

КАРТОТЕКАЛЫҚ ҚОРДЫ ЦИФРЛАНДЫРУ – ҚАЗАҚ ТІЛІНІҢ ҰЛТТЫҚ КОРПУСЫН ҚҰРУДЫҢ НЕГІЗІ

*Жаңабекова А.Ә.¹, Қожахметова А.Қ.², Бесиров Е.Б.³

¹Филология ғылымдарының докторы, А. Байтұрсынұлы атындағы Тіл білімі институты Қолданбалы лингвистика бөлімінің меңгерушісі, Алматы, Қазақстан, e-mail: aiman_miras@mail.ru, <https://orcid.org/0000-0002-6199-7444>

²Докторант, А. Байтұрсынұлы атындағы Тіл білімі институты Қолданбалы лингвистика бөлімінің кіші ғылыми қызметкері, Алматы, Қазақстан,

e-mail: akony_8484@mail.ru, <https://orcid.org/0000-0003-3783-3199>

³Магистр, А. Байтұрсынұлы атындағы Тіл білімі институты Қолданбалы лингвистика бөлімінің кіші ғылыми қызметкері, Алматы, Қазақстан, e-mail: erkin.besirov@mail.ru, <https://orcid.org/0000-0002-2291-9654>

Андатпа. Мақалада қазақ тіліндегі түсіндірме сөздіктер құрастырудың дереккөзі, иллюстрациялық базасы болып табылатын қазақ тілінің әртүрлі стильдері бойынша жинақталған мысалдар жиынтығы – картотекалық қорды цифрландырудың, автоматты сөздік базаға келтірудің ғылыми-практикалық мәселелері қарастырылады. Орыс тілінің картотекалық қорын жасау тәжірибесіне шолу жасалады. Картотекалық қорды жасаудың маңыздылығы туралы айтылған пікірлер беріледі. Қазақ тіліндегі ХІХ ғасырдың орта шенінде қол жұмысы арқылы басталған картотекалық қорды жасақтау тарихы баяндалады. Сонымен қатар ХХ ғасырдың басында профессор А.Жұбановтың басқаруымен жүзеге асқан «Тіл – қазына» атты деректер базасының жұмысына сипаттама беріледі. Автоматты, электронды картотекалық базаның жұмыс істеу механизмі көрсетіледі және сол арқылы пайдаланушы мүмкіндіктері айқындалады. Картотекалық қор – лексикографиялық база жасаудың, яғни түсіндірме сөздіктер құрастырудың негізгі шарты екендігі дәйектеледі. Қазақ тілінің картотекалық қоры негізінде жасалған түсіндірме сөздіктерге шолу жасалады.

Қазақ тілінің картотекалық қорын цифрландыру қажеттігін негіздеуде, осы бағыттағы әлемдік тіл біліміндегі қолданбалы бағыттағы жұмыстарды зерттеуде шолу, сипаттау, баяндау, аналитикалық талдау әдістері қолданылады. Картотекалық қорды автоматты сөздікке, электронды иллюстрациялық базаға айналдыру тетіктерін айқындауда алгоритмдік программалау әдісіне сүйенеміз.

Картотеканың дереккөздері нақтылығымен, стильдік жағынан алуан түрлілігімен, арнайы іріктеліп таңдалған және өңделген мәтіндер болғандықтан құндылығын ешқашан жоғалтпақ емес.

Картотека материалдары қазақ тілінің көптеген сөздіктері мен грамматикаларын дайындауда қолданылып келеді. Бұл картотеканы біздің еліміздің әр түрлі қалаларының және шетелден келген ғалымдар қазақ тіл білімінің әртүрлі мәселелерін зерттеу үшін пайдалана алады.

Тірек сөздер: картотека, автоматтандыру, цифрландыру, электронды база, картотекалық қор, корпус, лексикография, қолданбалы лингвистика

Мақала OR11465483 «Мемлекеттік тілді латынграфикалы ұлттық әліпбиге көшіруді қамтамасыз ететін жаңартылған нормативтік сөздіктер мен академиялық басылымдар сериясын әзірлеу» атты бағдарламалық-нысаналы зерттеу жұмысы аясында жарияланып отыр.

Негізгі ережелер

Картотекалық қор – ғалымдардың ғылыми-зерттеу қолжазбалары негізінде зерттеулерге қажетті сөздерді, сөз тіркестерін мысалдарымен, автормен, оқулық тақырыптарымен бірге жинақталған материалдар қоры. Сондықтан картотекалық қорды цифрландыру тақырыбы қазіргі қазақ тілінің салалық дамуында, қазақ тілін автоматтандыруда, қазақ тілінің ұлттық корпусын құрудың негізінде және цифрлі Қазақстан құру мәселесінде өзектілігін жоғалтпайды.

Кіріспе

Зерттеу мақаламыздың алғышарттары ретінде мынадай мәселелерді атап көрсетеміз. Бүгінгі таңда ғылым-білімнің барлық саласында ақпараттық технологияны кеңінен қолдану заман талабы болып отыр. Цифрлық Қазақстан бағдарламасын жүзеге асыру – еліміздің бәсекелестікке қабілетті ел болуының заманауи шарты, бұл – бір. Екіншіден, ақпараттық технологияның қарыштай дамуымен байланысты өркениетті елдерде кез келген қызмет түрін жеңілдетуге, тиімділігін арттыруға ерекше көңіл бөлініп отыр. Өйткені қайсібір өркениетті қоғам болмасын өз қызметін оның негізгі өкілі – Адамның өмір сүру ахуалының жақсаруына, сол арқылы қоғамның өркендеп, дамуына бағыттайды. Осы орайда тіл білімінің қазіргі зерттеу бағыты антропоөзектік парадигмаға қарай ойысты [1, 45 б.]. Үшіншіден, кез келген халықтың ұлттық болмысын танытатын жазба мұраларды сақтау және оны ұрпақтан-ұрпаққа жеткізу, әсіресе, жазу графикасын ауыстыру қарсаңында атқарылуға тиіс өте маңызды мәселелердің бірі болып отыр.

Қысқасы, цифрландыру, электронды форматқа көшіру, автоматтандыру, жұмысты жеңілдету, тиімділігін арттыру, қазақ сөзін келесі ұрпаққа жеткізу, әлемдік тіл білімінің дамуынан қалып қалмау, бәсекелестікке қабілетті болу сияқты талаптар бүгінгі күні қазақ тіл білімінде электронды форматта автоматты компьютерлік бағдарламалар арқылы жұмыс істейтін мәтіндік және сөздік базаларды жедел түрде жасақтау қажеттігін тудырады.

