

Қ.Қ. Рысберген, Н.М. Рсалиева
З.Б. Садырбаева, Д.М. Пашан, Ә.Б. Сейтбатқал*

А. Байтұрсынұлы атындағы Тіл білімі институты, Алматы, Қазақстан
*Байланыс үшін автор: arlen_98@mail.ru

Ономастикалық ішкорпусты жасау және әлемдік тәжірибе*

Аңдатпа. А. Байтұрсынұлы атындағы Тіл білімі институты Ұлттық корпусның дербес бір бөлігі саналатын ономастикалық ішкорпусты топонимика саласы бойынша құрастырып жатыр. Мақалада ономастикалық ішкорпусты құрастырудың бағыты, ұстанымдары талқыланады. Сол арқылы ономастикалық ішкорпусты құрастырудың негізгі критерийлері айқындалады. Қазақ тілінің Ұлттық корпусы мен әлемдегі ірі-ірі корпусардың тәжірибесі, ерекшеліктері сипатталады. Жұмыстың басты мақсаты ономастикалық белгіленімдердің түрлерін айқындауға қатысты. Ағылшын, орыс тілінің корпусарындағы топонимдердің белгіленімі, лингвистикалық және экстралингвистикалық ақпараттың берілу тәсілі, ономастикалық терминдердің кездесу жиілігі анықталады. Осы талдау нәтижесінде топонимдік белгіленімдер әзірлеуге қатысты салыстырулар жасалып, ұсыныстар айтылады. Шетелдік корпусар тәжірибесінен байқағанымыздай, неғұрлым ішкорпустар көп болса, соғұрлым ұлттық корпус кемелдене түседі. Ономастикалық ішкорпустың пайда болуы ономаст ғалымдардың зерттеулеріне көптеген жаңа мүмкіндіктер алып келеді. Өйткені корпусар жер-су аттары кездесетін мәтіндерді жылдам табуға мүмкіндік туады. Бұл тек іздеу жұмысын жеңілдетіп қана қоймай, ономастика саласындағы кешенді зерттеулерге де өз септігін тигізеді. Сонымен қатар, ішкорпустың көмегімен ономастикалық сөздіктердің жаңартылған нұсқасын басып шығару ісі де жылдам іске асырылатын болады. Ономастикалық ішкорпус географиялық атауларды тек бекіту және стандарттау тұрғысынан ғана емес, сонымен бірге топонимдерді жүйелеу тұрғысынан, оларды сақтау және ғылыми зерттеу мақсатында осы ұлттық қорға қол жеткізу және танымдық қызығушылықты қанағаттандыру тұрғысынан да сөзсіз қажет. Мақаладағы ғылыми-теориялық ұсыныстар ономастикалық ішкорпустың белгіленімдерін жетілдіруге ықпал етеді, сондай-ақ, лингвистикалық зерттеулерді жүргізудегі ономастикалық ішкорпустың рөлін айқындайды. Топонимдер қорымен басталған ономастикалық ішкорпустың ауқымы басқа да оним түрлерімен кеңейе түседі деп күтіледі.

Кілт сөздер: Корпустық лингвистика, корпусық ономастика, ономастикалық ішкорпус, топонимдер, qazcorpus.kz, қазақ тілін цифрлендіру

*Мақала ҚР БҒМ қаржыландырып отырған BR11765619 «Мемлекеттік тілдің ақпараттық-инновациялық базасы ретіндегі қазақ тілінің ұлттық корпусын әзірлеу: ғылыми-зерттеу және оқыту интернет-ресурсы» іргелі-зерттеу бағдарламасы аясында орындалды.

DOI: <https://doi.org/10.32523/2616-678X-2022-139-2-8-19>

Кіріспе

Қазіргі ғылым мен білім берудің т.б. қоғамдық өмірдің барлық салаларында ІТ технологиялардың қолданылуы күн озған сайын жеделдеп дамуда. Осыған орай заманауи технологиялардың жетістіктерін ғылыми зерттеулерде пайдалану және «Цифрлы Қазақстан» бағдарламасының талап қажеттіктері аясында мақалада қарастырып отырған қазақ тілін цифрлеу мәселесі өте өзекті болып отыр. Корпустық лингвистика – мәтіндік базаны түзумен, оны әзірлеумен және қолданумен айналысатын белсенді дамып келе жатқан отандық тіл білімінің жаңа бағыты. Қазіргі уақытта әртүрлі тілдерде көптеген корпустар бар, оларды қалыптастыру, белгілеу және статистикалық мәліметтерді шығару технологиясы жетілдіріліп келеді. Жалпы корпустық лингвистиканың негізгі өзегі әрі зерттеу объектісінің бірі ұлттық корпус болғандықтан, әр тілдегі корпус әзірлеу мәселелері өзекті болып отыр. Қазақ тілінің ұлттық корпусы да басқа тілдегі корпустармен салыстыра қарағанда өзіндік жолмен біршама дамып, лингвистикалық белгіленімдермен толығып келеді. Қазақ тілінің ұлттық корпусына жалқы есімдердің де базасы еніп, ономастикалық ішкорпус қажет белгіленімдермен толығыны анық. Ал корпустық ономастиканың қазіргі міндеті – әртүрлі жанрдағы мәтіндердің құрамындағы онимдерге белгіленім енгізу. Қазіргі таңда Қазақстан Республикасының ұлттық картографиялық-геодезиялық қоры жүйелеген 168000 (жүз алпыс сегіз мың) топонимдік атауды қамтитын microsoft access форматында сақталған электрондық база әзірленген [1]. Аталған база әр 6 ай сайын жаңа атаулармен жаңарып, ескі атаулары шегеріліп отырады. Алайда бұл база интернетте жүйеленбегендіктен, көпшілікке қолжетімсіз болып отыр. Корпустық ономастиканы дамыту міндеті тек топонимдерді ғана емес, сондай-ақ онимдердің басқа да түрлерін қамтитын, Қазақстанның ономастикалық қоры туралы жинақталған ақпаратты тек синхронды ғана емес, ретроспективті түрде де ұсыну үшін көп функциялы жалқы есімдер корпусын құруды

көздейді.

