeća očito promijenila jer se u *Gramatici hrvatskoga ili srpskog jezika* (Brabec i dr. 1952: 73–74), baš kao i kod suvremenih srpskih gramatičara, navodi kratkouslazni naglasak te se ističe kako kratkosilazni naglasak koji je uobičajen u nekim krajevima nije književan.

Tablica 1: Nenaglašeni u usporedbi s naglašenim oblicima ličnih zamjenica

| | nom. | gen. | | dat. | | ak. | |
|-----|----------------------|------------|--------------------|------------|--------------------|---------------|--------------------|
| jd. | <i>jâ</i> | <i>me</i> | <i>měne/mène</i> | <i>mi</i> | <i>měni/mèni</i> | <i>me</i> | <i>měne/mène</i> |
| | <i>tî</i> | <i>te</i> | <i>těbe/tèbe</i> | <i>ti</i> | <i>těbi/tèbi</i> | <i>te</i> | <i>těbe/tèbe</i> |
| | <i>ôn, ôno</i> | <i>ga</i> | <i>njěga/njèga</i> | <i>mu</i> | <i>njěmu/njèmu</i> | <i>ga, nj</i> | <i>njěga/njèga</i> |
| | <i>ôna</i> | <i>je</i> | <i>njê</i> | <i>joj</i> | <i>njôj</i> | <i>ju, je</i> | <i>njû</i> |
| mn. | <i>mî</i> | <i>nas</i> | <i>nâs</i> | <i>nam</i> | <i>nâma</i> | <i>nas</i> | <i>nâs</i> |
| | <i>vî</i> | <i>vas</i> | <i>vâs</i> | <i>vam</i> | <i>vâma</i> | <i>vas</i> | <i>vâs</i> |
| | <i>ôni, ôna, ône</i> | <i>ih</i> | <i>njîh</i> | <i>im</i> | <i>njîma</i> | <i>ih</i> | <i>njîh</i> |

Čak je i letimičan pogled na tablicu dovoljan da se uoče dvije vrlo važne činjenice. Kao prvo, sve lične zamjenice osim zamjenica za treće lice jednine muškoga, odnosno srednjega i ženskoga roda imaju iste zanaglasničke oblike u genitivu i akuzativu (usp. Mamić 1995: 187). Kao drugo, i lična zamjenica za treće lice jednine muškoga, odnosno srednjega roda baš kao i lična zamjenica za treće lice ženskoga roda u akuzativu, imaju po dvije varijante *ga* i *nj* te *ju* i *je* (usp. Raguž 1997: 78; Mrzović i Vukadinović 2009: 366).⁶

Zanaglasnica *nj* u sintaktičkome se smislu razlikuje od ostalih zanaglasnica jer obvezno slijedi prijedloge, ne grupira se s drugim zanaglasnicama i općenito nije vezana za drugo mjesto u rečenici (usp. Brabec i dr. 1952: 74; Težak i Babić 1994: 108; Raguž 1997: 78; Barić i dr. 1999: 177, 337; Ham 2007: 60; Hudeček i Mihaljević 2019: 78; Stevanović 1986: 306; Čavar i Wilder 1994: 445). Neki gramatičari (npr. Mrzović i Vukadinović 2009: 366; Piper i Klajn 2014: 96) napominju kako je spomenuta zanaglasnica u srpskoj gramatici arhaizam, dok je drugi (npr. Klajn 2005: 79; Čedić 2008: 111) čak i ne navode u zamjeničkoj paradigmi.⁷ U jezičnim savjetima *Govorimo hrvatski* (Dulčić 1997: 345) priznaje se kako zanaglasnica *nj* u hrvatskoj ima rubnu jezičnu uporabu, no preporučuje se njezino očuvanje, posebice nakon prijedloga *po, u, za, na, kroz, pred* i *uz*. I na kraju valja istaknuti da je raspodjela zanaglasničkih varijanata *ga* i *nj* dobro opisana te se deskriptivni i normativni opisi ne razlikuju od uzusa.⁸ Za razliku od toga, situacija je s akuzativnim zanaglasnicama *ju* i *je* ipak nešto složenija.

⁶ Alerić (1964) tvrdi kako za zamjenicu trećega lica jednine ženskoga roda postoje čak tri zanaglasnice: *ju, je* i *nju* te da i u genitivu osim zanaglasnice *je* postoji zanaglasnica *nje*. S obzirom na to da se zanaglasnički oblici *nju* i *nje* navode u samo nekim gramatikama, ovdje im se ne daje posebna pozornost.

⁷ Neki autori (npr. Jahić i dr. 2000: 250) samo navode spomenutu zanaglasnicu u paradigmi ne osvrćući se podrobnije na njezinu uporabu.

⁸ CQL upit [tag!="S."][word="nj"] pokazuje da se i u Riznici i u hrWaC-u konkordance u kojima se zanaglasnica *nj* ne javlja ispred prijedloga sa završnim dugim *â* mogu svesti na primjere s pogrešno označenim prijedlozima *kroza, nada, preda* i *poda*; primjere s jezičnim inovacijama u kojima genitivni

2.3. Akuzativne zanaglasnice *ju* i *je*

2.3.1. Kratka retrospektiva normativnih preporuka o uporabi akuzativnih zanaglasnica *ju* i *je*

U većini hrvatskih i dijelu bosanskih priručnika (usp. Brabec i dr. 1952: 74; Barić i dr. 1979: 354; Barić i dr. 1997: 209; Raguž 1997: 77; Hudeček i Mihaljević 2019: 78; Čedić 2008: 111) prvo se navodi zanaglasnički oblik *ju*, a ne *je*, iako je oblik *ju* sudeći prema u njima opisanim pravilima ograničeniji u uporabi od akuzativne zanaglasnice *je*. Za razliku od toga, većina srpskih te dio hrvatskih i bosanskih jezikoslovaca (Stanojčić i Popović 1992: 90; Klajn 2005: 79; Mrazović i Vukadinović 2009: 364; Piper i Klajn 2014: 97; Jahić i dr. 2000: 250; Težak i Babić 1994: 107; 2000: 126; Barić i dr. 1999: 177; Raguž 2010: 94) nakon naglašenoga oblika *nju* prvo navodi *je*, a potom *ju*.^{9,10} Mamić (1995: 187) pretpostavlja da razlike u redosljedju navođenja mogu proizlaziti iz djelomične podudarnosti naglašenoga oblika *nju* i nenaglašenoga oblika *ju*, ali i da bi mogle upućivati na otpor hrvatskih gramatičara spomenutome Maretićevev pravilu o ograničavanju uporabe zanaglasnice *ju*.

Naime, tijekom dvadesetoga stoljeća u postmaretićevskim se gramatikama hrvatskih autora nastojalo donekle proširiti uporabu zanaglasnice *ju* koja je trebala ne samo prethoditi glagolskoj zanaglasnici *je* nego i slijediti glagole i ostale riječi koje završavaju na *-je* (Mamić 1995: 187). *Gramatika hrvatskoga ili srpskoga jezika* (Brabec i dr. 1952: 74–75) jedan je od priručnika u kojemu se uporaba zanaglasnice *ju* ne ograničava samo na kontekste kada *ju* slijedi glagolska zanaglasnica *je* nego se, dapače, preporučuje njezino ostvarivanje i nakon riječi koje završavaju na slog *-je*, usp. primjere (2a–b).

(2a) *dajem je tebi* (Brabec i dr. 1952: 75)

(2b) *daje ju tebi* (Brabec i dr. 1952: 75)

Međutim, srpski su jezikoslovci uglavnom oštro kritizirali takve normativne preporuke. Stevanović (1964: 313; 1986: 306) primjerice ne popušta i ne odustaje od Maretićeva pravila da se zamjeničku zanaglasnicu *ju* smije rabiti samo u kombinaciji s glagolskom zanaglasnicom *je* te sve ostale uporabe zanaglasnice *ju* proglašava

prijedlog *od* (!) dobiva dugo *a* i postaje dvosložan (usp. Barić i dr. 1997: 209); primjere u kojima se raspravlja o grafemu *nj* i fonemu /ɲ/ te primjere u kojima je *nj* zapravo kratica za *njevog* u izrazu *njegovo veličanstvo*.

⁹ U *Srpskome jezičkome priručniku* (Ivić i dr. 2007: 161) pri pobrojavanju zanaglasnica uopće se ne navodi zamjenička zanaglasnica *ju*. Slično se opaža i u paradigmatičkim zamjenicama u gramatici *Bosnian for Foreigners* (Ridjanović 2012: 434).

¹⁰ Valja istaknuti kako usporedba natuknice *ona* otkriva da se i u hrvatskim i u srpskim rječnicima (Birtić i dr. 2012: 441; Jojić i dr. 2015: 956; Vujanić i dr. 2011: 861) prvo navodi akuzativna zanaglasnica *je*, a potom *ju*. U *Rječniku bosanskog jezika* (Čedić i dr. 2007: 539) ne navode se kosi oblici zamjenice za treće lice jednine.

Samo se u dijelu srpskih jezikoslovnih priručnika (npr. Klajn 2005: 79; Ivić i dr. 2007: 128; Mrazović i Vukadinović 2009: 366) objašnjava da niječni oblik *nije* uz koji se dopušta uporaba zamjeničke zanaglasnice *ju* dolazi ispred nje, dok se u drugima (npr. Klajn 2004: 102; Piper i Klajn 2014: 97) to ne precizira u obliku pravila, nego se samo navode primjeri u kojima *nije* stoji ispred akuzativne zanaglasnice *ju*. Međutim, sudeći prema objašnjenjima u *Srpskome jezičkome priručniku* (Ivić i dr. 2007: 128) uporaba se akuzativne zanaglasnice *je* nakon niječnoga oblika *nije* u srpskome standardu u međuvremenu počela smatrati prihvatljivom (usp. Antonić 1998: 9), usp. primjere (6) i (4).

(6) *Nije je stid.* (Ivić i dr. 2007: 128)

Većina bosanskih gramatika i jezičnih savjetnika (Jahić i dr. 2000: 249–250; Čedić 2008: 111–112; Čedić 2013) uopće ne precizira kada se rabi akuzativna zanaglasnica *ju*, a kada *je*. Iznimka je gramatika *Bosnian for Foreigners* (Ridjanović 2012: 434). Iako se u spomenutoj gramatici u paradigmi ne navodi zanaglasnica *ju*, nego samo *je*, u tekstu ispod paradigme ipak se spominje kako se akuzativna zanaglasnica *je* mijenja u *ju* ako ju slijedi glagolska zanaglasnica *je*, v. (7a).

(7a) *On ju je vidio.* (Ridjanović 2012: 434)

(7b) *On je vidio.* (Ridjanović 2012: 434)

Napominje se kako je to obilježje normiranoga idioma, dok se u razgovornome idiomu bosanskoga jezika rabi samo jedan kontrahirani *je*, v. (7b) (Ridjanović 2012: 434).

2.3.3. Zanaglasnice *ju* i *je* u suvremenome hrvatskome standardu

Kada je u pitanju izbor između varijanata *ju* i *je*, hrvatski jezikoslovni priručnici objavljeni u (kasnoj) drugoj polovici 20. stoljeća, no prije raspada Jugoslavije, uglavnom slijede tradiciju koju su postavili Brabec i dr. (1952: 74). Tako se i u *Jezičnome savjetniku s gramatikom* (Pavešić i dr. 1971: 112) i u *Jezičnim savjetima* (Vidović 1983: 11) navodi kako zanaglasnica *ju* dolazi samo u kombinaciji s glagolskom zanaglasnicom *je* te iza koje druge riječi na *-je* i to sve u svrhu blagoglasja, dok se inače rabi akuzativna zanaglasnica *je*. Štoviše, u potonjemu se priručniku izričito normativno diskvalificira rečenice sa zanaglasnicom *ju* koja nije ni u pretkontaktnome položaju s glagolskom zanaglasnicom *je* ni u postkontaktnome položaju s nekom drugom riječju, odnosno oblikom na *-je*, usp. (8a–b).

(8a) **Treba ju pažljivo slušati.* (Vidović 1983: 11)

(8b) *Treba je pažljivo slušati.* (Vidović 1983: 11)

Zanimljiv slučaj čini gotovo neprimjetna izmjena pravila u dvama izdanjima hrvatske gramatike autora s Instituta za hrvatski jezik i jezikoslovlje. Naime, u *Priručnoj gramatici hrvatskoga jezika* (Barić i dr. 1979: 123) pojašnjava se kako se zamjenička zanaglasnica *ju* rabi kada *ju* slijedi glagolska zanaglasnica *je* ili kada joj prethodi slog *-je* (v. primjere (9a) i (10a)), dok u svim ostalim slučajevima uvijek dolazi zanaglasnica *je* (v. primjere (9b) i (10b)).

- | | | |
|-------|---|-------------------------|
| (9a) | <i>Nije ju nikad vidio.</i> | (Barić i dr. 1979: 123) |
| (9b) | <i>Nisam je nikad vidio.</i> | (Barić i dr. 1979: 123) |
| (10a) | <i>Ne voli limunadu, ali pije ju kada je vruće.</i> | (Barić i dr. 1979: 123) |
| (10b) | <i>Ne voli limunadu, ali je pije kada je vruće.</i> | (Barić i dr. 1979: 123) |

Za razliku od toga, u gotovo dvadeset godina novijemu izdanju, točnije govoreći u *Hrvatskoj gramatici* (Barić i dr. 1997: 210) ponavlja se pravilo o obveznome ostvarivanju zanaglasnice *ju* u kontaktnome položaju s glagolskom zanaglasnicom *je* i s prethodnim slogom *-je* te se dodaje da se u svim drugim slučajevima obično rabi zanaglasnica *je*. Međutim, valja istaknuti kako prilog *obično* u formulaciji pravila *Hrvatske gramatike* (Barić i dr. 1997: 210) zapravo ne isključuje pojavljivanje zanaglasnice *ju* i izvan spomenutih obveznih konteksta, kao što je to bio slučaj u gotovo dvadeset godina starijemu izdanju, odnosno *Priručnoj gramatici hrvatskoga jezika* (Barić i dr. 1979: 123). Usporedbe radi, u *Hrvatskome jezičnome savjetniku* (Barić i dr. 1999: 173) ističe se kako je uporaba zanaglasnice *ju* obvezna kada uz nju stoji nenaglašeni prezent pomoćnoga glagola *biti* i kad se izbjegavaju istozvučni oblici.¹² Međutim, za potonje se, nažalost, ne navode nikakvi primjeri. Štoviše, spomenuti savjetnik ne govori ništa o tome dopušta li norma uporabu zanaglasnice *ju* u hrvatskome standardnome idiomu kada u njezinoj neposrednoj blizini nema ni glagolske zanaglasnice *je* ni tzv. istozvučnih oblika.

Babić i dr. (1991: 648) ističu kako akuzativne zanaglasnice trećega lica jednine ženskoga roda imaju komplementarnu raspodjelu: zanaglasnica *ju* rabi se ispred glagolske zanaglasnice *je* i iza riječi i oblika na *-je*, *-lje* i *-nje*,¹³ a u ostalim kontekstima dolazi *je*. Nadalje, u istoj se gramatici napominje i pokazuje kako su stariji hrvatski pisci¹⁴ rabili zanaglasnicu *ju* i izvan gore spomenutih konteksta (Babić i dr. 1991: 648).¹⁵

¹² Pravilo s tzv. istozvučnim oblicima nije baš najsretnije formulirano. Štoviše, čitatelju nelingvistu vrlo vjerojatno čak nije ni jasno s čime su ti oblici istozvučni: sa zanaglasnicom *ju* ili pak s nenaglašenim prezentom glagola *biti*? A koji je točno nenaglašeni oblik prezenta glagola *biti* u pitanju naslućuje se tek iz primjera (usp. Barić i dr. 1999: 173).

¹³ Valja istaknuti kako autori ne navode primjere u kojima zanaglasnica *ju* slijedi riječi i oblike na *-lje* i *-nje*, nego samo primjere u kojima se nalazi iza niječnoga oblika *nije* i ispred glagolske zanaglasnice *je* (usp. Babić i dr. 1991: 648).

¹⁴ Navode se primjeri iz tekstova A. G. Matoša, E. Kumičića, V. Novaka i F. Mažuranića.

¹⁵ Usp. „[...] u starijih pisaca češće se upotrebljava *ju* u svim okolinama“ (Matasović 2008: 239).

Korak dalje u prešutnome prihvaćanju znanaglasnice *ju* i izvan konteksta određenih strogim pravilima uočava se u *Praktičnoj gramatici hrvatskoga jezika* (Raguž 1997: 78) i *Gramatici hrvatskoga jezika* (Raguž 2010: 95). Istini za volju, u spomenutim se gramatikama ističe da je oblik *je* prvotni oblik znanaglasnice, dok se alternativni oblik *ju* rabi samo u posebnim slučajevima: kada ga slijedi glagolska znanaglasnica *je*. Međutim, isto se tako napominje kako se, iako rijetko, u standardnome idiomu oblik *ju* javlja i tolerira i izvan spomenutoga konteksta, za što se navode dva primjera (Raguž 1997: 78; Raguž 2010: 95), v. (11) i (12).

- (11) *Gleda ju.* (Raguž 1997: 78)
(12) *Ljuti ju stalno.* (Raguž 1997: 78)

Dvije godine nakon objavljivanja Mamićeva (1995) članka i godinu dana nakon što je *Praktična gramatika hrvatskoga jezika* (Raguž 1997) ugledala svjetlo dana, više hrvatskih jezičnih savjetnika (npr. Mamić 1997: 62; Zoričić 1998: 99) zauzima se za to da se akuzativni zamjenički oblici *ju* i *je* prepuste slobodnoj uporabi. To bi, tvrdi se, znatno pridonijelo prestanku prekida između suvremenoga hrvatskoga standarda i hrvatske književne tradicije (v. Zoričić 1998: 99). Nadalje, još krajem prošloga stoljeća u jezičnome savjetniku *Hrvatski u praksi* (Zoričić 1998: 99) predviđalo se kako znanaglasnica *ju* uskoro više neće biti na rubu gramatičke norme. Osim toga, u tome se jezičnome savjetniku preporučuje da lektori i učitelji više ne uklanjaju znanaglasnicu *ju* iz tekstova (v. Zoričić 1998: 99).

U nastavku, zanimljivu sliku svakako daje i *Jezični savjetnik – iz prakse za praksu* (Matković 2006: 76) u kojemu se ponavlja isto pravilo o izbjegavanju ponavljanja istovrsnih *-ju* odnosno *-je* slogova, a u slučaju riječi koje završavaju na *-e* dopušta se uporaba obiju akuzativnih znanaglasnica, usp. (13a–b).

- (13a) *Ne uznemirujte je.* (Matković 2006: 76)
(13b) *Ne uznemirujte ju.* (Matković 2006: 76)

Čini se kako za razliku od gore spomenutih autora hrvatskih gramatika Silić i Pranjković (2007: 121) zapravo ne ograničavaju uporabu akuzativne znanaglasnice *ju*, nego njezine varijante *je* tvrdeći da se potonja može rabiti svugdje, osim kada u njezinu neposrednu okruženju dolazi slog *-je*, usp. primjere (14)–(16).

- (14) *Vidio ju je.* (Silić i Pranjković 2007: 121)
(15) *Koje ju dijete zanima?* (Silić i Pranjković 2007: 121)
(16) *On ju jede kad god treba.* (Silić i Pranjković 2007: 121)

U svezi s tim valja skrenuti pozornost na činjenicu da se u *Gramatici hrvatskoga jezika* (Silić i Pranjković 2007: 121) ističe kako se zanaglasnicu *ju* može rabiti neovisno o tome nalazi li se u njezinu neposrednu okruženju slog *-je*, v. primjer (17).