Осы саладағы белгілі ғалымдардың пайымдауынша, компьютерлік тілдік қор – ғалымдардың өздерінің ғылыми зерттеу жұмыстарына жаңаша тұрғыда қарастыра алу мүмкіндігі. Осы сияқты компьютерлік тілдік қор неғұрлым қомақты болып келсе, соғұрлым қазақ тілі құрылысының сыры барынша тереңірек ашылады. Демек зерттеліп отырған нысан жөніндегі түсініктердің көлемі, адамзаттың білімі мен ғылыми өрісі кеңі түспек. Сол сияқты, зерттеуші ғалымның қабілет деңгейі әлденеше есе артып, шығармашылық қуат көздері ашыла түсетіні сөзсіз. Осылайша бұл жаңа мүмкіндіктер, қазақ тіл білімінің жүйелілік қасиеттерін жетілдіруге және тіл жүйесін мұқият тануға мүмкіндік береді.

Мақалада қазақ тіліндегі түсіндірме сөздіктер құрастырудың дереккөзі, иллюстрациялық базасы болып табылатын қазақ тілінің әртүрлі стильдері бойынша жинақталған мысалдар жиынтығы – картотекалық қорды

цифрландырудың, автоматты сөздік базаға келтірудің ғылыми-практикалық мәселелері қарастырылады.

Зерттеу материалы және әдістері

Мақалада қазақ тілінің картотекалық қорын цифрландыру қажеттігін негіздеуде, осы бағыттағы әлемдік тіл біліміндегі қолданбалы бағыттағы жұмыстарды зерттеуде шолу, сипаттау, баяндау, аналитикалық талдау әдістері қолданылады. Картотекалық қорды автоматты сөздікке, электронды иллюстрациялық базаға айналдыру тетіктерін айқындауда алгоритмдік программалау әдісіне сүйенеміз.

Зерттеу материалы А.Байтұрсынұлы атындағы Тіл білімі институтының қоржынындағы 5 миллионнан аса кәртішкеден тұратын картотекалық қордан алынды.

Әдебиетке шолу

ЭЕМ (Электронды есептеу машинасы) көмегімен лексикографиялық жұмыстарды автоматтандыру мәселесі 1973 жылдан бастап қолға алына бастады.

Отандық қолданбалы зерттеулердің негізін қалаған Қ.Бектаев, А.Жұбанов, А.Белботаев, А.Зекенова сынды тілші-ғалымдар алғаш қазақ тілінің жиілік сөздіктерін құрастырды.

1993 жылы «Қолданбалы лингвистика» бөлімінің қызметкерлері М.Әуезовтің 20 томдық толық шығармалар жинағын компьютер жадына түсіріп, әр түрлі жиілік сөздіктер шығарды.

Лексикография саласын автоматтандырудың әлем бойынша қысқаша тарихын сөз ететін болсақ, бұл бағыт осыдан 55 жыл бұрын Роберт Буздың жұмыстарынан бастама алды. Белгілі ғалым әлемдік тәжірибеде тұңғыш рет әдеби ескерткіштерді зерттеуге қажетті, дәлірек айтқанда, Фома Аквинскийдің шығармаларына сөзнұсқағыш пен конкорданс құрастыруда есептеуіш-перфорациялық мәшинені қолданған болатын [2, 85 б.].

Осындай лексикографиялық жұмыстарды автоматтандыру әрекеттері ұлттық тілдер, соның ішінде қазақ тілі қорларының автоматтанған жүйесін құру мәселесіне әкелген болатын.

Лексикография саласын автоматтандырудың қысқаша тарихына тоқтар болсақ, Францияның Нансий университетінде XVI-XX ғғ. француз тілінің автоматтандырылған сөздік картотекасының толық нұсқасы ақпарат сақтағыш магниттік таспада орналасқанын айтуға болады. Оның әдеттегі сөздік картотекасынан айырмашылығы тек техникалық орындалуында ғана емес, сонымен бірге құрылымдық жағында да болды. Мысалы, әрбір таңдалып алынған үлкен көлемдегі мәтіндер жиынын (шамамен алғанда 100 мың сөзқолданыс) өңдеу процесі жүйелі түрде жүргізіліп, нәтижесінде аз мезгіл аралығында (1963-1986 жж.) тек XIX-XX ғғ. тілінің 90 млн. сөзқолданыстан тұратын мәтіні өңдеуден өтіп, 175 мың лексемадан тұратын сөзтізбе түзілген. Сонымен бірге, жаңадан құрастырылып жатқан француз тілінің 15 томдық сөздігіне (*Treésor général de la Langue Francaise*) 70 мың сөздік бірлігі мен 1

млн. иллюстрациялық цитаталар таңдалып алынған. Мұндағы әрбір сөздің қолдануы 10 жылдық уақыт мерзімімен құжатталған деуге болады. Аталған сөздіктің бірінші томының 1972 жылы жарық көруі, ал 15 томның толық аяқталуы 1987 жылға жоспарлануы сөздік құрастыру жұмысының автоматты және жүйелі түрде жүргізілгендігінің куәсі. Шынында да, бұдан бұрын ешбір ұлттық лексикография мұндай қарқынмен және сапамен сөздік құрастырмағаны белгілі [3, 106-107 бб.].

Институттың көп жылдар бойы жинақталып, қол күшімен жазылып, сөз қоржынында сақталынған қазақ тілінің тіл байлығы деуге болатындай, қазақ тілінің картотекалық қоры (көлемі 5 млн. жуық) бар. Осы картотекалық жинақталған қорды саны мен сапасы жағынан сұрыпталып, компьютер жадына түсіріліп, автоматты түрде қолдану ісіне кірісу А.Байтұрсынұлы атындағы Тіл білімі институтында ертеден-ақ қолға алынған болатын. Аталған жұмыстың қажеттігі тек бүгінгі күнге ғана емес, оның келешегінің маңыздылығында екендігін ғалымдар сол кезде-ақ жақсы түсінді. Өйткені осындай қомақты көлемдегі мәнмәтін қорын автоматты жүйеге айналдырмайынша, мұны тілші қауымдардың ойдағыдай қолдану мүмкіндігі, сол күйінде сақтап тұруы немесе үнемі жаңартылып отыруы оңайлыққа түспейді.

Лексикографиядағы автоматтандыру мәселелері көпәспектiлi болғандықтан, сөздіктердің нақты түрлеріне және лексикографиялық жұмыс түрлеріне қатысты талқылаған жөн. Алайда, кез-келген сөздіктің негізі, оған қажетті лексикалық материал картотека екендігі жалпыға белгілі. Ұлттық тіл сөздіктерін құрастыру кезінде тек миллиондаған картотекалық қорымыз репрезентативті, фундаменталды база ретінде таптырмайтын материал болып табылады.

