

МЕТОДОЛОШКИ ОКВИР ЗА ИЗРАДУ КОРПУСА ПРАВНИХ ТЕКСТОВА 19. ВЕКА*

Предмет рада су нека методолошка разматрања везана за дизајн корпуса правних текстова на српском језику из 19. века. Уз уважавање репрезентативности и балансираности, издвојени су следећи критеријуми на основу којих корпус треба да буде формиран: (а) језички тип (унутрашњи/језички критеријум); (б) хронолошки критеријум; (в) функционалностилски, односно жанровски критеријум; (г) правна продукција (као спољашњи критеријум). На основу језичког критеријума корпус се дели на два поткорпуса: правни текстови из предстандардног периода и правни текстови из периода стабилизације савременог српског језика. Као секундарни критеријум узет је период од десет година, а као граничници узети су догађаји из политичке и културне историје Кнежевине, односно Краљевине Србије.

Кључне речи: историја језика, 19. век, правни стил, методологија формирања корпуса.

The focus of this paper lies in the methodological challenges connected to the design and creation of a corpus of Serbian 19th century legal texts. While following the principles of representativeness and balance, we highlight the following criteria that are employed in the creation of the corpus: (a) type of language (internal/linguistic criterion); (b) chronological criterion; (c) functional and stylistic criterion (genre criterion); (d) legal production (as a non-linguistic criterion). Based on the linguistic criterion, the corpus is divided into two sub-corpora (strata): legal texts from the pre-normative period and legal texts from the period of standardization of the modern Serbian language. As a secondary criterion, 10-year periods were taken into account, the limiting factors being the historical and political events that took place in the Principality / Kingdom of Serbia.

Key words: historical linguistics, 19th century, legal register, methodology of corpus design.

1. Увод

Циљ рада је разматрање основних методолошких питања везаних за процес креирања корпуса правних текстова на српском језику у 19. веку насталих на тлу Кнежевине, односно Краљевине Србије. Фокус рада је на критеријумима који ће се користити у процесу израде једног специјализованог статичног историјског корпуса¹, који ће превасходно бити конструисан да би се на њему

* Рад је рађен у оквиру пројекта *Динамика структура савременог српског језика*, ОИ 178014, који финансира Министарство просвете, науке и технолошког развоја Републике Србије. Рад је настао као припрема за израду корпуса текстова који ће ући у докторску дисертацију под насловом *Зависносложена реченица у административно-правном стилу српскога језика у XIX веку* (ментор проф. др Милош Ковачевић), а започет је за време студијског боравка аутора на Хумболтовом универзитету у Берлину у оквиру програма Еразмус +.

¹ О врстама корпуса в. McEnery, Hardie 2012: 3–21; Tognini-Bonelli, Sinclair 2005: 208–212; Utvić 2013: 24–31.

спровела синтаксичка и стилистичка истраживања². Пошто у раду разматрамо спољашње критеријуме везане за методологију формирања специјализованог историјског корпуса, циљ рада је и да се испита у којој мери методологија развијена за креирање корпуса савремених језика може бити примењена у дијахроној перспективи.

2. Теоријско-методолошка разматрања

Величина корпуса и његова структура првенствено зависе од намене самога корпуса и од истраживачког питања које усмерава његову креацију (Biber et al. 2006: 246; Koester 2010: 67–69, 71; Nelson 2010: 54; Xiao 2010: 148, 151). Из тог разлога у литератури се истиче да је први корак приликом формирања корпуса дефинисање истраживачког питања и дефинисање плана његове израде (Nelson 2010: 54; Reppen 2010: 31). Следећи важан корак у формирању сваког корпуса јесте дефинисање критеријума на којима ће бити заснован (Sinclair 2005: 8).

