

Александар М. МИЛАНОВИЋ\*  
Филолошки факултет  
Универзитета у Београду

## ТЕОРИЈСКО-МЕТОДОЛОШКИ ПРОБЛЕМИ СТВАРАЊА КОРПУСА СЛАВЕНОСРПСКИХ ТЕКСТОВА

У раду се расправља о теоријско-методолошким поставкама стварања корпуса славеносрпских текстова, намењеног изради *Речника славеносрпског језика* у Матици српској. Истакнути су најделикатнији проблеми: разграничење текстова писаних славеносрпским језиком од оних писаних рускословенским или руским језиком са једне, као и српским народним језиком са друге; обезбеђивање стилске и жанровске репрезентативности корпуса; одређивање временског оквира ексцерпираних текстова; одређивање територије на којој су текстови настали; решавање техничких проблема у вези са фонтом, лематизацијом корпуса и сл.

**Кључне речи:** славеносрпски језик, рускословенски језик, српски народни језик, речник, корпус, периодизација, стил, жанр.

1. У изради дијакхронијских корпуса различитих српских језичких идиома србистика заостаје за другим националним славистикама. Постојећи дијакхронијски корпуси, који нису дигитализовани, прављени су за потребе дијакхронијских речника српског језика и још увек су отворени. У питању су заправо два корпуса. Први је корпус текстова који се прави за потребе *Речника славеносрпског језика*, једног од типова књижевних језика код Срба, који је у разним својим видовима функционисао од друге половине 18. до половине 19. века. Други је корпус текстова који се израђује за потребе *Речника српског језика 12–18. века*, који би обухватио све текстове писане „народним језиком” кроз његову историју.

Оба речника, као и оба корпуса, израђују се под окриљем Матице српске. Први пројекат, речника славеносрпског језика, води доц. др Исидора Бјелакловић, а други, дијакхронијског речника српског народног језика, академик Јасмина Грковић-Мејџор. Први пројекат, покренут давне 1981. и оновљен 2010. године, ове године је дао свој први конкретан резултат у виду *огледне*

---

\*aleksandar.jus@gmail.com

свеске (ОС 2017), у којем су дате и многе теоријско-методолошке напомене о формирању корпуса славеносрпских текстова, циљ овога прилога биће допуна неких питања покренутих на наведеном месту, као и покретање нових у вези са физиономијом, величином, обимом, разноврсношћу и употребом корпуса.

2. Како је *корпус* скуп текстова селектираних према одређеним лингвистичким критеријумима, у ОС 2017 потрудили смо се да врло детаљно опишемо све теоријско-методолошке потешкоће на које су наилазили наши претходници на пројекту, а које су сачекале и нас при изради мерила за селекцију текстова и при изради тренутног списка дела која смо ексцерпирани и која ваља ексцерпирати. Будући да су од краја 2012. електронска обрада корпуса и израда огледне свеске постављени као примарни и нераздвајни задаци, вратили смо се критички још једном свим начелима која су утемељила претходна уредништва.

2.1. Као *славеносрпски језик* одредили смо „оне књижевне идиоме друге половине 18. и прве половине 19. века који претходе вуковском, а настали су после рускословенског утицаја” (ОС 2017: 9), дакле и *доситејевски језик*, како је позну фазу развоја славеносрпског језика именовао у више својих радова Павле Ивић (уп. нпр. Ивић 1990, Ивић 1991). То значи да је читава скала прелазних типова од рускословенског до српског „народног” језика у писаној форми наведене епохе подведена под славеносрпски језик као флуидну форму која је еволуирала у смеру константног посрбљавања, нарочито своје граматичке, али и лексичке структуре. Управо наведена особина овог типа српског књижевног језика условила је идеју међу члановима новог пројектног тима, да би речник могао носити и наслов *Речник славеносрпске епохе*, што је сасвим дискретно – у фусноти – наведено и у ОС (2017: 10, ф. 3), уз напомену да би то „безболније преместило питање корпуса”, али да је постојећи назив задржан због традиције и административних потешкоћа које би преименовање проузроковало.