А.Байтұрсынұлы атындағы Тіл білімі институтының картотекалық қорында түрлі сөздік және филологиялық жұмыстарды жүргізуге мүмкіндік беретін дәйексөздері бар көптеген сөздік карталар жиналып, жүйеленген, шамамен 5 миллион кәртiшке бар.

Картотеканың дереккөздері нақтылығымен, стильдік жағынан алуан түрлілігімен, арнайы іріктеліп таңдалған және өңделген мәтіндер болғандықтан құндылығын ешқашан жоғалтпақ емес.

Картотека материалдары қазақ тілінің көптеген сөздіктері мен грамматикаларын дайындауда қолданылып келеді. Бұл картотеканы біздің еліміздің әр түрлі қалаларының және шетелден келген ғалымдар қазақ тіл білімінің әртүрлі мәселелерін зерттеу үшін пайдалана алады.

Бүгінгі таңда оның көмегімен қазақ тілінің үлкен академиялық, түсіндірме сөздіктері және т. б. жасалды.

Ұлттық тілдің картотекалық қорын жасау – кез келген сөздік құрастырудың алғышарты. Орыс тілінде де Ресей Ғылым Академиясында жарық көрген түсіндірме сөздіктер құрастыруға 8 миллионнан тұратын картотекалық қор жинақталған [4, 22 б.]. Картотекалық қорды жасау ісі ХІХ ғасырда академик Я.К.Грот пен А.А.Шахматовтың жетекшілігімен басталған. Алғашқыда 1-кезеңде 2,5 миллион кәртiшке жинақталып, екінші кезеңде қор

тағы да толықтырылып отырған. Алғашқы кезең 1886-1968 жылдар аралығындағы мәтіндерді қамтыса, келесі кезеңде 1968-1994 жылдарды қамтиды. Бұдан басқа да әр кезеңдегі мәтіндер мен диалектілік, стильдік ерекшеліктер бойынша да кәртiшкі жинақтау жұмысы атқарылып отырған. Орыс халқының мұндай картотекалық қорға ерекше назар аударғанын 1986 және 2001 жылы осы мәселеге байланысты өткен ғылыми конференцияның ұйымдастырылуынан және баяндамашылардың сөзінен байқауға болады. Осы конференцияларда орыс тілінің картотекалық қоры мемлекет тарапынан қорғалуға тиісті бағалы мәдени ескерткіш екендігі атап көрсетіледі. Мәселен, В.П.Захаров картотеканың құндылығы туралы былай дейді: «Ең алдымен, үлкен академиялық картотека дерекқоры пайдаланушылардың картотекамен жұмысын жеңілдетіп, зерттеушілерге жаңа мүмкіндіктерді беруге арналған. Үлкен картотекалық қор – бұл ағаш сөрелерінде, каталог қораптарында сақталған көптеген қимақағаздардың жиынтығы. Соңғы уақыттарда үлкен академиялық картотека үш рет көшті. Сол себепті көп жылдық картотекалық қол еңбегі жұмыстарының, кәртiшкелердің әліпбилік ретімен бұзылуы, сөздердің жоғалуы және т.б. анықталады [5, 26 б.].

В.П.Захаровтың тұжырымдауынша картотеканың интеграцияланған дерекқорын құру және әртүрлі жүйелер бойынша жасалынған картотекалық дерекқорын реттеуге мүмкіндік береді. Демек қашан да картотекалық дерекқордың жұмыстарын реттеп, толықтырып, цифрландырып отыру ғылыми зерттеушілер үшін маңызды.

Зерттеуші М.Н. Приемышева картотеканы лексикографиялық және ғылыми-зерттеу жұмыстарының дереккөзі екенін ерекше атап көрсетеді [6, 23-28 бб.].

XIX ғасырдың екінші жартысында қазақ тілінің түсіндірме сөздіктерін құрастыру мәселесі алға қойылды. Сөздік құрастыру кезінде ең алдымен әр сөздің мағынасын дәйектейтін мысалдар қажет болды. Сонымен қатар әртүрлі стильдегі мәтіндерден (шығарма, газет т.б.) әрбір сөйлемді кәртiшкіге тіркеп, мысал беруде жаңа сөздер кездесіп отырды. Жаңа сөздер – қазақ тілінің сөздік қорын толықтырудың негізгі жолы еді. Кәртiшке жинақтау жұмысын үйлестіру кезінде кітаптарды кәртiшкеге түсіру жұмысына қалың көпшілік қатысқан. Әр кәртiшкеге тиын-тебен төлеп ынталандыру арқылы халықты жұмылдырған. Қазіргі Институт қоржынындағы 5 миллионнан аса картотекалық қор – осындай қалың көпшілік қауымның атсалысуымен, маңдай терімен жинақталған ұлттық қазынамыз.

Тіл білімі институтында осы 5 миллионнан аса кәртiшке мысалдары негізінде қазақ тілінің түсіндірме сөздіктері құрастырылды.

1968 жылы жарық көрген «Абай тілінің сөздігі» [7, 112 б.] мәтіні (6 мың сөз, 1 мыңдай сөз тіркес), «Қазақ тілінің 15 томдық түсіндірме сөздігі» [8].

1999 жылы баспада жарық көрген «Қазақ тілінің сөздігі» (50 мың сөз) және 15 томдық «Қазақ әдеби тілінің сөздігі», 1 томдық «Қазақ» сөздігі (106 000 сөз), сол сияқты салалық екітілді (қазақша-орысша, орысша-қазақша) терминологиялық сөздіктер мен қазақ тіліндегі салалық терминологиялық түсіндірме сөздіктері жарық көрді.

Нәтижелер және талқылау

Электронды компьютерлік қор жасаудың негізгі мақсаты – ғылыми зерттеуші тілшілердің жұмыс орнын толықтай автоматтандыру болып табылады. Осы автоматтандырудың негізінде зерттеушілердің, ғалымдардың уақыттарын мейлінше үнемдеу, яғни зерттеуге қажетті тілдік ақпараттарды іздеу жұмысы мен тілдік бірліктерді жинау жұмыстары жеңілдейді. Бұны тілші, ғалым – А.Жұбановтың: «Қазақ тілінің автоматты лексикографиялық базалары тілші-лексикографтарды немесе кез келген тұтынушыны қажетті деген ақпаратпен қамтамасыз ету және тілдік материалдарды іздеп табу әрекетін барынша жеңілдететіні сөзсіз» деген тұжырымдары дәйектей түседі [9, 41 б.].

Бұрынғы «қолмен» жиналған қазақ тілінің картотекалық қорын тұтынушы тілшілерге пайдалану аса ыңғайлы емес. Осындай қағаз күйіндегі мол қорды болашақ пайдаланушы жұртшылыққа сол қалпында сақтап немесе үнемі жаңартып тұру мәселелері де оңайлыққа түспейді.