Два основна критеријума на које се обраћа пажња приликом израде корпуса јесу репрезентативност и балансираност (McEnergy, Hardie 2012: 8–9, 10–11; Nelson 2010: 56–60; Sinclair 2005: 6, 12–14; Utvić 2013: 17–18, 20–21; Xiao 2010: 149–153)³. Неки аутори овоме додају и друге карактеристике као што је могућност поређења (McEnergy, Hardie 2012: 10–11). Репрезентативност према Дагласу Бајберу (Douglas Biber) зависи пре свега од популације⁴ текстова који ће чинити корпус. Када се одређује популација текстова обухваћених корпусом, по мишљењу овога аутора, битно је обратити пажњу на две ствари: на *границе популације* (који ће текстови бити укључени у корпус, а који ће бити из њега искључени) и на *хијерархију у оквиру популације* (дефинисати (пот)катеорије текстова у популацији и одредити њихове везе) (Biber 1993: 243). То омогућава да се формира *оквир узорковања / оквир узорка* (енг. *sampling frame*)⁵, што подразумева прављење списка дела која ће ући у корпус (Biber 1993: 243–244; McEnergy, Hardie 2012: 8; Utvić 2013: 16–17; Xiao 2010: 151–152).

Даглас Бајбер разликује две врсте репрезентативности (Biber 1993: 243–244; Koester 2010: 69; Xiao 2010: 149–150): (а) *спољашњу (екстралингвистичку, ситуациону) репрезентативност*, која подразумева да су различити типови текстова детерминисани спољашњим (социолингвистичким) критеријумима; (б) *унутрашњу (унутарјезичку, интрајезичку) репрезентативност*, која подразумева да различите језичке карактеристике буду равномерно распоређене у ок-

² Ово не искључује могућност да, уз одређене модификације, корпус буде искоришћен и за друге врсте историјскојезичких истраживања (ортографска, морфолошка, истраживања језичке промене итд.).

³ Репрезентативност и балансираност треба схватити пре као идеале којима треба тежити, а не као критеријуме које објективно можемо измерити (Biber et al. 2006: 246; McEnergy, Hardie 2012: 10–11; Nelson 2010: 60; Utvić 2013: 23; Xiao 2010: 147).

⁴ У раду термин *популација* користимо са значењем које има у корпусној лингвистици, где означава тип језика из кога потиче узорак који чини корпус (Biber 1993: 243; McEnergy, Hardie 2012: 8; Utvić 2013: 16).

⁵ Превод термина према Utvić 2013: 16.

виру корпуса. Корпуси се креирају према спољашњим критеријумима⁶ (Koester 2010: 69; Sinclair 1991: 17–18; Sinclair 2005: 5; Xiao 2010: 149) пошто је, без претходног истраживања, тешко утврдити како су језичке карактеристике дистрибуиране на нивоу корпуса (Xiao 2010: 150). Код специјализованих корпуса, као што је корпус правних текстова, примарност спољашњег критеријума је истакнута већ у самом називу корпуса (Koester 2010; Sinclair 2005: 7, 15).

Балансираност подразумева пропорционалну заступљеност одређених типова текстова у корпусу. Потребно је да у корпусу текстови буду заступљени у оној мери у којој су заступљени и у језику (Utvić 2013: 20–21; Xiao 2010: 151). Супротна тежња балансираности јесте потреба да се у корпусу прикаже што више варијација које постоје у језику, а које се не могу наћи у најзаступљенијим жанровима (Biber 1993: 247–248; Biber et al. 2006: 247–248). Такође, понекад је утицај неког жанра важнији од његове учесталости (Biber 1993: 247–248; Biber et al. 2006: 247–248; Utvić 2013: 20–21).

Репрезентативност и балансираност подразумевају и одабир узорка који ће ући у корпус (Biber 1993: 244). Обично се у литератури издвајају два начина узорковања преко којих се може добити балансирани корпус: *прости случајни узорак* и *стратификовани случајни узорак* (Biber 1993: 244; Biber et al. 2006: 247; Utvić 2013: 22; Xiao 2010: 151–152). Случајни узорак подразумева да постоји подједнака вероватноћа да ће се сваки текст наћи у корпусу. Међутим, методом случајног узорка могу се добити узорци који на неадекватан начин репрезентују популацију у целини и узорци који не укључују ретке језичке јединице. Из тог разлога боље решење у процесу формирања корпуса је стратификовани узорак, при чему се цела популација дели на *слојеве* (*стратуме*) да би се у оквиру њих текстови одабрали методом случајног узорка (Biber 1993: 244; Biber et al. 2006: 247–248; Utvić 2013: 22; Xiao 2010: 153).