Тіл білімінің барлық саласы компьютерлік құралдармен өңделіп жатқан тұста ономастикалық ішкорпусты да құрастырудың маңызы ерекше. Қазақ тілінің Ұлттық корпусында (qazcorpus.kz), тіл бірліктерінің барлығы көрініс тауып, арнайы шартты таңба беріліп, аннотацияланған. Онда, qazcorpus.kz-те, тіл бірліктеріне (сөз, сөз форма, сөз тіркесі, сөйлем, мәтін) морфологиялық, фонетикалық, синтаксистік, семантикалық белгіленімдер жасалғаны белгілі. Алдағы уақытта ономастикалық атауларды да арнайы шартты белгілермен сипаттап беру – Ұлттық корпусты кемелдендіріп қана қоймай, ономастиканың дамуына да серпін береді. Ономастикалық ішкорпус – электронды түрдегі ономастикалық мәліметтер қамтылған ақпараттық-анықтамалық жүйе. Онда Қазақстанның топонимдері туралы лингвистикалық ақпаратпен қамтамасыз етілген мәтіндер жинағы қамтылған. Ономастикалық корпус Ұлттық корпустың бір бөлігі болғандықтан, оның пайдасы мен үлесі өзара сәйкесімді. Мәселен, корпустар арқылы тілдің дамуын бақылауға болады. Сол сияқты жер-су атауларының өзгерісін, мәтіндегі қолданысы мен кездесу жиілігін айқындауға болады. Сондай-ақ, тілдегі жалқы есімдердің өзгерісі тіркеліп, олардың кезең-кезеңге тиесілі қоры қалыптасады.

Зерттеу әдістері

Зерттеу жұмысында әлемдік корпустарға шолу жасалынады, олардағы ономастикалық атаулардың қалай аннотацияланғанына назар аударылады. Қазақстан Республикасының 2 облысы, атап айтқанда, Ақмола және Алматы облысының топонимдері ономастикалық корпустың бастамасы ретінде жинақталды. Оларды сипаттау, талдау жұмыстары жасалды және де алынған материалдар ономастикалық ішкорпус құруға негіз болды. Материалдар Қазақстан Республикасының географиялық атауларының мемлекеттік каталогынан [2] алынды. Біз құрастырып отырған ономастикалық ішкорпус – Ұлттық корпустың тәуелсіз бір бөлігі. Сондықтан да біз Ұлттық

корпустың тәжірибесі мен ұстанымдарына сүйене отырып, ономастикалық ішкорпусты құрастырамыз. Мұнда топонимдердің этимологиясы, семантикасы, грамматикасы, географиялық орналасуы жөнінде ақпараттар жинақталды. Осы тектес ақпараттар Т.Жанұзақ, А.Әбдірахманов, Б.Бияров сынды ономаст ғалымдар мен топонимика ғылымына қатысты энциклопедия, сөздіктер, географиялық анықтамалықтардан алынды. Мәселен, әрбір топонимнің атау ретінде қойылу себебі, мағынасы, этимологиясы мәдени семантика бөлігіне кіріп, осы топтың өзегін құрайды. Бұдан бөлек әрбір атауға қатысты тарихи, әдеби және мәдени-экономикалық т.б. қосалқы мәліметтер қамтылады.

Талқылау

Ана тіліміздегі сөздіктердің, энциклопедиялардың т.б. мәліметтердің көп болуы, әрі олардың электронды түрде болуы корпусты құрастыруға үлкен септігін тигізеді. Өйткені корпус дегеннің өзі мәліметтер қоры. Тіл білімінің жаңа салаларының бірі, корпустық лингвистика осындай мәтіндік корпустарды әзірлеумен, құрумен және пайдаланумен айналысады. А.Қ.Жұбанов және А.Ә.Жаңабекова осы саладағы ғалымдардың пікірін жинақтай келе, тілдік корпусқа төмендегідей анықтама береді. «Корпус дегеніміз – әр тілдегі электронды пішінге келтірілген, яғни бір басқару орталығынан автоматты түрде жұмыс істейтін, лингвистикалық ақпараттарды қамтитын мәтіндер жинағы» [3, 13]. Э.Д.Сүлейменова да осыған ұқсас түсініктеме береді. «Корпус – бұл мәтіндерді электронды түрде жинауға негізделген ақпараттық-анықтамалық жүйе. Анықтамадан айқындалып тұрғандай, біріншіден, корпус – бұл мәтіндерді электронды (цифрлендірілген) түрде жинау. Екіншіден, корпус анықтамалық-ақпараттық жүйе ретінде қолданыла алады.