«Қолмен» жасақталған 5 миллионнан аса картотека материалдарын 2010 жылдары профессор А.Жұбанов «Тіл – қазына» атты деректер базасына салуды жоспарлаған болатын. Жұмыс барысы туралы ғалым былай деп жазады: «Өткен ғасырдың 90-шы жылдарынан бастап назар аударылып келген қазақ тілінің «Қазына» атты цифрландырылған картотекалық қордың маңызы айтарлықтай зор деп есептеуге болады. 1991-1993 жылдар аралығында «Қолданбалы лингвистика» бөлімінің қызметкерлерінің қатысуымен «Қазына» картотекалық қорына 900 мың кәртiшкі-мәнмәтiн iшiнен 70 мыңдай кәртiшкенi сұрыптап алып, оларды компьютердiң жадына енгiзген болатын».

Бұрынғы «Қазына» қоры кейіннен жаңа нұсқада қайта өңделіп, «Тіл – қазына» деп аталды. «Тіл – қазына» бағдарламасына әліпбилік ретпен «А» әрпінен басталып, «С» әрпіне дейінгі кәртiшкелер сұрыпталып, енгiзiлiп отырды. Кәртiшкелер қайта өңделген «Тiл – қазына» компьютерлiк базасына енгiзiлiп, лингвист-филологтар үшiн керектi деген ақпаратты iздеу, тiлдiк бiрлiктердi табу жұмыстары бұрынғыға қарағанда әлдеқайда қолжетiмдi болды.

А.Жұбанов өзі үйлестіріп, атаған «Тіл – қазына» атты қазақ сөзінің картотекалық компьютерлік қорының негізгі мүмкіндіктеріне төмендегідей қысқаша сипаттама берген. «Олар:

– қажетті деген сөзге немесе сөзформаға қатысты кәртiшкі-мәнмәтiндердi лезде тауып алып, олармен компьютердiң экраны арқылы немесе қағаз бетiне басып шығарып та танысу мүмкiндiгi;

– әрбiр әрiпке байланысты реестрлiк сөздiк түзiлгендiктен, қажет деген сөздер мен оларға қатысты кәртiшкі-мәнмәтiндердi лезде iздеп тауып алу мүмкiндiгi. Мұндай әрекеттi кәртiшкі-мәнмәтiн iшiндегi сөзқолданыстар бойынша да жүзеге асыру алдын ала қарастырылған;

– қажетті деген сөзге немесе сөзформаға (олардың бастапқы, ортаңғы, соңғы бөліктеріне) байланысты кәртiшкі-мәнмәтiндердi авторына, шығармасына, баспа аттары мен шыққан жылдарына қатысты іздеп табу мүмкiндiгi;

– «Тiл – қазына» атты қазақ сөзiнiң деректер базасындағы кәртiшкі-мәнмәтiндердi өндеуге, өшiруге немесе жаңадан қосуға да мол мүмкiндiк бар» [9, 52 б.].

Ғалым А.Жұбанов «Тiл – қазына» деп атаған компьютерлiк қоры Қазақ тiлiнiң 15 томдық түсiндiрме сөздiгiн құрастыру кезiнде әзiрлемешiлердi мысалдар және бұрынғы 10 томдық түсiндiрме сөздiкте жоқ жана сөздердi картотекалық базадан тауып беру сияқты жұмыстармен қамтамасыз еттi. Алайда, өкiнiшке орай, бұл картотекалық қордың электронды базасы абайсызда техникалық себептермен жойылып кетедi. Бiрақ осы картотекалық қорды электронды базаға салуда ерен еңбек еткен Асқар Жұбанов оны қалпына келтiруге өмiрiнiң соңына дейiн күш салды, Министрлiкке, Ғылым комитетiне, Институт әкiмшiлiгiне т.б. өкiлеттi органдарға картотеканы цифрландыру кезек күттiрмейтiн мәселе екенiн, оған арнайы қаражат бөлу керектiгiн жеткiзуге тырысты. 2020-2022 жылдарға арналған «Мемлекеттiк тiлдi латынграфикалы ұлттық әлiпбиге көшiрудi қамтамасыз ететiн жаңартылған нормативтiк сөздiктер мен академиялық басылымдар сериясын әзiрлеу» жобасы аясында Институт директоры А.Фазылжанованың басқаруымен картотекалық қорды цифрландыру iсi жалғасын табуда.

Картотекалық қорды цифрландыру – Қазақ тiлiнiң ұлттық корпусының сөздiк және мәтiндiк базасын толықтырудың бiр көзi болып табылады. Өйткенi тiл бiлiмiнде қазiргi инновациялық база – Ұлттық корпус тар осы мақалада арнайы сөз болып отырған картотекалық қордың жетiлген түрi болып табылады. Екеуi де – ғылыми-зерттеу жүргiзудiң, сөздiктер құрастырудың дереккөзi болып табылады.

Бұрындары, тiптi қазақ тiлiнде қазiрге дейiн қолмен көшiрiлiп алынған кәртiшкелердi тiлдiк ғылыми зерттеу жұмыстарында зерттеушiлер пайдаланып келгенi көпшiлiкке мәлiм.

Мұндай тiлдiк мәлiметтердiң жиынтығын кәртiшкелiк қор деп атап, тiлдi зерттеушiлер тiлдiк бiрлiктердi жинап, өздерiнiң ғылыми зерттеу еңбектерiнде пайдаланып келген. Зерттеушi ғалымдар зерттеу тақырыптарына байланысты уақытының көбiн қажет деген мысалдарды жинауға және оларды кәртiшкеге көшiрiп жазып, мысал алынған авторы мен кiтаптардың атын, кiтаптың бетiн, баспаның атын, шыққан жылын және тағы басқа да ақпараттарды көрсетiп, керектi деген сұрыптаудан өткiзгеннен кейiн, зерттеушi оны өздерiнiң қажеттiлiгiне пайдалануға жұмсайтын.

Жаңа технологиялық мүмкiндiктер сөздiктердiң автоматтанған картотекасы тұрғысынан жаңаша құрастыру қажеттiгiне әкелдi. Мәселен, сөздiк картотекасына қатысты практикалық қажеттiктен туындаған мынадай әрекеттер орын алады: сөзтiзбенi дайындау, контекстердi тандап алу мен iздеу, сөзқолданыстарды (бiр немесе бiр топ сөзге қатысты) салыстырмалы талдау,

сөздік жобасына байланысты иллюстрациялық контекстер мен цитаталарды іріктеу, сөзге қатысты стильдік таңба жүйесін таңдау және т.б. [10, 14-19 бб.].

Корпус мәселесі тілшілерге таныс дәстүрлі картотека қорын жасау мен қолдану әрекетінің заңды жалғасы деуге болады. XX ғасырда картотекалар компьютер жадына енгізіліп, электрондық қорға айналып, көпшілік қолдануына мүмкіндік жасалды.