3. Критеријуми за изградњу корпуса правних текстова на српском језику у 19. веку

У нашем случају циљ израде корпуса јесте истраживање инхерентних, преваходно синтаксичких, карактеристика административно-правних текстова насталих на територији Кнежевине, односно Краљевине Србије у 19. веку. Секундарно истраживачко питање јесте праћење развоја административно-правног стила у току 19. века, те је приликом прављења плана израде корпуса важно узети у обзир и временску (развојну) димензију⁷. За креирање корпуса овог типа треба користити метод стратификованог узорка.

Ми ћемо се у раду осврнути на најважније спољашње критеријуме везане за формирање корпуса правних текстова на српском језику у 19. веку. Фокус ће бити на критеријума који треба да буду примењени за поделу корпуса на

⁶ Као примарни критеријум за изградњу корпуса Џон Синклер (John Sinclair) издваја комуникативну функцију текстова (Sinclair 2005: 5).

⁷ Развојна димензија карактерише се следећим општим тенденцијама: традиционализмом, интернационализацијом, пуризмом, поларизацијом, унификацијом, интелектуализацијом и специјализацијом средстава (Nehring 2009: 43).

поткорпусе (стратуме) у оквиру којих треба вршити узорковање. Као најширу радну дефиницију популације правних текстова⁸ у 19. веку узећемо правну продукцију насталу у оквиру рада државне администрације у Кнежевини, односно Краљевини Србији.

Најважнији критеријуми које треба узети у обзир приликом дизајна корпуса правних текстова на српском језику у 19. веку јесу:

- (а) језички тип (предстандардни језички типови – вуковски језички тип);
- (б) хронолошки критеријум (као оријентир су узети периоди од десет година);
- (в) функционалностилски, односно жанровски критеријум;
- (г) правна продукција (вањезички фактор).

Као додатни критеријум узима се значај документа⁹, па се у корпус уносе документи као што су устави (*Сретењски устав* из 1835, *Турски устав* из 1838, *Устав за Књажевство Србију* из 1869, *Устав за Краљевину Србију* из 1888, *Устав за Краљевину Србију* из 1901, *Грађански законик за Књажевство Србију* из 1847. итд.).

Узети су у разматрање текстови од *Сретењског*, односно *Турског устава* (1835, односно 1838) до *Устава за Краљевину Србију* из 1901. Три су разлога зашто је укључен и овај *Устав* иако је донет почетком 20. века: (а) ово је последњи устав донет за време владавине династије Обреновића; (б) израда овог устава почела је већ у 19. веку (Луковић 1994: 71); (в) природа правног језика је таква да он по правилу боље чува архаичније конструкције него што рефлектује иновације које се јављају у време настанка документа (Грковић-Мејџор 2007: 41–44; Visković 1989).

Важно је скренути пажњу да различити аутори узимају различите спољашње критеријуме за омеђавање корпуса правних текстова 19. и почетка 20. века. Тако Милош Луковић у својој књизи *Развитак српског правног језика: прилог историји језика и историји права* анализира текстове од 1868 (година званичне Вукове победе) до 1918 (стварање Краљевине Срба, Хрвата и Словенаца и престанак правне продукције у оквиру Србије као самосталне државе) (Луковић 1994: 73–76). Тодор Подгорац разматра језичке особине српских устава од 1835. до 1903 (Подгорац 1990). Потребно је нагласити да су спољашњи критеријуми за разграничење корпуса нужно арбитарни, док се језик развија у континуитету и у њему коегзистирају елементи настали у различитим етапама (Грковић-Мејџор 2007: 19–20, 28–30; Павловић 2013: 5–6).

⁸ Досадашња истраживања српског правног језика 19. века спорадична су и везују се углавном за рад Милоша Луковића (Луковић 1988/90, 1994, 2006, 2010). Срећу се и истраживања других аутора (на пример Јовић 2011; Марковић 1994; Невекловски 2008; Подгорац 1990). О статусу српског правног језика ван територије Кнежевине, односно Краљевине Србије в. на пример Луковић 2009; Окука 1990: 32–36; Окука 1996; Остојић 2000.

⁹ У специјализованим корпусима текстови се обично укључују не због њихове лингвистичке репрезентативности, већ због њиховог културно-историјског значаја: “It is the uniqueness of their language choices, and not the universality of them, that cause them to be collected” (Tognini-Bonelli, Sinclair 2005: 210).