Поменута недоумица заслужује и додатне коментаре, за које није било места у ОС, а која се тичу дистинктивних црта славеносрпског језика у односу на коегзистентне идиоме. Наиме, још од времена круцијалних радова Ирене Грицкат, од пре више од пола века, више се инсистирало на дистинкцији између текстова писаних руским или рускословенским језиком с једне и славеносрпским језиком с друге стране, уз савет да првенствено треба ексцерпирати оригинална дела српских писаца, тј. дела „која не представљају преводе с руског и која нису теолошке природе” (Грицкат 1966: 66). Наведени цитат указује да су се у првобитним поставкама анализе текстова писаних славеносрпским језиком теоријско-методолошке замке препознавале пре свега у њиховој корелацији са текстовима српских писаца на руском и рускословенском, док је корелација са текстовима писаним српским народним језиком доживљавана као мање упитна.

Новија истраживања писаног језика у другој половини 18. и првој половини 19. века, а нарочито покретање пројекта *Речника српског језика 12–18. века*, ову теоријско-методолошку матрицу довели су у питање. Сматрамо, на-

име, да питање раздвајања славеносрпских текстова од рускословенских или руских теоријски није ништа мање компликовано од питања разграничавања са текстовима писаних језиком који се најчешће одређује као „српски народни”. Примера за различито одређивање једног истог српског текста, нарочито из друге половине 18. века, има посвуда у литератури, и није редак случај да се о његовој руској, рускословенској или славеносрпској природи споре чак и највећи научни ауторитети у овој области. Са друге стране, и поједина дела писана језиком ближим свакодневном, неформалном изразу српског становништва у наведеном периоду оцењивана су час као дела писана „народним”, час као дела писана славеносрпским језиком. Таква судбина није заобишла чак ни нека од најзначајнијих дела тадашње српске филологије, попут научних радова Саве Мркаља. То је уочио Божо Ћорић, те је покушао одредити којим типом језика су писани Мркаљеви филолошки радови *Сало дебелог јера либо азбукопротрес* и *Палинодија либо обрана дебелог њ* и закључио да је ова тема подразумевала „два истраживачка пута: а) анализу релевантних радова о Сави Мркаљу; б) анализу самих филолошких радова” (Ћорић 2012: 211).

За нашу тему битно је раздвојити два дела Ћорићевог закључка. У првом констатује: „Први истраживачки пут није дао очекивани резултат: прецизног одговора на постављено питање није било.” У наставку текста, пак, пише: „Други истраживачки пут био је плодотворнији: утврдили смо да је Мркаљ своје филолошке радове написао *славеносрпским језиком* (или *србизираним славеносрпским језиком*, по терминологији Александра Младеновића” (Ћорић 2012: 211).<sup>1</sup> Први део закључка указује да некритичко ослањање на постојећу литературу може бити веома опасно: у њој је одређење типа језика којим је поједини текст писан понекад магловито, недовољно прецизно, само описно или – у најгорем случају – погрешно. Други део Ћорићевог закључка упућује на једини исправан методолошки пут: да би једно дело ушло у корпус славеносрпских текстова (а не у корпус текстова писаних руским, рускословенским или српским народним језиком), претходно мора бити изведена ваљана лингвистичка анализа његовог текста.

Потпуно идентичан теоријско-методолошки проблем разграничења славеносрпског (тј. доситејевског) и народног језика, као при анализи Мркаљевог језика, имамо и при анализи писаног израза великог броја српских аутора с почетка 19. века, било да су у питању књижевници, или публицисти. Као ваљан пример могли би послужити писани опуси Јована Дошеновића, Павла Соларића, Димитрија Давидовића и др. Мислимо ту, уосталом, и на писани израз самог Вука Стефановића Карацића у почетној фази стваралаштва, током које је имао толерантан став према лексичким славенизмима па их стога и сам често активирао и мимо термилошке службе (Милановић 2015).

