

Стагоддзе на далоні

Каб дасканала пазнаёміцца з беларускай літаратурай XIX стагоддзя ці напісаць бліскучае даследаванне па мове гэтага перыяду, ужо не трэба збіраць стосы кніг і тыднямі сядзець у бібліятэцы. У выдавецтве «Беларуская навука» пабачыў свет электронны дыск «Канкарданс беларускай мовы XIX стагоддзя», які змяшчае не толькі грунтоўны даведнік, але і поўны корпус твораў азначанага перыяду (497 тэкстаў). З некалькіх клікаў дыск прапануе карыстальніку агульны канкарданс, чатыры аўтарскія канкардансы, частотны слоўнік, адваротны слоўнік, спіс твораў і тэкстаў. Дык што ж такое канкарданс? Вось як тлумачыць аўтар даведніка — малодшы навуковы супрацоўнік Інстытута мовазнаўства Акадэміі навук Беларусі Ніна Міхайлаўна Сянкевіч.

— Канкарданс — даведнік, які складаецца на аснове корпуса тэкстаў — пэўнага перыяду альбо пэўнага аўтара. Спецыфіка ў тым, што ён уключае ўсе словаўжыванні тэксту. Гэта поўны пералік словаўжыванняў з кантэкстамі і адраснымі характарыстыкамі, да якога можа дадавацца і іншая інфармацыя. Як правіла, на гэтым матэрыяле робяцца частотныя і адваротныя слоўнікі, паказальнік слоў, якія ёсць і на нашым дыску. Канкарданс разлічаны на спецыялістаў, але змяшчае ўсе тэксты — таму цікавы і для шырокага чытача. Бо да гэтага часу тэксты XIX стагоддзя не выдаваліся ў поўным аб'ёме, за выключэннем хіба кніг серыі «Залатая калекцыя беларускай літаратуры».

— Якую гісторыю маюць гэтыя глабальныя даведнікі?

— Гісторыя пачынаецца больш як сем стагоддзяў таму. Лічыцца, што першы канкарданс быў створаны ў 1262 годзе па Бібліі. Эра сучасных канкардансаў пачалася ў сярэдзіне мінулага стагоддзя. На працягу некалькіх дзесяцігоддзяў былі створаны канкардансы па творах многіх паэтаў і пісьменнікаў — перш за ўсё англамоўных. Найбольш вядомым можна лічыць шасцітомны канкарданс да твораў Уільяма Шэкспіра, выдадзены ў 1968 годзе. Сёння гэта справа квітнее ў Расіі — там працуюць вялікія калектывы, ствараюцца новыя канкардансы на больш высокім узроўні.

У айчыннай лексікаграфіі канкардансы не былі распаўсюджаны, і можна сказаць, што наша група пад кіраўніцтвам прафесара Віктара Мартынава была адной з першых, хто ўзяўся за гэту справу ў канцы 70-х гадоў мінулага стагоддзя. Калі я выступала на канферэнцыях з дакладамі, было заўсёды шмат пытанняў, мовазнаўцы цікавіліся гэтай справай. У той час канкардансы задумваліся як альтэрнатыва картатэкам, патрэбным для стварэння новых слоўнікаў. Картатэкі складаюцца з адзінкавых выпадкаў, а ў канкардансе змешчаны поўны пералік кантэкстаў, што дазваляе лёгка сфармуляваць значэнне слова. Гэта зручны інструмент для даследавання семантыкі.

— Ці вялася праца над іншымі канкардансамі?

— Праца распачалася ў 1979 годзе. Першым планаваўся канкарданс беларускай мовы XIX стагоддзя, бо гэта вельмі значны перыяд у развіцці беларускай мовы. Менавіта тады пачалося станаўленне яе літаратурнай формы. Але гэты перыяд быў малавывучаны, бо даследчыкі не мелі доступу да арыгіналаў літаратурных твораў, раскіданых па розных бібліятэках

і архівах. Беларуская мова XIX стагоддзя і сёння застаецца не вельмі даследаванай...

„Даследаванне частотнасці лексічных адзінак дазваляе зрабіць высновы пра асаблівасці мовы і стылю пісьменніка, у прыватнасці, знайсці яго «ўлюбёныя» словы, выявіць суадносіны паміж часта- і рэдкаўжывальнымі, прасачыць, наколькі блізкая па тэматыцы творчасць розных аўтараў.“

У 1985 годзе мы падрыхтавалі першы канкарданс. Потым вырашылі зрабіць серыю канкардансаў беларускіх пісьменнікаў:



Ніна Сянкевіч.

аўтарская лексікаграфія ў Беларусі ў той час амаль не развівалася, у адрозненне ад іншых краін.

Канкардансы планаваліся як аснова для распрацоўкі слоўнікаў мовы асобных пісьменнікаў, але гэта задача засталася нявыкананай. Было вызначана зрабіць канкардансы мовы Максіма Багдановіча, Кузьмы Чорнага, Кандрата Крапівы. Мы

пачалі ўводзіць тэксты Кузьмы Чорнага, але потым я на шэсць год сышла ў адпачынак, і за гэты час змянілася тэхніка. Мы запісвалі збор твораў на магнітную стужку для электронна-вылічальных машын ЕС. А калі я вярнулася, замест гэтых машын з'явіліся персанальныя камп'ютары, і магнітныя стужкі нельга было прачытаць, таму работу над канкардансам мовы Чорнага давялося спыніць. Канкарданс мовы Багдановіча быў амаль завершаны, але цяпер пры ўзнаўленні яго трэба дапоўніць творами, не ўлічанымі ў той час. Канкарданс мовы Кандрата Крапівы быў завершаны ў 1996 годзе — і з таго часу праца спынілася. Выдаць даведнікі не атрымалася, бо яны вялікія па аб'ёме. Мы раздрукавалі «Канкарданс беларускай мовы XIX стагоддзя» ў адзіным



Вокладка электроннага дыска.