Цифрлендірілген мәтіндердің базасы корпус белгісінің көмегімен арнайы өңделеді. Дәл осы белгі корпусты түбегейлі тілдің қолданылуы мен өзгеруі жайлы ақпараттың жаңа және объективті дереккөзіне айналдырады.

Белгі – мәтіндердің және олардың бірліктеріне тән қасиеттер жайлы арнайы лингвистикалық ақпарат беретін корпустың негізгі сипаттамасы. Корпустың белгісі дәлірек және әрқилырақ болған сайын оның әлеуметтік, ғылыми және оқыту құндылығы жоғарырақ болады». [4, 371] Ал The Oxford English Dictionary сөздігі корпусқа «Лингвистикалық талдауға негізделген жазбаша немесе ауызша материалдардың жиынтығы» деген анықтама береді [5].

Әлемдік лингвистикадағы ұлттық корпустардың ерекшеліктеріне қысқаша шолу жасаймыз. Ең жетілдірілген, әлеуеті жоғары әлемдік танымал корпустар қатарына Британ корпусы, чех тілінің корпусы жатады, орыс тілінің корпусы да едәуір жетілген деуге болады. Төменде ірі көлемді ұлттық корпустармен қатар әр алуан тілдердегі арнайы қызметтегі функционалды корпус түрлері бар. Корпустық лингвистика негізінен мәтіндік корпус жиынтығы болса, әлемде одан өзге алуан түрлі мультимедиялық (сериялдар, фильмдер, комикстер, мультфильмдер, түрлі сипаттағы телебағдарламалардың видеожазбалары және т.б. атай беруге болады) корпусты қамтиды. Мәселен, Колорадо университеті кітапханасынның көрсетуінше, тілдік, мәтіндік корпустардың ішінде Brown University Corpus of American English корпусы компьютермен қолданысқа арналған алғашқы цифрлендірілген корпус. Браун корпусы 1964 жылдан басталып жасалған.

Ал британдық ұлттық корпус (British National Corpus-BNC) – XX ғасырдың аяғынан бастап британдық ағылшын тілінің кең спектрін ауызша да, жазбаша да көрсетуге арналған көптеген көздерден алынған жазбаша және ауызша сөйлеу үлгілерінің 100 миллион сөздің жиынтығы.

Ағылшын тілді академиялық ортада қолданылатын өзге де көптеген ішкорпус түрлері бар, мәселен, BYU Corpora – Brigham Young University Corpora. CU – BYU Corpora тіркелген кітапхана болып табылады. Айрықша нұсқаулары мыналар: толық мазмұнға қол жеткізу үшін жүйеге кіру қажет. Тегін жазба үшін colorado.edu электрондық пошта мекенжайын қолдану арқылы тіркелуге болады. Дерек-

тер жиынтығы жүктеу үшін қолжетімді емес. Бұл дерекқорға кампустан тыс кіру үшін VPN пайдалану қажет. Corpora-ға мыналар кіреді: ғаламтордағы жаңалықтар (NOW), ғаламдық веб-ағылшыны (GloWbE), Уикипедия корпусы (Wikipedia), Хансард корпусы (Hansard), интернеттегі ерте ағылшын тілі кітаптары (Early English Books Online), қазіргі америкалық ағылшын тілі корпусы (Coca), америкалық ағылшын тілінің тарихы корпусы (COHA), АҚШ Жоғарғы Сотының пікірлер корпусы (Corpus of US Supreme Court Opinions), TIME журналының корпусы, америкалық сериалдар корпусы, Британдық ұлттық корпус (BYU-BNC), Страти корпусы (Канада), CORE корпусы, Португал тілінің корпусы.

CHILDES – (Child Language Data Exchange System) корпусы – бұл балалардың бірінші тілді меңгеруіне бағытталған дыбыстық, бейне және мәтіндік файлдарды қамтитын мәліметтер базасының жиынтығы.

Келесі қытай-ағылшын параллельді корпустарының Chinese-English Parallel Corpora жүктемелерін TranslateX зерттеушілері мен тілшілері көпшілікке пайдалану үшін жасаған. Корпустар Гонконгтың қаржы және заң салаларын қамтитын сапалы екі тілді мәтіндерден тұратын сөйлемдердің жұптарынан тұрады. Дереккөздерге үкіметтік заңдар мен нормативтік актілер, қор биржасының хабарландырулары, қаржылық ұсыныстар туралы құжаттар, нормативтік құжаттар, нормативтік нұсқаулықтар, корпоративтік құрылтай құжаттары және басқалар кіреді. Барлық мәтіндер Гонконг қор биржасынан, Гонконгтың бағалы қағаздар және фьючерстер комиссиясынан және Гонконг үкіметінің веб-сайттарынан алынған. Жинақ дереккөз түріне сәйкес ұйымдастырылған және .tsv файл форматына сәйкес келеді.

COCA – Corpus of Contemporary American English. 1990-2015 жылдар аралығында 520 миллионнан астам сөз жиналған, корпустар ағылшын әдет-ғұрыптарын ауызша сөйлеу, көркем әдебиет, жаңалықтар көздері мен ғылыми көздерден алынған [5]. Корпус 485 202 мәтінде қолданылған 1 миллиардтан астам сөзден тұрады, оған 1990-2019 жж. аралығында

жыл сайын 24-25 миллион сөз қосылып отырған. Корпус жыл сайын теледидар бағдарламалары мен фильмдер субтитрлері, ауызша тіл, көркем әдебиет, журналдар, газеттер және академиялық журналдар жанрлары арасында бірдей пропорцияны сақтап отырады [6].