Овакво разграничавање славеносрпског и српског народног језика потенцијално може оштетити *Речник славеносрпског језика* управо у лексичком фонду који је и најзанимљивији, а то су *славенизми*. Могуће је, наиме, да

<sup>1</sup> По терминологији Павла Ивића овај идиом би био именован као *доситејевски*.

ће поједини облици славенизама, нарочито они хибридни (тј. славеносрбизми). Исто важи и за богат фонд *индивидуалних неологизама*. Вратимо се за тренутак Мркаљевог језику: ако би његов језик био проглашен народним, кованице *азбукопротрес* и *језикопротрес* нашле би се ван корица *Речника славеносрпског језика*.<sup>2</sup>

3. Уколико се вратимо кључним проблемима, при изради корпуса текстова за *Речник славеносрпског језика* ваљало је, ипак, пре свега разграничити текстове писане овим језиком од оних који су писани рускословенским или руским књижевним језиком, који су у српској средини настајали напоредо, неретко са минималним дистинктивним језичким цртама у односу на славеносрпске текстове. Упитно је, наравно, и да ли су сами аутори при писању текстова били свесни ових филолошких проблема којима ми данас придајемо толико пажње, а о *језичкој свести* тадашњих аутора, нажалост, има јако мало поузданих сведочанстава.

Иако се традиционална статистичка метода Александра Младеновића, коју су прихватили и његови бројни ученици и следбеници (Јован Јерковић, Јован Кашић, Александар Албијанић/Албин, Питер Херити, а у извесној мери и Љиљана Суботић и Јухани/Јуси Нуорлуото) показала као доста поуздана за разграничење ова три типа књижевних језика, првобитни пројектни тим овога речника (управо са академиком Младеновићем на челу) ексцерпирao је својевремено за речник и нека дела која језички не припадају славеносрпском корпусу, на шта смо скренули пажњу у ОС.

Наравно, ваљало је сада узети у обзир и сва теоријско-методолошка упозорења на недостатке Младеновићевог метода упитника, на које је пажњу скретала прво Ана Кречмер (1990), а потом, између осталих, и Александар Милановић и Исидора Бјелаковић (Милановић/Бјелаковић 2016). Основна замерка односила се на занемаривање стила којим је текст писан и жанра којем припада при статистичкој обради селектираних црта из упитника. У свим методолошким напоменама у вези са корпусом за речник, међутим, Младеновић (1982: 137) је истицао нужност његове стилске и жанровске раслојености, што је било у складу са исправним саветима већег броја лингвиста (И. Грицкат, В. Михајловића и др.) да акценат у славеносрпској епохи ваља ставити на текстове световног садржаја и свакодневне намене: календаре, алманахе, часописе, дневне новине и сл. То се у одређеној мери косило са појединим Кашићевим напоменама (Кашић 1984: 92–94), који је приоритет при формирању корпуса давао ексцерпирању књижевних дела, што је наишло на критику новог пројектног тима, који наглашава неоправданост таквог априорног закључивања, као и чињеницу да је овај лингвиста потпуно превидео значај дневних новина и часописа (ОС 2017: 12).

4. Подједнако велики проблем за формирање репрезентативног корпуса представља и временско омеђење славеносрпске епохе, које је драстично различито у постојећој литератури – нарочито од Ивићевог увођења терми-

<sup>2</sup> На значај индивидуалних неологизама за *Речник славеносрпског језика* скретао је пажњу и Јован Кашић (1984: 93), што је посебно назначено и у ОС 2017: 13.

на *доситејевски књижевни језик*. Са овим проблемом први су се суочили историчари језика који су покушавали створити периодизације – пре свих Борис Унбегаун (1995), а потом и Никита Иљич Толстој (2004). И док је код Унбегауна (1995: 21) славеносрпска епоха доста мутно одређена јер обухвата, кроз формулацију о конкурисању рускословенском језику, и други период (од 1740. до 1780), али углавном трећи период (од 1780. до почетка 19. века), Толстојева периодизација је прецизнија јер под „славеносрпским добом” подразумева период 1760–1780, док период 1780–1818. одређује као период „конкуренције норми”.