тэксты, якія трэба было дадаць. Бо асноўная вартасць канкарданса — паўната і аб'ектыўнасць пададзенай інфармацыі — забяспечваецца апрацоўкай максімальна поўнага тэкставага матэрыялу.

— Якія лінгвістычныя даследаванні сталі магчымыя дзякуючы з'яўленню канкарданса? На якія актуальныя пытанні можа даць адказ гэты даведнік?

— Мы пачалі ўжо працу над двума слоўнікамі на аснове канкарданса. Гэта «Кароткі слоўнік беларускай мовы XIX стагоддзя», дзе будуць прыведзены значэнні спецыфічных слоў, і «Анамастычны слоўнік беларускай мовы XIX стагоддзя» — такога слоўніка няма, хаця ў творах гэтага перыяду сустракаецца шмат адмысловых тапонімаў, уласных імёнаў.

Карыснымі будуць і іншыя слоўнікі, якія змяшчае канкарданс. Нават простае даследаванне на аснове частотных слоўнікаў дае прыемныя вынікі. Мне было надзвычай цікава паглядзець, якія словы часцей за ўсё ўжывалі аўтары твораў у XIX стагоддзі. Аказалася, што самае частотнае слова — *пан*. А вось спіс дзесяці найбольш ужывальных слоў: *пан, Бог, людзі, дзеці, мужык, брат, цар, хата, народ, зямля*. Калі я параўнала самыя частотныя словы з агульнага і аўтарскіх канкардансаў, то выявілася, што гэтыя спісы супадаюць 1:1 у Вінцэнта Дуніна-Марцінкевіча і практычна супадаюць у Аляксандра Ельскага і Каруся Каганца. А вось у Кастуся Каліноўскага найбольш частотнае слова — *мужык*. Даследаванне частотнасці лексічных адзінак дазваляе зрабіць высновы пра асаблівасці мовы і стылю пісьменніка, у прыватнасці, знайсці яго «ўлюбёныя» словы, выявіць суадносіны паміж часта- і рэдкаўжывальнымі, прасачыць, наколькі блізкая па тэматыцы творчасць розных аўтараў.

— Якія рэдкія тэксты змешчаны на дыску?

— Ёсць невялічкі вершаваны твор «Разговор Лянокіцкіх крэсцян» — назва сустракаецца ў даследаваннях, але бачыць тэкст не даводзілася. Яго знайшлі ў Віленскім архіве, і я сама расчытвала рукапіс,

што было няпростай справай: старая папера, неразборлівы почырк. Упершыню змешчаны і ўрываек прэзідэнтскага твора Адама Гурыновіча, вядомага як вершатворца. Уключаны тэксты гутарак, адносна якіх няма звестак.

Тэксты шэрагу твораў, што ўжо былі апублікаваныя, маюць досыць істотныя адрозненні ад аўтарскіх. Нярэдка ў тэксты ўносіліся сур'ёзныя праўкі, яны публікаваліся са значнымі скарачэннямі. Узятая з першакрыніц, тэксты ў корпусе пададзены без купюр. Вершаваныя творы «Старшына» А. Абуховіча, «Саўсім не тое, што было» Ф. Тапчэўскага, «Езусу Цвецінскаму скарга на даўмантаў» І. Тамашэвіча,

«Літвінам, запісаўшымся ў мой альбом, на пажагнанне» А. Вярыгі-Дарэўскага, ананімны твор «Дзякуй Богу і гасудару...». Некаторыя вершы К. Каганца, Л. Вітана-Дубейкаўскага з рукапіснага зборніка «Мае думы і мыслі» ўвайшлі ў корпус у поўным аб'ёме ў адрозненне ад апублікаваных тэкстаў.

— Ці мае сэнс перавыдаваць творы беларускай літаратуры XIX стагоддзя, калі ўсе яны змешчаны на дыску?

— Гэта чытабельныя тэксты, але... Калі мы пачыналі працу над канкардансамі, праграма было зручнай апрацоўваць тэксты, надрукаваныя вялікімі літарамі. Таму ўсе тэксты ўведзены

„Канкарданс разлічаны на спецыялістаў, але змяшчае ўсе тэксты — таму цікавы і для шырокага чытача. Бо да гэтага часу тэксты XIX стагоддзя не выдаваліся ў поўным аб'ёме.“

менавіта так, і чытаць іх не вельмі зручна, бо нельга, напрыклад, вылучыць уласныя імёны. Мы рабілі мінімальнае рэдагаванне: замянілі літары, якія не існуюць у сучасным беларускім алфавіце (так, Карусь Каганец распрацоўваў уласны алфавіт, і яго сімвалы з рукапісных шрыткаў мы замянілі), і правялі транслітарцыю тэкстаў з лацінкі на кірыліцу.

Агульнавядома, што ў той час не існавала правілаў граматыкі і правапісу, адсюль — мноства розных графічных варыянтаў аднаго слова (напрыклад, у канкардансе зарэгістравана 14 варыянтаў слова *яшчэ*). Дарэчы, знайсці ўсе варыянты аднаго слова дапамагае адваротны слоўнік. У тэкстах досыць часта сустракаецца злітнае напісанне слоў, асабліва часціц і прыназоўнікаў. Мы не ставілі знакаў прыпынку, пакідалі аўтарскія, якія не заўсёды ставіліся нават у канцы сказаў. Усё гэта ўскладняе ўспрыманне. Таму чытаць мастацкія творы, на мой погляд, зручней у адрэдагаваным варыянце.

Алеся ЛАПІЦКАЯ