1810 жылдан 2009 жылға дейінгі кезеңде жиналған 400 миллионнан астам сөзі бар корпус (COHA – Corpus of Historical American English) тарихи ағылшын қолданыстары мен грамматикалық құрылымдарды іздеуге мүмкіндік береді.

AnCora Catalan and Spanish Corpus. AnCora деген Catalan (AnCora-CA) және Spanish (испан) корпусынан (AnCora-ES) тұрады, олардың әрқайсысы 500 000 сөзді құрайды. Корпустарға әртүрлі деңгейде аннотациялар берілген.

1200-1900 жылдар аралығында испан тіліндегі мәтіндерді Corpus del Español-дан іздеуге болады. Жақында корпус кеңейіп, 1900-2000 жылдардағы мәтіндерді қамтитын болды.

CREA – Corpus de Referencia del Espanol Actual – Корольдік испан академиясының бұл мәтіндік корпусында бүкіл испан тілді әлемнің кітаптары мен мақалалары бар.

Ал CU Linguistics' Verb Corpora Server CU тіл білімі бөлімінде өз корпустары үшін интернетте «тілі бойынша», «аты бойынша» іздеу тетігі бар сервер. Кейбір деректер дискілерде немесе басқа форматтарда сақталған.

Соңғы екі онжылдықтағы Париж тұрғындарынан алынған сұхбаттар жиынтығы. Аудиофайлдар мен стенограммалар жүктеп алушылар үшін қолжетімді корпус Discours sur la ville Corpus de Français Parlé Parisien des années 2000 (Corpus of Parisian French from the 2000's) деп аталады.

Корпустарды іздестіру қызметін көздейтін орталықтандырылған сервистік корпус түрі ELRA Catalogue of Language Resources деп аталады. ELRA каталогы тілдік корпустың баржоғын анықтауға және оларға қол жеткізуге көмектесетін тілдік ресурстар қорын ұсынады. Бұл корпустардың кейбіреулері қосымша тіркелуді қажет етеді.

Француз тіліндегі корпустар (French Language Corpus Collection), мәтіндік және

аудио файлдар, ересектер мен балалардың өзара сөйлесім қарым-қатынасына, сондай-ақ ересектердің өзара коммуникациясына ерекше назар аударады.

Ағылшын тілді ортақ халықаралық интернет-ресурс International Corpus of English (ICE) деп аталады және бұл жерде орналасқан әрбір ICE корпусы 1989 жылдан кейін Канада, Шығыс Африка, Ұлыбритания, Гонконг, Үндістан, Ирландия және Австралия, Ямайка, Жаңа Зеландия, Нигерия, Филиппин, Сингапур, Шри-Ланка және АҚШ корпустарымен жиналған 1 миллион жазбаша және ауызша ағылшын тілінің қорынан тұрады.

Мичиган университетінде академиялық спектакльдердің транскриптітерінің онлайн жиынтығы (сонымен қатар кейбір сөйлеу аудиожазбаларын қамтуы мүмкін) ғана арналған мынадай MICASE - Michigan Corpus of Academic Spoken English корпус түрі де бар екен. Бұл корпусты ғылым тілінде сөйлеу сипаттамаларын зерттеу үшін ғылыми жобаларда пайдалануға арналған.

Тарихи ағылшын тілін зерттеу корпусын Penn Parsed Corpora of Historical English ұлттық корпустың тарихи ішкорпусы деп те бағалауға болады, оның ішінде орта ғасырдағы ағылшын тілін талдау бойынша Пенн-Хельсинки корпусы, екінші басылым (PPCME2), ерте немесе заманауи ағылшын тілін талдау бойынша Пенн-Хельсинки корпусы (PPCEME) және қазіргі Британдық ағылшын талдау бойынша Пенн корпусы, екінші басылым (PPCMBE2), Британдық ағылшын прозасының мәтіндері мен мәтіндерінің үлгілері оның бүкіл тарихында – ең ерте орта ағылшын тілінен Бірінші дүниежүзілік соғысқа дейінгі құжаттар қамтылған. Мәтіндер үш формада ұсынылған: қарапайым мәтін, сөз таптары мен синтаксиске арналған мәтіндер.

Лос-Анджелес қаласындағы Калифорния университетінің фонетика зертханасы ғалымдары ана тілінде сөйлеушілер мен тіл үйренушілер үшін құнды болып табылатын фонетикалық және фонологиялық зерттеулер үшін әлемнің жүздеген тілдерінің жазбаларын жинаған. Осы сайттағы материалдарда 200-ден астам тілде фонетикалық құрылымдарды

суреттейтін аудио жазбалар UCLA Phonetics Lab Archive корпусына жинақталған [5].

TS Corpus: online corpus of Turkish – TS Corpus: түрік тілінің онлайн корпусы. Түрік корпусын құру жобасы. Құрамында миллиардтан астам POST тарға негізделген жинақ. TS Corpus құрылысын (Taner мен Turker Sezer) Танер мен Түркер Сезер жүргізеді [7].