(17) Često **ju** pita za savjet. (Silić i Pranjković 2007: 121)

Slično se mišljenje o standardnosti zanaglasnice *ju* i izvan kontaktnoga konteksta s glagolskom zanaglasnicom *je*, niječnicom *nije* te drugim riječima i oblicima koji završavaju na *-je*¹⁶ zastupa u knjigama *Hrvatski jezik i jezična kultura* (Frančić i Petrović 2013: 143) i *Hrvatski jezični savjeti* (Ham i dr. 2014: 74), v. primjer (1a). Štoviše, u *Hrvatskim jezičnim savjetima* (Ham i dr. 2014: 74) ističe se kako je zanaglasnica *ju*, kojoj je uporaba tijekom prošloga stoljeća slijedom unitarističke jezične politike bila značajno sužena, vraćena u standardni idiom hrvatskoga jezika. Ipak, u obama priručnicima uporaba zanaglasnice *ju* nije potpuno neograničena zbog pravila o blagoglasju, odnosno o izbjegavanju ponavljanja istoga sloga. Naime, kako bi se izbjeglo ponavljanje dvaju istih slogova, oba priručnika (usp. Frančić i Petrović 2013: 143; Ham i dr. 2014: 74) preporučuju: ako prethodna riječ završava na *-ju*, treba uporabiti akuzativnu zanaglasnicu *je*, a ako pak završava na *-je*, savjetuje se uporaba zamjeničke zanaglasnice *ju*, usp. primjere (18a–b).

(18a) *Voda je studena i piju je kada je vruće.* (Ham i dr. 2014: 74)

(18b) **Voda je studena i piju ju kada je vruće.* (Ham i dr. 2014: 74)

I na kraju, valja istaknuti kako Frančić i Petrović (2013: 143) napominju da izvorni govornici, kada se nastoje koristiti standardnim idiomom hrvatskoga jezika, zapravo vrlo često odstupaju od spomenutoga pravila o blagoglasju.

2.3.4. Zanaglasnice *ju* i *je* u školskim gramatikama hrvatskoga jezika

U kontekstu izbora između dviju varijanata nenaglašenoga akuzativnoga oblika zamjenice trećega lica jednine ženskoga roda posebno zanimljivu sliku daju usporedbe hrvatskih školskih gramatika. U *Gramatici hrvatskoga jezika* (Težak i Babić 1994: 108) potpuno u skladu s tradicijom koju su postavili Brabec i dr. (1952: 74–75) uporaba zanaglasnice *ju* ograničava se na kontaktni kontekst s glagolskom zanaglasnicom *je* te drugim riječima i oblicima koji imaju završni slog *-je* (Težak i Babić 1994: 108), v. primjere (19)–(21).

(19) *Uzela ju je.* (Težak i Babić 1994: 108)

(20) *Sova iz duplja izađe svoga, vuka i lisice nije ju strah.* (Težak i Babić 1994: 108)

(21) *Kupuje ju.* (Težak i Babić 1994: 108)

¹⁶ U *Hrvatskim jezičnim savjetima* (Ham i dr. 2014: 74) spomenuto se pravilo proglašava starom normom koju suvremena hrvatska norma već desetljeće i pol ostavlja po strani.

Međutim, iako se u 12. prerađenoj i dopunjenoj izdanju *Gramatike hrvatskoga jezika* (Težak i Babić 2000: 126) u paradigmi uz naglašeni akuzativni oblik *nju* prvo navodi *je*, a tek potom *ju*,¹⁷ u objašnjenju se ističe da zbog sustavnosti kao izvorni akuzativni oblik prednost ipak ima oblik *ju*, a ne *je* koji je preuzet iz genitiva. Težak i Babić (2000: 127) ograničavaju uporabu oblika *je* te savjetuju da ga se rabi kako bi se izbjeglo ponavljanje sloga *-ju* koji se pak može nalaziti ispred ili iza akuzativne zamjenice trećega lica jednine ženskoga roda, v. primjere (22) i (23).

(22) *Poštuju je.* (Težak i Babić 2000: 127)

(23) *Vidjeli smo je jutros.* (Težak i Babić 2000: 127)

Na ovome mjestu valja dodati kako su Težak i Babić (2000: 127) i Silić i Pranjko-
vić (2007: 121) jedini autori koji upućuju na važnost prvoga sloga susljedne riječi pri izboru između dviju varijanata zamjeničke zanaglasnice trećega lica jednine ženskoga roda u akuzativu, v. primjere (16) i (23).

Sličan se stav o teoretskoj ravnopravnosti obiju varijanata zastupa i u *Školskoj gramatici hrvatskoga jezika* (Ham 2007: 60). Međutim, i u toj se gramatici, baš kao i kod Težaka i Babića (2000: 126) ističe kako se prednost ipak treba dati zanaglasnici *ju*. Potom se pravilo blagoglasja oprimjeruje raznim kontekstima u kojima se ograničava uporaba i jedne i druge varijante (Ham 2007: 60).

I na kraju valja istaknuti kako se u najnovijoj *Hrvatskoj školskoj gramatici* (Hudeček i Mihaljević 2019: 78) uporaba zanaglasnice *ju* u standardnome idiomu hrvatskoga jezika slično kao i u standardnome idiomu srpskoga jezika ograničava samo na kontaktne kontekste s glagolskom zanaglasnicom *je* i riječima koje završavaju slogom *-je*.

Nenaglašeni oblik akuzativa zamjenice ona glasi *je*, a oblik *ju* upotrebljava se ispred oblika *je* pomoćnoga glagola *biti*, npr. *On ju je volio.* i kada je ispred njega slog *je*, npr. *Nije ju volio.*

(Hudeček i Mihaljević 2019: 78)

Autorice *Hrvatske školske gramatike* (Hudeček i Mihaljević 2019: 78) ostaju, dakle, vjerne tradiciji koju su postavili Brabec i dr. (1952: 74–75) još sredinom prošloga stoljeća čime se jasno razlikuju od ostalih hrvatskih gramatika i jezičnih savjetnika objavljenih od 90-ih godina prošloga stoljeća u kojima se između redaka nazire ili pak jasno uočava zaokret prema potpunome prihvaćanju kontekstno neuvjetovane uporabe zanaglasnice *ju*.

¹⁷ Pri pripremi novoga izdanja autori su vrlo vjerojatno zaboravili ispraviti redoslijed zanaglasnica u paradigmi.

3. Istraživačka pitanja

Detaljna kontrastivna analiza suvremenih bosanskih, hrvatskih i srpskih gramatika i jezičnih savjetnika upućuje na moguću varijaciju ne samo među trima standardima nego i među njihovim inim idiomima. Štoviše, razlike su moguće i među neorganskim i organskim idiomima unutar jednoga jezika. Empirijski dio ovoga rada donosi odgovore na sljedeća istraživačka pitanja:

P1: Kakva je distribucija akuzativnih zamjeničkih zanaglasnih varijanata *ju* i *je* u korpusima koji zrcale standardne idiome bosanskoga, hrvatskoga i srpskoga jezika? Odražava li spomenuta korpusna građa razlike opisane u analiziranoj literaturi (usporedba standardni idiom – standardni idiom)?¹⁸

P2: Kakva je distribucija akuzativnih zamjeničkih zanaglasnih varijanata *ju* i *je* u korpusima koji predstavljaju razgovorne idiome bosanskoga, hrvatskoga i srpskoga jezika (usporedba razgovorni idiom – razgovorni idiom)?

P3: Kakva je distribucija akuzativnih zamjeničkih zanaglasnih varijanata *ju* i *je* u okvirima jednoga jezika (bosanskoga, hrvatskoga i srpskoga)? Uočavaju li se razlike u uporabi akuzativnih zamjeničkih zanaglasnih varijanata *ju* i *je* u standardnome i razgovornome idiomu jednoga jezika (usporedba standardni idiom – razgovorni idiom)?

P4: Utječe li završetak riječi koja prethodi zanaglasnici na izbor između dviju akuzativnih zanaglasnih varijanata *ju* i *je* kako se to pretpostavlja u predstavljenoj literaturi, (v. razdjelje 2.3.1 – 2.3.4)?

U sljedećemu razdjelu detaljno se prikazuje istraživački nacrt i način prikupljanja empirijskih podataka u svrhu iznalaženja odgovora na istraživačka pitanja.

4. Metodologija i prikupljanje jezične građe

4.1. Istraživački nacrt i operacionalizacija varijabli

Istraživački nacrt uključuje pet kategoričkih varijabli: jednu zavisnu i četiri nezavisne varijable. Varijanta nenaglašenoga akuzativnoga oblika zamjenice trećega lica jednine ženskoga roda s dvjema razinama, *ju* i *je*, zavisna je varijabla. Iako je riječ o kategoričkoj varijabli, u svrhe ovoga istraživanja i statističke analize kodirana je kao binarna varijabla: oblik *ju* uzimao se kao vrijednost 1, dok je varijanta *je* odgovarala vrijednosti 0 (usp. Janda, Nessel i Baayen 2010: 32). Prva nezavisna varijabla

¹⁸ Razdjel 4.1 donosi informaciju o tome što se u radu smatra standardnim, a što razgovornim idiomom, odnosno kratko se komentira kako je nezavisna varijabla idiom operacionalizirana i zašto je to učinjeno na baš takav način, dok se u razdjelu 4.2 opisuju izvori jezične građe koji odražavaju spomenute idiome.

riječ – koja prethodi zanaglasnici akuzativnoga oblika zamjenice trećega lica jednine ženskoga roda ima tri razine (v. Tablicu 2): riječ koja završava na *-(j)u*, riječ koja završava na *-(j)e* te riječ s ostalim završecima.

Kao što je opisano u razdjelima 2.3.3 i 2.3.4, u dvjema hrvatskim gramatikama (Silić i Pranjaković 2007: 121; Težak i Babić 2000: 127) napominje se kako i početak riječi koja slijedi akuzativnu zanaglasnicu trećega lica jednine ženskoga roda ima utjecaja na to koja bi od dviju varijanata trebala biti uporabljena. Međutim, u ovo istraživanje nije bilo moguće uključiti nezavisnu varijablu početak riječi koja slijedi zanaglasnicu s trima razinama: riječ koja počinje na *(j)u-*, riječ koja počinje na *(j)e-*, riječ s ostalim počecima. Naime, u prikupljenoj je građi samo potonja razina spomenute varijable bila dobro potvrđena, dok su prve dvije razine bile iznimno slabo potvrđene zbog čega pak statistički model ne bi dobro radio.¹⁹ Iako su podatci bili označeni i za tu varijablu, ona u daljnjemu istraživanju nije bila praćena zbog prirode same jezične građe.

Umjesto upravo spomenute varijable, u istraživanje je kao druga nezavisna varijabla uključena riječ koja slijedi akuzativnu zanaglasnicu trećega lica jednine ženskoga roda s dvjema razinama: glagolski oblik *je* te ostale riječi. Jezik kao treća nezavisna varijabla ima tri razine: bosanski, hrvatski i srpski. Idiom, četvrta je i posljednja nezavisna varijabla s dvjema razinama: standardni i razgovorni. Tablica 2 donosi sažet prikaz varijabli i njihov opis, odnosno podatke o njihovoj vrsti i tipu, razinama i kodiranju.

¹⁹ Da bi se prikupio dovoljan broj primjera za riječi koje slijede akuzativnu zanaglasnicu *ju* ili *je* i počinju s *(j)e-* ili *(j)u-*, trebalo bi oblikovati CQL upite kojima se ciljano u korpusima traže riječi s tim počecima dok se izbor između dviju spomenutih varijanata akuzativne zanaglasnice ostavlja slobodnim. Tada bi pak postojala opasnost da zanaglasnica *je* bude prezastupljena u podacima, odnosno da postoji premalo podataka sa zanaglasnicom *ju*, posebice u korpusima s bosanskim i srpskim idiomima. Osim toga, CQL upiti primjenjivi su samo u nekim korpusima koji su u ovome istraživanju poslužili kao izvor jezične građe, v. sljedeći razdjel. Podatke bi se, naravno, alternativno moglo prikupljati i s pomoću psiholingvističkih metoda kao što su procjena prihvatljivosti ili ispunjavanje praznina (za više informacija v. Kolaković i dr. 2022: 59–66).

Tablica 2: Popis varijabli

| Vrsta | | Varijabla | Tip | Razina | Kod |
|-----------|----|--|-------------|---|-----|
| ZAVISNA | 1. | varijanta akuzativne zanaglasnice trećega lica jednine ženskoga roda | binarna | ju (1) je (0) | CL |
| | | | | | |
| NEZAVISNA | 1. | riječi koja prethodi zanaglasnici | kategorička | riječ koja završava na <i>-(j)u</i> riječ koja završava na <i>-(j)e</i> riječ s ostalim završcima | |
| | 2. | riječi koja slijedi zanaglasnicu | kategorička | glagolski oblik <i>je</i> ostale riječi | |
| | 3. | jezik | kategorička | bosanski hrvatski srpski | |
| | 4. | idiom | kategorička | standardni razgovorni | |

Kao što se može pretpostaviti, nezavisnu varijablu idiom zasigurno je najteže operacionalizirati na potpuno zadovoljavajuć i pouzdan način, a glavni su uzrok tomu prijepori u svezi s tim što čini i što bi trebalo činiti eksplicitnu kodifikacijsku normu i koja je uloga uzusa u tome. Iako se dobar dio stručne javnosti zalaže za to da kodifikacijskonormativni kriterij treba biti visoka uporabna čestota (usp. Kordić 2006: 195–196; Vuković 2016: 228–229; Starčević i dr. 2019; Belaj 2022: 17–18; Kapović 2022: 23),²⁰ u Hrvatskoj primjerice procvat jezičnosavjetodavne djelatnosti ne jenjava, štoviše uz uobičajene tiskane priručnike s jezičnim savjetima, kolumne i blogove, u posljednje vrijeme mogu se pratiti i *podcasti* s jezičnim savjetima. To da postoji nesuglasje oko toga treba li kodificirana norma zapravo odražavati uzus, nije nešto što se može pratiti samo u Hrvatskoj. S time je usko povezano i pitanje je li „razgovorni stil“ jedan od pojava oblika „standardnoga jezika“ koje se nalazi u središtu mnogobrojnih znanstvenih radova.²¹ Iako to pitanje nije jedno od središnjih istraživačkih pitanja ovoga rada, jedno je od ključnih problema koji se izravno tiču metodologije i izbora izvora jezične građe.

Iako dobar dio hrvatskih jezikoslovaca zastupa stajalište da je „razgovorni stil“ dio „standardnoga jezika“, u konačnici, baš kao i u svezi s mnogobrojnim drugim jezikoslovnim problemima, i u slučaju spomenutoga pitanja u lingvistici nema jedno-

²⁰ Dok dio autora smatra da uzor normi treba biti uzus razgovornoga idioma, valja istaknuti kako Belaj (2023) argumentira da se pri kodifikaciji moraju posebno slijediti visokofrekventne konstrukcije u svakome pojedinome funkcionalnome stilu. Drugim riječima, ono što je uzus, tj. uobičajeno, ili često u primjerice administrativnome stilu ne mora ujedno biti i u ostalim funkcionalnim stilovima. Posljedično, pojedini se funkcionalni stilovi samo jednim dijelom podudaraju u svojoj kodifikacijskoj normi.

²¹ U ovome se radu rabi termin *standardni idiom*, a ne *standardni jezik* čime se neizravno pokazuje da standard ni po čemu nije superiorniji u odnosu na ostale idiome (usp. Langston i Peti-Stantić 2014: 31; Starčević i dr. 2019: 10), v. termine *standardni dijalekt* (Starčević i dr. 2019: 10–11) i *jezični varijetet* (engl. *language variety* u Stefanotwisch 2020: 6).

glasnoga odgovora (usp. Silić 2006: 108; Opačić 2006; Horvat 2007: 37; Škarić 2009: 20–21; Kryžan-Stanojević 2010: 240; Udier i Gulešić Machata 2011: 334; Jelaska i Bjedov 2015: 229, 247). Štoviše, radovi u kojima se polazi od pretpostavke da je „razgovorni stil“ dio „standardnoga jezika“ često joj izravno ili pak neizravno proturječe:

Razgovorni je jezik ušao svugdje gdje je do sada vladao književni, odnosno standardni jezik. Pojavio se u medijima, ušao u državne ustanove, obrazovne institucije, u Sabor. U književnosti se pojavio kao jezik naracije. Nije ni čudo da se pojavio i u udžbenicima. [...] Vladalo je uvjerenje da ćemo kada naučimo standardni jezik lako savladati i sve njegove varijante i stilove od razgovornoga do sociolekata i jezika subkultura.

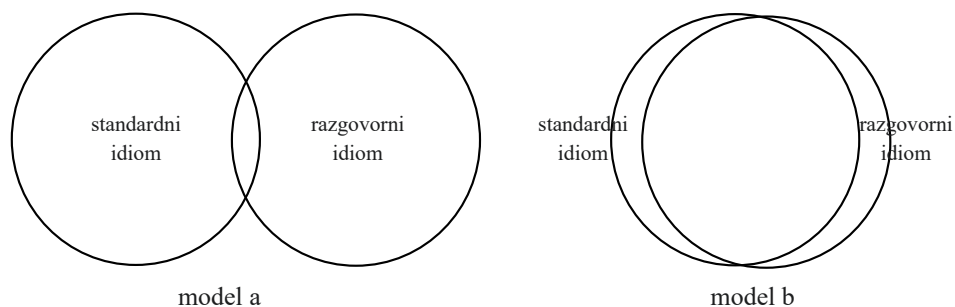
Kryžan-Stanojević (2010: 240)

Prethodni citat pokazuje kako se s jedne strane „razgovorni stil“ drži dijelom „standardnoga jezika“. Međutim, s druge strane isto je tako razvidno da se „razgovorni jezik“ ipak ne smatra jednim od pet jasno odvojenih funkcionalnih stilova, nego nečim što se proteže kroz ostale funkcionalne stilove, od administrativnoga preko publicističkoga i književnoumjetičkoga do znanstvenoga i ulazeći u njihove pore izravno konkurira „standardnomu jeziku“ potiskujući njegovu uporabu u spomenutim domenama (usp. Hudeček i Mihaljević 2009: 20–22). Spomenuto proturječe vjerojatno proizlazi iz činjenice da iako se „razgovorni stil“ „standardnoga jezika“ redovito navodi u literaturi, pokazuje se da je njegovo postojanje u praksi upitno slijede li se pravila standarda kako je kodificiran u postojećim referentnim priručnicima. Drugim riječima, „razgovorni [se] stil standardnoga jezika rijetko ostvaruje i rijetko čuje“ (Frančić i dr. 2005: 242).

Iako se na prvu ruku odluka o razdvajanju standardnoga i razgovornoga idioma možda može činiti donekle oprečnom dijelu hrvatske teorijske tradicije u okviru koje se „razgovorni stil“ smatra jednom od realizacija „standardnoga jezika“, ona je ipak i teorijski i metodološki utemeljena. Kao prvo, standardni se idiom, neovisno o tome je li riječ o pismenome ili usmenome mediju njegove realizacije, definira kao kodificirani varijetet (v. Brozović 1970: 28). Za razliku od njega, za razgovorni se idiom tvrdi da se, iako umnogome dijeli gramatičku strukturu i leksik sa standardnim idiomom, od njega razlikuje nedostatkom eksplicitne kodifikacije te ga kao takvoga karakterizira mnogo veći stupanj varijacije uključujući i sadržavanje jezičnih elemenata svojstvenih inim idiomima (usp. Mićanović 2008: 14; Langston i Peti-Stantić 2014: 30; Brozović 2016: 39, 385–402; Pasini 2018: 587–593; Bošnjak Botica i dr. 2022: 29). Drugim riječima, razgovorni idiom, neovisno o pismenoj ili usmenoj realizaciji (usp. Silić i Pranjković 2007: 387; Pasini 2018: 592), sadrži obilježja i koja jesu dio kodificirane norme i ona koja to nisu. Iako se potpuna kodifikacija u konačnici nikada ne može postići te je standard više teoretski apstraktno zamišljen koncept nego realnost i iako je između ostaloga zbog jezičnih promjena dijelom neispravno svoditi standardni idiom na popis ponekad i proturječnih pravi-

la zabilježenih u normativnim priručnicima (usp. Milroy i Milroy 1985: 22–23; Joseph 1987: 160–161; Ammon 1987: 328), za potrebe ovoga empirijskoga istraživanja standardni je idiom bilo najprikladnije operacionalizirati upravo kao aproksimaciju skupa pravila zabilježenih u referentnim priručnicima (gramatikama, jezičnim savjetnicima i rječnicima). S obzirom na to da se spomenuta pravila u stvarnoj jezičnoj uporabi mogu ostvariti u većoj (\approx standardni idiom) ili pak manjoj (\approx razgovorni idiom) mjeri, kao izvor jezične građe za ovo je istraživanje bilo potrebno odabrati i prikladne korpuse (v. sljedeći razdjel).