3.1. Критеријум језичког типа

У току 19. века разликују се два важна периода у развоју српског књижевног језика, а као година разграничења узима се 1868 (Ивић 2014: 218–227; Луковић 1994: 27–42, 61–72; Луковић 2006: 1540–1545) као година званичног усвајања Вукове језичке и правописне реформе: (а) предстандардни период, у коме се јавља језичка диглосија и триглосија (паралелна употреба рускословенског, славеносрпског и српског народног језика); (б) период стабилизације језика вуковског типа (Ивић 2014: 159–305). Језички тип коришћен у правним текстовима насталим на тлу Кнежевине Србије пре 1868. оцењен је као „модернизовани *доситејевски*“ (Луковић 1994: 51, курзив у оригиналу), док се у текстовима после 1868. године користи „вуковски језички тип, али са специфичностима србијанске, првенствено београдске културне средине“ (Луковић 1994: 72). Ако се узме у обзир досад речено, одређивање језичког типа за текстове из предстандардног периода треба извршити након анализе језика сваког (типа/жанра) документа понаособ.

На основу критеријума језичког типа корпус треба бити подељен на два поткорпуса: (а) поткорпус текстова предстандардног периода (настали до 1868); (б) поткорпус текстова из периода стабилизације савременог стандардног српског језика (текстови настали после 1868). Ми оба периода сагледавамо у дијахроној вертикали као две фазе у континуираном развоју стандардног српског језика (Јерковић 1997; Младеновић 2008: 147–152). Док се нагли прекид на плану коришћених идиома може пратити на површинском графематско-ортографском плану (уп. на пример Ивић 2014: 159–255; Милановић 2014: 29–90; Суботић 1989: 14–97), као и на лексичком плану, који најбрже региструје ванјезичке промене (Драгићевић 2018; Ивић 2014: 44–52, 100–104, 150–154, 250–254, 256–259), слабији је утицај ванјезичких околности на морфолошки и синтаксички план¹⁰. Разлог за одабир синтаксичких и стилистичких особина јесте тај што се њиховим сагледавањем најбоље уочавају инхерентне одлике одређеног функционалног стила (у овом случају жанра правних текстова) (уп. Луковић 1994: 73). Поред тога, очекује се да су синтаксичке и стилистичке особине, с обзиром на високу формализованост административно-правног стила, стабилније у односу на графематско-ортографске и лексичке језичке црте, те да се на основу њих могу одредити инхерентне карактеристике језика права¹¹ у току целог 19. века без обзира на језички тип којем документ припада. Овакво полазиште подржавају досадашња истраживања конкретних синтаксичких феномена на функционалностилски разноврсном корпусу текстова из предстандардног периода, где је утврђено да је функционални стил, а не језички тип, пресудан за репартицију синтаксичких црта у специјалним стиловима (на пример в. Бјелаковић 2008). У оквиру књижевних идиома треба узети у обзир и „тенденције и варијабилности“ (Nehring 1999: 217) њихових (експлицитних или имплицитних) норми.

¹⁰ На синтаксичке промене утиче садејство унутарјезичких и ванјезичких фактора (Грковић-Мејџор 2007: 31–44). Утицај ванјезичких фактора на морфолошки и синтаксички план у 19. веку најбоље илуструје повлачење партиципа (Суботић 1984).

¹¹ Ово се не односи на типичне синтаксичке славенизме као што је употреба партиципа (в. Суботић 1984).

Са методолошке тачке гледишта важно је питање балансираности текстова из два периода (в. одељак 3.4).

3.2. Хронолошки критеријум

Код специјализованих корпуса постоји чвршћа веза између текстова и контекста у којима су настали (Koester 2010: 66–67). Љиљана Суботић истиче да се „[п]ериодизација тзв. *спољашње* историје језика обично [...] заснива на догађајима социополитичке историје језика“ (Суботић 2004: 166). С обзиром на релевантне догађаје у друштвеном, правном и културном животу Србије (в. Ивић 2014: 159–255, 256–305; Луковић 1994: 27–72; Суботић 2004), два по-торпуса делимо на следеће стратуме (референти период од десет година¹²):

(1) Предстандардни период:

(а) Први потпериод од 1835 (односно 1838) до 1858, који обухвата прву владавину Милоша Обреновића, прву владавину Михаила Обреновића и владавину Уставобранитеља и који делимо на два стратума:

(а1) Период од 1835/8. до 1847/8, при чему је граничник изабран јер је 1847. година незваничне победе Вукове језичке и правописне реформе и пошто те године Јован Хаџић, један од најважнијих кодификатора Кнежевине Србије, одлази из Србије за Нови Сад.