Ағылшын тілінің корпусына дүние жүзіндегі жүздеген университеттер мен мыңдаған жеке тұлғалар, сондай-ақ Германияның зерттеу қоры сияқты ұйымдар қолдау көрсетеді. Корпус деректерін (мысалы, толық мәтінді, сөз жиілігі жөніндегі ақпаратты) Amazon, Google, Microsoft, IBM, Sony, Disney, Intel, Adobe, Samsung деген сияқты көптеген компаниялар, бірнеше (AI) жасанды интелект компаниялары, Merriam-Webster, Dictionary.com, Grammarly, Duolingo, TurnItIn, Oxford University Press, Sketch Engine сияқты тікелей тілге қатысты компаниялар, т.б. пайдаланып келеді.

Бұл корпус төмендегідей бірнеше субкорпустардан тұрады:

- News on the Web (NOW) Ғаламтордағы жаңалықтар
- iWeb: The Intelligent Web-based Corpus – Зияткерлік вебке негізделген корпус
- Global Web-Based English (GloWbE) – Ағылшын тіліндегі ғаламдық веб-сайт
- Wikipedia Corpus – Уикипедия корпусы
- Coronavirus Corpus - Коронавирустық корпус
- Corpus of Contemporary American English (COCA) – Қазіргі америкалық ағылшын тілі корпусы (COCA)
- Corpus of Historical American English (COHA) – Тарихи америкалық ағылшын тілі корпусы (COHA)
- The TV Corpus – Теледидар корпусы
- The Movie Corpus – Көркем фильмдер корпусы, диалектілер: 6 ел
- Corpus of American Soap Operas – Америкалық сериалдар корпусы
- Hansard Corpus – Хансард корпусы
- Early English Books Online – Ерте ағылшын тілінің онлайн кітаптары

- Corpus of US Supreme Court Opinions – АҚШ Жоғарғы соты пікірлерінің корпусы
- TIME Magazine Corpus – TIME журналының корпусы
- British National Corpus (BNC) – Британ ұлттық корпусы (BNC)
- Strathy Corpus (Canada) – Страси корпусы (Канада)
- CORE Corpus – CORE корпусы
- American English – Америка ағылшыны
- British English – Британ ағылшыны [8].

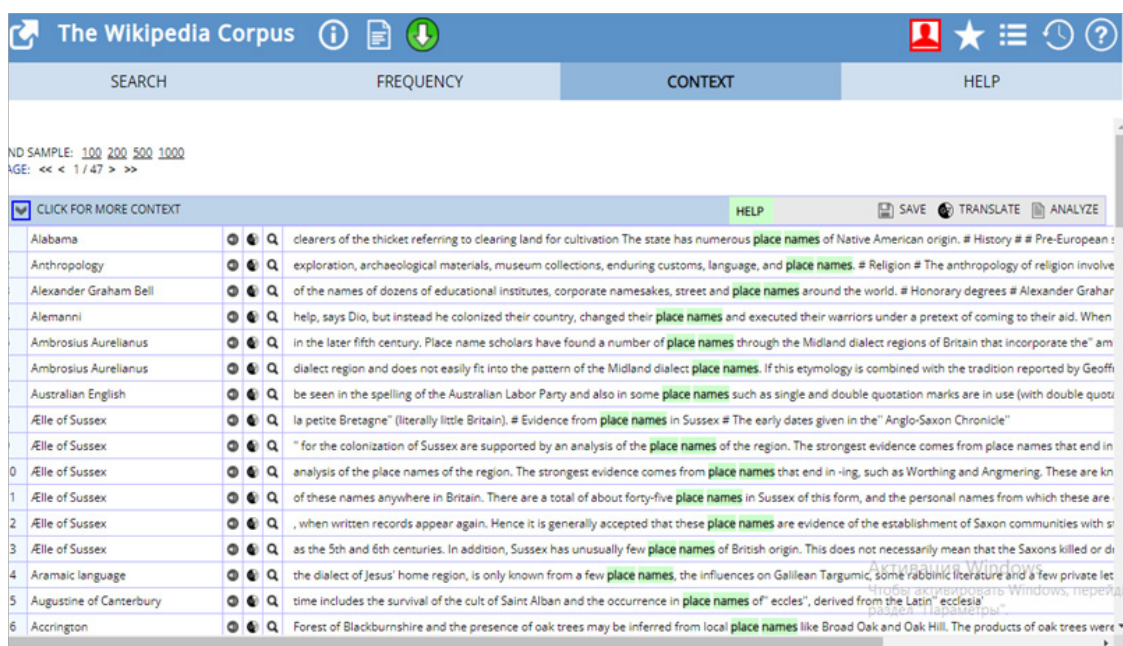
Осы корпусардың бірнешеуіне ономастика саласына қатысты 'onomastics', 'geographical names', 'place names', 'exonym', 'endonym' терминдерін енгізгенде, төмендегідей нәтижелер шықты:

- 1) WIKIPEDIA CORPUS-нда (1-сурет):
 - 'onomastics' – 198
 - 'geographical names' – 645
 - 'place names' – 4667
 - 'exonym' – 491
 - 'endonym' – 383
- 2) COCA Қазіргі америкалық ағылшын тілі корпусында:
 - 'onomastics' – 19
 - 'geographical names' – 3
 - 'place names' – 317

