

УДК 811.161.3'0

І. В. Наўроцкая

Беларускі дзяржаўны ўніверсітэт інфарматыкі і радыёэлектронікі

**ДА ПЫТАННЯ АБ УКЛАДАННІ КОРПУСА
БЕЛАРУСКІХ НАВУКОВЫХ ТЭКСТАЎ 1920–1930-х ГГ.**

Выкарыстанне інфармацыйных тэхналогій у мовазнаўстве дазваляе ствараць вялікія масівы тэкставых даных, распрацоўваць адпаведнае праграмае забеспячэнне, што адкрывае новыя магчымасці ў лінгвістычных корпусных даследаваннях. Гістарычная значнасць навуковага стылю 1920–1930-х гг. выступае аб'ектыўнай і актуальнай задачай для стварэння дыяхранічнага корпуса беларускіх навуковых тэкстаў. Шырокая тэматычная і жанравая разнастайнасць друкаванай перакладной і аўтарскай навуковай літаратуры 1920–1930-х гг. адкрываюць перспектывы для стварэння адпаведнага корпуса навуковых тэкстаў, які можа быць як аднамоўным беларускім, так і двухмоўным – з прыцягненнем арыгінальных тэкстаў, з якіх рабіліся пераклады. Да актуальных задач па распрацоўцы корпуса беларускіх навуковых тэкстаў 1920–1930-х гг. адносяцца стварэнне анатаванага спіса крыніц для выбару найбольш рэпрэзентатыўных адзінак, апрацоўка PDF-файлаў навуковых выданняў і выяўленне ў іх тэкставага слоя. Стварэнне дыяхранічнага корпуса навуковай літаратуры дазволіць якасна прааналізаваць значны аб'ём інфармацыі і атрымаць колькасныя паказчыкі для абгрунтаваных высноў, выявіць факты дынамікі моўных зменаў.

Ключавыя словы: гісторыя беларускай літаратурнай мовы, навуковы стыль, дыяхранічны корпус беларускіх навуковых тэкстаў, беларуская мова 1920–1930-х гг., корпусная лінгвістыка.

Для цытавання: Наўроцкая І. В. Да пытання аб укладанні корпуса беларускіх навуковых тэкстаў 1920–1930-х гг. // Труды БГТУ. Сер. 4, Принт- и медиатехнологии. 2021. № 1 (243). С. 74–78.

I. V. Naurotskaya

Belarusian State University of Informatics and Radioelectronics

**ON THE ISSUE OF CORPORA CREATION
OF BELARUSIAN SCIENTIFIC TEXTS OF THE 1920S–1930S**

The use of information technology in linguistics allows creating large arrays of text data, developing the relevant software that provides new opportunities in linguistic research. The historical significance of the scientific style in the 1920s–1930s acts as an objective and urgent task for the creation of a diachronic corpus of Belarusian scientific texts. Extensive thematic and genre variety of printed translated and authored scientific literature of the 1920s–1930s offers the prospect for creating the corpus of scientific texts, which can be both monolingual Belarusian and bilingual – with the involvement of original texts from which translations were made. Developing the corpus of Belarusian scientific texts from the 1920s–1930s includes devising an annotated list of sources to select the most representative ones, processing PDF files of scientific journals and identifying a text layer in them. The creation of the diachronic corpus of scientific literature will allow to qualitatively analyze a significant amount of information, obtain quantitative indicators for reasoned conclusions, reveal the facts of the language changes dynamics.

Key words: history of the Belarusian literary language, scientific style, diachronic corpus of Belarusian scientific texts, Belarusian language of the 1920s–1930s, corpus linguistics.

For citation: Naurotskaya I. V. On the issue of corpora creation of Belarusian scientific texts of the 1920s–1930s. *Proceedings of BSTU, issue 4, Print- and Mediatechnologies*, 2021, no 1 (243), pp. 74–78 (In Belarusian).