Технологиялық құралдар ретінде заманауи компьютерлік технологияларды қолданатын тіл білімінің басқа да салаларымен қатар, корпуслық лингвистика тілдік материалдарды жинау мен сақтаудың дәстүрлі әдіс-тәсілдеріне – жазбаша мәтіндерді қол күшімен өңдеуден өткізіп, лұғаттық кәртішкелерді жинау мен сақтау және тағы басқа да түрлеріне тікелей әсер етті. Дәстүрлі тәсілдердің кемшілік жақтарына жататындар: аса көп уақыт шығыны, зерттеу нысанының көлем жағынан шектелуі, жаңартудың күрделілігі, статистикалық өңдеудің қолайсыздығы, қашықтық жағдайындағы (дистанционный) ақпарат алмасу мүмкіндігінің болмауы. Тілдік зерттеулерге қажетті материалдарды жинау мен сақтауға қатысты компьютерлік технологияны пайдалану жаңа тәсілдің артықшылығын толығымен пайдалануға мүмкіндік тудырды. Атап айтқанда, тілдік материалға «қол жеткізу» жедел түрде іске асады, интернет желісін пайдалану ақпарат алмасу қашықтығын жояды, ақпарат көлемі талай мәрте өседі, деректерді сақтайтын және өңдейтін компьютерлік бағдарламалар арқылы қажет деген ақпаратпен танысу мен оны өңдеу әрекеттері барынша тез жүзеге асады [11, 24 б.].

Сол сияқты, корпуслық әдістің қалыптасуына интернет желісі айтарлықтай рөл атқарды. Мысалы, корпус жадына енгізілетін мәтіндердің көлемін және әртүрлі лингвистикалық зерттеулерге қажетті мәтін түрлерін арттыруға мүмкіндік туды. Аса кең электронды ресурстарды пайдалану мүмкіндігі лингвистикалық зерттеулердегі қажетті материалдарды жинау әрекетін айтарлықтай жеңілдетті [12, 9 б.].

Корпуста пайдаланушы қажеттігіне қарай метамәтіндік және лингвистикалық белгіленімдер қойылады. Ол белгіленімдер пайдалану кезінде автоматты іздеу бағдарламасына салынады. Метабелгіленім дегеніміз – мәтіннің авторы, тақырыбы, жарық көрген уақыты, т.б. ондаған параметр бойынша жасалады. Бұл параметрлер бойынша пайдаланушылар өздеріне қажетті мәтінді, ондағы сөзді аталған параметрлер (авторы, жылы, тақырыбы, стилі, жанры т.б.) бойынша іздеу жасай алады. Бұл зерттеушінің не сөздік құрастырушының материалды тез тауып алуына мүмкіндік береді.

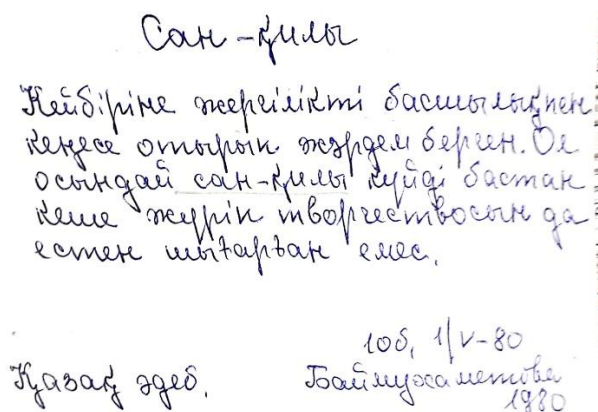
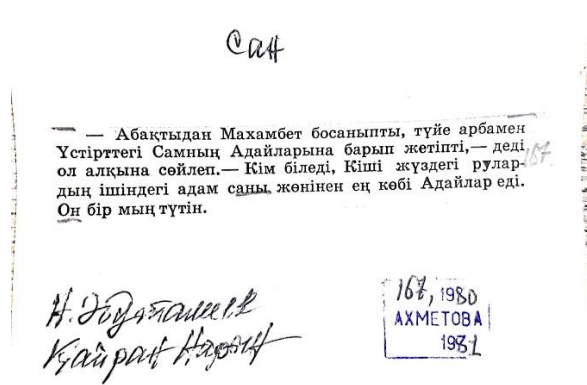
Картотекалық қор – қазіргі уақытта Ұлттық корпустар базасының мәтіндік базасын толықтырушы мысалдар массиві болып табылады. Картотекалық қорды да электронды пішінге келтіріп, автоматтандыру қазіргі Ұлттық корпустардың жұмыс істеу механизмімен өте ұқсас. Десек те картотекалық қорды цифрландыру Ұлттық корпус құрастыру ісімен қатар жүре береді. Өйткені картотекалық қор – қазақ тілінде әр кезеңде, әртүрлі стильде жарық көрген мәтіндер. Ол тарихи дереккөз болады. Қазіргі Ұлттық корпустар жасалып жатқан дамыған технология заманында картотекалық қорды корпустардағы іздеу жүйесімен (поисковик) жетілдіру қажет. Осы ретте

қазіргі таңда А.Байтұрсынұлы атындағы Тіл білімі институтында бұл игілікті жұмысты мына кезеңдер бойынша атқару жоспарланған:

- 1) Картотекалық қорды санитарлық тазалаудан өткізу;
- 2) Кәртiшкелердi сұрыптау;
- 3) Кәртiшкелердi сканерлеу;
- 4) Сканерленген картотекалық қордың электронды (word) нұсқасын әзірлеу;
- 4) Картотекалық базаның іздеу жүйесін жасау;
- 5) Картотекалық деректер қорының сайтыን әзірлеу;
- 6) Ақпаратқа еркін қол жеткізуді қамтамасыз ету.

Институттағы 5 миллионнан асатын картотекалық қордан алынған кәртiшкелердi 1-суреттен көруге болады.

Сурет - 1. Сөздік кәртiшкелердiң үлгілері



Бұл суретте көріп отырғанымыздай, ең алғашқы кәртiшкелер осылай қол күшімен жасалған. Бұдан өте үлкен еңбекпен жасалған анықтамалық кәртiшке қоры жиналған.

Қазіргі таңда осы 5 миллионға жуық картотекалық сөздік қорды цифрландыру мәселелері талқылану үстінде. Картотеканы цифрландыру жұмысы – лексикографиялық сөздіктер жасауға, сонымен қатар басқа да ғылыми зерттеу жұмыстарына, арнайы корпустарды құруда да орны мен рөлі айрықша. Осы ретте онлайн режимінде ашық картотеканы құру және жүргізу идеясы туындайды.