- 'exonym' – 0
- 'endonym' – 0
- 3) BRITISH NATIONAL CORPUS Британ ұлттық корпусында (BNC):
 - 'onomastics' – 2
 - 'geographical names' – 1
 - 'place names' – 63
 - 'exonym' – 2
 - 'endonym' – 5
- 4) NOW CORPUS Ғаламтордағы жаңалықтар корпусында:
 - 'onomastics' – 48
 - 'geographical names' – 451
 - 'place names' – 4353
 - 'exonym' – 22
 - 'endonym' – 14
- 5) THE TV CORPUS Теледидар корпусында:
 - 'onomastics' – 0
 - 'geographical names' – 0
 - 'place names' – 36
 - 'exonym' – 0
 - 'endonym' – 0

Бұдан көретініміз, ағылшын тілінің ішкорпусарында топонимикалық мәліметтердің қамтылу аясы әрқелкі. Демек, ономастика саласы терминдерінің мұндай әр корпусқа түрлі жиілікте кездесу нәтижесі, тиісінше, әр

1-сурет



корпустың мүмкіндігіне қарай өзінің тақырыптық аясына байланысты жұмыс істейтінін көрсетеді.

Келесі көлемді корпустың бірі Орыс тілінің ұлттық корпусы. Орыс тілінің Ұлттық корпусы – 500 миллионнан астам сөздерді қамтитын мәтіндердің ең үлкен электронды жинағы. Бұл орыс тілін зерттейтін лингвистердің негізгі көздерінің бірі. Жоба 9 жыл бұрын, 2004 жылдың сәуірінде іске қосылды. Корпус көптеген ұйымдардың, соның ішінде РҒА В.В. Виноградов атындағы Орыс тілі институты, РҒА ақпарат беру проблемалары институты, ММУ және басқалардың күшімен толықтырылып, қолдау көрсетіледі.

Орыс тілінің Ұлттық корпусы қазіргі кезде келесі шағын корпустардан (субкорпустардан) тұрады:

- әрбір сөйлемге толық сипаттағы морфологиялық және синтаксистік құрылымдары (тәуелділік сызбасы) (деревозависимость) салынған терең аннотацияланған корпус;
- орыс немесе ағылшын тіліндегі белгілі бір сөздің не сөз тіркесінің барлық аудармаларын іздеп табуға мүмкіндік беретін мәтіндердің орысша-ағылшынша параллель корпусы;
- Ресейдің әртүрлі аймақтарының диалектілік сөйлеу тілінің жазбасын жадында сақтайтын және олардың грамматикалық ерекшеліктерін ескере отырып, диалектілік морфология бойынша арнайы іздеу үдерісі қарастырылған диалектілік мәтіндер корпусы;
- тек лексикалық және грамматикалық белгілер ғана емес, сонымен бірге өлеңдерге тән белгілер бойынша іздеу (мысалы, сонета, эпиграмма, өлеңдер және т.б. белгілі тіркестерді іздеу) мүмкіндіктері қарастырылған поэтикалық мәтіндер корпусы;
- орыс тілінің мектеп бағдарламасына бағытталған белгіленімі бар, омонимдер қатары ажыратылған корпус – орыс тілінің оқыту корпусы;
- көпшілік және жекелік сөйлеу тілінің, сонымен бірге 1930-2000 жылдардағы кинофильмдер транскриптерінің магнитофондық жазбаларының мағынасы ашылған орыс тілінің сөйлеу тілі корпусы [3, 46].

Бұл корпуста ономастикалық атаулар лингвистикалық ақпаратпен қамтамасыз етілген деуге болады. Біздің корпусымыз топонимдер бойынша құрастырылып жатқандықтан, орыс корпусындағы топонимдердің аннотациялануына назар аудардық. Бірінші, табиғи нысан жайлы геоақпаратты, атаудың мәдени семантикасы мен этимологиясын, орфографиясын көрсетілген сілтеме арқылы өтіп, academic.ru сөздігінен көруге болады. Мысалы, Волга жалқы есімі жайлы сөздіктен 200 түсініктеме табылды. Ол дереккөздерден атаудың семантикасына кіретін ақпараттың кең көлемін табуға болады. Екінші, жалқы есімнің қай түрі болуына байланысты, жер-су атауы болса t:toron, r:prorn деп, ал кісі аты болса t:persn, r:prorn, t:hum деген шартты белгілермен таңбаланған. Орыс корпусында топонимдерінің басым бөлігі аннотацияланғанына көз жеткіздік.

Нәтижелер

Біз, ең алдымен, ономастикалық/топонимдік ішкорпусты құрастыруда топонимдерді сипаттау критерийлерін әзірледік. Мәселен, жер-су атауларының корпусын құру әр түрлі жанрдағы көптеген мәтіндермен, көркем әдебиеттермен, публицистикамен, бұқаралық ақпарат құралдарындағы мәтінмен жұмыс жасауды қамтитыны белгілі. Мұндай жұмыс атаулардың контекстік ортасын бекіту және оның семантикалық көлемін сипаттайтын барлық мүмкін жағдайларды анықтау үшін маңызды саналады. Мұндай мәтіндерді жүйелеудің міндетті шарты – олардың төлқұжатын құру, яғни басылымның толық деректерін, автордың жасын, жынысын көрсету және т.б.