Odluka o razdvajanju standardnoga i razgovornoga idioma proizlazi i iz gore spomenute činjenice da ne postoji konsenzus o tome je li „razgovorni stil“ dio „standardnoga jezika“ ili ne. Predostrožnosti radi metodološki je i ispravnije i pouzdanije podatke koji bi mogli pripadati standardnomu idiomu i podatke koji bi mogli biti dio razgovornoga idioma isprva promatrati odvojeno – jedino se na taj način može utvrditi omjer sličnosti i razlika među njima, v. Sliku 1.



Slika 1: Modeli s vrlo različitim i vrlo sličnim standardnim i razgovornim idiomima

Neovisno o tome pada li odnos između standardnih i razgovornih idioma u ovome radu istraživanih jezika bliže jednomu ili drugomu polu prikazanih krajnjih vrijednosti (Slika 1) ili pak negdje u sredinu kontinuuma, vjerujemo da je kontrastivan empirijski varijacionistički pristup ne samo moguć nego i nužan. Iako razgovorni idiom ne podrazumijeva isključivo obilježja suprotna kodificiranoj normi i iako držimo da bi, kao i dio spomenutih autora, kodificirana norma trebala biti u većoj korelaciji sa stvarnom jezičnom uporabom izvornih govornika, vjerujemo da bi zadatak varijacionističkih istraživanja, između ostaloga, trebao biti kvalitativno i kvantitativno utvrditi i opisati i sličnosti i razlike između standardnoga idioma (onakvoga kakav je (pr)opisan u normativnim priručnicima) i razgovornoga idioma (onakvoga kakav je često demoniziran u jezičnim savjetnicima i inoj literaturi) ne samo na svim jezičnim razinama, nego i s obzirom na pragmatiku uzimajući u obzir ne samo jezične nego i izvanjezične čimbenike varijacije. Za razliku od postojećih normativnih priručnika koji se bave inventarizacijom pravila svojstvenih standar-

du i odstupanja od njega, varijacionistička bi istraživanja trebala dati podatke i o čestoći ostvarivanja tih pravila i o čestoći odstupanja od njih, kao i o čimbenicima koji stoje iza spomenutih procesa. Drugim riječima, jasni empirijski uvidi u to koliko su standardni i razgovorni idiom bliski, odnosno različiti ono je čemu se, među ostalim, u ovome radu teži. Rezultati pak takvih istraživanja zasigurno mogu biti korisni (socio)lingvistima koji se bave (re)standardizacijom i pitanjem kakvo jezično planiranje trebamo.

Sljedeći podrazdjeli detaljno prikazuju izvore jezične građe, uzorkovanje i prikupljanje podataka te njihovu anotaciju i čišćenje prije kvantitativne i kvalitativne analize.

4.2. Izvori jezične građe: korpusi bosanskoga, hrvatskoga i srpskoga jezika

S obzirom na to da se istraživačkim pitanjima ne nastoji utvrditi samo razlika među trima standardima nego i razgovornim idiomima, potrebna je jezična građa koja dobrim dijelom odražava jezičnu uporabu u ciljanim jezicima pod izravnim utjecajem propisanih nomativnih pravila kao i jezična građa koja dobrim dijelom izlazi iz okvira spomenutih pravila. Kao što je već rečeno, zbog razloga spomenutih u prethodnome razdjelu operacionalizacija varijable idiom kako god se provela nikada neće biti sasvim zadovoljavajuća, a isto vrijedi i za izbor vrela, odnosno izvore jezične građe. Jezična je uporaba vrlo heterogena, a varijacija ovisi o mnogim čimbenicima od kojih dio čini sam kontekst i svrha uporabe (stilovi, žanrovi, teme...). Neovisno o samoj veličini i nastojanju da budu balansirani, jezični su korpusi u konačnici uvijek nepotpun odrazi idioma koje bi trebali odražavati (usp. Kilgariff & Grefenstette 2003: 9; Stefanowitsch 2020: 6–7).

S obzirom na to da ne postoje javno dostupni i elektronički pretraživi usmeni korpusi istraživanih idioma koji bi ujedno zadovoljavali i kriterij usporedivosti, u ovome se istraživanju građa crpila samo iz korpusa pisanih idioma. Kao izvor standardne jezične građe poslužili su Oslo korpus bosanskih tekstova (Santos 1998, nadalje Oslo korpus), Hrvatska jezična riznica (Brozović Rončević i dr. 2018, nadalje Riznica), Korpus savremenog srpskog jezika (Utvić 2011, nadalje KSSJ) i InterCorp (Čermák i Rosen 2012). Riznica je odabrana kao izvor hrvatskoga standardnoga idioma jer kao i KSSJ sadrži ne samo recentnije, nego i tekstove iz cijeloga 20. stoljeća i jer je automatski provedena morfosintaktička anotacija točnije provedena nego u Hrvatskome nacionalnome korpusu (Tadić 1998, 2002, 2009; Agić i dr. 2008: 449–450; Ljubešić i dr. 2016: 4269).

S obzirom na to da anotacija vrste riječi u KSSJ-u (Utvić 2011) nije dovoljno kvalitetno provedena,²² kao dodatni izvor za standardni idiom srpskoga jezika poslužio je InterCorp (Čermák i Rosen 2012) koji je morfosintaktički anotiran. U istraživanju je korištena 12. verzija češkoga paralelnoga korpusa InterCorp koja sadrži dijelove 13 izvornih srpskih tekstova desetorice autora (Adamovičová i Vavřín 2019). Riječ je o suvremenim književnim djelima napisanim nakon druge polovice prošloga stoljeća koji broje 40 568 726 riječi.^{23, 24} Nadalje, valja istaknuti kako je riječ o tekstovima koji čine središnji dio InterCorpa iz čega proizlazi da je automatska anotacija ručno provjeravana. Korpusu se pristupa putem sučelja *KonText* (Machálek 2014, 2020), a MSD anotacija identična je kao i u *NoSketchEngineu* (Rychlý 2007; Kilgariff 2014) što omogućuje primjenu istoga CQL upita kao u drugim MSD anotiranim korpusima korištenima u ovome istraživanju. Osim toga valja napomenuti kako je InterCorp poslužio kao komplementarni izvor podataka i zbog posrednoga cilja ovoga istraživanja. Naime, željelo se utvrditi postoje li bitne razlike u podacima iz KSSJ-a (Utvić 2011) i InterCorpa (Čermák i Rosen 2012), odnosno provjeriti bi li InterCorp (Čermák i Rosen 2012) u budućim istraživanjima mogao poslužiti kao primarni izvor jezične građe koja odražava suvremeni standardni idiom srpskoga jezika.²⁵

Mrežni bsWaC, hrWaC i srWaC korpusi (Ljubešić & Klubička 2014, nadalje {bs,hr,sr}WaC) nisu samo najveći nego su i jedini veliki elektronički dostupni korpusi bosanskoga, hrvatskoga i srpskoga koji sadrže tekstove ne samo sa značajkama norminarnoga nego i inih idioma,²⁶ i to u prvome redu u sadržajima kao što su komentari, blogovi ili forumi koje stvaraju sami korisnici interneta.²⁷ Prema procjeni koju Klubička i Ljubešić (2014: 64–65) iznose na temelju provedenoga eksperimen-

²² Nijedan oblik riječi *je* nije anotiran kao zamjenica, svi su anotirani kao glagol. Slični se problemi vrlo vjerojatno javljaju i u svezi s drugim homonimnim, odnosno homografnim oblicima, no u okviru ovoga istraživanja nisu bili provjeravani.

²³ Češki paralelni korpus InterCorp neprestano se nadograđuje. Međutim, najnovija, 14. verzija sadrži isti broj izvornih srpskih tekstova.

²⁴ Ne uzimaju li se interpunkcijski znakovi u obzir, broj je riječi u srpskoj dijelu InterCorpa 32 444 479.

²⁵ Jurkiewicz-Rohrbacher i dr. (2017b) također rabe InterCorp kao izvor podataka za srpski standardni idiom.

²⁶ Za razliku od standardnoga i razgovornoga idioma u spomenutim su korpusima dijalektalni idiomu vrlo vjerojatno izrazito slabo zastupljeni. Ta se pretpostavka temelji na opažanju za hrvatske dijalekte iznesenome u Polančec i Mihaljević (2016: 444).

²⁷ Valja napomenuti da govornici hrvatskoga jezika mogu ostavljati komentare na svojem materinskom jeziku ne samo na stranicama s .hr, nego i na stranicama s .bs i .rs domenom. U svezi s tim Ljubešić i Klubička (2014: 32–33) tvrde da je većina dokumenata u spomenutim mrežnim korpusima točno automatski klasificirana, no upozoravaju da se u pojedinome korpusu i dalje mogu pojaviti ulomci na drugome blisko srodnome zapadnojužnoslavenskome jeziku. Tome je problemu u ovome istraživanju, isto kao i jezičnoj građi nastaloj strojnim prevodenjem, doskočeno prilikom ručne provjere empirijskih podataka, odnosno rečenica u uzorcima.

ta, udio standardnoga idioma u hrWaC-u ne iznosi više od 57%.²⁸ Spomenuta količina i raznolikost podataka te dostupnost metainformacija koje omogućuju praćenje podrijetla, odnosno pravoga izvora teksta, velika su prednost {bs,hr,sr}WaC korpusa (Jurkiewicz-Rohrbacher i dr. 2017a: 53–54) zbog čega su korišteni kao primarni izvor jezične građe razgovornih idioma. Budući da je jedan od ciljeva ovoga istraživanja dubinska analiza razlika između standardnih i razgovornih idioma kada je u pitanju izbor između dviju varijanata nenaglašenoga akuzativnoga oblika zamjenice trećega lica jednine ženskoga roda *ju* i *je*, metodološki bi zapravo bilo ispravnije ne koristiti cijele {bs,hr,sr}WaC korpuse (Ljubešić & Klubička 2014), nego samo njihove dijelove s tekstovima za koje se sa sigurnošću može reći da ih lektori nisu prilagođavali i usklađivali sa službenom normom kao što je npr. Forum potkorpus hrWaC korpusa (v. Bošnjak Botica i dr. 2022; Kolaković i dr. 2019). Nažalost, bsWaC i srWaC trenutačno nemaju slične potkorpuse, a ručna identifikacija internetskih stranica na kojima se nalaze slični sadržaji, dakle isključivo komentari, blogovi ili forumi, bila bi iznimno mukotrpana i dugotrajna. Upravo zbog toga, unatoč činjenici da spomenuti korpusi sadrže i tekstove pisane standardnim idiomima, radi ostvarivanja što bolje kontrastivnosti samoga istraživanja kao usporedivi izvori razgovornih idioma poslužili su {bs,hr,sr}WaC korpusi u cjelini, a ne samo neki njihovi dijelovi (usp. Kolaković i dr. 2022: 70, 86). Valjda dodati i to da će svaki (pot)korpus kako god pomno odabran bio, uvijek biti heterogen s obzirom na standardnost. Čak i u okviru foruma, blogova i korisničkih komentara dok se neki ljudi radije služe razgovornim idiomom, drugi se zasigurno odlučuju držati se norme. Isto tako i referentni su korpusi uvijek dijelom heterogeni s obzirom na standardnost tekstova koje sadrže – tome posebice pridonose književni tekstovi i razne novinarske kolumne koji često sadrže elemente inih idioma (Kolaković i dr. 2019: 512).

Tablica 3 donosi najvažnije podatke o svim korištenim korpusima u usporednome prikazu.

²⁸ Treba dodati da su procjenitelji standardnosti bili studenti studija Informatologije na Filozofskome fakultetu u Zagrebu, a ne autori referentnih normativnih priručnika. Da su potonji sudjelovali u eksperimentu, vrlo bi vjerojatno u hrWaC-u udio tekstova pisanih na standardnome idiomu bio procijenjen kao niži.

Tablica 3: Osnovna obilježja korištenih korpusa

| Korpus | veličina | tekstovi | | lematizacija i anotacija | | |
|--------------|---------------|---|---|--------------------------|--------------|-----|
| | br. riječi | starost | vrste | lema | vrsta riječi | MSD |
| Oslo | 1 500 000 | 1989. – 1997. | fikcija, dječja književnost, eseji, novine, vjerski i pravni tekstovi, folklorni tekstovi | da | ne | ne |
| Riznica | 85 273 724 | ² / ₂ 19. – 21. st. | fikcija, znanstvene publikacije, (mrežni) časopisi i novine | da | da | da |
| KSSJ | 122 000 000 | 20. – 21. st. | književni, znanstveni, publicistički, administrativni i ini tekstovi | da | da | ne |
| InterCorp Sr | 40 568 726 | 20. – 21. st. | fikcija (izvorna i prijevodna), podnatpisi | da | da | da |
| bsWaC | 248 478 730 | – | .ba | da | da | da |
| hrWaC | 1 210 021 198 | – | .hr | da | da | da |
| srWaC | 476 888 297 | – | .sr | da | da | da |

Iako građa u odabranim izvorima nije savršen odraz razgovornih i standardnih idioma analiziranih jezika nego prije njihova aproksimacija, ta se činjenica ne bi smjela koristiti kao glavni argument protiv uporabe korpusa jer svako je prikupljanje podataka ne samo u lingvistici nego i u drugim znanostima uvijek nepotpuno. Znanstveni rad počiva na izgradnji modela od širokodostupnih podataka te njihovu postupnoj nadogradnji tijekom vremena, tako kako veći broj podataka biva sve dostupnijim (Stefanowitsch 2020: 6–7).

4.3. Uzorci jezične građe

4.3.1. Određivanje veličine uzoraka

Sljedeći korak u istraživanju bio je utvrđivanje veličine cijele populacije zanaglasnih varijanata akuzativnoga oblika zamjenice trećega lica jednine ženskoga roda u odabranim korpusima. Kako bi se dobili podatci o veličini cijele populacije akuzativnih zanaglasnih varijanata *ju* i *je*, u morfosintaktički anotiranim korpusima (Riznica, InterCorp, {bs,hr,sr}WaC) primijenjeni su sljedeći CQL upiti: [tag="Pp.*a" & word="ju"] i [tag="Pp.*a" & word="je"], a u ostalim korpusima (Oslo, KSSJ) spomenute su zanaglasnice pretraživane kao točno ti oblici riječi, odnosno kao "ju" i kao "je". U potonjemu slučaju poseban je problem točno utvrđivanje veličine populacije akuzativnih zanaglasnih varijanata *ju* i *je* tim više što se pretragom bez morfosintaktičke anotacije ne dobivaju samo primjeri s akuzativnim nego i genitivnim oblikom zamjenice trećega lica jednine ženskoga roda te primjeri s nesvršenim nenaglašenim oblikom prezenta glagola *biti* u trećemu licu jednine. U tome slučaju veličina populacije procjenjivana je posredno, o čemu se govori nešto više niže u tekstu. Kao

rezultat pretrage dobiveni su podatci o veličini populacije, odnosno o apsolutnoj čestoti prikazani u Tablici 4.²⁹

Tablica 4: Čestota akuzativnih znanaglasnih varijanata *ju* i *je* u odabranim korpusima

| Korpus | apsolutna čestota | | odnos uzorka i cijele populacije | |
|--------------|-------------------|--------------------|----------------------------------|---------------|
| | <i>ju</i> | <i>je</i> | <i>ju</i> | <i>je</i> |
| Oslo | 249 | 52 919 (10 000) | 249 = 100 % | 300 = 3 % |
| Riznica | 15 526 | 13 808 | 300 = 1,93 % | 300 = 2,17 % |
| KSSJ | 14 326 (10 200) | 4 109 019 (59 169) | 300 = 2,94 % | 300 = 0,51 % |
| InterCorp Sr | 67 | 485 | 67 = 100 % | 300 = 61,86 % |
| bsWaC | 27 433 | 33 305 | 300 = 1,09 % | 300 = 0,90 % |
| hrWaC | 323 981 | 147 951 | 300 = 0,09 % | 300 = 0,20 % |
| srWaC | 28 658 | 61 593 | 300 = 1,05 % | 300 = 0,49 % |

Usporedni podatci iz Tablice 4, posebice kada su u pitanju morfosintaktički anotirani korpusi (Riznica, InterCorp, {bs,hr,sr}WaC), već na prvi pogled pružaju iznimno zanimljivu sliku. Tako se, primjerice, uočava da su u Riznici, korpusu koji predstavlja standardni idiom hrvatskoga jezika u pisanome mediju, varijante nenaglašena akuzativnoga oblika zamjenice trećega lica jednine ženskoga roda *ju* i *je* gotovo podjednako česte, uz neznatno veću prevagu znanaglasnice *ju*. Štoviše, u hrWaC-u, koji je u ovome istraživanju poslužio kao izvor razgovornoga idioma hrvatskoga jezika, nenaglašeni akuzativni oblik zamjenice trećega lica jednine ženskoga roda *ju* više je nego dvostruko češći od akuzativne znanaglasnice *je*. U svim ostalim istraženim idiomima, odnosno korpusima, vrijedi obrnuto: nenaglašeni akuzativni oblik zamjenice trećega lica jednine ženskoga roda *je* češći je od oblika *ju* koji mu se jedino u bsWaC-u, korpusu koji je u ovome istraživanju poslužio kao izvor razgovornoga idioma bosanskoga jezika, donekle uspijeva svojom čestotom „približiti“.³⁰

U svezi s navedenim podacima potrebno je istaknuti da ni u jednome od pretraženih korpusa pretraga zasigurno nije rezultirala prikazivanjem cijele populacije u njezinoj potpunosti što je posljedica pogrešaka u automatskoj obradi podataka: segmentaciji teksta, lematizaciji i na kraju morfosintaktičkoj anotaciji.³¹ Valja imati

²⁹ Pretpostavljena (približna) veličina cijele populacije akuzativne zamjeničke znanaglasnice *je* u Oslo i KSSJ korpusu navodi se u zgradama.

³⁰ Potonje bi moglo ukazivati na prijelazni položaj razgovornoga idioma bosanskoga jezika, no isto tako moglo bi biti i posljedica relativno velikoga broja dokumenata s hrvatskim jezičnim materijalom koji su pogrešno automatski klasificirani kao bosanski i koji su kao takvi ostali u bsWaC-u, v. prethodni razdjel.

³¹ Svi korpusi zasigurno sadrže pogreške u segmentaciji i lematizaciji teksta. S obzirom na to da se morfosintaktička anotacija u korpusima provodi automatski, i morfosintaktički anotirani korpusi ne

na umu da je riječ, dakle, o relativno točnome prikazu apsolutne čestote. Osim toga, dok veličina nekih korpusa, primjerice Oslo i InterCorp korpusa omogućuje ručnu provjeru svih dobivenih primjera sa zamjeničkom zanaglasnicom *ju*, ručna bi provjera cijele populacije akuzativne zamjeničke zanaglasnice *je* bila (donekle) izvediva jedino u slučaju InterCorp korpusa. U svim ostalim slučajevima ručno pregledavanje svih primjera prekoračivalo bi ljudske i ine resurse koji su stajali na raspolaganju za ovo istraživanje. Upravo zbog toga odlučeno je prikupiti uzorke rečenica s objema varijantama iz svih odabranih korpusa. S obzirom na to da su korpusi korišteni u istraživanju različite veličine (v. Tablicu 3), veličine populacija zamjeničkih zanaglasnica *ju* i *je*, odnosno njihove čestote također su vrlo različite. Kako bi obje zamjeničke zanaglasničke varijante bile jednako zastupljene u svim ciljanim idiomima, odlučeno je da u konačnu statističku analizu uđu uzorci od 300 rečenica po zanaglasničkoj varijanti i po idiomu, što bi u konačnici rezultiralo građom od 3600 ručno anotiranih rečenica. A s obzirom na to da su kao građa za srpski jezik poslužila tri, a ne dva korpusa (v. Tablicu 3), broj je analiziranih rečenica zapravo bio nešto veći. Riječ je o dodatnih 367 rečenica iz InterCorpa koje su poslužile kao komplementarni podatci za standardni idiom srpskoga jezika.