Уводзіны. Аб'ектыўная асаблівасць сучаснай парадэгмы ведаў – далейшае пашырэнне інфарматызацыі грамадства. З развіццём інфармацыйных тэхналогій узрастае якасць падрыхтоўкі спецыялістаў, якія могуць за адносна кароткі тэрмін апрацоўваць вялікія масівы даных. У гэтым плане лінгвістыка не з'яўляецца выключэннем і набывае інтэрдысцыплінарны характар. Як адзначаюць навукоўцы, ідэя ствараць корпус тэкстаў даўно вядомая ў мовазнаўстве. Так, у тлумачальных слоўніках выкарыстоўваецца практыка раскрываць значэнне слова праз

кантэкст [1, с. 34]. Аднак менавіта выкарыстанне інфармацыйных тэхналогій у лінгвістычных мэтах абумовіла з'яўленне новых магчымасцей для даследавання: ствараць вялікія масівы тэкставых даных і распрацоўваць адпаведнае праграмае забеспячэнне.

Асноўная частка. Пад паняццем корпуса тэкстаў у мовазнаўстве звычайна разумеецца база тэкстаў, якія адабраны па крытэрыях рэпрэзентатыўнасці пэўнай мовы [2], маюць электронны выгляд, утрымліваюць метамоўную лінгвістычную разметку і, як адзначаюць даследчыкі,

пажадана даступныя праз інтэрнэт для шырокага карыстальніка і павелічэння «паўнаwartаснай функцыянальнасці корпуса» [1, с. 35].

Дасягненні беларускай корпуснай лінгвістыкі прадстаўлены ўдзельнікамі праектаў па пытаннях метадыкі і распрацоўкі корпусаў нацыянальнай мовы, апісваецца гісторыя стварэння корпусаў англійскай, рускай, польскай і інш. моў [1; 2; 3; 4; 5; 6; 7], праведзены супастаўляльны аналіз даступных у інтэрнэце корпусаў беларускай мовы [8].

Першай спробай асобнага корпуса навуковых тэкстаў з'яўляецца распрацаваны Corpus Albaruthenicum прыкладна на 350 тыс. словаўжыванняў, створаны з выкарыстаннем 74 тэкстаў. Пошук пэўнай словаформы (камбінацыі слоў) па корпусе ажыццяўляецца з дапамогай анлайн-навага сэрвісу <http://grid.bntu.by/corpus/> [9]. Навуковы стыль у яго вучэбнай разнавіднасці выкарыстоўваецца для стварэння паралельнага корпуса беларускіх і рускіх падручнікаў [4]. Важна адзначыць, што пералічаныя корпусы прадстаўляюць сінхронны зрэз моўнага развіцця. Тым не менш распрацоўшчыкі нацыянальнага корпуса беларускай мовы (Беларускі N-корпус) прызнаюць, што «корпус можа мець дыяхранічна-сінхронны характар і ўключаць матэрыялы па-чынаючы з 20-х гадоў, асабліва гэты тычыцца падкорпусу «*Спецыяльныя тэксты*» (*курсіў наш – І. Н.*)» [2]. Такім чынам, гістарычная значнасць навуковага стылю 1920–1930-х гг. абгрунтавана выступае аб'ектыўнай і актуальнай задачай для распрацоўкі дыяхранічнай часткі нацыянальнага корпуса беларускай мовы. Аднак акрэсленае пытанне яшчэ чакае свайго вырашэння метадамі корпуснай лінгвістыкі, паколькі пры распрацоўцы беларускага N-корпуса «рэкамендавалася аддаць перавагу храналагічна больш вузкаму перыяду» [2]. Праца над нацыянальным корпусам беларускай мовы пастаянна вядзецца, таму аптымістычна ўспрымаюцца словы пра «стварэнне сінхроннага корпусу з паступовым яго рэтраспектыўным пашырэннем» [2].

Даследчыкамі па гісторыі мовы даўно адзначана неабходнасць выкарыстання інфармацыйных тэхналогій для апрацоўкі вялікай колькасці тэкстаў [10, с. 174; 11, с. 300]. Зрэшты, спецыялісты па корпуснай лінгвістыцы адзначаюць, што гатовыя корпусы дазваляюць выходзіць за межы вузкіх лінгвістычных задач і маюць вялікую практычную значнасць [1, с. 35].