(а2) Период од 1847/8. до 1858, у коме се осећа највећи утицај Срба из Аустроугарске на настанак правних текстова (уп. Луковић 1994: 32–53). Као граничник је изабрана 1858. година, када долази до смене власти у Србији.

(б) Период од 1858. до 1868 (друга владавина Милоша Обреновића и друга владавина Михаила Обреновића), у коме слаби утицај Срба из Аустроугарске и јача утицај младе интелигенције школоване у иностранству, пре свега у Француској, и у којем се постепено припрема терен за прихватање Вукове реформе (уп. Луковић 1994: 53–59).

Година 1868. узета је као граничник из следећих разлога: (а) то је година престанка владавине Михаила Обреновића; (б) то је година званичног прихватања Вукове реформисане азбуке у Кнежевини Србији; (в) наредне 1869. године донет је нов *Намеснички устав* за Кнежевину Србију.

(2) Корпус који чине текстови из периода стабилизације савременог српског језика делимо на три стратума:

(а) Период од 1868/9. до 1878, који почиње доласком на власт Милана Обреновића, односно од доношења *Устава за Књажевство Србију*, и траје до стицања независности и проглашења Кнежевине Србије за Краљевину.

¹² Период од десет година узет је из практичних разлога да би се лакше пратила фреквенција конкретних језичких црта и да би могао да се стекне статистички увид у њихов развој. Могућа је, а у неким истраживањима и методолошки оправданија, алтернативна подела текстова на основу конкретних политичких прилика у Кнежевини, односно Краљевини Србији (на пример, текстови настали за време прве владавине Милоша Обреновића, текстови донети за време Уставобранитеља, текстови настали након стицања независности итд.).

(б) Период од 1879. до 1888. обухвата време од стицања независности до доношења *Устава за Краљевину Србију*.

(в) Период од 1888. до 1898. године (односно до 1901) обухвата време од доношења *Устава* до краја периода, с тим што је нагласак на 1898. години као години у којој је донето преко 700 правних прописа (Луковић 1994: 68).

С методолошке стране, важно је балансирати број текстова по периоду да би се могао пратити континуитет развоја правних текстова на српском језику (в. одељак 3.4).

Оваква подела корпуса на поткорпусе (стратуме) има предности и недостатака. Предност је то што се прегледно пружа увид у развој конструкција у периоду од десет година. Међутим, овај приступ има следеће недостатке: (1) пошто су „граничници“ издвојени на бази екстралингвистичких критеријума, они се могу узети само условно; (2) правни језик је архаичан, те садржи и конструкције из ранијих фаза језичког развоја; (3) разликују се моменат настанка текста и моменат штампања документа, те је питање шта узети за референту тачку приликом формирања корпуса.

У завршној фази формирања и публикавања корпуса потребно је образложити кораке учињене у његовом формирању пошто они касније имају утицаја на интерпретацију добијених резултата (Sinclair 2005: 11).

3.3. Функционалностилски (жанровски критеријум)

Жанр као спољашњи критеријум за формирање корпуса један је од најбитнијих у корпусној лингвистици (Biber 1993: 244–247; Utvić 2013: 18–20). Жанровски критеријум као важан параметар за дефинисање корпуса текстова средњовековне пословноправне писмености издвојио је Слободан Павловић (Павловић 2009: 6)¹³.