Картотекалық қорды цифрландырудың алғашқы нұсқасы «Тіл – қазына» бағдарламасы арқылы жүргізілді. Бұл бағдарламаның басқы терезесіне жаңа сөзді базаға енгіземіз, одан кейін қажетті мәліметтерді толтырғаннан кейін, оларды «сақтау» батырмасын басу арқылы сақтап отырамыз. Енгізілген мәліметтерді тұтынушылар өз қалауына қарай оны қағаз бетіне шығара алады.

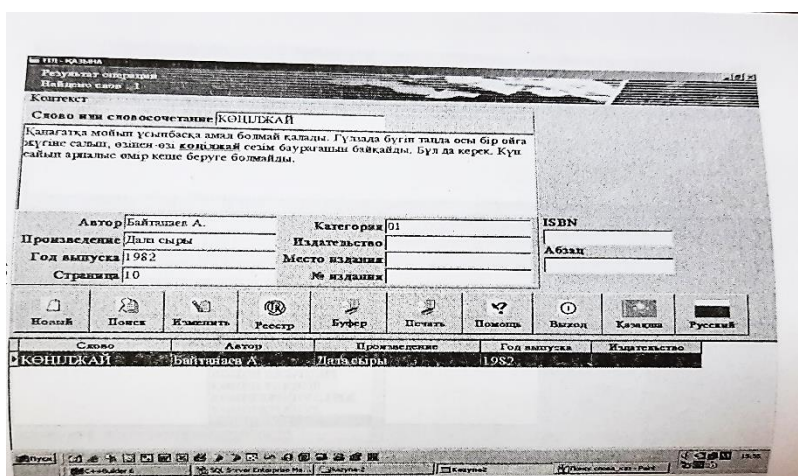
Ақпаратты (кәртiшкенi) iздестiру оның бiрнеше параметрлерi арқылы жүргiзуге болады:

- сөз немесе сөз тіркесі бойынша;
- мәнмәтін бойынша;
- авторы бойынша;
- шығарманың аты бойынша;
- сөздің категориясы бойынша;
- баспа аты бойынша;
- басылымның жылы бойынша.

Деректерді (кәртiшкелердi) базадан iздестiру «iздеу» атты батырманы басу арқылы жүзеге асады [10, 14 б.].

Бұл бағдарламада iздестiрiлген сөз немесе сөз тiркесi база iшiнен табылған соң, оған қатысты деректер мынадай кәртiшке түрiнде экран бетiне шығады (сурет-2).

2-сурет. Тіл – қазына бағдарламасы



Бұл бағдарламаға деректер базасын енгізу кезінде алдымен кәртiшкелердi стильдiк ерекшелiктерiне, мысалдың жазылған жылына, қолданыс аясына қарай сұрыптап, iрiктеп алып, енгiзiлiп отырды.

Қазiргi таңда картотекалық деректер базасын цифрландырудың тағы бiр жолы –кәртiшкелердi сканерлеу жұмыстары жүргiзiлiп отыр. Картотекалық деректер базасын «Тiл – қазына» автоматтандыру бағдарламасында iрiктелiп, сұрыпталып алынып отырса, бұл цифрландыруда жаппай бүкiл кәртiшкелер сканерленiп жатыр. Алдағы уақытта бұл сканерленген кәртiшкелердi онлайн режимiнде «iздеу» жүйесi арқылы сурет күйiнде қажеттi деген сөз немесе сөз тiркесiн iздеп табатын арнайы бағдарлама дайындау көзделiп отыр. Iздеген сөздi немесе сөз тiркесiн базадан тек сурет форматында, яғни кәртiшке түрiнде ғана экранға шығарып бере алатын бағдарлама жасауды қолға алу

жоспарланып отыр. Кәртiшкенi тек экран арқылы көзбен көрiп және қолмен қағазға жазып ала аламыз. Автоматтандырудың бұл әдiсi де өз кезегiнде зерттеушiлер үшiн қызмет атқарады. Алайда болашақта картотекалық деректер базасын цифрландыруды уақыт өте келе жетiлдiру басты мiндеттердiң қатарына енiп отыр.

Картотекалық қорды цифрландыру оңай жұмыс емес, өйткенi ерте заманнан берi тұрған кәртiшкiлер мен жәшiктер шаң басып, санитарлық талаптарға сай келмейдi. Ол үшiн, Ұлттық кiтапханадағыдай қағаз кәртiшкелердi арнайы шаңнан және т.б. зиянкестерден үздiксiз залалсыздандырып, дезинфекциялап, қайта өңдеп отыру, қағазға қаламмен жазылған әртүрлi сиялардың өшiп қалмауын қамтамасыз етiп отыратын арнайы сақтау орындары болу қажет.

Оның үстiне бiр сөздiң өзiне жүздеген мысал сөйлем берiлген кәртiшкелердi *мағынасын, мысал авторына, стилiне* т.б. сұрыптау өте көп уақытты алады және бiлiктiлiктi талап етедi. Бұл болашақта атқарылатын уақыттың еншiсiнде. Осы ретте әзiрге кәртiшкелердi сұрыптамай, жиналған мысалдарды түгелдей электронды базаға салу қажеттiгi тұр. Десек те, қазiргi уақытта алдымен сканерлеу арқылы базаға енгiзу жұмыстары атқарылуда. Қазақ тiлiнiң ұлттық корпусы жобасы аясында сканер кәртiшкелердi алдағы уақытта «уорд» нұсқаға келтiру жағы да ойластырылып жатыр. Сонымен қатар сканер суреттердi модельдеу, «уорд» нұсқа жасалса, түрлi программалық iздеу интерфейсi әзiрленуде.

Осындай үлкен көлемдегi картотекалық мәнмәтiн қорын цифрландырмайынша, тiлшi қауымның ойдағыдай пайдалану мүмкiндiгi оңайлыққа түспейдi. Мiне, осындай себептерге байланысты сонау жылдардан бастап қолға алынған картотекалық қорды автоматтандырудың маңызы айтарлықтай зор. Мұндай компьютерлiк қорды цифрландырудың негiзгi мақсаты – зерттеушiлердiң уақыттарын мейлiнше үнемдеу, яғни қажеттi тiлдiк ақпаратты iздеу жұмысы мен тiлдiк бiрлiктердi реттеу т.б. iс-әрекеттердiң жеңiлдейтiнi сөзсiз.

Қорытынды

Электронды картотека технологиясы талдауға алынатын мәтiндегi сөзқолданыстардың тұтастығын сақтауда және сөздердi iрiктеуде жүйелiлiк пен толықтықты қамтамасыз етедi. Басқару жүйесiн *деректердiң мәтiндiк базасына* (ДМБ) қолдану мен оларға терминал дисплейi арқылы «кiру» мүмкiндiктерi базаны автоматтанған лексика-статистикалық анықтағыш ретiнде әрекет етуiне жағдай жасайды. Автоматтанған лексика-статистикалық анықтағыш – ол әрi сөзнұсқағыш, әрi жиiлiк сөздiк, әрi компьютер жадында тұрақты сақталатын мәтiндердiң конкордансы болады.