Сондай-ақ, топонимдердің төлқұжатына төмендегідей белгіленімдер енгізіледі (2-сурет):

- 1) Географиялық нысанның типі;
- 2) Топонимнің орфографиялық нормасы мен бейнормасы. Мәселен, Қарағаш/Қарағаш, Сарыарқа/Сарарқа және т.б.;
- 3) Топонимнің сөзжасамдық ерекшеліктері;

А	В	С	Д	Е	Ғ	Г	Н	І
1 Онім түрі Белкайкақ	Шартты белгісі t.toron	Топоним түрі ойконим	Шартты белгісі t.oikon	Орфографиялық норма Белкайкақ	Орфограмма (Бейнорма) жоқ	Сөзжасамы Бел+қайқақ (зат есім + зат есім) (*Жер-су аттарының сөзжасамдық үлгілері* 261 б.)	Геоаппарат (кітап) БЕЛҚАЙҚАҚ - қоныс. Итжон үстіртінің оңт. бөлігінде, Талғар ауданда (*Жерінің аты, елінің хаты* 160 б.)	Геоаппарат (интернет) https://kk.wikipedia.org/wiki/%D0%98%D1%8B%D0%BE%D1%8A%D0%98%D1%8A
2 Белкайың	t.toron	Ороним	t.toron	Белкайың	Белкайың, белкайың	Бел+қайың (зат есім+зат есім) (*Жер-су аттарының сөзжасамдық үлгілері* 261 б.)	БЕЛҚАЙЫҢ - жота. Тастан өз нін он жағалауында, Алақол ауданда. (*Жерінің аты, елінің хаты* 160 б.)	https://kk.wikipedia.org/wiki/%D0%91%D0%9B%D0%B9%D1%8F%D0%B9%D1%8F A3 https://stom.tilimejta-shi-oleshop-ami-januarlardan-elik-arari-a.html?w
3 Белқұдық	t.toron	гидроним	t.gidron	Белқұдық	өлқұдық, белқұдық, белқұдық	Бел+құдық (зат есім+зат есім) (*Жер-су аттарының	Белқұдық - Бақанас аумағының солтүстік-шығысқа, Бірлік аумағының	деректер жоқ

- 4) Топонимнің хронологиялық мәліметі;
- 5) Ареалы мен географиялық орналасуы;
- 6) Этимологиялық және лексика-семантикалық анықтамасы

Қазақ тілінің ұлттық корпусының ономастикалық ішкорпусын жасауда осында (2-суретте) көрсетілген белгіленімдер топонимдер төңірегіндегі ең маңызды ақпарат ретінде танылып, біздің ішкорпус жасаудағы басты ұстанымдарымызды құрап отыр.

Көпшілік ғаламтор желісін ақпарат алу үшін қолданатыны белгілі. Әдетте адамдар бір немесе бірнеше тірек сөзді тереді, компьютер осы сөздерге байланысты бар ақпарат көздерін немесе арнайы сайттарды тауып береді. Бұл ақпарат алудың бір тәсілі. Берілген сөзге қатысты ақпаратты табу үшін лингвистика-

лық сараптама қолданылады. Онда тірексөздердің қолданылу жиілігіне, морфологиялық формаларына, мағыналас сөздеріне, семантикалық жақтан сәйкес ұғымдарға байланысты ғаламтор сайттар жауап береді. Сол арқылы керекті мәліметті ұсынады» [9, 519-520]. Алайда ғаламтордағы мәтіндер мен тілдік корпустың мәтіндердің айырмашылығы бар. Бұл туралы Ә.Д.Сүлейменова «Цифрлендірілген мәтіндердің базасы корпус белгісінің көмегімен арнайы өңделеді, ал электронды кітапханалар мұндай өңдеуден өтпейді. Дәл осы белгі корпусты түбегейлі тілдің қолданылуы мен өзгеруі жайлы ақпараттың жаңа және объективті дереккөзіне айналдырады. Белгі (аннотация, код) – мәтіндердің және олардың бірліктеріне тән қасиеттер жайлы ар-

1-кесте

Топонимнің түрлері	Түсініктемесі	Шартты белгілері
ойконим	Елдімекен атаулары: қала, ауыл, аудан, қоныс, қыстау, кент, мекен,	t.oikon
ороним	Жер бетінен жоғары және төмен жатқан жер бедері атаулары: Тау, төбе, кезең, жыра, сай, мүйіс, шоқы,	t.toron
гидроним	Су нысандарының атаулары: өзен, көл, бұлақ, тоған, канал, құдық т.б.	t.gidron
хороним	Аймақ, өлке атауы	t.horon
некроним	Жерлеу орындары: мола, бейіт, мазар, қорым	t.nekron
дримоним	Орман, алқап, тоғай атаулары.	t.drimon
дромоним	Кез келген байланыс, қатынас жолының, жол тораптарының атауы	t.dromon

найы лингвистикалық ақпарат беретін корпустың негізгі сипаттамасы. Корпустың белгісі дәлірек және әрқилырақ болған сайын оның әлеуметтік, ғылыми және оқыту құндылығы жоғарырақ болады. Кейбір корпустар аз ғана белгімен қанағаттанса, кейбіреулерінде белгі ұдайы жетілдіріліп отырады» [10, 77]. Мәселен, Google-ге Жетісу деп жазсақ ол туралы қыруар материалдар шығады. Алайда ол жерде Жетісу атауы туралы лингвистикалық (ономастикалық) ақпарат берілмейді. Ономастикалық ішкорпустағы белгіленімдердің түрлеріне қарай төмендегіше түсіндірме беруге болады:

а) топонимнің түрі: ороним, гидроним, ойконим, некроним т.б. байланысты шартты белгілер енгізіледі (1-кесте).