За све текстове који улазе у корпус треба одредити (под)жанр административно-правног стила коме припадају. У оквиру административно-правног стила издвајају се усмена и писмена варијанта, у оквиру које постоји и више подстилова (Тошовић 2002: 355–358). Одређивање типа текста с обзиром на жанр, односно поджанр, треба да буде урађено у складу са ванјезичким критеријумима (било с обзиром на експлицитни назив документа у наслову и садржају или на његов реконструисани назив). Главни критеријум који се користи је медијум (писмена – усмена варијанта правног језика). Следећи поткритеријум јесте конкретни (под)жанр. У оквиру писаних жанрова језгро корпуса чине уставни и текстови штампани у *Зборницима закона и уредба и уредбених указа за Кнежевину, односно Краљевину Србију*. Усмену варијанту чине *Стенографске белешке са седница Народне скупштине Краљевине Србије, с нагласком на Стенографским белешкама о седницама Народне скупштине Краљевине Србије за 1898. годину*. Корпус треба бити формиран тако да је могуће касније додавање и других (под)жанрова административно-правних текстова у 19. веку (као што су протоколи, записи са суђења, устаничка преписка итд.).

¹³ Поред жанровског Павловић користи и хронолошки и територијални критеријум (Павловић 2009: 6).

3.4. Критеријум правне продукције

Најважнији спољашњи критеријум важан за формирање корпуса правних текстова у 19. веку јесте сам процес правне продукције. Како истиче Милош Луковић (Луковић 1994: 68–72), значајно се разликује правна продукција у два издвојена периода. У предстандардном корпусу преовладава индивидуални стил аутора законског решења (Луковић 1994: 71). С друге стране, у периоду стабилизације већ се постепено устаљује норма језика вуковског типа (Ивић 2014: 256–305) и правна продукција прелази из руку појединаца (као што су Јован Хаџић, Јован Стерија Поповић, Димитрије Давидовић, Аврам Петронијевић и други) на организованији рад у оквиру државних институција (Луковић 1994: 71–72).

Узевши претходно речено у обзир, текстовима из првог и другог потпериода прилази се на различити начин. Текстови предстандардног периода узимају се у целини, док се из другог периода узимају само репрезентативни текстови у целини (устави, значајни закони), док се врши избор других текстова. Избор текстова врши се тако да у корпус улазе краћи текстови, а да се од дужих текстова узима почетак, средина и крај¹⁴.

Из свега реченог следи да се у анализи корпуса може очекивати: (а) већи утицај индивидуалног стила аутора у текстовима предстандардног периода; (б) равнoмернија дистрибуција језичких конструкција у текстовима другог периода.

4. Закључци

У раду смо се бавили методолошким проблемима везаним за процес формирања корпуса правних текстова на српском језику у 19. веку. Уважавајући критеријуме балансираности и репрезентативности, издвојили смо четири главна критеријума важна за изградњу овог специјализованог историјског корпуса: (а) језички тип (предстандардни идиоми – вуковски језички тип); (б) хронолошки критеријум (период од десет година); (в) функционалностилски, односно жанровски критеријум; (г) језичка продукција (као најважнији екстралингвистички критеријум). На основу језичког критеријума корпус треба поделити на два поткорпуса: поткорпус текстова предстандардног периода и поткорпус текстова стабилизације савременог српског књижевног језика. Као поткритеријум за даљу поделу узели смо период од десет година, а као граничнике узели смо значајне културне и политичке догађаје из историје Србије у 19. веку.

С методолошке тачке гледишта потребно је обратити пажњу на следеће аспекте дизајна корпуса правних текстова у 19. веку:

(1) Мањи број текстова је штампан у предстандардном периоду, те ови текстови треба да буду узети у целини.

¹⁴ Постоје два приступа у одабиру текстова за корпус. Први је приступ заснован на узорку, обично од 2000 речи (McENERY, HARDIE 2012: 9–10). Предност овог приступа је тај што се смањује утицај индивидуалног аутора или појединачног текста на укупне резултате претраге и анализе. Други приступ заступа познати корпусни лингвиста Џон Синклер, који истиче да документ треба донети у целости (Sinclair 1991: 19–20; Sinclair 2005: 11).

(2) Организованији рад на правној кодификацији у време Милана Обреновића утицао је на повећање правне продукције, те текстове из овог периода по правилу треба донети у изводу.

(3) Треба укључити што већи број различитих жанрова административно-правног стила тако да се обухвати и његова писмена и његова усмена варијанта.

(4) Треба пратити цео период са потпериодима (стратумима) да би се добила хронолошка слика развоја језичких феномена који се проучавају.

У будућим истраживањима требало би испитати и евентуални утицај страних језика и страних кодификација на језик правних прописа у Србији у 19. веку.