Перыяд 1920–1930-х гг. асабліва паказальны ў плане дынамікі моўнага функцыянавання: беларуская мова імкліва развівалася дзякуючы пашырэнню камунікатыўных функцый беларускай мовы на ўсе сферы тагачаснага грамадскага жыцця. У якасці аб'ектыўнага паказчыка можна

разглядаць кнігавыдавецкую справу, якая адлюстроўвае вынік шырокага ўключэння беларускай мовы ў сацыяльныя працэсы. Так, на працягу 1920–1930-х гг. выдавалася вялікая колькасць беларускамоўнай літаратуры. Да параўнання прывядзём лічбы беларускай друкаванай кнігі на тэрыторыі БССР (СССР) паводле бібліяграфічных даных: у 1922 г. выпушчана 26 назваў, праз 10 гадоў, у 1932 г. пабачыла свет 1300 назваў беларускамоўных выданняў [12, с. 211]. Адзначана ў тым ліку шырокая тэматычная і жанравая разнастайнасць навуковай кнігі [13]. Наяўнасць вялікай колькасці беларускіх перакладных і аўтарскіх работ, прадстаўленых у трох падстылях навуковай функцыянальнай разнавіднасці (уласна навуковы, навучальны і навукова-папулярны), тэматычнае багацце беларускамоўнай навуковай літаратуры (матэматыка, фізіка, хімія, геадэзія, радыётэхніка, мінералогія, геаграфія, астраномія, медыцына, эканоміка, тэхніка, ваенная справа, сельская гаспадарка, гісторыя, філалогія, краязнаўства і інш.) адкрываюць перспектывы для стварэння корпуса навуковых тэкстаў 1920–1930-х гг., які можа быць як аднамоўным беларускім, так і двухмоўным – з прыцягненнем арыгінальных тэкстаў, з якіх рабіліся пераклады. Тым больш што ў даследаваннях гісторыкаў мовы абгрунтоўваецца, што супастаўленне арыгінальнага і перакладнага тэкстаў дазваляе выявіць глыбінныя заканамернасці моўнага развіцця, не заўважныя пры карыстанні выключна перакладам [14, с. 14]. Істотна пры ўкладанні корпуса мець на ўвазе, што некаторыя навуковыя выданні перавыдаваліся за 1920–1930-я гг., таму параўнанне некалькіх версій аднаго тэксту, якія выходзілі ў розны час, можа служыць надзейным паказчыкам моўных зменаў [10, с. 174; 11, с. 300]. Такім чынам, стварэнне дыяхранічнага корпуса навуковай літаратуры дазволіць якасна прааналізаваць значны аб'ём інфармацыі і атрымаць колькасныя паказчыкі для абгрунтаваных высноў, выявіць факты дынамікі моўных зменаў.

У Нацыянальнай бібліятэцы Рэспублікі Беларусь праводзіцца вялікая работа па аблічбоўцы захаваных выданняў 20–30-х гг. ХХ ст., што дазваляе ўводзіць іх у навуковы ўжытак. Аднак неабходна заўважыць, што на пачатак 2021 г., якраз амаль праз стагоддзе, калі пачалі шырока выдавацца беларускамоўныя кнігі, поўны каталог друкаваных выданняў Беларусі 1920–1930-х гг., створаны з выкарыстаннем інфармацыйных тэхналогій, яшчэ толькі чакае свайго рэалізацыі. Адпаведная электронная калекцыя «Кніга Беларусі 1517–1917 гг. Зводны каталог» распрацавана (уключае анатаваны спіс крыніц, лічбавыя копіі асобных тэкстаў) і даступная для карыстальніка ў лакальнай сетцы Нацыянальнай бібліятэкі. Падрыхтоўка другой часткі каталога

беларускай савецкай кнігі з 1917 да 1950-х гг. была запланавана яшчэ ў 1974 г. [15, с. 120–121]. Магчыма, у бліжэйшы перыяд будзе створана электронная бібліяграфія кніг, выдадзеных у 1920–1930-я гг., што дазволіць найбольш поўна прадставіць рэпертуар нацыянальнай кнігі і закласці трывалыя асновы для дыяхранічных даследаванняў як па гісторыі мовы, так і па гісторыі навукі Беларусі. Тым больш што станючыя крокі ў гэтым накірунку ўжо зробленыя на матэрыяле перыядычнага друку. Так, Нацыянальнай бібліятэкай прадстаўлена электронная калекцыя «Беларуская літаратурная спадчына», якая ўключае лічбавыя копіі больш за 400 нумароў 270 назваў часопісаў 1919–1939 гг., падзеленых паводле тэматыкі (адукацыя, медыцына, ваенная, выдавецкая справы, геофізіка, жывёлагадоўля, лясная гаспадарска, мастацтва, транспарт, эканоміка і інш.). Выданні на літаратурную тэматыку падзяляюцца ў залежнасці ад месца выдання: часопісы БССР, Заходняй Беларусі і беларускага замежжа. Калекцыя аблічбаваных копіі часопісаў даступная па адрасе <http://digital.nlb.by/collections/show/9>.