I na kraju treba nešto dodati i o „proizvoljnosti“ veličine uzorka u ovome istraživanju. Naime, prema standardima u društvenim znanostima, smatra se kako je u slučaju malih populacija (do 1000) potreban uzorak od 30 %, za umjereno velike populacije (do 10 000) za jednaku točnost nužan je uzorak od oko 10 %, za velike populacije (više od 150 000) uzorci od 1 % smatraju se vrlo točnima, dok se u slučaju golemih populacija (više od 10 000 000) velika preciznost može postići i s iznimno malim uzorcima od samo 0,025 % (usp. Neuman 2007: 162; Buchstaller & Khattab 2013: 82). Kao što se iz Tablice 4 može iščitati, veličina uzoraka u ovome istraživanju svojevrsni je kompromis između metodološki preporučljivoga i tehnički izvedivoga.

jamče apsolutnu točnost anotacije traženih oblika. Drugim riječima, može se očekivati da se među konkordancama koje se dobivaju za akuzativnu zamjeničku zanaglasnicu *je* kriju i primjeri s genitivnim oblikom iste zamjenice, odnosno s nesvršenim nenaglašenim oblikom glagola *biti* u trećemu licu jednine u prezentu. Ručnom provjerom uzoraka doskočeno je problemu tzv. lažnih pozitivnih vrijednosti (engl. *false positive*, v. Stefanowitsch 2020: 112), na taj su način iz uzoraka uklonjeni svi primjeri s genitivnim oblikom zamjenice trećega lica jednine ženskoga roda i primjeri s nesvršenim enklitičkim oblikom glagola *biti* u trećemu licu jednine u prezentu. Međutim, valja istaknuti kako je isto tako moguće da je akuzativni oblik zanaglasnice *je* u nekim slučajevima bio krivo anotiran kao genitiv iste zamjenice ili kao nesvršeni enklitički oblik glagola *biti* u trećemu licu jednine u prezentu. Taj je problem lažnih negativnih vrijednosti (engl. *false negative*, v. Stefanowitsch 2020: 112) gotovo nerješiv. Alternativno, primjere bi se moglo tražiti na temelju točno određenih oblika riječi kao i u korpusima koji nisu morfosintaktički anotirani, no razlika u točnosti koju bi se moglo dobiti bila bi neznatna s obzirom na cijenu koju bi se za to moralo platiti, odnosno s obzirom na vrijeme koje bi se u tako nešto moralo uložiti. U svezi s tim valja dodati kako je procijenjena točnost morfosintaktičke anotacije za {bs,hr,sr}WaC i Riznicu iznosi 92,5 % (Ljubešić i dr. 2016: 4269; osobna komunikacija N. Ljubešić).

4.3.2. Prikupljanje jezičnih uzoraka za analizu

Od konkordanci dobivenih s pomoću CQL upita navedenoga u prethodnome razdjelu u korpusima pretraživima preko *NoSketchEngine* sučelja ($\{bs,hr,sr\}$ WaC, Riznica) u prvome koraku filtrirani su samo prvi primjeri iz pojedinoga dokumenta (funkcija *Ist hit in doc*) kako bi se u građi izbjeglo gomilanje primjera iz istoga dokumenta, odnosno istoga autora. U sljedećemu su koraku konkordance promiješane (funkcija *Shuffle*) čime se osiguralo da se rečenice prikazuju nasumičnim redoslijedom te su potom oblikovani nasumični uzorci (funkcija *Sample*) s po 400 konkordanci (usp. Kolaković 2018: 210; Kolaković i dr. 2022: 295). Uzorci su pohranjeni na lokalno računalo (funkcija *Save*) u obliku .csv datoteka.

Iako InterCorp sadrži i originalne tekstove i prijevode, primjeri sa zamjeničkim zanaglasnicama *ju* i *je* preuzimani su samo iz originalnih tekstova na srpskoj jeziku, prijevodni tekstovi nisu bili uključeni jer nije bilo sasvim jasno jesu li prevoditelji hrvatski ili srpski izvorni govornici, odnosno jesu li se eventualno srpski prevoditelji pri prevođenju pomagali već gotovim hrvatskim prijevodima. Nadalje, podnaslovi također nisu bili uključeni u građu. Za razliku od građe iz $\{bs,hr,sr\}$ WaC-a i Riznice, u slučaju InterCorpa nije primijenjeno filtriranje samo prvih primjera iz pojedinih dokumenata jer bi u tome slučaju broj primjera s *ju* zanaglasnicom bio samo devet, dok bi za akuzativnu zamjeničku zanaglasnicu *je* na taj način bilo prikupljeno tek 13 primjera. Za akuzativnu zamjeničku zanaglasnicu *ju* prikupljeno je ukupno 67 podataka, dok je za njezinu varijantu *je* pronađeno 485 primjera, od čega je nakon miješanja s pomoću funkcije *Shuffle* oblikovan slučajni uzorak s 400 primjera koji su, baš kao i u slučaju podataka iz $\{bs,hr,sr\}$ WaC-a i Riznice, radi daljnje ručne anotacije pohranjeni na lokalno računalo u obliku .csv datoteka.

KSSJ kao korpus pokazao se višestruko problematičnim u procesu prikupljanja građe jer ne omogućuje automatsko pohranjivanje svih dobivenih primjera čime bi se kasnije osigurali uvjeti za nasumičan izbor primjera u Excel tablici. U KSSJ-u postoji funkcija kojom se treba prikazivati svaki n -ti primjer, no to zapravo nije isto što i nasumično prikazivanje primjera. Osim toga, kada se u polje predviđeno za tu funkciju pokuša upisati ili broj nula, čime bi se trebao osigrati prikaz svih primjera, ili bilo koji broj iznad 2000, dobiva se povratna poruka da tako nešto nije moguće. Budući da bi ručno kopiranje svih stranica s dobivenim primjerima u Excel dokument iziskivalo previše vremena, za prikupljanje podataka s akuzativnom zanaglasnicom *je* izabrana je sljedeća kompromisna strategija: pohranjivani su samo primjeri sa svake druge stranice. Pri tome se pazilo da broj primjera iz jednoga djela, odnosno jednoga autora ne bude veći od pet tako da se dobivene stranice s primjerima listalo sve dok se ne bi došlo do prikaza primjera iz djela sljedećega autora. S obzirom na to da, kao što je već rečeno u prethodnome razdjelu, u KSSJ-u anotacija vrste riječi u slučaju oblika riječi *je* nije provedena točno, prije pohranjivanja primjera, provjeralo se odnosi li se spomenuti oblik riječi na zamjeničku zanaglasnicu trećega lica

jednine ženskoga roda u akuzativu ili genitivu, ili je pak riječ o nenaglašenome prezentu trećega lica jednine *je*. U Excel tablicu na lokalnome računalo pohranjivani su samo primjeri koji sadrže oblik riječi *je* koji se odnosi na zamjeničku zanaglasnicu trećega lica jednine ženskoga roda u akuzativu.

S obzirom na to da oblik riječi *ju* nema homonima, odnosno homografa, on je u KSSJ-u puno rjeđi pa je za akuzativnu zamjeničku zanaglasnicu *ju* izabrana nešto drukčija strategija. Budući da se od 9100 primjera javljaju gotovo isključivo primjeri u kojima je oblik riječi *ju* zapravo dio neke riječi pogrešno rastavljene na slogove, u Excel tablicu ručno su kopirani primjeri sa svake desete stranice³² sve do stranice na kojima se nalaze primjeri od 9001 do 9100. Nakon toga brzim su se automatskim pretraživanjem s pomoću funkcije nađi i zamijeni (*find and replace*) te naprednim označavanjem bojom iz uzorka pokušali ukloniti svi neodgovarajući podaci odnosno primjeri s oblikom riječi *ju* koji se nisu odnosili na akuzativnu zamjeničku zanaglasnicu nego na *ju es open*, *ju tjub*, *aj lav ju*, *ju* ? (potonje u pravilu sadrži pogrešno rastavljene riječi) i slične izraze. Nakon takvoga automatskoga čišćenja ostalo je 778 primjera koji su potom s pomoću Excelove funkcije *=rand()* nasumično poredani te je za ručno označavanje pripremljeno prvih 400 primjera.

Iako ni Oslo korpus nije morfosintaktički anotiran, podaci su u njemu bili lakše prikupljeni nego u KSSJ-u. Svi su primjeri s oblicima riječi *ju* (256)³³ i *je* (52.919) pohranjeni na računalo, a potom su u Excel tablici nasumično poredani i pripremljeni za daljnje ručno označavanje. Ručna anotacija primjera s oblikom riječi *je* obuhvaćala je odabir samo onih primjera koji se odnose na akuzativnu zamjeničku zanaglasnicu *je* i isključivanje svih primjera koji sadrže genitivnu zamjeničku zanaglasnicu *je* i nesvršeni nenaglašeni oblik prezenta trećega lica jednine glagola *biti*.

4.3.3. Anotacija i čišćenje podataka

Na računalo u Excel dokumentu svi su primjeri iz uzoraka opisanih u prethodnome razdjelu dalje ručno provjeravani i označavani s obzirom na varijable naznačene u razdjelu 4.1. Kako bi se postigla što veća točnost, svaki su primjer provjerile i označile dvije anotatorice (usp. Stefanowitsch 2020: 128–129).³⁴

Tijekom ručne anotacije dio je podataka uklonjen jer su u primjerima bili upotrijebljeni oblici riječi *ju* i *je* koji se nisu odnosili na akuzativnu zamjeničku zanaglasnicu trećega lica jednine ženskoga roda. U takvim su se primjerima nalazili nesvršeni

³² U slučaju oblika riječi *ju* primjena funkcije svaki *n*-ti primjer u konkordance uvijek je unosila previše neodgovarajućih primjera kao što su primjerice *ju es open*, *ju tjub*, *aj lav ju*, *ju* ? itd.

³³ U Tablici 4 navodi se da se u Oslo korpusu nalazi 249 primjera s akuzativnom zanaglasnicom *ju*. Prilikom ručne anotacije sedam je neodgovarajućih primjera bilo uklonjeno iz građe.

³⁴ „We can (and must) try to minimize errors in our data and our classification, but we can never get rid of them completely (this is true not only in corpus-linguistics but in any discipline).“ (Stefanowitsch 2020: 74)

nenaglašeni prezent trećega lica jednine glagola *biti* (24), genitivna zanaglasnica trećega lica jednine ženskoga roda (25), riječi rastavljene na slogove (26), fonetski zapisane engleske riječi (27), strojno prevođenje ili pak pogreške koje su nastale pri brzome oblikovanju teksta (28).

- (24) *Vrsta saobraćaja koji želite je onaj koji čine ljudi [...].* [srWaC]
(25) *Bila je to moja duševna hrana i nisam je se namjeravala odreći.* [hrWaC]
(26) *Uneo je pu no no vi na u is tra ži va nje ki ne ske ci vi li za ci je, in si sti ra ju ci na ve zi i te snom prožimanju po li ti ke.* [srWaC]
(27) *Ju mejk maj hart sing.* [InterCorp]
(28) *[...] da naftni poslovi u Iraku, zemlji koju ju je nazvao naftaškim rajem [...].* [Riznica]

Osim toga, primjeri su isključivani iz građe kada je bilo očito da se odnose na idiome koji nisu trebali biti u tome korpusu. Tako je primjer (29) bio isključen iz građe jer ekavski refleks jata upućuje na to da primjer ne pripada razgovornom idiomu bosanskoga jezika. Na sličan način jezična obilježja u (30) upućuju na to da primjer ne pripada razgovornom idiomu srpskoga jezika. Primjer (31) također je isključen iz građe jer je jasno da ne pripada razgovornom idiomu bosanskoga jezika.

- (29) *Znam da sam je negde videla, ali se nisam mogla setiti.* [bsWaC]
(30) *Ohladenu kremu izvadite iz hladnjaka pa ju izvažite. (Slatko vrhnje istucite u šlag da dobijete jednu trećinu težine kreme.)* [srWaC]
(31) *Upoznal sam neku plavušu, pa sam ju pozval doma na kavu.* [bsWaC]

Neki od primjera isključeni su jer redosljed zanaglasnica u njima odstupa od kodificirane norme, v. primjer u (32). Iako takva i slična odstupanja nisu neuobičajena u idiomima na bosanskoj, hrvatskoj i srpskoj jezičnoj području (v. Kolaković i dr. 2022: 147–151, 193–194), u ovome su istraživanju takvi podatci isključivani iz građe jer su bili malobrojni i jer bi zahtijevali uvođenje posebne kategorije, odnosno varijable u anotacijsku shemu i jer njihova malobrojnost ne bi dopuštala valjanu statističku analizu.

- (32) *[...] već je ju povuklo iz procedure dok se ne konstituirala [...].* [bsWaC]
(33) *[...] kršio sam sva pravila — ljubio sam je.* [InterCorp]
(34) *[...] đavo je odneo u bestragiju.* [KSSJ]
(35) *Nadi-beg ju ostavio samu.* [Oslo]

Iz građe su isključivani i primjeri poput (33) u kojima zamjeničke zanaglasnice *ju* i *je* stoje ispred interpunkcijskih znakova jer istraživački nacrt opisan u razdjelu 4.1

zahtijeva da se spomenute zanaglasnice nalaze između riječi. I na kraju, iz građe su isključivani i primjeri poput (34) i (35) u kojima je ispuštena ili zanaglasnica trećega lica jednine ženskoga roda u akuzativu (v. Antonić 1998) ili oblik nesvršenoga nena-
 glašenoga prezenta trećega lica jednine glagola *biti*.

5. Rezultati

Kako bi se odgovorilo na istraživačka pitanja o mogućim razlikama u distribuciji dviju varijanata akuzativnih zamjeničkih zanaglasnica *ju* i *je* u bosanskim, hrvatskim i srpskim idiomima, ukupno je prikupljeno i ručno označeno 3916 primjera iz elektronički dostupnih korpusa, a nakon dodatnoga čišćenja podataka preostalo je 1749 primjera s akuzativnom zamjeničkom zanaglasnicom *ju* i 1800 s njezinom varijantom *je*.³⁵ Distribucija spomenutih zamjeničkih zanaglasnica statistički je analizirana s pomoću programa *R* (R Core Team 2016). Tablica 5 daje detaljan uvid u distribuciju zamjeničkih zanaglasnica *ju* i *je* s obzirom na nezavisne varijable (riječi koje prethode i slijede zanaglasnicu te idiome).

Tablica 5: Distribucija akuzativnih zanaglasnica *ju* i *je* u analiziranoj građi

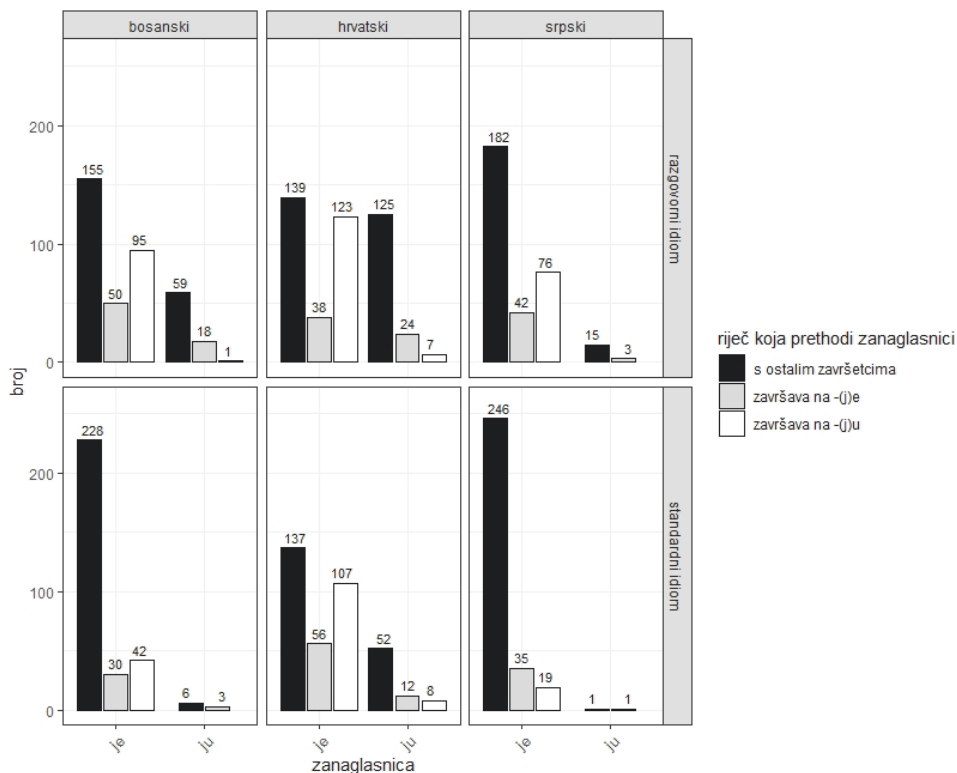
| <i>ju</i> | | riječ koja prethodi zanaglasnici | | | riječ koja slijedi zanaglasnicu | |
|--------------------|------------|----------------------------------|-------|--------|---------------------------------|--------|
| jezik | idiom | -(j)u | -(j)e | ostalo | glagolski oblik <i>je</i> | ostalo |
| bosanski | standardni | 2 | 7 | 240 | 240 | 9 |
| | razgovorni | 6 | 34 | 260 | 222 | 78 |
| hrvatski | standardni | 12 | 29 | 259 | 228 | 72 |
| | razgovorni | 7 | 35 | 258 | 144 | 156 |
| srpski | standardni | 4 | 6 | 290 | 298 | 2 |
| | razgovorni | 1 | 20 | 279 | 282 | 18 |
| Σ <i>ju</i> | | 32 | 131 | 1586 | 1414 | 335 |
| <i>je</i> | | | | | | |
| bosanski | standardni | 42 | 30 | 228 | 0 | 300 |
| | razgovorni | 95 | 50 | 155 | 0 | 300 |
| hrvatski | standardni | 107 | 56 | 137 | 0 | 300 |
| | razgovorni | 123 | 38 | 139 | 0 | 300 |
| srpski | standardni | 19 | 35 | 246 | 0 | 300 |
| | razgovorni | 76 | 42 | 182 | 0 | 300 |
| Σ <i>je</i> | | 462 | 251 | 1087 | 0 | 1800 |

³⁵ U građi prikupljenoj iz korpusa InterCorp, koji sadrži srpsku standardnojezičnu građu, zanaglasnica *ju* javlja se isključivo u kombinaciji s glagolskom zanaglasnicom *je*, dakle u 100 % slučajeva. Kako bi se izbjeglo gomilanje podataka za standardni idiom srpskoga jezika, prikupljeni i anotirani podatci iz InterCorpa isključeni su iz daljnjega istraživanja.

Podatci u Tablici 5 daju odgovor na prvo istraživačko pitanje. Zamjenička se zanaglasnica *ju* u analiziranoj standardnojezičnoj građi u većini slučajeva rabi kada *ju* slijedi glagolska zanaglasnica *je*. U korpusu KSSJ, koji sadrži srpsku standardnojezičnu građu, to se odnosi na 99 % slučajeva. U korpusu s bosanskom standardnojezičnom građom to se odnosi na 96 %, dok se u korpusu s hrvatskom standardnojezičnom građom to odnosi na 76 % slučajeva uporabe akuzativne zanaglasnice *ju*. Ti su podatci u skladu sa stanjem opisanim u gramatikama bosanskoga, hrvatskoga i srpskoga jezika, v. razdjela 2.3.2–2.3.4.