Првобитни временски оквир који је формиран у Младеновићевој фази рада на речнику обухватио је временски период од сто година, тј. 1750–1850. У њему се, очито, тежило „округлим” годинама, али је потом ширен списак дела која треба ексцерпирати, а са њим је истовремено прошириван и временски оквир, који је потом обухватио период 1748–1899 (ОС 2017: 11).

Нови пројектни тим тренутно стоји на позицији у којој је као почетна година издања дела која су ушла у корпус узета 1748, а као крајња 1850 ОС 2017: 17), иако је потпуно свестан чињенице да има текстова који су и после овог периода задржали лексичке особине доситејевског облика славеносрпског језика:

Обухваћена је продукција од 1748. до 1850. године, при чему у најмлађе изворе спада периодика. Будући да се репрезентативност речника може остварити једино заступљеношћу свих функционалних стилова (или боље – жанрова), постојећи корпус би требало допунити новим изворима, као што су новине, пословно-правна документа, приватна преписка и др. (ОС 2017: 17).

Овакав временски захват, барем у почеткој тачки временског оквира, у доброј мери одговара периодизацији коју установио Никита И. Толстој, модификујући старију периодизацију Бориса Унбегауна.

5. За одређивање репрезентативног корпуса славеносрпских текстова поред времена њиховог настанка дуго се као релевантан критеријум наводио и простор са којег је аутор, при чему се по правилу истицала екавска територија шумадијско-војвођанског дијалекта (уп. нпр. Младеновић 1989: 136 или Младеновић 2008: 195–196).

Лингвистички анализирајући текстове административно-правне и медицинске намене из Црне Горе и Босне, Исидора Бјелаковић (2017а, Бјелаковић 2017б) је у најновијим радовима утврдила све репрезентативне одлике славеносрпских текстова и на ијекавској територији, сасвим удаљеној од оне која је по правилу истицана као епицентар стварања хибридног (мешовитог) језика, што су у најужем и најбуквалнијем смислу били Сремски Карловци.

Будући да и друга, још непубликована истраживања текстова са тла Црне Горе, Босне и Херцеговине, као и Баније и Кордуна, упућују на чињеницу да се и на овим просторима чињенично може бранити став о егзистенцији славеносрпског језика, то значи да се при формирању коначног корпуса славеносрпских текстова за речник мора водити рачуна и о проширењу територије на којој су текстови штампани или написани.

6. Проблем штампе и рукописа даље нас води до новог проблема у потрази за репрезентативним корпусом, а то је проблем писма. Дешавало се у прошлости да се поједини текстови а priori оцене као текстови писани рускословенским, а не славеносрпским језиком само зато што су штампани црквеном руском ћирилицом, као што је био случај са првим српским дневним новинама браће Пуљо.<sup>3</sup> Анализе језика тих новина, међутим, показале су да је у питању славеносрпски језик (Ивић 1998, Милановић 2013), чак србизованији од језика у неким од потоњих новина које су штампане грађанском ћирилицом (Бјелаковић 2012).

Иако се избор црквене ћирилице заиста махом функционално везивао за специфичну тематику (теолошке природе), а сходно са тим и за рускословенски тип књижевног језика, бројни изузеци терају на опрез, на шта нас посебно подсећа чак и Стеријина збирка *Даворје*. Када су у питању рукописи, нарочито нпр. они везани за црквену администрацију, овај опрез мора бити још већи.

7. Појава коегзистенције грађанске и црквене ћирилице у славеносрпској епохи, иако су биле углавном функционално диференциране, условила је и решавање проблема одговарајућег фонта за израду корпуса славеносрпских текстова. За ОС 2017 стварање корпуса је подразумевало уношење грађе у фонту Times New Roman, док ће корпус за сам речник, који ће бити неупоредиво обимнији и графематски разноврснији, подразумевати нови фонт, креиран посебно за сврху двају дијахронијских речника покренутих у Матици српској.