Для ўкладання корпуса беларускіх навуковых тэкстаў 1920–1930-х гг. неабходна вырашыць наступную задачу: скласці максімальна поўны анатаваны спіс навуковых выданняў адзначанага перыяду, які пакуль уяўляе сабой разрозненыя бібліяграфічныя звесткі. Напрыклад, навуковыя выданні на беларускай мове прыводзяцца ў «Летапісе беларускага друку» (1924–1931 гг.), «Летапісе друку Беларускай ССР» (з 1932 г.). Асобныя выданні даступныя ў электронных каталогах бібліятэк Рэспублікі Беларусь, пытанні навуковай літаратуры разглядаюцца ў літаратуры па кнігавыдавецкай дзейнасці [12 і інш.]. Складзены анатаваны спіс дазволіць выбраць найбольш рэпрэзентатыўныя адзінкі паводле падстыляў, наяўнасці перакладаў і перавыданняў навуковай літаратуры. Другой актуальнай задачай для ўкладання корпуса навуковых тэкстаў 1920–1930-х гг. з’яўляецца атрыманне лічбавай копіі тэксту-крыніцы, што дазволіць ажыццяўляць аўтаматычны пошук па тэксце. З выкарыстаннем інфармацыйных тэхналогій тэкст павінен распазнавацца камп’ютарнай праграмай як тэкставы код і адпаведным чынам адлюстроўвацца ў гэтай жа праграме. Так, у Нацыянальнай бібліятэцы Рэспублікі Беларусь некаторыя друкаваныя выданні 1920–1930-х гг. прадстаўлены ў выглядзе лічбавай копіі ў фармаце PDF-файла. Гэта дазваляе вывучаць крыніцу ў праграме-праглядальніку, выкарыстоўваць зручны машаб, захоўваць копію дакумента на карту флэш-памяці. Паколькі аблічбаваная кніга захоўваецца ў выглядзе фармату PDF, у якім тэкставы слой не распазнаны,

а пададзены як гатовы малюнак, што абгрунтавана прадстаўленнем у электронным асяроддзі надрукаванай кнігі з захаваннем усіх элементаў яе афармлення, то ў даследчыка для далейшай працы ўзнікае неабходнасць распазнавання тэксту. У сваю чаргу гэта дазваляе выкарыстоўваць распазнаны тэкст у праграмах, якія аўтаматызуюць пошук адпаведнай словаформы, параўноўваюць некалькі выданняў аднаго твора або арыгінал і яго пераклад, адлюстроўваюць вынікі параўнання ў асобным дакуменце. Да найбольш вядомых праграм такога тыпу адносіцца тэкставы рэдактар Word. Напрыклад, супастаўленне некалькіх версій аднаго тэксту магчыма прывесці, націснушы на ўкладку «*Сэрвіс*» і выбраўшы срод спіса каманд «*Параўнаць і аб’яднаць*».

З распазнаваннем тэкставага слоя PDF-файла звязаны некаторыя цяжкасці. Па-першае, паліграфічнае афармленне кніг 1920-х гг. не заўсёды без памылак распазнаецца праграмай як тэкст, гатовы варыянт таксама патрабуе карэкціровак і зверкі з арыгіналам, тым больш што мовы пачатку XX ст. і пачатку XXI ст. адрозніваюцца ў тым ліку арфаграфіяй, лексікай. Па-другое, не ва ўсіх праграмах распазнавання ўбудавана магчымасць выбару беларускай мовы, што таксама зніжае якасць гатовага тэксту і вымагае большую колькасць правак. Сучасныя інфармацыйныя тэхналогіі дазваляюць апрацоўваць файлы як з усталёўкай дадатковых праграм на персанальны камп’ютар, так і з выкарыстаннем рэсурсаў сусветнага сеціва або прымяняюцца два спосабы ў сукупнасці. Так, распазнаванне тэксту можа праводзіцца праз усталяваную на камп’ютары праграму ABBYY FineReader, якая ў пераліку інструментаў утрымлівае беларускую мову. Да тыпу інтэрнет-рэсурсаў адносяцца сэрвісы <https://convertio.co/ru/ocr/>, <https://tools.pdf24.org/ru/ocr-pdf>, на якіх магчыма праводзіць разнастайныя аперацыі з PDF-файламі: выяўленне тэкставага слоя з выкарыстаннем беларускай мовы, канвертацыю ў тэкставы фармат. Тэкставы слой таксама магчыма распазнаваць наступным чынам: спачатку утыліта **Lightshot** дапамагае рабіць хуткія фотаздымкі старонкі кнігі, захоўваць іх, а потым трэба загрузаць малюнак на старонку перакладчыка <https://translate.yandex.by/ocr>, дзе праводзіцца распазнаванне тэксту. Параўнанне некалькіх PDF-дакументаў у тым ліку прадстаўлена на наступных сэрвісах:

<https://draftable.com/compare>;

<https://www.diffchecker.com/pdf-diff>;

<https://www.inetsoftware.de/products/pdf-content-comparer> і інш.

Некаторыя з іх маюць ліміт загрузкі па мегабайтах і старонках. Для аўтаматызаванага якаснага параўнання PDF-файлаў неабходна мець версію дакумента з тэкставым слоём, які можна

атрымаць, карыстаючыся згаданым вышэй адрасам <https://tools.pdf24.org/ru/ocr-pdf>.

У межах дысертацыйнага даследавання пра станаўленне навуковага стылю намі праводзіцца падбор матэрыялу і вычстка некаторых распазнаных PDF-дакументаў, папярэдне калекцыя навуковых тэкстаў 1920–1930-х гг. складаецца з 85 тэкстаў і амаль з 900 тыс. словаўжыванняў. Тэксты прыкладна раўнамерна размеркаваны паводле трох падстыляў навуковага стылю. Калекцыя апрацаваных тэкстаў адкрывае перспектывы для пераўтварэння ў корпус шляхам дадання граматычнай разметкі і распрацоўкі адпаведнага праграмнага забеспячэння.

Заклучэнне. Такім чынам, стварэнне корпуса (калекцыі, базы) навуковых тэкстаў 1920–1930-х гг., а таксама распрацоўка якаснага праграмнага забеспячэння з’яўляюцца неабходнымі для правядзення даследаванняў у галіне дыяхранічнага мовазнаўства, што дазволіць найбольш поўна апісаць гісторыю беларускай мовы, асабліваці складання яе нормаў, стыляў, падстыляў навуковай разнавіднасці і асэнсаваць спецыфіку функцыянавання на сучасным этапе. Тым больш што дыяхранічныя даныя выступаюць неад’емнай часткай лінгвістычных даследаванняў і істотна дапаўняюць аспекты сінхрані моўнага развіцця.