Na temelju podataka u Tablici 5 dobiva se odgovor i na drugo istraživačko pitanje. Akuzativna se zanaglasnica *ju* u srWaC-u, korpusu s jezičnom građom koja predstavlja razgovorni idiom srpskoga jezika, samo u 6 % slučajeva ne rabi zajedno s glagolskom zanaglasnicom *je*. Za razliku od toga, uporaba spomenute zanaglasnice u bsWaC-u, korpusu s jezičnom građom koja predstavlja razgovorni idiom bosanskoga jezika, samo u 74 % slučajeva ovisi o neposrednoj blizini glagolske zanaglasnice *je*. Distribucija je, dakle, akuzativne zanaglasnice *ju* u bsWaC-u donekle slična njezinoj distribuciji u Riznici, dok je u hrWaC-u, korpusu koji predstavlja razgovorni idiom hrvatskoga jezika, njezina uporaba u 52 % slučajeva neovisna o direktnome kontaktu s glagolskom zanaglasnicom *je*. Razlike u uporabi akuzativne zanaglasnice *ju* u korpusima koji sadrže jezičnu građu s razgovornim idiomima bosanskoga, hrvatskoga i srpskoga jezika na neki način preslikavaju razlike koje su uočene i u standardnim idiomima bosanskoga, hrvatskoga i srpskoga s tim da je u razgovornim idiomima uporaba spomenute zanaglasnice općenito manje ograničena na izravni kontakt s glagolskom zanaglasnicom *je*.

Kako se čini, vjerojatnost pojavljivanja akuzativne zanaglasnice *ju* koju ne slijedi glagolska zanaglasnica *je* najmanja je u standardnim idiomima srpskoga i bosanskoga jezika, a najveća je u razgovornim idiomima hrvatskoga i bosanskoga jezika. Štoviše, razvidno je da u slučaju akuzativne zanaglasnice *ju* u građi prevladavaju primjeri u kojima *ju* slijedi glagolska zanaglasnica *je*. Radi daljnje statističke analize iz građe su uklonjeni svi takvi primjeri. Grafikon 1 daje uvid u distribuciju akuzativnih zanaglasnica *ju* i *je* s obzirom na završetak riječi koja im prethodi nakon izuzimanja svih primjera koji sadrže zanaglasničku skupinu *ju je*.



Grafikon 1: Distribucija akuzativnih zanaglasnica *ju* i *je* s obzirom na završetak prethodne riječi

Kako bi se provjerilo koje od nezavisnih varijabli statistički značajno utječu na distribuciju akuzativne zanaglasnice *ju* u analiziranoj građi, podatci su analizirani s pomoću linearnoga regresijskoga modela.³⁶ Njime je, v. Tablicu 6, kao što je već rečeno, nakon što su iz podataka uklonjeni primjeri u kojima se zanaglasnica *ju* javlja u kombinaciji s glagolskom zanaglasnicom *je*, provjeravano utječu li nezavisne varijable jezik, idiom te završetak riječi koja prethodi akuzativnoj zanaglasnici na vjerojatnost pojavljivanja zanaglasnice *ju*. Pojavljivanje zanaglasnice *ju* nakon riječi s ostalim završetcima u razgovornome idiomu bosanskoga jezika (v. Intercept), odnosno, u podatcima prikupljenima iz bsWaC-a, uspoređivano je s ostalim podatcima. Prema rezultatima, jezik kao nezavisna varijabla statistički je značajan čimbenik. Naime, pojavljivanje zanaglasnice *ju* u pozitivnoj je korelaciji s hrvatskim, dok je u negativnoj korelaciji sa srpskim.

³⁶ Uporabljena je glm funkcija u R-u: `glm(formula = enklitika ~ idiom * jezik * riječ koja prethodi zanaglasnici, family = binomial(link = "logit"), data = ohneje)`.

Tablica 6: Rezultati statističke analize čimbenika koji utječu na pojavljivanje zanaglasnice *ju*

```

Deviance Residuals:
Min                1Q            Median            3Q                Max
-1.1327           -0.6231          -0.3203          -0.0901           3.3195

Coefficients:
                Estimate      Std. Error  z value  Pr(>|z|)
Intercept (bosanski:razgovorni:s
ostalim završetcima)
standardni          -2.67170    0.44097    -6.059    1.37e-09    ***
hrvatski            0.85973    0.19646     4.376    1.21e-05    ***
srpski             -1.53007    0.30913    -4.950    7.44e-07    ***
završava na -(j)e    -0.05576    0.31457    -0.177    0.859298
završava na -(j)u  -3.58799    1.01682    -3.529    0.000418    ***
standardni:hrvatski  1.80912    0.48598     3.723    0.000197    ***
standardni:srpski    -0.33768    1.12724    -0.300    0.764513
standardni:završava na -(j)e  1.39076    0.79792     1.743    0.081336    .
standardni:završava na -(j)u  -10.34049    610.45286    -0.017    0.986485
hrvatski:završava na -(j)e    -0.29761    0.42677    -0.697    0.485584
srpski:završava na -(j)e    -0.08734    0.72682    -0.120    0.904353
hrvatski:završava na -(j)u    0.82788    1.09549     0.756    0.449824
srpski:završava na -(j)u    -11.48212    453.80623    -0.025    0.979814
standard:hrvatski: završava na
-(j)e          -1.60910    0.92064    -1.748    0.080497    .
standard:srpski:
završava na -(j)e    -13.30840    668.71808    -0.020    0.984122
standard:hrvatski: završava na
-(j)u    11.47596    610.45313     0.019    0.985001
standard:srpski: završava na
-(j)u    27.97150    760.65354     0.037    0.970666
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 1855.3 on 2134 degrees of freedom
Residual deviance: 1413.1 on 2117 degrees of freedom
AIC: 1449.1
Number of Fisher Scoring iterations: 16
    
```

Sve tri nezavisne varijable (vrsta završetka riječi ispred zanaglasnice, jezik i idiom) pomažu u objašnjavanju vjerojatnosti pojavljivanja akuzativne zanaglasnice *ju*. Općenito govoreći postoji statistički značajna razlika u uporabi zanaglasnice *ju* u

bosanskoj, srpskoj i hrvatskoj jeziku. Naime, najvjerojatnija je njezina uporaba u hrvatskoj jeziku: 2,36 puta veća je od vjerojatnosti njezina pojavljivanja u bosanskoj jeziku (pod istim uvjetima, odnosno u istim kontekstnim konfiguracijama). Drugim riječima, relativni je rizik 136 % veći nego u bosanskoj jeziku, dok je za srpski jezik relativni rizik 79 % manji nego u bosanskoj.³⁷

Osim toga, postoji općenita statistički značajna razlika u uporabi zanaglasnice *ju* u razgovornim i standardnim idiomima. Točnije govoreći, vjerojatnost pojavljivanja zanaglasnice *ju* veća je u razgovornim nego u standardnim idiomima. Međutim, valja istaknuti kako se uočava razlika između bosanskoga i srpskoga jezika s jedne strane i hrvatskoga s druge strane. U bosanskoj i srpskoj jeziku vjerojatnost pojavljivanja zanaglasnice *ju* u standardnome idiomu jednaka je vjerojatnosti da se zanaglasnica *ju* pojavi u razgovornome idiomu pomnožena s 0,069. Drugim riječima, relativni je rizik za pojavljivanje zanaglasnice *ju* u standardnome idiomu bosanskoga i srpskoga jezika 93 % manji nego u njihovim razgovornim idiomima. U razgovornome je idiomu hrvatskoga jezika pak taj relativni rizik 58 % niži nego u standardnome idiomu hrvatskoga jezika.

Kako statistička analiza pokazuje, vrsta završetka riječi koja prethodi zamjeničkoj zanaglasnici trećega lica ženskoga roda u akuzativu samo djelomično igra značajnu ulogu u pojavljivanju zanaglasničke varijante *ju* ili *je*. Međutim, kako usporedba različitih statističkih modela pokazuje, uključivanje prediktora, odnosno razine nezavisne varijable „riječ koja prethodi zanaglasnici i završava *-(j)u*“ opravdano je s obzirom na to da primijenjeni model objašnjava više varijance u podacima od modela bez varijable „riječ koja prethodi zanaglasnici“ (vrijednost AIC viša je u slučaju izostavljanja spomenute nezavisne varijable i niža u slučaju njezina uključivanja $AIC\ 1592.1 > AIC\ 1449.1$).³⁸

Ako riječ koja prethodi zamjeničkoj zanaglasnici trećega lica ženskoga roda u akuzativu završava na *-(j)u*, vjerojatnost se pojavljivanja zanaglasnice *ju* u svim analiziranim idiomima (standardnim i razgovornim) smanjuje: 0,03, odnosno relativni je rizik 97 %. Štoviše, u srpskoj razgovornome idiomu i u standardnome idiomu bosanskoga jezika u okviru analizirane građe nije uočen nijedan primjer u kojemu ispred zanaglasnice *ju* stoji riječ koja završava na *-(j)u*, v. razdjel 6.1. Za razliku od toga, u srpskoj standardnome i u razgovornome idiomu bosanskoga jezika u okviru je analizirane građe potvrđen samo jedan takav primjer.³⁹

³⁷ Engl. *relative risk*, kratica *RR*, „[o]mjer vjerojatnosti za određeni ishod između skupine izložene određenom čimbeniku i skupine koja nije bila izložena tomu čimbeniku“ (v. Brač i dr. 2015; <http://struna.ihj.hr/naziv/relativni-rizik/35644/#naziv>).

³⁸ enklitika ~ idiom * jezik nasuprot formuli enklitika ~ idiom * jezik * riječ koja prethodi zanaglasnici.

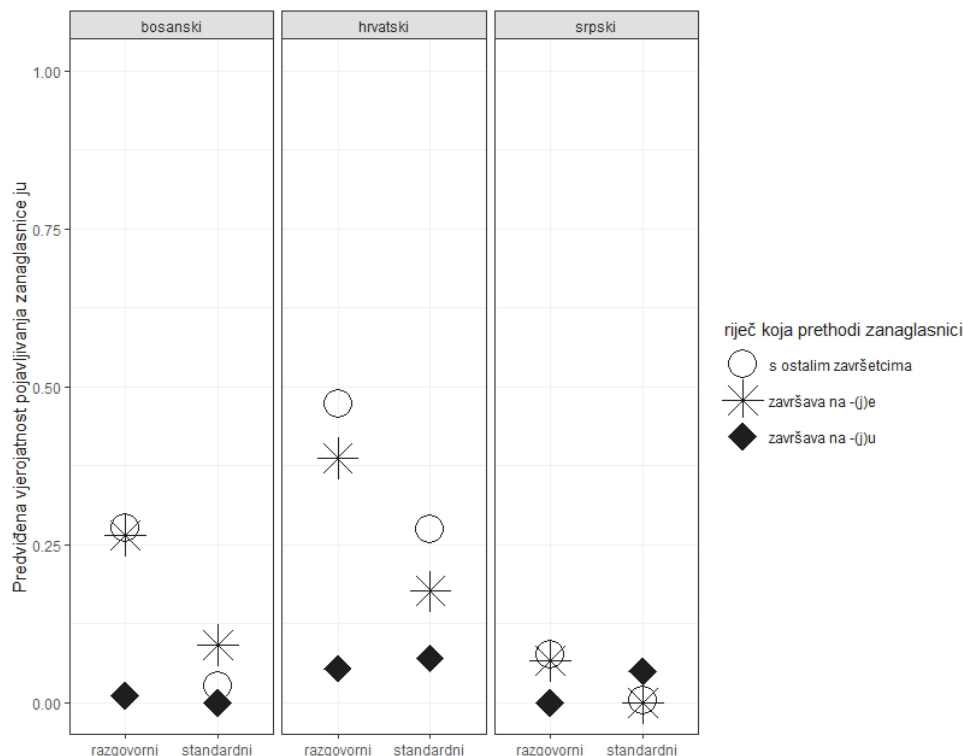
³⁹ Nedostatak primjera u kojima riječ koja završava na *-(j)u* prethodi zanaglasnici *ju* dijelom bi se mogao objasniti općenitom težnjom da se u srpskoj jeziku izostavljaju zanaglasnice pomoćnih gla-

Prema analizi, nema statistički značajne razlike između riječi koje prethode zamjeničkoj zanaglasnici trećega lica ženskoga roda u akuzativu i završavaju na *-(j)e* ili na što drugo.⁴⁰ Valja, međutim, napomenuti da se u slučaju standardnih idioma ta razlika približava statističkoj značajnosti (p vrijednost = 0,08; v. Tablicu 6). U standardnome idiomu bosanskoga i srpskoga jezika postoji blaga tendencija za češće pojavljivanje zanaglasnice *ju* nakon riječi koje završavaju na *-(j)e* u usporedbi s riječima koje imaju ostale završetke (RR 301 %). Za razliku od toga, u standardnome idiomu hrvatskoga jezika riječi koje završavaju na *-(j)e* i dolaze neposredno ispred zanaglasnice trećega lica ženskoga roda u akuzativu imaju negativan utjecaj na ostvarivanje zanaglasnice *ju* (80 % niži RR). S obzirom na ukupnu nisku učestalost zanaglasnice *ju* u bosanskoj i srpskoj poduzorku, trebalo bi prikupiti veći skup podataka da bi se istražilo ovo pitanje.

Na temelju linearnoga regresijskoga modela procijenjena je vjerojatnost pojavljivanja zanaglasnice *ju* u analiziranim idiomima bosanskoga, hrvatskoga i srpskoga jezika, v. Grafikon 2. Za razgovorni idiom bosanskoga jezika u slučaju riječi koja prethodi zanaglasnici trećega lica ženskoga roda u akuzativu i ne završava na *-(j)u* vjerojatnost je pojavljivanja zanaglasnice *ju* 0,28, dok je u slučaju riječi sa završetkom *-(j)e* još niža i iznosi 0,01. U standardnome idiomu bosanskoga jezika ukupna je vjerojatnost pojavljivanja zanaglasnice *ju* izrazito niska, s obzirom na to da iznosi 0,02 u slučaju riječi koje prethode zanaglasnici trećega lica ženskoga roda u akuzativu i završavaju na bilo koji nastavak osim na *-(j)u*, dok je u slučaju riječi koje prethode zanaglasnici i završavaju na *-(j)u* još niža i iznosi tek 0,0007. U srpskoj jeziku vjerojatnost je uporabe zanaglasnice *ju* izrazito mala u svim konfiguracijama, no u razgovornome idiomu nakon riječi koje završavaju na *-(j)e* ili ostalim završetcima (izuzev završetka *-(j)u*) iznosi 0,07. I na kraju, u hrvatskoj jeziku vjerojatnost je pojavljivanja zanaglasnice *ju* nakon riječi s nastavcima koji nisu *-(j)u* 0,47 za razgovorni, a 0,27 za standardni idiom. Vjerojatnost pojavljivanja zanaglasnice *ju* nakon riječi koje završavaju na *-(j)u* nešto je veća nego u srpskim i bosanskim idiomima: za razgovorni iznosi 0,02, a standardni 0,01. Očigledno je da je u usporedbi s vjerojatnošću pojavljivanja akuzativne zanaglasnice *je*, vjerojatnost pojavljivanja zanaglasnice *ju* niža. Valja, međutim, istaknuti da u razgovornome idiomu hrvatskoga jezika akuzativna zanaglasnica *ju* posebno snažno konkurira zanaglasnici *je*. Naime, ako prethodna riječ ne završava ni na *-(j)u* ni na *-(j)e* vjerojatnost je pojavljivanja zanaglasnice *ju* gotovo 0,5, odnosno 50 %.

gola, za više o toj pojavi v. Sonnenhauser i Meermann (2016). Budući da u ovome istraživanju vrsta prethodne riječi i njezin točan oblik nisu bili sustavno ručno označavani, u ovome se trenutku o tome ne mogu dati nikakvi konkretni zaključci. Valja, međutim, istaknuti kako je među riječima koje prethode zanaglasnici *ju* i završavaju na *-(j)u* u našoj građi najčešća bila upravo glagolska zanaglasnica *su*.

⁴⁰ Valja imati na umu da u potonje ne ulazi završetak riječi *-(j)u* koji je posebno analiziran.



Grafikon 2: Vjerojatnost pojavljivanja zanaglasnice *ju* s obzirom na završetak prethodne riječi

Kao što je razvidno iz prethodnih redaka, na temelju provedene statističke analize odgovoreno je na sva istraživačka pitanja (v. razdjel 3).

O1: Distribucija akuzativnih zamjeničkih zanaglasnih varijanata *ju* i *je* u korpusima koji predstavljaju bosanski, hrvatski i srpski standardni idiom potvrđuje razlike opisane na temelju analizirane literature (odnos standardni idiom – standardni idiom). Vjerojatnost je pojavljivanja zanaglasnice *ju* statistički značajno veća u standardnome idiomu hrvatskoga jezika nego u standardnim idiomima bosanskoga i srpskoga jezika. Kako statistička analiza pokazuje, nema značajne razlike u distribuciji zanaglasnice *ju* u standardnim idiomima bosanskoga i srpskoga jezika, što je u skladu i sa stanjem opisanim u literaturi, v. razdjel 2.3.2.

O2: Distribucija je akuzativne zamjeničke zanaglasne varijanate *ju* u korpusima koji predstavljaju razgovorni idiom bosanskoga, hrvatskoga i srpskoga standardnoga jezika statistički značajno različita (odnos razgovorni idiom – razgovorni idiom). Neovisno o tome koji se od triju razgovornih idioma međusobno uspoređuju, razlike su u distribuciji zanaglasnice *ju* statistički značajne.

O3: Distribucija akuzativne zamjeničke zanaglasne varijante *ju* u promatranim okvirima jednoga jezika (bosanskoga, hrvatskoga i srpskoga) pokazuje statistički značajne različite obrasce u standardnome i razgovornome idiomu (odnos standardni idiom – razgovorni idiom). Drugim riječima, distribucija je zanaglasnice *ju* značajno različita uspoređuju li se standardni i razgovorni idiom u okviru jednoga jezika.

O4: Na temelju statističke analize pokazalo se da su samo riječi koje prethode zanaglasnici i završavaju na *-(j)u* u negativnoj korelaciji s pojavljivanjem akuzative zanaglasne varijante *ju*, što je i u skladu s teorijom, odnosno opisom u hrvatskoj gramatikografiji, v. razdjele 2.3.3 – 2.3.4. Za razliku od toga, iako se približava statističkoj značajnosti u standardnim idiomima, utjecaj riječi koja završava na *-(j)e* i prethodi akuzativnoj zanaglasnici ipak nije statistički značajan. Iako neki hrvatski jezikoslovci ističu da bi riječi koje završavaju na *-(j)e* trebala slijediti akuzativna zanaglasna varijanta *ju*, a ne *je* (v. razdjele 2.3.3 – 2.3.4), podatci iz Riznice, odnosno iz referentnoga korpusa s tekstovima koji bi trebali zrcaliti standardni idiom hrvatskoga jezika govore nešto suprotno: riječi koje završavaju na *-(j)e* i dolaze neposredno ispred zanaglasnice trećega lica ženskoga roda u akuzativu u negativnoj su korelaciji s pojavljivanjem zanaglasne varijante *ju*.

6. Rasprava

6.1. Riječi s *-(j)u* završetkom ispred akuzativne zanaglasnice

U prethodnome razdjelu s pomoću statističke analize pokazano je da je vjerojatnost pojavljivanja zanaglasnice *ju* u svim analiziranim idiomima manja ako joj prethode riječi koje završavaju na *-(j)u*. Unatoč tomu u analiziranoj su građi u okviru razgovornoga idioma hrvatskoga jezika pronađeni primjeri u kojima riječ koja završava na *-ju* stoji ispred zanaglasnice *ju*, v. (36). Osim toga, u razgovornome idiomu bosanskoga i hrvatskoga, te standardnome idiomu hrvatskoga i srpskoga jezika u okviru analizirane građe pronađene su potvrde zanaglasnice *ju* koja se naslanja na riječi koje završavaju na *-u*, v. primjere (37)–(39) i (65).

- (36) *na kraju ju ovi u americi ipak nisu htjeli [...].* [hrWaC]
(37) *[...] a u ranom srednjem vijeku benediktinci su ju donijeli u srednju Europu.* [bsWaC]
(38) *[...] u sporu sa BiH, a u kojemu ju BiH optužuje za genocid, objavio je beogradski B92.* [hrWaC]
(39) *Množina ideja koje su je zaokupljale i namjera da ih realizira održavale su ju u životu.* [Riznica]

Međutim, spomenuti su primjeri iznimka jer je, sudeći prema rezultatima statističke analize (v. Tablicu 6), nakon riječi s *-(j)u* završetkom veća vjerojatnost pojavljivanja akuzativne zanaglasnice *je* nego *ju*, v. primjere (40)–(45).