Цитирана литература

- Бјелаковић, Исидора. Партиципски пасив у књижевном језику код Срба у XIX веку. Нови Сад: Филозофски факултет, 2008.
- Грковић-Мејдор, Јасмина. Списи из историјске лингвистике. Сремски Карловци – Нови Сад: Издавачка књижарница Зорана Стојановића, 2007.
- Драгићевић, Рајна. Српска лексика у прошлости и данас. Нови Сад: Матица српска, 2018.
- Ивић, Павле. Преглед историје српског језика. Сремски Карловци – Нови Сад: Издавачка књижарница Зорана Стојановића, 2014.
- Јерковић, Јован. „О континуитету развоја екавске верзије савременог српског књижевног језика“. Зборник Матице српске за филологију и лингвистику XL/2, 1997: 132–149.
- Јовић, Надежда. „О лексици *Зборника закона и уредба* из 1867. године“. [У:] М. Ковачевић (ур.) Српски језик, књижевност, уметност: зборник радова са V међународног научног скупа одржаног на Филолошко-уметничком факултету у Крагујевцу (29-30. X 2010). Књ. 1. Књижевни (стандардни) језик и језик књижевности. Крагујевац: Филолошко-уметнички факултет, 2011, 387–397.
- Луковић, Милош. „Настанак и синтаксичка својства 'Устава за Књажество Србију' од 1869. године“. Прилози проучавању језика 24/26, 1988/90: 87–97.
- Луковић, Милош. Развој српскога правног стила: прилог историји језика и права у Србији. Београд: Службени гласник, 1994.
- Луковић, Милош. „Српски правни језик“. Архив за правне и друштвене науке 3/4, 2006: 1533–1554.
- Луковић, Милош. Богишићев законик: припрема и језичко обликовање. Београд: Балканолошки институт САНУ, 2009.
- Луковић, Милош. „Језички израз Уставâ Кнежевине/Краљевине Србије“. [У:] А. Фира, Р. Марковић (ур.) Два века српске уставности: зборник радова са научног скупа одржаног 11–12. марта 2010. Београд: Српска академија наука и уметности – Правни факултет Универзитета у Београду, 2010, 159–174.
- Марковић, Жељко. „Из језичке проблематике *Сретењског устава*“. Зборник Матице српске за филологију и лингвистику XXXVII, 1994: 329–337.
- Милановић, Александар. Језик Јована Суботића. Београд: Филолошки факултет, 2014.

- Младеновић, Александар. Историја српског језика: одабрани радови. Београд: Чигоја штампа, 2008.
- Невекловски, Герхард. „Дивулзацијска лексика у српском и хрватском језику у XIX веку (на примеру неких економских и банкарских термина)“. Јужнословенски филолог LXIV, 2008: 247–260.
- Окука, Милош. „Српски језик у Аустро-угарској монархији средином 19. вијека“. Научни састанак слависта у Вукове дане 25/2, 1996: 165–176.
- Остојић, Бранислав. „Језичка структура црногорских законика 19. вијека с посебним освртом на њихову лексичку слојевитост“. Српски језик V/1–2, 2000: 239–249.
- Павловић, Слободан. Старосрпска зависна реченица од XII до XV века. Сремски Карловци – Нови Сад: Издавачка књижевница Зорана Стојановића, 2009.
- Павловић, Слободан. Узроци и механизми синтаксичких промена у српском језику. Сремски Карловци – Нови Сад: Издавачка књижевница Зорана Стојановића, 2013.
- Подгорац, Тодор. „Језик и стил Устава Кнежевине и Краљевине Србије 1835–1903“. [У:] М. Јовичић (ур.) Уставни развитак Србије у XIX и почетком XX века: зборник радова. Београд: Српска академија наука и уметности – Научна књига, 1990, 185–191.
- Суботић, Љиљана. „Судбина партиципа у књижевном језику код Срба у 19. веку“. Прилози проучавању језика 20, 1984: 5–81.
- Суботић, Љиљана. Језик Јована Хаџића. Нови Сад: Матица српска, 1989.
- Суботић, Љиљана. „Из историје књижевног језика: ’питање језика’“. [У:] Предавања из историје језика. Лингвистичке свеске. Књ. 4. Нови Сад: Филозофски факултет, 2004, 145–191.
- Тошовић, Бранко. Функционални стилови. Београд: Београдска књига, 2002.
- Biber, Douglas. “Representativeness in Corpus Design”. *Literary and Linguistic Computing* 8/4, 1993: 243–257.
- Biber, Douglas, Susan Conrad, Randi Reppen. *Corpus linguistics: Investigating language structure and use*. Cambridge – New York: Cambridge University Press, 2006.
- Koester, Almut. “Building small specialized corpora”. [In:] A. O’Keeffe, M. McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics*. London – New York: Routledge, 2010, 66–79.
- McEnery, Tony, Andrew Hardie. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press, 2012.
- Nehring, Dieter. „(Граматицке) категорије и питања промена књижевног језика“. Научни састанак слависта у Вукове дане 28/2, 1999: 203–217.
- Nehring, Gerd-Dieter. „О дескриптивном приступу описивању повезаности између језика и културе“. Научни састанак слависта у Вукове дане 39/1, 2009: 41–54.
- Nelson, Mike. “Building a written corpus: What are the basics?”. [In:] A. O’Keeffe, M. McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics*. London – New York: Routledge, 2010, 53–65.
- Okuka, Miloš. *Ogledi o našem književnom jeziku*. Nikšić: Univerzitetaska riječ, 1990.
- Reppen, Randi. “Building a corpus: What are the key considerations?”. [In:] A. O’Keeffe, M. McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics*. London – New York: Routledge, 2010, 31–37.
- Sinclair, John. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.