Спіс літаратуры

1. Барковіч А. А. Корпусная лінгвістыка і Інтэрнэт // Роднае слова. 2014. № 9. С. 33–36.
2. Кошчанка У. Нацыянальны корпус беларускай мовы ў кантэксте корпуснай лінгвістыкі славыянскіх краін [Электронны рэсурс]. URL: https://bnkorporus.info/articles/Artykul_Koscanki_2018.html. (дата звароту: 10.01.2021).
3. Волчек О. А., Порицкий В. В. Экспериментальный корпус белорусского языка: текущее состояние и перспективы развития // Труды Междунар. конф. «Корпусная лингвистика – 2013». Санкт-Петербург: Изд-во С.-Петерб. ун-та, 2013. С. 231–239.
4. Зубов А. В. О задачах создания подкорпуса текстов русских и белорусских учебников по языку и литературе // Карповские научные чтения: сб. науч. ст. Вып. 8: в 2 ч. / редкол.: А. И. Головная (отв. ред.) [и др.]. Минск: Белорусский Дом печати, 2014. Ч. 1. С. 13–15.
5. Зубов А. В. Создание большого корпуса текстов белорусского языка и его использование для изучения белорусского языка и его связи с другими языками Европы // Труды Междунар. конф. «Корпусная лингвистика – 2019». СПб: Изд-во С.-Петерб. ун-та, 2019. С. 23–29.
6. Калета Р., Капылоў І., Кошчанка У. Корпусная лінгвістыка як перспектыўны напрамак развіцця беларуска-польскіх моўных сувязей // Acta Albaruthenica. Т. 20. Warszawa: Katedra Białorutenistyki Uniwersytetu Warszawskiego, 2020. С. 185–198.
7. Рычкова Л. В. Регионально распределённая полнотекстовая база данных в исследовании речевого облика районных газет Беларуси // Славянский мир и национальная речевая культура в современной коммуникации: сб. науч. тр. / гл. ред. М. И. Конюшкевич. Гродно: ГрГУ им. Янки Купалы, 2018. С. 79–82.
8. Станкевич А. Ю. Программные инструменты создания конкордансов для белорусскоязычных опорных единиц // Вестник МГЛУ Сер. 1, Филология. 2016. № 1 (80). С. 108–115.
9. Кошчанка У., Капылоў І., Міклашэвіч І. Corpus Albaruthenicum як частка міжнароднага праекта «BalticGrid-II» // Наука и инновации. № 10 (92). 2010. С. 30–31.
10. Запрудскі С. М. Да пытання аб дынаміцы узусу беларускай літаратурнай мовы ў 1920–1930-х гг. // Язык и социум: материалы V Междунар. науч. конф., Минск, 6–7 дек. 2002 г.: в 2 ч. Минск: РИВШ БГУ, 2003. Ч. 2. С. 173–177.
11. Запрудскі С. М. Беларускае мовазнаўства і развіццё беларускай літаратурнай мовы: 1920-я – 1930-я гады. Мінск: БДУ, 2013. 367 с.
12. Гісторыя беларускай кнігі. Т. 2: Кніжнасць новай Беларусі (XIX–XXI стст.) / рэд. В. В. Антонай, М. В. Нікалаеў. Мінск: Бел. Энцыкл. імя Петруся Броўкі, 2011. 435 с.
13. Наўроцкая І. В. Крыніцы навуковага стылю беларускай літаратурнай мовы 1920-х гг. // Мова і літаратура: матэрыялы 72-й навук. канф. студэнтаў і аспірантаў філал. фак. БДУ, Мінск, 28 крас. 2015 / пад рэд. К. А. Тананушкі. Мінск: РИВШ, 2015. С. 82–87.
14. Кулеш Г. І. Мова беларускага заканадаўства XX стагоддзя: генезіс і эвалюцыя. Мінск: БДУ, 2015. 303 с.
15. Киреева Г. В. Сводный каталог «Книга Беларуси»: к созданию полного библиографического репертуара белорусских изданий XVI–XVII вв. // Інфармацыйныя рэсурсы Нацыянальнай бібліятэкі Беларусі: праблемы фарміравання і выкарыстання: зб. арт. / склад. Т. В. Кузьмінч. Мінск: Нац. б-ка Беларусі, 2008. С. 120–129.

References

1. Barkovich A. A. Corpus linguistics and Internet. *Rodnaye slova* [Native word], 2014, no 9, pp. 33–36 (In Belarusian).
2. Koshchanka U. *Natsyyanal’ny korpys belaruskay movy u kantekstse korpusunay lingvistyki slavyanskikh krain* [National Corpus of the Belarusian Language in the context of Slavic countries

corpus linguistics]. Available at: https://bnkorus.info/articles/Artykul_Koscanki_2018.html. (accessed 10.01.2021).

3. Volchek O. A., Poritski V. V. An experimental corpus of Belarusian: its present and future. *Trudy Mezhdunarodnoy konferentsii "Korpusnaya lingvistika – 2013"* [Proceedings of the international conference «Corpus linguistics – 2013». St. Petersburg, 2013, pp. 231–239 (In Russian).

4. Zubov A. V. On the tasks of creating a subcorpus of texts of Russian and Belarusian textbooks on language and literature. *Karpovskiy nauchnyye chteniya: sbornik nauchnykh statey* [Karpov Scientific Readings: Collection of Scientific Articles], issue 8: in 2 vol. Minsk. Vol 1. 2014, pp. 13–15 (In Russian).

5. Zubov A. V. Creation of a large corpus of texts of Belarusian language and the use of it for investigation of the Belarusian language and its connection with the different languages of European. *Trudy Mezhdunarodnoy konferentsii "Korpusnaya lingvistika – 2019"* [Proceedings of the international conference «Corpus linguistics – 2019». St. Petersburg, 2019, pp. 23–29 (In Russian).