- (40) *Slušaju je s pomnjom.* [Oslo]
(41) *Neki su je upozorili da pobjedom ne podržava sve aspekte [...].* [bsWaC]
(42) *[...] za selektivni pristup rušenju i pitaju je zašto ne sankcionira [...]* [Riznica]
(43) *Porodili su je ukućani i to u surovim uvjetima [...].* [hrWaC]
(44) *Smatraju je veoma lepom mada se meni ne sviđa.* [KSSJ]
(45) *[...] koji su je ubacili u priču o elitnoj prostituciji.* [srWaC]

6.2. Niječni oblik nije i druge riječi s *-(j)e* završetkom ispred akuzativne zanaglasnice

Kako je objašnjeno u prethodnome razdjelu, rezultati statističke analize provedene na temelju prikupljene građe pokazuju da *-(j)e* kao završetak riječi ispred akuzativne zanaglasnice ni u jednome istraživanome idiomu nema utjecaja na veću vjerojatnost pojavljivanja zanaglasnice *ju*. Jezična je uporaba, dakle, u svim istraživanim idiomima bliža opisima u srpskim jezičnim priručnicima, usp. preglede u razdjelima 2.3.2–2.3.4. Valja istaknuti kako su u svim istraživanim idiomima, dakle ne samo u razgovornim nego i u standardnim, nakon riječi koje završavaju na *-(j)e* zabilježeni primjeri s akuzativnom zanaglasnicom *je*, v. (46)–(51).

- (46) *Nije je poznavao, nije je ljubio [...].* [Oslo]
(47) *[...] i gdje je mogu nabaviti?* [bsWaC]
(48) *[...] da strankama koje je sačinjavaju i na sljedećim izborima daju povjerenje.* [Riznica]
(49) *[...] i gdje je pokušavam svesti na nešto što je meni jasno.* [hrWaC]
(50) *Stisnula mi je ruku i nije je odmah pustila.* [KSSJ]
(51) *[...] nelagodnost koju oseća i signalizuje je onima koji se o njemu brinu [...].* [srWaC]

Sudeći prema primjerima u građi, analizirani bi se idiomi mogli razlikovati i u tome koja vrsta riječi ispred akuzativne zanaglasnice *je* s *-je*, odnosno *-e* završetkom može imati ulogu domaćina. Naime, u okviru analizirane građe u standardnome i razgovornome idiomu hrvatskoga jezika ispred zanaglasnice *je* potvrđene su isključivo riječi sa završetkom *-je* koje nisu glagoli (usp. 48 i 49 s 50 i 51).⁴¹

⁴¹ Uzme li se pak u obzir jezična građa cijele Riznice i cijeloga hrWaC-a moguće je naći primjere u kojima prezent punoznačnoga glagola na *-je* stoji ispred akuzativne zanaglasnice *je*. Spomenuti primjeri vrlo vjerojatno nisu dio nasumičnih uzoraka analizirane građe iz standardnoga i razgovornoga idioma

Kako primjeri (46) i (50) pokazuju, u standardnim idiomima bosanskoga i srpskoga jezika zabilježena je uporaba akuzativne zanaglasnice *je* i nakon niječnoga oblika nesvršenoga prezenta trećega lica jednine glagola *biti*. Štoviše, u okviru uzoraka analizirane građe ni u standardnome ni u razgovornome idiomu srpskoga jezika nije zabilježena uporaba zanaglasnice *ju* nakon niječnoga oblika *nije*. Ti su rezultati u potpunosti u skladu s normom srpskoga standarda kako ju opisuju Ivić i dr. (2007: 128), v. razdjel 2.3.2. Nadalje, uporaba zanaglasnice *ju* nakon niječnoga oblika *nije* nije zabilježena ni u uzorku razgovornoga idioma bosanskoga, ali je zabilježena u standardnim idiomima bosanskoga i hrvatskoga te u razgovornome idiomu hrvatskoga jezika, v. primjere (52)–(54).

- (52) *Nije ju niko vidio.* [Oslo]
(53) *Nije ju uvrstila ni nova intendantica Mani Gotovac.* [Riznica]
(54) *[...] nije ju prosljedila nadležnom sudu, što je bila dužna učiniti.* [hrWaC]

Naime, uporaba se akuzativne zamjeničke zanaglasnice *je* nakon niječnoga oblika *nije* u suvremenome standardnome idiomu srpskoga jezika s normativnoga stajališta više ne smatra pogrešnom (Ivić i dr. 2007: 128), v. razdjel 2.3.2. Nadalje, kao što je već rečeno u razdjelu 2.3.2, samo se u dijelu srpskih jezikoslovnih priručnika (npr. Klajn 2005: 79; Ivić i dr. 2007: 128; Mrazović i Vukadinović 2009: 366) objašnjava da niječni oblik *nije* koji uvjetuje uporabu zamjeničke zanaglasnice *ju* dolazi ispred nje. U tome kontekstu zanimljivo je istaknuti kako je u okviru analizirane građe, odnosno analiziranih uzoraka jedino u standardnim idiomima bosanskoga i srpskoga jezika potvrđena uporaba akuzativne zamjeničke zanaglasnice *je* ispred niječnoga oblika *nije*, v. primjere u (55) i (56).⁴²

- (55) *Lina je nije ni pogledala [...].* [Oslo]
(56) *[...] ali je nije zadavio.* [KSS]

Valja, međutim, dodati da je, iako rjeđe, u standardnome idiomu bosanskoga jezika ispred niječnoga oblika *nije* zabilježena i akuzativna zanaglasnica *ju*, baš kao i u standardnome idiomu hrvatskoga te u razgovornim idiomima bosanskoga i hrvatskoga jezika, v. primjere u (57)–(60).

- (57) *On ju nije razumio.* [Oslo]

hrvatskoga jezika jer su u njima puno rjeđi nego u srpskoj razgovornome idiomu u kojemu su potvrđeni u sklopu analizirane građe, odnosno uzoraka.

⁴² U Riznici i hrWaC-u mogu se naći potvrde s akuzativnom zanaglasnicom *je* ispred naglašenoga niječnoga oblika *nije*. Primjeri sa spomenutim *je nije* slijedom nisu ušli u naše nasumične uzorke s gradom iz standardnoga i razgovornoga idioma hrvatskoga jezika vrlo vjerojatno jer su u njima puno rjeđi nego u standardnome bosanskoj i standardnome srpskoj idiomu u kojima su potvrđeni u sklopu analizirane građe, odnosno uzoraka.

- (58) *Iako ju nije mogla patentirati ni od nje zaraditi novac [...].* [bsWaC]
(59) *[...] ali iako ju nije dodirnuo, uspio je omesti gostujućeg [...].* [Riznica]
(60) *[...] u Hrvatskoj i Splitu nije bilo uloge a da ju nije tumačio Genda [...].* [hrWaC]

6.3. Kontekstno neuvjetovana uporaba akuzativne zanaglasnice *ju*

U raspravi se posebna pozornost pridaje kontekstno neuvjetovanoj uporabi akuzativne zamjeničke zanaglasnice *ju*, odnosno slučajevima kada ni lijevi ni desni kontekst ne sadrže riječi koje završavaju, odnosno počinju na (-)(j)e(-), uključujući i naglašeni niječni prezentski oblik glagola *nije* te susljednu uporabu glagolske zanaglasnice *je*. U korpusima bosanskoga jezika takvih je primjera relativno malo, u Oslo korpusu tek šest od 249 analiziranih primjera s akuzativnom zamjeničkom zanaglasnicom *ju* nije kontekstno uvjetovano, riječ je o četiri različita teksta u kojima se tako nešto javlja, v. primjer u (61).

- (61) *[...] da ju nitko ne može usrećiti [...].* [Oslo]

Spomenuti je primjer iz djela autora A. H. Bjelevca za kojega se u biografijama navodi kako je bio bošnjački i hrvatski književnik koji je posljednjih 30 godina svoga života proveo u Zagrebu. Osim kontekstno neuvjetovane uporabe akuzativne zanaglasnice *ju*, u navedenoj se rečenici i na primjeru uporabe neodređene zamjenice *nitko* vidi utjecaj hrvatskoga jezika. Nadalje, valja istaknuti da je jedan od tih četiriju tekstova iz Oslo korpusa u kojima se javlja kontekstno neuvjetovana uporaba akuzativne zanaglasnice *ju* napisao M. Jergović, pisac čiji se idiolekt teško može jednoznačno odrediti kao bosanski standardni idiom.

Kontekstno neuvjetovana uporaba akuzativne zamjeničke zanaglasnice *ju* u bsWaC-u javlja se u 60 od 300 analiziranih primjera koji su zabilježeni na 35 različitih internetskih stranica, v. primjer u (62).

- (62) *Tek sam ju počela čitati, no već su prve tri stranice bile [...].* [bsWaC]

Valja, međutim, naglasiti da je jedna od stranica na kojima se zamjenička zanaglasnica *ju* rabi iako nije u kontaktu s riječima koje završavaju, odnosno počinju na (-)(j)e(-), uključujući i glagolsku zanaglasnicu *je*, Croportal.ba.

U analiziranoj je građi iz Riznice kontekstno neuvjetovana uporaba akuzativne zanaglasnice *ju* također zabilježena u 60 od 300 primjera, v. primjer u (63). Treba dodati kako su od toga tri primjera s kraja 19., odnosno početka 20. stoljeća,⁴³ je-

⁴³ U Riznici uz Kovačićeve Feljtone i članke u metapodatcima stoji 1952. kao godina objave, no tekst je zapravo nastao puno prije.

dan je iz 50-ih godina 20. stoljeća, jedan je iz 1997. godine, dok je svih ostalih 55 iz razdoblja nakon 2000. godine. Ti su podatci, dakle, u skladu s promjenama u normi, odnosno u hrvatskim jezičnim priručnicima opisanima u razdjelima 2.3.3 i 2.3.4.

(63) [...] *primjerice, jedan metar kad ju razvučemo na dvostruku duljinu.* [Riznica]

Kontekstno neuvjetovana uporaba akuzativne znanaglasnice *ju* u analiziranoj je građi najčešća u podacima prikupljenima iz hrWaC-a, korpusa koji predstavlja razgovorni idiom hrvatskoga jezika. U 132 od 300 ručno anotiranih primjera s akuzativnom znanaglasnicom *ju* njezina se uporaba ne može pripisati ni glagolskoj znanaglasnici *je*, kao ni prethodnim ni susljednim riječima koje počinju, odnosno završavaju na (-)(j)e(-), v. primjer u (64).

(64) [...] *a tema prve radionice bila je Riva i kako ju djeca doživljavaju.* [hrWaC]

Kao što je već rečeno, u analiziranoj građi prikupljenoj iz srpskoga dijela InterCorpa (v. razdjel 4.2) koji predstavlja standardni idiom srpskoga jezika, kontekstno neuvjetovana uporaba znanaglasnice *ju* uopće nije potvrđena. Za razliku od toga, u KSSJ-u, koji je također u ovome istraživanju poslužio kao izvor građe za standardni idiom srpskoga jezika, dva su primjera od 300 analiziranih s akuzativnom znanaglasnicom *ju* kontekstno neuvjetovana, v. primjere (65) i (66).

(65) [...] *a mnogi su ju smatrali privlačnom i nakon što je prešla 60. godinu.* [KSSJ]

(66) [...] *i trebalo bi da ju izudaram prvi put kada čujem njenu reč.* [KSSJ]

Primjer (65) ušao je u KSSJ kao tekst objavljen na stranicama Radio-televizije Srbije. Valja, međutim, dodati kako se cijela (gotovo) identična rečenica nalazi i na srpskim i na srpskohrvatskim i na hrvatskim stranicama Wikipedije.⁴⁴ Drugi je primjer iz djela R. Petrovića napisanoga 1927. godine.

⁴⁴ Hrvatske stranice Wikipedije (i), srpskohrvatske stranice Wikipedije (ii), srpske stranice Wikipedije (iii), u originalu i transliterirane s ćirilice na latinicu.

(i) *Empire časopis ju je 1995. prozvaao 25. najseksi zvijezdom u filmskoj povijesti, a mnogi su ju smatrali privlačnom i nakon što je prešla 60. godinu.* [https://hr.wikipedia.org/wiki/Sophia_Loren]

(ii) *Časopis Empire ju je 1995. prozvaao 25. najseksi zvijezdom u filmskoj povijesti, a mnogi su ju smatrali privlačnom i nakon što je prešla 60. godinu.* [https://sh.wikipedia.org/wiki/Sophia_Loren]

(iii) *Часопис Емпajer ју је 1995. године прозвао 25. најсекси звездом у филмској историји, а многи су ју сматрали привлачном и након што је прешла 60. годину.* [https://sr.wikipedia.org/wiki/Софија_Лорен]

Časopis Empajer ju je 1995. godine prozvaao 25. najseksi zvezdom u filmskoj industriji, a mnogi su ju smatrali privlačnom i nakon što je prešla 60. godinu.

Valja istaknuti da je tekst u hrvatskoj Wikipediji najstariji i datira iz 2006., dok je tekst unesen u srpsku Wikipediju nastao 2007., a onaj unesen u srpskohrvatsku Wikipediju još je noviji i datira iz 2009. godine. Na temelju toga može se pretpostaviti da je kontekstno neuvjetovana uporaba znanaglasnice *ju* potvrđena u KSSJ-u, korpusu koji predstavlja standardni idiom srpskoga jezika doista, kako ju Piper i

U analiziranoj građi prikupljenoj iz srWaC-a koji predstavlja razgovorni idiom srpskoga jezika, kontekstno je neuvjetovana uporaba zanaglasnice *ju* potvrđena u 15 od 300 primjera, v. (67). Spomenuti su primjeri zabilježeni na 11 različitih internetskih stranica.

(67) [...] *za koju napravis SVE da ju realizujes.* [srWaC]

Dok je u analiziranim idiomima srpskoga jezika kontekstno neuvjetovana uporaba akuzativne zanaglasnice *ju* iznimno rijetka (u 17 od 667 analiziranih primjera), u hrvatskim je idiomima znatno češća (u 192 od 600 analiziranih primjera), dok se bosanski idiomi nalaze negdje u sredini (u 66 od 549 analiziranih primjera). Nadalje, posebno je zanimljiva činjenica da je kontekstno neuvjetovana uporaba akuzativne zanaglasnice *ju* češća u razgovornim nego u standardnim idiomima u okviru analiziranoga južnoslavenskoga jezičnoga područja.

I na kraju, stave li se u žarište pozornosti standardni idiomi, posebno se zanimljiva slika dobiva na temelju podataka iz Riznice, korpusa s tekstovima koji predstavljaju standardni idiom hrvatskoga jezika. Kao što je već rečeno, većina primjera za kontekstno neuvjetovanu uporabu zamjeničke zanaglasnice *ju* potječe iz razdoblja nakon 2000. godine. Osim toga, valja istaknuti kako je većina primjera iz publicističkoga registra, no to može biti i u izravnoj svezi s jezičnom građom u korpusu, odnosno s njegovim ustrojem. Naime, u Riznici u razdoblju od 2000. godine prevladavaju publicistički tekstovi. Tri su primjera iz književnoga registra, dva su iz djela koja su nastala prije razdoblja Kraljevine Srba, Hrvata i Slovenaca, dakle iz djela nastalih neposredno prije ili neposredno nakon *Gramatike i stilistike hrvatskoga ili srpskoga književnog jezika* (Maretić 1899). U analiziranoj je građi, i to u okviru književnoga registra, potvrđen samo jedan primjer kontekstno neuvjetovane uporabe zamjeničke zanaglasnice *ju* u djelu objavljenome u vrijeme Socijalističke Federativne Republike Jugoslavije. Čini se kako spomenuti podatci indirektno potvrđuju Mamićeve (1995: 188) pretpostavke da se kontekstno neuvjetovana uporaba zanaglasnice *ju*, koja je u postmaretićevskoj eri bila izrazito potiskivana, polako vraća u uporabu u hrvatskome jeziku (ne samo u njegov razgovorni nego i u njegov standardni idiom).

7. Zaključci i smjernice za buduća istraživanja

Još krajem prošloga stoljeća Mamić (1995: 188) uočava postupne promjene u hrvatskoj normi⁴⁵ te tvrdi da zanaglasnica *ju* polako postaje prihvaćenija nego što je to

Klajn (2014: 97) nazivaju, strana konstrukcija (v. razdjel 2.3.2) unesena prilikom kopiranja teksta napisanoga na hrvatskome.

⁴⁵ Nije zapravo jasno misli li Mamić (1995: 198) na noviju normu standardnoga idioma hrvatskoga jezika koja se opisuje u gramatikama nastalim nakon 90-ih (npr. Raguž 1997, 2010; Silić i Pranjković

bio slučaj tijekom 20. stoljeća. Međutim, iako mnogi hrvatski jezikoslovci odobravaju slobodnu uporabu zanaglasničkoga oblika *ju* i njegovu ravnopravnost s oblikom *je* ili mu čak daju prednost, detaljna kontrastivna analiza suvremenih hrvatskih gramatika i jezičnih savjetika pokazuje da zamjenička zanaglasnica *ju* još uvijek nije potpuno istisnula iz uporabe zanaglasnicu *je* u standardnome idiomu hrvatskoga jezika (usp. Mamić 1995: 188). Jezikoslovci koji daju prednost zanaglasnici *ju* kao argumente obično navode jezik starijih hrvatskih pisaca, dijakronijsku prvotnost akuzativoga oblika *ju* i njegovu podudarnost s naglašenim oblikom *nju* te hrvatskim dijalektima uključujući i kajkavsko narječje. Kontrastivna pak analiza suvremenih bosanskih i srpskih gramatika i jezičnih savjetika pokazuje da je uporaba zanaglasnice *ju* u standardnim idiomima bosanskoga i srpskoga jezika, iako nešto slobodnija nego tijekom 20. stoljeća, i dalje znatno ograničenija nego u standardnome idiomu hrvatskoga jezika i uglavnom svodiva na kontekst u kojemu prethodi glagolskoj zanaglasnici *je*.

Spomenute su razlike bile poticaj da se uporaba akuzativnih zanaglasnica *ju* i *je* provjeri u korpusima s jezičnom građom koja priližno predstavlja standardne idio-
me s jedne strane i jezičnom građom koja u dobroj mjeri odgovara razgovornim idiomima bosanskoga, hrvatskoga i srpskoga jezika s druge strane. U tu je svrhu u {bs,hr,sr}WaC-u, Rizinici, KSSJ-u, srpskoj dijelu InterCorpa i Oslo korpusu prikupljena i ručno označena građa s 3916 rečenica u kojima su uporabljene zanaglasnice *ju* i *je* kako bi se provjerilo koje varijable utječu na vjerojatnost pojavljivanja zanaglasnice *ju*. Svi primjeri u kojima akuzativna zanaglasnica *ju* dolazi isključivo uz glagolsku zanaglasnicu *je* bili su isključeni iz daljnega istraživanja jer slijed akuzativne zanaglasnice *je* i glagolske zanaglasnice *je* teoretski moguć u nestandardnim idiomima, kako je bilo i očekivano, nije bio potvrđen u građi. Rezultati uopćenoga regresijskoga modela pokazuju da je vjerojatnost pojavljivanja akuzativne zanaglasnice *ju* veća u razgovornim nego u standardnim idiomima analiziranih jezika. Osim toga, statistička analiza upućuje na to da su razlike između uporabe akuzativne zanaglasnice *ju* u standardnome idiomu hrvatskoga jezika s jedne strane i standardnim idiomima bosanskoga i srpskoga jezika s druge strane statistički značajne: uporaba je zanaglasne varijante *ju* vjerojatnija u standardnome idiomu hrvatskoga jezika. I na kraju statistička analiza potvrđuje, kako se u literaturi i pretpostavlja, da riječi sa završetkom *-(j)u* koje stoje ispred akuzativne zanaglasnice imaju negativan utjecaj na pojavljivanje varijante *ju*. Drugim riječima, nakon tih je riječi vjerojatnija uporaba akuzativne zanaglasnice *je* u svim analiziranim idiomima.