- Sinclair, John. "Corpus and text – basic principles". [In:] M. Wynne (ed.) *Developing Linguistic Corpora: a Guide to Good Practice*, <<https://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm>> 16.2.2019.
- Tognini-Bonelli, Elena, John Sinclair. "Corpora". [In:] K. Brown (ed.) *The Encyclopedia of Language and Linguistics*. Second Edition. Vol 2. Oxford: Elsevier, 2005, 206–219.
- Utvić, Miloš. *Izgradnja referentnog korpusa savremenog srpskog jezika*. Neobjavljena doktorska disertacija. Beograd: Filološki fakultet, 2013 <<https://fedorabg.bg.ac.rs/fedora/get/o:10061/bdef:Content/get>> 16.2.2019.
- Visković, Nikola. *Jezik prava*. Zagreb: Naprijed, 1989.
- Xiao, Richard. "Corpus creation". [In:] N. Indurkha, F. J. Damerou (eds.) *Handbook of Natural Language Processing*. Second Edition. Boca Raton – London – New York: CRC Press, Taylor & Francis Group, 2010, 147–165.

Jelena M. Pavlović

METHODOLOGICAL FRAME FOR DEVELOPING A CORPUS OF 19th CENTURY SERBIAN LEGAL TEXTS

Summary

The focus of the paper lies in the methodological challenges tied to the process of the creation of a corpus of Serbian 19th century legal texts. While following the principles of balance and representativeness, we have emphasized four major criteria that are vital to the development of such a static specialized historical corpus: (a) type of language; (b) chronological criterion (10-year period); (c) functional and linguistic criterion (genre criterion); (d) legal production (as the most relevant non-linguistic criterion). Based on the linguistic criterion, the corpus is divided into two sub-corpora: legal texts of the pre-normative period and legal texts from the period of standardization of the modern Serbian language. As a sub-criterion for further categorization, we took into account time spans of 10 years, while the limiting factors were momentous cultural and political events in Serbian history of the 19th century. From a methodological point of view, we have come to the conclusion that one must bear in mind the following when creating a corpus of 19th century legal texts: (1) A small amount of text was printed in the pre-normative period, thus all the texts should be taken into account. (2) A more organized approach to legal codification (during the rule of Milan Obrenović) influenced the expansion of legal production, thus the texts from this period will be represented as samples, not as whole texts. (3) A large number of genres within the legal register are observed, including oral and written variations, thus genre annotation will be performed on selected texts. (4) The whole period will be studied, divided into 10-year time spans, so that it is possible to develop a chronological picture of the development of certain syntactic phenomena.

Key words: historical linguistics, 19th century, legal register, methodology of corpus design.