6. Kaleta R., Kapylou I., Koshchanka U. Corpus linguistics as a perspective direction for the development of Belarusian-Polish linguistic relations [*Acta Albaruthenica*], 2020, vol. 20, pp. 185–198. DOI: 10.32612/aw.18988091.2020.20.

7. Rychkova L. V. Regionally distributed full-text database in the study of the speech appearance of regional newspapers in Belarus. *Slavyanskiy mir i natsional'naya rechevaya kul'tura v sovremennoy kommunikatsii: sbornik nauchnykh trudov* [Slavic world and national speech culture in modern communication: collection of scientific papers], Grodno, 2018, pp. 79–82 (In Russian).

8. Stankevich A. Yu. Concordancing software for Belarusian-speaking reference units. *Vestnik MGLU* [MSLU Bulletin], series 1, Philology, 2016, no 1 (80), pp. 108–115 (In Russian).

9. Koshchanka U., Kapylou I., Miklashevich I. Corpus Albaruthenicum as part of the international project “BalticGrid-II”. *Nauka i innovatsii* [Science and innovation], 2010, no 10 (92), pp. 30–31 (In Belarusian).

10. Zaprudski S. On the question of the dynamics of the use of the Belarusian literary language in 1920–1930s. *Yazyk i sotsium. Materialy V Mezhdunarodnoy nauchnoy konferentsii: 6–7 dekabrya 2002 g., Minsk* [Language and society: materials of the V international scientific conference, Minsk, December 6–7, 2002]. Minsk, 2003, vol. 2, pp. 173–177 (In Belarusian).

11. Zaprudski S. *Belaruskaye movaznaustva i razvitsyio belaruskay litaraturnay movy: 1920–1930-ya gady* [Belarusian linguistics and Belarusian Literary Language development: 1920–1930s]. Minsk, BDU Publ., 2013. 367 p.

12. *Gistoryya belaruskay knigi. T. 2: Knizhnasts' novay Belarusi (XIX–XXI stst.)* [History of the Belarusian book. Vol. 2: The literature of the new Belarus (XIX–XXI centuries)]. Minsk, Belaruskaya Entsi-klapedyya imya Petrusya Brouki Publ., 2011. 435 p.

13. Naurotskaya I. V. Sources of the scientific style of the Belarusian literary language of the 1920s. *Mova i litaratura: materyaly 72-y navukovay kanferentsyi studentau i aspirantau filalagichnaga fakul'teta BDU, Minsk, 28 krasavika 2015* [Language and Literature: Proceedings of the 72nd Scientific Conference of Students and Postgraduates of the Philological Faculty of BSU, Minsk, April 28, 2015], Minsk, 2015, pp. 82–87 (In Belarusian).

14. Kulesh G. I. *Mova belaruskaga zakanadaustwa XX ctagoddzya: genesis i evalutsyya* [The language of Belarusian legislation of the twentieth century: genesis and evolution]. Minsk, BDU Publ., 2015. 303 p. Kulesh H. I. Language Belarusian legislation of the twentieth century: genesis and evolution. Minsk, Belarusian State University, 2015. 303 p. (In Belarusian).

15. Kireeva G. V. Consolidated catalog “The Book of Belarus”: towards the creation of a complete bibliographic repertoire of Belarusian editions of the 16th – 17th centuries. *Infarmatsyynyya resursy Natsyyanal'nyy bibliyateki Belarusi: prablemy farmiravannya i vykarystannya: zbornik artykulau* [Information resources of the National Library of Belarus: problems of formation and use: a collection of articles]. Minsk, 2008, pp. 120–129 (In Belarusian).

Інфармацыя пра аўтара

Наўроцкая Ірына Валянцінаўна – выкладчык кафедры агульнаадукацыйных дысцыплін. Беларускі дзяржаўны ўніверсітэт інфарматыкі і радыёэлектронікі (220013, г. Мінск, вул. П. Броўкі, 6, Рэспубліка Беларусь). E-mail: navrotskaya@bsuir.by

Information about the author

Naurotskaya Iryna Valyantsinauna – lecturer, the Department of General Subjects. Belarusian State University of Informatics and Radioelectronics (6, P. Brovki str., 220013, Minsk, Republic of Belarus). E-mail: navrotskaya@bsuir.by

Пасмыніў 11.02.2021