Uopćeni linearni regresijski model na temelju analiziranih podataka nije potvrdio statističku značajnost *-(j)e* završetka riječi koja prethodi akuzativnoj zanaglasnici na pojavljivanje varijante *ju* kako se to pretpostavlja u (uglavnom) hrvatskoj

2007; Težak i Babić 2000; Ham i dr. 2014) ili misli na uzus, odnosno na način na koji govornici hrvatskoga jezika općenito govore i pišu nakon 90-ih.

dijelu predstavljene literature. Na ovome mjestu treba svakako istaknuti i ograničenja same građe, odnosno analiziranih uzoraka. Kao prvo, u analiziranoj se građi nalazi relativno malo primjera u kojima *-(j)e* riječ prethodi zanaglasnici trećega lica jednine ženskoga roda u akuzativu i većina se odnosi na *-e*, a ne na *-je* završetak.⁴⁶ Osim toga, iako relativno malobrojni, primjeri u građi upućuju na to da bi se analizirani idiomi mogli razlikovati i u tome koja vrsta riječi sa završetkom *-je*, odnosno *-e* može stajati ispred akuzativne zanaglasnice *je*. U analiziranim se primjerima u okviru prikupljene građe iz standardnoga i razgovornoga idioma hrvatskoga jezika ispred zamjeničke zanaglasnice *je* nalaze isključivo neglagolske riječi sa završetkom *-je* (usp. 48 i 49 s 50 i 51). Sve spomenuto govori u prilog tome da su potrebni i dodatni podaci s detaljnijim razinama anotacije za završetke riječi koje prethode i početke riječi koje slijede zanaglasnicu i nove analize. Potonje je posebice važno tim više što ovo istraživanje zbog prirode samih podataka (nedovoljnoga broja primjera u kojima akuzativnu zanaglasnicu trećega lica jednine ženskoga roda slijedi riječ koja počinje na *(j)u-* odnosno *(j)e-*) u statistički model i analizu nije moglo obuhvatiti i početak susljedne riječi kao potencijalno značajnu nezavisnu varijablu koja utječe na ostvarivanje *ju* ili *je* varijante. Iako se pretpostavlja da zbog same zanaglasničke, odnosno naslanjačke prirode zamjenice trećega lica jednine ženskoga roda u akuzativu važniju ulogu igra završetak riječi koja prethodi, a ne početak riječi koja slijedi, hipotezu je ipak nužno empirijski potvrditi. Zanaglasnička, odnosno naslanjačka priroda vjerojatno stoji i iza dijelom potvrđenoga pravila o blagoglasju, odnosno izbjegavanja ponavljanja *-(j)u ju* sloga i davanja prednosti slijedu *-(j)u je*. S obzirom na to da ih ta naslanjačka priroda čini bliskima afiksima, ostvarivanje varijante *je* umjesto varijante *ju* nakon riječi koje završavaju na *-(j)u* slično je razjednačavanjima koja se događaju u okviru fleksije tipa *Bečom* (**Bečem*). Međutim, ako je to tako i dalje je otvoreno pitanje zašto eufonija, odnosno razjednačavanje nije potvrđeno kao statistički značajan čimbenik i u slučaju slijeda *-(j)e ju*. Kao što je već napomenuto, u svakome slučaju potrebna su daljnja istraživanja jer bi jedan od razloga zašto je to tako mogao ležati i u količini i vrsti, odnosno strukturi analiziranih podataka.

Kvalitativna analiza predstavljena u okviru rasprave pokazuje da je, unatoč tome što se u dijelu predstavljene literature tvrdi suprotno, uporaba akuzativne zanaglasnice *ju* moguća i izvan konteksta sa susljednom glagolskom zanaglasnicom *je* i prethodnom riječju koja završava na *-(j)e* i to u svim analiziranim idiomima. Spomenuta je kontekstno neuvjetovana uporaba akuzativne zanaglasnice *ju* iznimno rijetka u standardnome i razgovornome idiomu srpskoga jezika te standardnome idiomu bosanskoga, nešto manje rijetka u razgovornome idiomu bosanskoga i standardnome idiomu hrvatskoga, dok je najčešća u razgovornome idiomu hrvatskoga

⁴⁶ Upravo zbog maloga broja primjera, riječi koje završavaju i na *-u* i na *-ju* bile su anotirane u okviru *-(j)u* završetaka, a isto vrijedi i za *-e* i *-je*, odnosno *-(j)e* završetke riječi koje prethode zanaglasnici trećega lica jednine ženskoga roda u akuzativu.

jezika. S obzirom na to da je zanaglasnica *ju* potvrđena ne samo u hrvatskim nego i u srpskim i bosanskim dijalektima, njezina relativno slaba kontekstno neuvjetovana zastupljenost u standardnim i razgovornim idiomima bosanskoga i srpskoga jezika vrlo je vjerojatno dijelom odraz jezičnoga planiranja kao i društvenoga statusa bosanskih i srpskih idioma s kontekstno neuvjetovanom uporabom zanaglasnice *ju*. Drugim riječima, na primjeru uporabe zanaglasnice *ju* moglo bi se reći da su hrvatski i standardni i razgovorni idiom otvoreniji utjecajima iz inih idioma hrvatskoga jezika nego što su to standardni i razgovorni idiomi bosanskoga i srpskoga jezika.

Analizirana građa pokazuje kako su promjene i u referentnim normativnim priručnicima i u građi koja približno odražava standardni idiom hrvatskoga jezika posebno zanimljive zbog čega bi mogle biti temom sljedećih istraživanja. Naime, u okviru analizirane građe kontekstno je neuvjetovana uporaba akuzativne zanaglasnice *ju* češća u primjerima nastalima nakon 2000. godine čime se potvrđuju u Mamićevu (1995: 188) članku nagoviještene mijene u hrvatskoj normi. Međutim, i dalje ostaje dijelom nejasno kako se točno uporaba akuzativnih zanaglasnica *ju* i *je* mijenjala tijekom povijesti i u predstandardnizacijskome razdoblju i u standardnome idiomu hrvatskoga jezika i koji su sve točno jezični, ali i izvanjezični čimbenici imali utjecaja na te promjene. Sljedeća bi dijakronijska i sociolingvistička istraživanja trebala dati odgovore i na ta i dalje otvorena pitanja.

8. Popis literature

- Agić, Željko; Dovedan, Zdravko; Tadić, Marko. 2008. Improving Part-of-Speech tagging accuracy for Croatian by morphological analysis. *Informatica* 32/4. 445–451.
- Alerić, Danijel. 1964. O enklitikama *nju* i *nje*. *Jezik* 12/2. 47–51.
- Alexander, Ronelle. 2008. Rhythmic structure constituents and clitic placement in Bosnian, Croatian, Serbian. *American contributions to the 14th International Congress of Slavists. Volume 1: Linguistics*. Ur. Bethin, Christina Y. Slavica. Bloomington. 1–19.
- Aljović, Nadira. 2005. On clitic climbing in Bosnian/Croatian/Serbian. *Lingvistički vidici* 34/5. 58–84.
- Ammon, Ulrich. 1987. Language – variety/standard variety – dialect. *Sociolinguistics: An international handbook of the science of language and society*. Ur. Ammon, Ulrich; Dittmar, Norbert; Mattheier, Klaus J. Walter de Gruyter. Berlin. 316–35.
- Antonić, Ivana. 1998. Zaboravljena enklitika *ju*. *Jezik danas: glasilo Matice srpske za kulturu usmene i pisane reči* 6. 8–11.
- Babić, Stjepan. 1963. Enklitika između imena i prezimena. *Jezik* 11/2. 63–64.
- Babić, Stjepan. 1964a. Smije li se upotrebljavati zamjenički *si*. *Jezik* 12/4. 118–119.
- Babić, Stjepan. 1964b. Lingvistika i logika Ivana Brabeca. *Jezik* 12/5. 151–159.
- Babić, Stjepan; Brozović, Dalibor; Moguš, Milan; Pavešić, Slavko; Škarić, Ivo; Težak, Stjepko. 1991. *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika*. Nakladni zavod Globus. Zagreb.
- Barac-Grum, Vida; Malić, Dragica; Pavešić, Slavko; Vince, Zlatko. 1971. *Jezični savjetnik s gramatikom*. Matica hrvatska. Zagreb.

- Barić, Eugenija; Lončarić, Mijo; Malić, Dragica; Pavešić, Slavko; Peti, Mirko; Zečević, Vesna; Znika, Marija. 1979. *Priručna gramatika hrvatskoga književnog jezika*. Školska knjiga. Zagreb.
- Barić, Eugenija; Lončarić, Mijo; Malić, Dragica; Pavešić, Slavko; Peti, Mirko; Zečević, Vesna; Znika, Marija. 1997. *Hrvatska gramatika*. Školska knjiga. Zagreb.
- Barić, Eugenija; Hudeček, Lana; Koharović, Nebojša; Lončarić, Mijo; Lukenda, Marko; Mamić, Mile; Mihaljević, Milica; Šarić, Ljiljana; Švačko, Vanja; Vukojević, Luka; Zečević, Vesna; Žagar, Mateo. 1999. *Hrvatski jezični savjetnik*. Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- Belaj, Branimir. 2022. O standardnom idiomu iz kognitivne perspektive. *Standardni i nestandardni idiomi (zbornik sažetaka)*. Ur. Glušac, Maja; Mikić Čolić, Ana. Filozofski fakultet Osijek. Osijek. 17–19.
- Belaj, Branimir. 2023. *Kognitivna lingvistika i hrvatski jezik*. Filozofski fakultet. Osijek.
- Birtić, Matea; Blagus Bartolec, Goranka; Hudeček, Lana; Jojić, Ljiljana; Kovačević, Barbara; Lewis, Kristian; Matas Ivanković, Ivana; Mihaljević, Milica; Miloš, Irena; Ramadanović, Ermina; Vidović, Domagoj. 2012. Školski rječnik hrvatskoga jezika. Školska knjiga. Zagreb.
- Bošković, Željko. 2004. Clitic placement in South Slavic. *Journal of Slavic Linguistics* 12/1. 37–90.
- Bošnjak Botica, Tomislava; Polančec, Jurica; Sviben, Robert. 2022. Korpusno istraživanje hrvatskih imenica s dugom i kratkom množinom. *Jezikoslovlje* 23/1. 35–74. <https://doi.org/10.29162/jez.2022.2>.
- Brabec, Ivan. 1964. Enklitika – šta je to. *Jezik* 12/5. 143–151.
- Brabec, Ivan; Hraste, Mate; Živković, Sreten. 1952. *Gramatika hrvatskoga ili srpskog jezika*. Školska knjiga. Zagreb.
- Brač, Ivana; Bratanić, Maja; Ostroški Anić, Ana. 2015. Hrvatsko nazivlje i nazivoslovlje od Šuleka do Strune – hrvatski jezik i terminološko planiranje. *Od Šuleka do Schengena: terminološki, terminografski i prijevodni aspekti jezika struke*. Ur. Bratanić, Maja; Brač, Ivana; Pritchard, Boris. Institut za hrvatski jezik i Pomorski fakultet u Rijeci. Zagreb. 3–26.
- Brozović, Dalibor. 1970. *Standardni jezik*. Matica hrvatska. Zagreb.
- Brozović, Dalibor. 2016. *Jezik današnji: 1965–1968*. Ur. Krešimir Mićanović. Disput. Zagreb.
- Browne, Wayles. 1968. Srpskohrvatske enklitike i teorija transformacione gramatike. *Zbornik za filologiju i lingvistiku Matice srpske* 11. 25–29.
- Browne, Wayles. 1974. On the problem of enclitic placement in Serbo-Croatian. *Slavic transformational syntax. Volume 10*. Ur. Brecht, Richard; Chvany, Catherine. Michigan Slavic Publications. University of Michigan. Ann Arbor. 36–52.
- Browne, Wayles. 2003. Razlike u redu riječi u zavisnoj rečenici. *Wiener Slawistischer Almanach* 57. 39–44.
- Browne, Wayles. 2004. Serbo-Croatian Enclitics for English-Speaking Learners. *Journal of Slavic Linguistics* 12/1. 249–283. (Reprint from 1975).
- Browne, Wayles. 2014. Groups of Clitics in West and South Slavic Languages. *Slavic and German in Contact: Studies from Areal and Contrastive Linguistics. Slavic Eurasian Studies* 26. Ur. Kaczmarzka, Elżbieta; Nomachi, Motoki. Slavic Research Center. Hokkaido. 81–96.
- Brozović Rončević, Dunja; Čavar, Damir; Čavar, Małgorzata; Stojanov, Tomislav; Štrkalj Despot, Kristina; Ljubešić, Nikola; Erjavec, Tomaž. 2018. Croatian language corpus Riznica 0.1. *Slovenian language resource repository CLARIN.SI*. <http://hdl.handle.net/11356/1180>.

- Buchstaller, Isabelle; Khattab, Ghada. 2013. Population samples. *Research Methods in Linguistics*. Ur. Podesva, Robert J.; Sharma, Devyani. Cambridge University Press. Cambridge. 74–95.
- Bukumirić, Mileta. 2003. Govori severne Metohije. *Srpski dijalektološki zbornik. Volume 50*. Ur. Mladenović, Aleksandar. Srpska akademija nauka i umetnosti i Institut za srpski jezik SANU. Beograd. 3–356.
- Čedić, Ibrahim. 2008. *Osnovi gramatike bosanskog jezika*. Institut za jezik u Sarajevu. Sarajevo.
- Čedić, Ibrahim. 2013. *Jezički savjetnik*. Institut za jezik u Sarajevu. Sarajevo.
- Čedić, Ibrahim; Hajdarević, Hadžem; Kadić, Safet; Kršo, Aida; Valjevac, Naila. 2007. *Rječnik bosanskog jezika*. Institut za jezik. Sarajevo.
- Čermák, František; Rosen, Alexandr. 2012. The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics* 17/3. 411–427.
- Čavar, Damir; Wilder, Chris. 1994. Clitic third in Croatian. *Linguistics in Potsdam* 1. 25–63.
- Čavar, Damir; Brozović-Rončević, Dunja. 2012. Riznica: The Croatian Language Corpus. *Prace filologiczne* 63. 51–65.
- Ćurković, Dijana. 2014. *Govor Bitelića*. Doktorski rad. Filozofski fakultet Sveučilišta u Rijeci.
- Diesing, Molly; Filipović Đurđević, Dušica; Zec, Draga. 2009. Clitic placement in Serbian: Corpus and experimental evidence. *The Fruits of Empirical Linguistics. Volume 2*. Ur. Winkler, Susanne; Featherston, Sam. Mouton de Gruyter. Berlin. 59–73. <https://doi.org/10.1515/9783110216158>.
- Dulčić, Mihovil (ur.). 1997. *Govorimo hrvatski: jezični savjeti*. Hrvatski radio. Zagreb.
- Fekete, Egon; Čupić, Drago; Terzić, Bogdan. 2005. *Srpski jezički savjetnik*. Službeni list Srbije i Crne Gore. Beograd.
- Frančić, Anđela; Hudeček, Lana; Mihaljević, Milica. 2005. *Normativnost i višefunkcionalnost u hrvatskoj standardnoj jeziku*. Hrvatska sveučilišna naklada. Zagreb.
- Frančić, Anđela; Petrović, Bernardina. 2013. *Hrvatski jezik i jezična kultura*. Visoka škola za poslovanje i upravljanje „Baltazar Adam Krčelić“. Zaprešić.
- Franks, Steven; Progovac, Ljiljana. 1994. On the Placement of Serbo-Croatian Clitics. *Indiana Slavic Studies* 7. 69–78.
- Franks, Steven; Peti-Stantić, Anita. 2006. Splitting puzzles in South Slavic. Presented at FDSL 6.5. University of Nova Gorica. Online dostupno na: https://www.academia.edu/2438162/Splitting_puzzles_in_South_Slavic_talk_with_Steven_Franks_. Posljednji pristup 30.12.2020.
- Franks, Steven. 2010. Clitics in Slavic. *Contemporary Issues in Slavic Linguistics* 10. 1–157.
- Franks, Steven; King, Tracy Holloway. 2000. *A Handbook of Slavic clitics*. Oxford University Press. Oxford.
- Gortan, Veljko. 1956. O mjestu enklitike u nezavisnoj rečenici. *Jezik* 5/1. 14–16.
- Gračanin-Yukse, Martina. 2019. Clitics that Don't Climb. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 45/2. 401–416. <https://doi.org/10.31724/rihjj.45.2.9>.
- Ham, Sanda. 2006. O mjestu znanaglasnice i o lektorima. *Jezik* 53/2. 70–72.
- Ham, Sanda. 2007. *Školska gramatika hrvatskoga jezika*. Školska knjiga. Zagreb.
- Ham, Sanda; Mlikota, Jadranka; Baraban, Borko; Orlić, Alen. 2014. *Hrvatski jezični savjeti*, Školska knjiga. Zagreb.
- Hansen, Björn; Kolaković, Zrinka; Jurkiewicz-Rohrbacher, Edyta. 2018. Clitic climbing and stacked infinitives in Bosnian, Croatian and Serbian – A corpus-driven study. *Grammar*

- and Corpora* 2016. Ur. Fuß, Eric; Konopka, Marek; Trawiński, Beata; Waßner, Ulrich H. Heidelberg University Publishing (heiUP). Heidelberg. 259–268. <https://doi.org/10.17885/heiup.361.509>.
- Horvat, Marijana. 2007. Novi standardnojezični priručnik. *Jezičnik* 54/1. 36–39.
- Hoyt, Alexander D. 2012. *Hrvatski jezik u Zagrebu: sociolingvistički pogled*. Hrvatska sveučilišna naklada. Zagreb.
- Hudeček, Lana; Mihaljević, Milica. 2009. *Jezičnik medija – publicistički funkcionalni stil*. Hrvatska sveučilišna naklada. Zagreb.
- Hudeček, Lana; Matković, Maja; Čutuk, Igor. 2012. *Jezični priručnik Coca-Cole HBC Hrvatska – Hrvatski jezik u poslovnoj komunikaciji Hrvatski jezik u poslovnoj komunikaciji*. Coca-Cola HBC Hrvatska. Zagreb.
- Hudeček, Lana; Mihaljević, Milica. 2019. *Hrvatska školska gramatika*. Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- Ivić, Pavle; Klajn, Ivan; Pešikan, Mitar; Brborić, Branislav. 2007. *Srpski jezički priručnik*. Beogradska knjiga. Beogradska knjiga.
- Jahić, Dževad; Halilović, Senahid; Palić, Ismail. 2000. *Gramatika bosanskoga jezika*. Dom štampe. Zenica.
- Janda, Laura Alexis; Nessel, Tore; Baayen, Harald R. 2010. Capturing correlational structure in Russian paradigms: A case study in logistic mixed-effects modeling. *Cognitive Linguistics and Linguistic Theory* 6/1. 29–48.
- Jelaska, Zrinka; Bjedov, Vesna. 2015. Pogreške ili promjene – ovladanost odabranim hrvatskim morfosintaktičkim sadržajima učenika završnoga razreda osnovne škole. *Jezičnik* 16/2–3. 227–252.
- Jojić, Ljiljana; Nakić, Anuška; Vajs Vinja, Nada; Zečević, Vesna. (ur.) 2015. *Veliki rječnik hrvatskoga standardnog jezika*. Školska knjiga. Zagreb.
- Jonke, Ljudevit. 1953. O pomicanju enklitike prema početku rečenice. *Jezičnik* 2/5. 149–150.
- Joseph, John Earl. 1987. *Eloquence and power: The rise of language standards and standard languages*. Blackwell. New York.
- Jurkiewicz-Rohrbacher, Edyta; Kolaković, Zrinka, Hansen, Björn. 2017a. Web Corpora – the best possible solution for tracking rare phenomena in underresourced languages: clitics in Bosnian, Croatian and Serbian. *CMLC-5+BigNLP 2017*. Ur. Bański, Piotr; Kupietz, Marc; Lungen, Harald; Rayson, Paul; Biber, Hanno; Breiteneder, Evelyn; Clematide, Simon; Mariani, John; Stevenson, Mark; Sick, Theresa. Institut für Deutsche Sprache. Mannheim. 49–55.
- Jurkiewicz-Rohrbacher, Edyta; Hansen, Björn; Kolaković, Zrinka. 2017b. Clitic climbing, finiteness and the raising-control distinction. A corpus-based study. *Jazykovedný časopis* 68/2. 179–190.
- Kapović, Mate. 2022. Standard, kvazistandard i nestandard: tragikomedija zabunā. *Standardni i nestandardni idiomi (zbornik sažetaka)*. Ur. Glušac, Maja; Mikić Čolić, Ana. Filozofski fakultet Osijek. Osijek. 23–24.
- Kilgarriff, Adam; Grefenstette, Gregory. 2003. Introduction to the special issue to the Web as Corpus. *Computational Linguistics* 29/3. 333–347. <https://doi.org/10.1162/089120103322711569>.
- Kilgarriff, Adam; Baisa, Vít; Bušta, Jan; Jakubíček, Miloš; Kovář, Vojtěch; Michelfeit, Jan; Rychlý, Pavel; Suchomel, Vít. 2014. The Sketch Engine: Ten Years on. *Lexicography* 1/1. 7–36.
- Klajn, Ivan. 2004. *Rečnik jezičkih nedoumica*. Srpska školska knjiga. Beograd.