Нажалост, због ограничених материјалних средстава, а самим тим и људских ресурса, корпус није дигитализован у правом смислу, тј. није извршена лематизација, као ни означавање морфосинтаксичких категорија и сл. У питању су проблеми који нарастају када се у обзир узме да су многе појаве у славеносрпском језику управо на међи лексике и морфосинтаксе, попут облика руске и рускословенске компарације или облика партиципа. Ови морфосинтаксички облици су с временом лексикализовани у доситејевској фази развоја славеносрпског језика, те захтевају специфичан граматички и лексикографски третман, који се мора одредити на основу контекста који чини *текст*, његова тематика, жанр којем припада и стил којим је писан. Дobar пример за наведене теоријско-методолошке потешкоће при потенцијалној дигитализацији могла би бити компарација придева у титулацији у оквиру приватне преписке, а такви облици придева почетком 19. века несумњиво су били лексикализована категорија, о чему сведочанства налазимо чак и у познијим граматицама, попут Суботићеве *Српске граматике* из 1847. године.

7. У раду смо се задржали на теоријско-методолошким и конкретним проблемима у вези са стварањем корпуса славеносрпских текстова, насталог за потребе *Речника славеносрпског језика*. Далеко су већи проблеми везани за корпус неопходан при будућој изради историјског речника српског народног

<sup>3</sup> Наведени пропуст направили су, додуше, историчари друштва, а не историчари књижевног језика.

језика, пројекта на чијем се челу налази академик Јасмина Грковић Мејџор, а који се односе, између осталог, и на разграничење српске и несрпске писане баштине, а потом и дефинисања текста који је писан „народним” језиком, који кроз историју – као што знамо – готово ни у једном тексту није био чист.

Истакли смо и најделикатније проблеме, а то су: а) разграничење текстова писаних славеносрпским језиком од оних писаних рускословенским или руским језиком са једне, као и српским народним језиком са друге; б) обезбеђивање стилске и жанровске репрезентативности корпуса; в) одређивање временског оквира ексцерпираних текстова; г) одређивање територије на којој су текстови настали; д) решавање техничких проблема у вези са фонтом, лематизацијом корпуса и сл.

## ЛИТЕРАТУРА

- Бјелаковић 2012:** Исидора Бјелаковић, „Језик новина код Срба крајем XVIII и почетком XIX века”, у: *An den Anfängen der serbischen Philologie. На почецима српске филологије. Salo debeloga jera libo azbukoprotres von Sava Mrkalj (1810–2010). Сало дебелого јера јера либо азбукопротрес Саве Мркаља. Philologica Slavica Vindobonensia, Band 1, Frankfurt am Main – Bern – Bruxelles – New York – Oxford – Wien: Peter Lang, 49–80.*
- Бјелаковић 2017а:** Исидора Бјелаковић, „Језик докумената Архива манастира Савине”, у: *Српско језичко и књижевно наслеђе на простору данашње Црне Горе / Српски језик и књижевност данас*, Подгорица [у штампи].
- Бјелаковић 2017б:** Исидора Бјелаковић, „Медицинска терминологија у трима лекарушама из Босне (18. и 19. век)”, у: *Српско писано наслеђе и историја средњовековне Босне и Хума*, Вишеград [у штампи].
- Грицкат 1966:** Ирена Грицкат, „У чему је значај и какве су специфичности славеносрпског периода у развоју српскохрватског језика (Поводом иницијативе Матице српске да се састави Речник књижевног језика славеносрпског периода)”, *Зборник Матице српске за филологију и лингвистику*, IX, 61–66.
- Ивић 1990:** Павле Ивић, „Доситејевски књижевни језик између славеносрпског и вуковског”, *Научни састанак слависта у Вукове дане*, 19/2, 5–14.
- Ивић 1991:** Павле Ивић, *Из историје српскохрватског језика*, Изабрани огледи II, Ниш: Просвета.
- Ивић 1998:** Павле Ивић, *Преглед историје српског језика*, Сремски Карловци – Нови Сад: Издавачка књижарница Зорана Стојановића.
- Кречмер 1990:** Ана Кречмер, „Неколико напомена поводом 30-годишњег јубилеја истраживања славеносрпског доба”, *Зборник Матице српске за филологију и лингвистику*, XXXIII, 221–231.