Алғашында тiрнектеп жинақталған картотекалық қорды цифрландыру болашақта тiлшiлердiң, тарихшылар мен әлеуметтанушылардың зерттеу нысандарына аса құнды материал болуы мүмкiн. Қазiргi таңда анықтағыштарды, сөздiктердi, энциклопедияларды, оқулықтар мен оқу құралдарын, жинақтарды, әдеби және ғылыми-көпшiлiк шығармаларын т.б.

мәтіндерді компьютер жадында енгізе отырып сақтау технологиясын меңгеру өте маңызды болып отыр. Жаппай цифрландыру кезеңінде қоғамның әрбір мүшесі белгілі бір дәрежеде, компьютерді пайдаланушы, қолданушы болып саналатыны белгілі. Сондықтан қазақ тіл білімінің әлеуетін көтеруге, компьютер мүмкіндіктерін толығымен қолдану егемен еліміздің барынша дамуына қосқан үлесі зор болар еді. Қазіргі кезеңдегі ғылым мен технологияның түрлі салаларында ғылыми және өндірістік ақпараттармен алмасудың айтарлықтай ұлғаюына байланысты мейлінше кең ауқымды қамтитын көптілді сөздіктер шығару қажеттілігі туындауда.

Тұтынушы мәтіндік база мен бағдарламалық қамтамасыз етуді өз қалауынша өзгертіп, толықтырып отыруына және қажетті деген материалды немесе конкорданстың толық түрін (кері-әліпбилі сөздікті де) қажетті реттілікпен қағаз бетіне басып шығара алады. Сонымен бірге ол тілдік деректерді редакциялау мен түзетулер жүргізудің жүйелік құралдарын пайдалану арқылы бастапқы, аралық және соңғы сатыдағы мәтіндерге кез келген түзетулер енгізуіне немесе оларды жаңадан туындатуына мүмкіндігі мол.

ӘДЕБИЕТ

- [1] Смолян Г.Л. Человек и компьютер: Социально-философские аспекты автоматизации управления и обработки информации. –М.: Политиздат, 1981. –192 с.
- [2] Busa R. Index Thomisticus: Sancti Thomae Aquinatis Operum Omnium Indices et Concordantiae. Section Secunda. Concordantiae Operum Thomisticorum – Concordantia Prima. Stuttgart, 1974. Book
- [3] Вертель В.А., Вертель Е.В., Рогожникова Р.П. К вопросу об автоматизации лексикографических работ (некоторые результаты применения ЭВМ) // –ВЯ. –1980. –№ 1. – 208 с.
- [4] Большой академический словарь русского языка. Т. 1-6. –СПб.: 2004-2006.
- [5] Захаров В.П. Компьютерная модель Большой словарной картотеки Института лингвистических исследований РАН // Технологии информационного общества - Интернет и современное общество: Труды VIII Всероссийской объединенной конференции (IST/IMS-2005). –СПб., 2005. –С. 33-34.).
- [6] Приемышева М.Н. Картотека как источник лексикографической и научной работы. // Журнал Лингвистической Петрополитаны. Труды Института лингвистических исследований / Отв. редактор Н.Н. Казанский. – СПб.: Наука, 2003. –С. 23-28.
- [7] Абай тілі сөздігі. –Алматы, «Ғылым» баспасы, 1968. – 729 б.
- [8] Қазақ тілінің 15 томдық түсіндірме сөздігі. – Алматы, 2011.
- [9] Жұбанов А.Қ. Қазақ сөзінің «Тіл – қазына» деректер базасы және оның теориялық негіздері. – Алматы: «Арыс» баспасы, 2009. – 304 бет.
- [10] Рогожникова Р.П. Большой Словарной картотеке 100 лет // Практическая лексикография. 100 лет словарной картотеке. –М., 1989. –С. 14-19.
- [11] Жұбанов А., Жаңабекова А. Корпустық лингвистика. – Алматы, 2016. – 336 б.
- [12] Перцов Н.В. О роли корпусов в лингвистических исследованиях // Труды международной конференции «Корпусная лингвистика – 2006». – СПб.: Изд-во С.-Петербур. ун-та, 2006. –С. 318-331.

REFERENCES

- [1] Smolyan G.L. Chelovek i kompiyter: Sosialno-filosovskie aspekty avtomatizatsii upravleniya i obrabotki informatsii (Man and computer: Socio-philosophical aspects of automation of management and information processing). M., Politizdat, 1981. 192 s. [In Rus.]
- [2] Busa R. Index Thomisticus: Sancti Thomae Aquinatis Operum Omnium Indices et Concordantiae. Section Secunda. Concordantiae Operum Thomisticorum – Concordantia Prima. Stuttgart, 1974.
- [3] Vertel V.A., Vertel E.V., Rogojnikova R.P. K voprosu ob avtomatizatsii leksikograficheskikh rabot (nekotorye rezultaty primeneniya EVM) (On the issue of automation of lexicographic works (some results of computer application) // VІA. 1980. № 1. 208 s. [In Rus.]
- [4] Bolshoi akademicheskii slovar russkogo yazyka (A large academic dictionary of the Russian language). T. 1-6. SPb.: 2004-2006. [In Rus.]
- [5] Zaharov V.P. Kompiyternaya model Bolshoi slovarnoi kartoteki Instituta lingvisticheskikh issledovaniy RAN (Computer model of a large dictionary file of the Institute of Linguistic Research of the Russian Academy of Sciences) // Tehnologii informatsionnogo obshestva - internet i sovremennoe obshestvo: Trudy VIII Vserossiiskoi obedinnoi konferentsii (IST/IMS-2005). SPb.: 2005. S. 33-34. [In Rus.]
- [6] Priemysheva M.N. Kartoteka kak istochnik leksikograficheskoi i nauchnoi raboty (The card file as a source of lexicographic and scientific work). // Zhurnal Lingvistika Petropolitana. Trudy Instituta lingvisticheskikh issledovaniy / Otv. redaktor N.N. Kazanskii. SPb.: Nauka, 2003. S. 23-28. [In Rus.]
- [7] Abai tılı sözdığı (Dictionary of the Abay language.). Almaty, «Ėylym» baspasy, 1968. 729 b. [In Kaz.]
- [8] «Qazaq tılınıń 15 tomdyq túsındırme sözdığı» («Explanatory dictionary of the Kazakh language in 15 volumes»). Almaty, 2011. [In Kaz.]
- [9] Jūbanov A.Q. Qazaq söziniń «Tıl – qazyna» derekter bazasy jáne onyń teoriatyq negizderi. (The database of the Kazakh word "Til-kazyna" and its theoretical foundations.). Almaty: «Arys» baspasy, 2009. 304 bet. [In Kaz.]
- [10] Rogozhnikova R.P. Bolshoi Slovarnoi kartoteka 100 let (A large Dictionary card index of 100 years). // Prakticheskaya leksikografiya. 100 let slovarnoi kartoteka. M.: 1989. S. 14-19. [In Rus.]
- [11] Jūbanov A., Jańabekova A. Korpustyq lingvistika (Corpus linguistics). Almaty, 2016. 336 b. [In Kaz.]
- [12] Persov N.V. O roli korpusov v lingvisticheskikh issledovaniyah (On the role of corpora in linguistic research) // Trudy mezhdynarodnoi konferentsii «Korpysnaiya lingvistika – 2006». SPb.: Izd-vo S. Peterb. un-ta, 2006. S. 318-331. [In Rus.]