- Klubička, Filip; Ljubešić, Nikola. 2014. Using crowdsourcing in building a morphosyntactically annotated and lemmatized silver standard corpus of Croatian. *Language Technologies: Proceedings of the 17th International Multiconference Information Society – IS 2014. Volume G*. Ur. Erjavec, Tomaž; Žganec Gros, Jerneja. Institut Jožef Stefan. Ljubljana. 62–68.
- Kolaković, Zrinka. 2018. *Dvoaspektni glagoli – razlike između (p)opisa u priručnicama i stanja u korpusu s posebnim osvrtom na uporabu izvornih govornika*. Doktorski rad. Sveučilište u Regensburgu/Sveučilište u Zagrebu. Regensburg/Zagreb.
- Kolaković, Zrinka; Jurkiewicz-Rohrbacher, Edyta; Hansen, Björn. 2019. Clitic Climbing, the Raising-Control Dichotomy and Diaphasic Variation in Croatian. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 45/2. 505–522. <https://doi.org/10.31724/rihjj.45.2.13>
- Kolaković, Zrinka; Jurkiewicz-Rohrbacher, Edyta; Hansen, Björn; Filipović Đurđević, Dušica; Fritz, Nataša. 2022. *Clitics in the wild. Empirical studies on the microvariation of the pronominal, reflexive and verbal clitics in Bosnian, Croatian and Serbian*. Berlin. Language Science Press.
- Kolaković, Zrinka; Horvat, Marijana; Perić Gavrančić, Sanja. 2022. Jezično planiranje i sudbina znanaglasnice *ju* pod retrospektivnim svjetlom dijakronijskih podataka. *Međunarodni znanstveni skup Hrvatske dopreporodne gramatike u europskoj kontekstu*. https://retrogram.jezik.hr/wp-content/uploads/2022/11/RETROGRAM_knjizica-sazetaka.pdf. Posljednji pristup 11.4.2023.
- Kolaković, Zrinka; Blagus Bartolec, Goranka. 2022. Pronominal accusative clitics *ju* and *je* in 20th and 21st century Croatian. *14. Deutscher Slavistiktag*. https://slavtag2022.blogs.ruhr-uni-bochum.de/wp-content/uploads/2022/09/Abstractband_2.pdf. Posljednji pristup 11.4.2023.
- Kordić, Snježana. 2006. Upute jezičnim cenzorima. *Književna republika* 4/5–6. 194–208.
- Kryžan-Stanojević, Barbara. 2010. Norma i uzus u novim udžbenicima stranoga jezika. *Lahor* 2/10. 233–246.
- Langston, Keith; Peti-Stantić, Anita. 2014. *Language planning and national identity in Croatia*. Palgrave Macmillan. New York.
- Lisac, Josip. 2003. *Hrvatska dijalektologija 1. Hrvatski dijalekti i govori štokavskoga narječja i hrvatski govori torlačkoga narječja*. Golden marketing – Tehnička knjiga. Zagreb.
- Ljubešić, Nikola; Klubička, Filip. 2014. {bs,hr,sr}WaC — Web corpora of Bosnian, Croatian and Serbian. *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*. Ur. Bildhauer, Felix; Schäfer, Roland. Association for Computational Linguistics. Gothenburg. 29–35. <https://doi.org/10.3115/v1/W14-04>
- Ljubešić, Nikola; Klubička, Filip; Agić, Željko; Jazbec, Ivo-Pavao. 2016. New Inflectional Lexicons and Training Corpora for Improved Morphosyntactic Annotation of Croatian and Serbian. *Proceedings of the tenth international conference on language resources and evaluation (LREC' 2016)*. Ur. Calzolari, Nicoletta; Choukri, Khalid; Declerck, Thierry; Goggi, Sara; Grobelnik, Marko; Maegaard, Bente; Mariani, Joseph; Mazo, Helene; Moreno, Asuncion; Odijk, Jan; Piperidis, Stelios. ELRA. Paris. 4264–4270.
- Machálek, Tomáš. 2014. KonText – Corpus Query Interface. Filozofická fakulta Univerzity Karlovy. Praha. <http://kontext.korpus.cz/>. Posljednji pristup 18.2.2022.
- Machálek, Tomáš. 2020. KonText: Advanced and Flexible Corpus Query Interface. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*. Ur. Calzolari, Nicoletta; Béchet, Frédéric; Blache, Philippe; Choukri, Khalid; Cieri, Christopher; De-

- clerck, Thierry; Goggi, Sara; Isahara, Hitoshi; Maegaard, Bente; Mariani, Joseph; Mazo, Hélène; Moreno, Asuncion; Odijk, Jan; Piperidis, Stelios. ELRA. Marseille. 7005–7010.
- Mamić, Mile. 1995. Upotreba akuzativnih oblika zamjenice *ona*. *Jezik* 43/5. 183–188.
- Mamić, Mile. 1997. *Jezični savjeti*. Vlastita naklada. Zadar.
- Maretić, Tomislav. 1899. *Gramatika i stilistika hrvatskoga ili srpskoga književnog jezika*. Hartman. Zagreb.
- Maretić, Tomislav. 1924. *Hrvatski ili srpski jezični savjetnik za sve one, koji žele dobro govoriti i pisati našim jezikom*. Nadbiskupska štamparija. Zagreb.
- Matasović, Ranko. 2008. *Poredbenopovijesna gramatika hrvatskoga jezika*. Matica hrvatska. Zagreb.
- Matković, Maja. 2006. *Jezični savjetnik – iz prakse za praksu*. Škorpion. Zagreb.
- Mićanović, Krešimir. 2006. *Hrvatski s naglaskom: standard i jezični varijeteti*. Disput. Zagreb.
- Milićević, Jasmina. 2007. Co-occurrence of Serbian second-position clitics: syntactic and morphological constraints. *Morphologie à toulouse. Actes du colloque international de morphologie 4emes Décembrettes*. Ur. Hathout, Nabil; Montermini, Fabio. Lincom Europa. München. 99–120.
- Milićević, Jasmina. 2009. Linear placement of Serbian clitics in a syntactic dependency framework. *Dependency in Linguistic Description*. Ur. Polguère, Alain; Mel'čuk, Igor. Benjamins. Amsterdam. 235–276.
- Milićević, Jasmina. 2023. *Serbian Clitics*. Benjamins. Amsterdam. <https://doi.org/10.1075/slcs.229>.
- Milroy, James; Milroy, Lesley. 1985. *Authority in language: Investigating language prescription and standardisation*. Routledge & Kegan Paul. London.
- Mišeska-Tomić, Olga. 1996. The Balkan Slavic Nominal Clitics. *Approaching second: second position clitics and related phenomena*, Ur. Halpern, Aaron; Zwicky, Arnold M. Stanford University Press. Stanford. 511–535.
- Mladenović, Radivoje. 2010. *Zamenice u govorima jugozapadnog dela Kosova i Metohije*. Beograd: Institut za srpski jezik SANU.
- Mrazović, Pavica; Vukadinović, Zora. 2009. *Gramatika srpskog jezika za strance*. Izdavačka knjižarnica Zorana Stojanovića. Sremski Karlovci.
- Mulić, Malik. 1964. Da zaista ne bude zbrke oko enklitike. *Jezik* 12/1. 23–27.
- Neuman, W. Lawrence. 2007. *Basics of social research: Qualitative and quantitative approaches*. Pearson. Boston.
- Okuka, Miloš. 2008. *Srpski dijalekti*. SKD Prosvjeta. Zagreb.
- Opačić, Nives. 2006. Što je standardni jezik? *Vijenac* 14/315.
- Pasini, Dinka. 2018. Supstandardni varijeteti hrvatskoga jezika u 20. stoljeću. *Povijest hrvatskoga jezika. Knj. 5., 20. stoljeće: prvi dio*. Ur. Bičanić, Ante; Samardžija, Marko; Pranjković, Ivo. Društvo za promicanje hrvatske kulture i znanosti CROATICA. Zagreb. 580–629.
- Peco, Asim. 2007. *Izabrana djela I–VI. Knjiga I: govori istočne i centralne Hercegovine*. Bosansko filološko društvo. Sarajevo.
- Peti-Stantić, Anita. 2002. Stilistika klitika u tumačenjima hrvatskih gramatičara. *Važno je imati stila*. Ur. Bagić, Krešimir. Disput. Zagreb. 165–179.
- Peti-Stantić, Anita. 2007. Wackernagelovo pravilo – norma ili mogućnost. *Sintaktičke kategorije*. Ur. Kuna, Branko. Institut za hrvatski jezik i jezikoslovlje. Zagreb. 173–187.

- Pešikan, Mitar. 1959. O mestu enklitike u rečenici. *Nаш jezik* 9/7–10, 305–311.
- Piper, Predrag; Klajn, Ivan. 2014. *Normativna gramatika srpskog jezika*. Matica srpska. Novi Sad.
- Polančec, Jurica; Mihaljević, Ana. 2016. Croatian non-standard invariant relativizers in European perspective. *Jezikoslovlje* 17/1–2. 429–451.
- Popović, Ljubomir. 1979. Antizipative und postverbale Stellung der Enklitika im Serbokroatischen. *Zeitschrift für Slawistik* 24/JG. 678–695. <https://doi.org/10.1524/slav.1979.24.jg.678>.
- Popović, Ljubomir. 2004. *Red reči u rečenici*. Društvo za srpski jezik i književnost Srbije. Beograd.
- Progovac, Ljiljana. 1996. Clitics in Serbian/Croatian: comp as the second position. *Approaching second: second position clitics and related phenomena*, Ur. Halpern, Aaron; Zwicky, Arnold M. Stanford University Press. Stanford. 411–428.
- Radanović-Kocić, Vesna. 1988. *The grammar of Serbo-Croatian clitics: A synchronic and diachronic perspective*. PhD Thesis. University of Illinois at Urbana-Champaign.
- Radanović-Kocić, Vesna. 1996. The placement of Serbo-Croatian clitics: A prosodic Approach. *Approaching second: second position clitics and related phenomena*, Ur. Halpern, Aaron; Zwicky, Arnold M. Stanford University Press. Stanford. 429–445.
- R Core Team. 2016. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna. <https://www.R-project.org/>. Posljednji pristup 19.1.2022.
- Raguž, Dragutin. 1997. *Praktična hrvatska gramatika*. Medicinska naklada. Zagreb.
- Raguž, Dragutin. 2010. *Gramatika hrvatskoga jezika*. Vlastito izdanje. Zagreb.
- Reinkowski, Ljiljana. 2001. *Syntaktischer Wandel im Kroatischen am Beispiel der Enklitika*. Otto Sagner. München.
- Ridjanović, Midhat. 2012. *Bosnian for foreigners with a comprehensive grammar*. Rabic Publishing Company. Sarajevo.
- Rychlý, Pavel. 2007. Manatee/Bonito – A Modular Corpus Manager. *First Workshop on Recent Advances in Slavonic Natural Languages Processing*. RASLAN 2007. Ur. Sojka, Petr; Horák, Aleš. Masaryk University. Brno. 65–70.
- Santos, Diana. 1998. Providing access to language resources through the World Wide Web: the Oslo Corpus of Bosnian Texts. *Proceedings of The First International Conference on Language Resources and Evaluation*. ELRA. 475–481.
- Schütze, Carson T. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. *MIT working papers in linguistics. Papers on phonology and morphology. Volume 21*. Ur. Carnie, Andrew; Harley, Heidi; Bures, Tony. 373–473.
- Sekereš, Stjepan. 1977. Govor Hrvata u južnoj Baranji. *Hrvatski dijalektološki zbornik. Volume 4*. Ur. Finka, Božidar. Jugoslavenska akademija znanosti i umjetnosti. Zagreb. 323–484.
- Silić, Josip. 2006. *Funkcionalni stilovi hrvatskoga jezika*. Disput. Zagreb.
- Silić, Josip; Pranjković, Ivo. 2007. *Gramatika hrvatskoga jezika za gimnazije i visoka učilišta*. Školska knjiga. Zagreb.
- Sonnenhauser, Barbara; Meermann, Anastasia. 2016. Das Perfekt im Serbischen zwischen slavischer und balkanischer Entwicklung. *Beiträge zur Slavistik. XXII. JungslavistInnen-Treffen in München. 12.-14. September 2013*. Ur. Bazhutkina, Alena; Sonnenhauser, Barbara. Bibliion Media. Leipzig. 83–110.
- Stefanowitsch, Anatol. 2020. *Corpus linguistics (Textbooks in Language Sciences 7)*. Language Science Press. Berlin. DOI: 10.5281/zenodo.3735822.

- Stanojčić, Živojin; Popović, Ljubomir. 1992. *Gramatika srpskoga jezika. Udžbenik za I, II, III i IV razred srednje škole*. Zavod za udžbenike i nastavna sredstva. Beograd.
- Starčević, Anđel; Kapović, Mate; Sarić, Daliborka. 2019. *Jeziku je svejedno*. Sandorf. Zagreb.
- Stevanović, Mihailo. 1950. Đakovački govor. Srpski dijalektološki zbornik. *Volume 11*. Ur. Belić, Aleksandar. Naučna knjiga. Beograd. 1–152.
- Stevanović, Mihajlo. 1964. *Savremeni srpskohrvatski jezik. Gramatički sistemi i književnojezička norma. I*. Naučno delo. Beograd.
- Stevanović, Mihajlo. 1986. *Savremeni srpskohrvatski jezik. Gramatički sistemi i književnojezička norma. I*. Naučna knjiga. Beograd.
- Stjepanović, Sandra. 2004. Clitic Climbing and Restructuring with “Finite Clause” and Infinitive Complements. *Journal of Slavic Linguistics* 12/1. 173–212.
- Škarić, Ivo. 2009. *Hrvatski izgovor*. Nakladni zavod Globus. Zagreb.
- Tadić, Marko. 1998. Raspon, opseg i sastav korpusa suvremenoga hrvatskoga jezika. *Filologija* 30–31. 337–347.
- Tadić, Marko. 2002. Building the Croatian National Corpus. *Proceedings of the third international Conference on Language Resources and Evaluation (LREC'02)*. Knj. 2. Ur. González Rodríguez, Manuel; Suarez Araujo, Carmen Paz. ELRA. Las Palmas. 441–446.
- Tadić, Marko. 2009. New version of the Croatian National Corpus. *After Half a Century of Slavonic Natural Language Processing*. Ur. Hlaváčková, Dana; Horák, Aleš; Osolsobě, Klára; Rychlý, Pavel. Masaryk University. Brno. 199–205.
- Težak, Stjepko; Babić, Stjepan. 1994. *Gramatika hrvatskoga jezika – priručnik za osnovno jezično obrazovanje*. Školska knjiga. Zagreb.
- Težak, Stjepko; Babić, Stjepan. 2000. *Gramatika hrvatskoga jezika – priručnik za osnovno jezično obrazovanje*. Školska knjiga. Zagreb.
- Udier, Sanda Lucija; Gulešić Machata, Milvia. 2011. Kakvom jeziku poučavati neizvorne govornike? Viši i niži varijetet u poučavanju hrvatskoga kao inog jezika. *Nova Croatica* 5/5. 329–349.
- Utvić, Miloš. 2011. Annotating the corpus of contemporary Serbian. *INFOtheca: Journal of Informatics & Librarianship* 12/2. 36–47.
- Verber, Adolfo. 1873. *Slovnica hrvatska za srednja učilišta*. Dionička tiskara u Zagrebu. Zagreb.
- Vidović, Radovan. 1983. *Jezični savjeti*. Logos. Split.
- Vujančić, Milica; Gortan-Premk, Darinka; Dešić, Milorad; Dragičević, Rajna; Nikolić, Miroslav; Nogo, Ljiljana; Pavković, Vasa; Ramić, Nikola; Stijović, Rada; Radović-Tešić, Milica; Feckete, Egon. 2011. *Rečnik srpskoga jezika*. Matica srpska. Novi Sad.
- Vuković, Petar. 2016. Vijeće za normu i teorija upravljanja jezikom. *Suvremena lingvistika*. 42/82. 219–235.
- Zoričić, Ivan. 1998. *Hrvatski u praksi*. Zavičajna naklada „Žakan Juri“. Pula.
- Wilder, Chris; Čavar, Damir. 1994. Long Head Movement? Verb movement and cliticization in Croatian *Lingua* 93/1. 1–58. [https://doi.org/10.1016/0024-3841\(94\)90352-2](https://doi.org/10.1016/0024-3841(94)90352-2).
- Wonisch, Arno. 2012. *Das Pronominalsystem Des Bosnischen/Bosniakischen, Kroatischen Und Serbischen*. Lit Verlag. Wien.

ACCUSATIVE PRONOMINAL CLITICS *JU* AND *JE* IN BOSNIAN, CROATIAN AND SERBIAN: A CONTRASTIVE CORPUS STUDY OF STANDARD AND COLLOQUIAL VARIETIES

Empirical studies on variation in clitic use are almost entirely lacking. However, a careful contrastive analysis of both BCS grammar books and language advisor handbooks indicates differences in the usage of clitics not only between BCS standard varieties but also within their colloquial varieties. This paper deals specifically with variation in usage of the third-person feminine accusative CL *ju* and *je* 'her'. According to the previous literature, the usage of the CL variant *ju* is more limited in Bosnian and Serbian than in Croatian standard (cf. Silić & Pranjković 2007; Mrazović & Vukadinović 2009; Ridjanović 2012; Piper & Klajn 2014).

To test the range of this variation, we turn not only to traditionally compiled reference corpora with language material that should reflect standard BCS varieties (Santos 1998; Utvić 2011; Čermák & Rosen 2012; Brozović Rončević et al. 2018) but also to {bs,hr,sr}WaC, three massive web corpora (Ljubešić & Klubička 2014) available via a unified, functional interface NoSketchEngine, with morphosyntactic annotation based on the common tagset. We manually annotated 3.916 data points with CL variants *ju* and *je*. This study shows how language corpora can give us valuable insights into the range of variation not only in the BCS standard varieties but also in the real language use, i.e., their colloquial varieties through the frequency of forms and patterns which are partially discarded by the BCS normativists.

The generalized linear regression model revealed statistically significant differences in the usage of the CL *ju* between the standard Croatian variety on the one hand, and standard Bosnian and Serbian on the other. It also showed that BCS colloquial varieties differ significantly concerning the distribution of the CL variant *ju*. Furthermore, our data showed that the distribution of the CL variant *ju* is significantly different in standard and colloquial varieties of analyzed languages. Finally, based on our generalized linear regression model, we showed that the *-(j)u* ending of the CL host has a significant impact on the choice between the CL variants *ju* and *je*: it decreases the probability of the *ju* usage.

Keywords: Bosnian, Croatian, Serbian, standard and colloquial varieties, variation, corpora, third-person singular feminine clitic in the accusative

Adrese autorica:

Zrinka Kolaković

Alpen-Adria-Universität Klagenfurt
Institut für Slawistik
Universitätsstraße 65/67
9020 Klagenfurt
Österreich
Zrinka.Kolakovic@aau.at

Edyta Jurkiewicz-Rohrbacher

Universität Regensburg
Institut für Slavistik
Universitätsstraße 31
93053 Regensburg
Deutschland
Edyta.Jurkiewicz-Rohrbacher@ur.de

Jasmin Denise Gradischnig

Alpen-Adria-Universität Klagenfurt
Institut für Slawistik
Universitätsstraße 65/67
9020 Klagenfurt
Österreich
jasmindenigr@edu.aau.at