- Милановић 2013:** Александар Милановић, *Језик весма полезан*, Београд: Друштво за српски језик и књижевност Србије.
- Милановић 2015:** Александар Милановић, „Рана Вукова језичка свест и статус славенизама (1814–1818), у: *Вук Стефановић Караџић (1787–1864–2014)* (уредник Нада Милошевић-Ђорђевић), Београд: САНУ, 145–157.
- Милановић 2016:** Александар Милановић, „Филолошки текстови из прве половине 19. века као ресурс за истраживање српског језика”, *Научни састанак слависта у Вукове дане*, 45/3, 135–147.
- Милановић/Бјелаковић 2016:** Александар Милановић и Исидора Бјелаковић, „Толстојева периодизација и питање жанра”, у: *Slavische Geisteskultur: Ethnolinguistische und philologische Forschungen. Teil 2 / Славянскія духовная культура: етнолингвистические и филологические исследования. Часть 2* (ред. А. А. Алексеев, Н. П. Антропов, А. Г. Кречмер, Ф. Б. Поляков, С. М. Толстой), Frankfurt am Main – Bern – Bruxelles – New York – Oxford – Warszawa – Wien: Peter Lang, 197–217.
- Младеновић 1982:** Александар Младеновић, „Напомене о раду на речнику славеносрпског типа књижевног језика”, у: *Лексикографија и лексикологија. Зборник радова*, Београд – Нови Сад: САНУ – Институт за јужнословенске језике Филозофског факултета у Новом Саду – Матица српска – Филолошки факултет у Београду, 137–140.
- Младеновић 1989:** Александар Младеновић, *Славеносрпски језик*, Нови Сад – Горњи Милановац: Књижевна заједница Новог Сада – Дечје новине.
- Младеновић 2008:** Александар Младеновић, *Историја српског језика : одабрани радови*, Београд: Чигоја штампа.
- ОС 2017:** *Речник славеносрпског језика : огледна свеска* (приредили Исидора Бјелаковић, Ирена Цветковић Теофиловић и Александар Милановић), Нови Сад: Матица српска.
- Толстој 2004:** Никита И. Толстој, *Студије и чланци из историје српског књижевног језика*, Београд – Нови Сад: Завод за уџбенике и наставна средства – Вукова задужбина – Матица српска.
- Ђорић 2012:** Божо Ђорић, „О језику Мркаљевих филолошких радова”, у: *An den Anfängen der serbischen Philologie. На почецима српске филологије. Salo debeloga jera libo azbukoprotres von Sava Mrkalj (1810–2010). Сало дебелого јера јера либо азбукопротрес Саве Мркаља. Philologica Slavica Vindobonensia*, Band 1, Peter Lang, 189–215.
- Унбегаун 1995:** Борис Унбегаун, *Почеци књижевног језика код Срба*, Београд – Нови Сад: Вукова задужбина – Матица српска – Орфелин.



Aleksandar. M. Milanović

CORPUS OF SLAVONIC-SERBIAN TEXTS; THEORETICAL AND METHODOLOGICAL  
ISSUES

Summary

The paper discusses the theoretical and methodological tenets related to the task of collecting the corpus of Slavonic-Serbian texts for the purpose of compiling *Rečnik slavenosrpskog jezika (Dictionary of Slavonic-Serbian)* by Matica srpska. The paper highlights some of the burning issues such as: a) distinguishing texts written in Slavonic -Serbian from those written in Russian or Russian (Church) Slavonic on the one hand, and from the those written in Serbian language spoken by the people, on the other; b) ensuring that the texts included in the corpus be representative in style and genre; c) identifying the collected texts according to when and where they were written; d) technical matters such as the usage of fonts, corpus lemmatization, etc.