ЦИФРОВИЗАЦИЯ КАРТОТЕЧНОГО ФОНДА – ОСНОВА СОЗДАНИЯ НАЦИОНАЛЬНОГО КОРПУСА КАЗАХСКОГО ЯЗЫКА

*Жанабекова А.А.¹, Кожахметова А.К.², Бесиров Е.Б.³

¹Доктор филологических наук, Институт языкознания имени А.Байтурсынулы, заведующий отделом Прикладной лингвистики, Алматы, Казахстан, e-mail: aiman_miras@mail.ru, <https://orcid.org/0000-0002-6199-7444>

²Докторант, Институт языкознания имени А.Байтурсынулы, младший научный сотрудник отдела Прикладной лингвистики, Алматы, Казахстан, e-mail: akony_8484@mail.ru, <https://orcid.org/0000-0003-3783-3199>

³Магистр, Институт языкознания имени А.Байтурсынулы, младший научный сотрудник отдела Прикладной лингвистики, Алматы, Казахстан, e-mail: erkin.besirov@mail.ru, <https://orcid.org/0000-0002-2291-9654>

Аннотация. В статье рассматриваются научно-практические вопросы оцифровки картотечного фонда, приведения его в автоматическую словарную базу-совокупность примеров по различным стилям казахского языка, являющуюся источником, иллюстративной базой для составления толковых словарей на казахском языке. Дается обзор практики создания картотечного фонда русского языка. Приводятся мнения о важности создания картотечного фонда. Рассказывается история формирования картотечного фонда на казахском языке, начавшаяся в середине XIX века путем ручной работы. Также дается характеристика работы над базой данных «Тіл – қазына», которая реализовалась под руководством профессора А. Жубанова в начале XX века. Указывается механизм функционирования автоматической, электронной картотечной базы, и тем самым определяются возможности пользователя. Подтверждается, что картотечный фонд является основой создания лексикографической базы, т. е. составления толковых словарей. Дан обзор толковых словарей, составленных на основе картотечного фонда казахского языка.

При обосновании необходимости цифровизации картотечного фонда казахского языка, при изучении работ прикладной направленности в мировом языкознании используются методы обзорного, описательного, повествовательного, аналитического анализа. При определении механизмов преобразования картотечного фонда в автоматический словарь, в электронную иллюстрированную базу реализован метод алгоритмического программирования.

Источники картотеки никогда не утратят ценности благодаря своей точности, стилистическому разнообразию, специально подобранным и обработанным текстам.

Картотека используется при составлении многих словарей и грамматик казахского языка. Данная картотека может быть использована казахстанскими учеными и учеными зарубежья для изучения различных вопросов казахского языкознания.

Ключевые слова: картотека, автоматизация, оцифровка, электронная база, картотечный фонд, корпус, лексикография, прикладная лингвистика

Статья опубликована в рамках программно-целевой исследовательской работы OR11465483 «Разработка серии обновленных нормативных словарей и академических изданий, обеспечивающих перевод государственного языка на латинографический национальный алфавит».

DIGITALIZATION OF THE CARD FILE FUND – THE BASICS OF CREATING THE NATIONAL CORPUS OF THE KAZAKH LANGUAGE

*Zhanabekova AA.¹, Kozhahmetova A.², Bessirov Y.B.³

*¹Doctor of Philology, Institute of Linguistics named after A. Baitursynuly, Head of the Department of Applied Linguistics, Almaty, Kazakhstan, e-mail: aiman_miras@mail.ru, <https://orcid.org/0000-0002-6199-7444>

²Doctoral student, Institute of Linguistics named after A. Baitursynuly, Junior Researcher of the Department of Applied Linguistics, Almaty, Kazakhstan,

e-mail: akony_8484@mail.ru, <https://orcid.org/0000-0003-3783-3199>

³Master's degree, Institute of Linguistics named after A. Baitursynuly, Junior Researcher of the Department of Applied Linguistics, Almaty, Kazakhstan,

e-mail: erkin.besirov@mail.ru, <https://orcid.org/0000-0002-2291-9654>

Abstract. The article deals with the scientific and practical issues of digitizing the card stock, bringing it into an automatic dictionary database—a set of examples on various styles of the Kazakh language, which are the source, illustrative base for compiling explanatory dictionaries in the Kazakh language. The review of the practice of creating a file fund of the Russian language is given. Opinions are given on the importance of creating a file fund. The history of the formation of the card file fund, which began in the middle of the nineteenth century with manual work, is told in the Kazakh language. The work of the database "Til – kazyna" under the guidance of Professor A. Zhubanov at the beginning of the twentieth century is also characterized. The mechanism of functioning of the automatic, electronic file database is indicated and thereby the user's capabilities are determined. It is confirmed that the card stock is the main condition for the creation of a lexicographic base, i.e. the compilation of explanatory dictionaries. Review of explanatory dictionaries compiled on the basis of the Kazakh language file fund.

When justifying the need for digitalization of the Kazakh language file fund, when studying works of applied orientation in world linguistics in this direction, methods of overview, descriptive, narrative, analytical analysis are used. When determining the mechanisms for converting the card stock into an automatic dictionary, an electronic illustrated database, we rely on the method of algorithmic programming.

The sources of the card file will never lose their value due to their accuracy, stylistic diversity, specially selected and processed texts.

Card files are used in the compilation of many dictionaries and grammars of the Kazakh language. This card file can be used by scientists from different cities of our country and abroad to study various issues of Kazakh linguistics.

Keywords: card file, automation, digitization, electronic database, card file fund, corpus, lexicography, applied linguistics

The article was published as part of the program-targeted research work OR11465483 "Development of a series of updated normative dictionaries and academic publications providing translation of the state language into the Latin national alphabet".

Статья поступила 22.07.2022