ә) атаудың құрылымына байланысты сөзжасамдық белгіленімдер енгізіледі. Топонимдердің сөзжасамдық ерекшеліктері қарастырылады. Әрбір атаудың ең алдымен қандай сөзден немесе сөздерден тұратыны, оның құрамы мен қай сөз табынан жасалғаны көрсетіледі. Мысалы, Сарыарқа – Сары + арқа, күрделі жалқы есім: біріккен сөз – сын есім + зат есім. Зат есімдердің белгіленімдерін орнатудың үлгісін А.Жұбанов пен А.Жаңабекова төмендегідей көрсетеді. «1. Түбір морфема зат есім; 2. Туынды зат есім (түбір + жұрнақ); 3. Күрделі зат есім: а) біріккен сөз; ә) қосарланған сөз; б) құрама сөз; қысқарған сөз» [3, 112]. Б.Бияров жер-су атауларының сөзжасамдық үлгілерін зерттеген еңбегінде топонимдердің жасалу модельдері бойынша топтайды. 1. Түр-түс мәнді сын есім + зат есім үлгісімен жасалған біріккен топонимдер (Ақайдар, Алакөл, Бозарал, Жиренкөл, Көкбел, Қарақұдық, Қоңыртөбе, Қызылағаш); 2. Сапа белгісін білдіретін сын есім + зат есім үлгісімен жасалған біріккен топонимдер (Айырбидайық, Арықбалық, Байсала, Баянауыл, Беріктас, Бособа, Ескікеңсе, Жақсықұм, Жалтыржон, Жаңадала); 3. Көлемін және бағытын білдіретін сын есімдер + зат есім (Бертоған, Биіккөл, Дәутөбе, Дөңгелеккөл, Жалпақбас, Жіңішкебұлақ, Кішкенесай, Келтетал, Кішікемер, Қысқабұлақ); 4. Зат есім + Зат есім үлгісімен жасалған біріккен топонимдер (Ке-

мербастау, Кепесай, Қабанасу, Науатоған; 5) Сан есім + зат есім үлгісімен жасалған біріккен топонимдер (Алпыссала, Бесарық, Алтыағаш, Егізтөбе, Жартыкөл, Жүзқарағай, Тоғызақ) [11].

б) орфографиялық нормасы мен бейнормасы беріледі. Әлеуметтік желі де, түрлі баспасөзде, тағы да басқа дереккөздерде жер-су атауларын жазуда орфографиялық ала-құдалық туындап жатады. Бұл белгіленім түрі топонимдерді біріздендіруде үлкен рөл атқарады.

Қорытынды

Ономастикалық ішкорпусты құрастырудың маңызы жоғары және оны құрастыру өте ауқымды жұмыс екеніне көз жеткіздік. Мақаламызда әлемдік корпустардағы ономастикалық атаулардың белгілену ерекшеліктерін анықтап, өз тәжірибемізде ұтымды пайдалану жолдарын қарастырдық. Орыс тілінің ұлттық корпусындағы мәліметтерді талдай отырып, олардағы электронды дереккөздердің мол екенін байқадық. Бұл өз кезегінде корпустың көлемді, толық болуындағы негізгі шешуші фактор болып табылады. Сондықтан қазақ тілінде жарық көрген барлық энциклопедия, сөздік, анықтамалықтардың электронды түрде қолжетімді болуы Қазақ тілінің ұлттық корпусын құрастыруға үлкен септігін тигізетін болады.

Сонымен қазақтың ежелгі тарихынан бүгінгі күнге дейін келе жатқан жер-су атауларымыздың корпуста көрініс табуы, олардың метабелгіленімдерінің айқындалуы ғылыми-көпшілік қауымға пайдасы үлкен болады. Ономастикалық ішкорпусты географ, тілші, тарихшы, этнограф мамандар, саяхатшылар, университет және мектеп оқытушылары мен оқушылары, сондай-ақ көпшілік қауым пайдалана алады. Өйткені кез келген ізденуші ономастикалық атаулардың үлкен қоймасына тап болады. Онда ақпараттар топтастырылып, жүйелі түрде берілуі ізденушіге керек мәліметті жылдам табуға көмектеседі әрі жан-жақты білім беретін болады. Алдағы уақытта ҚР барлық облыстарының топонимдері ономастикалық ішкорпуста енгізіледі. Жай

ғана енгізіліп қоймай, әр жер-су атауының мәдени семантикасы, мүмкіндігінше этимологиясы да беріліп отырады. Яғни бұл ұлт игілігіне қызмет ететін ауқымды заманауи дерекқор көзі болмақ. Сонымен бірге ономастикалық белгіленімдер де кеңейіп, жетілдіріліп отыратын болады және келешекте әр атауды интернет желісіне нақты географиялық нысанның 2GIS іздеу-орналасуын көрсететін сервисі арқылы геоақпараттық сипаты да берілмек.

Пайдаланған әдебиеттер

1. Ұлттық картографиялық-геодезиялық қоры [Электронды ресурс]. URL: <https://nkgf.kz/> (жарияланған уақыты: 2021)
2. Қазақстан Республикасының географиялық атауларының мемлекеттік каталогы: 14 том. – Алматы: 2012. – Т.1. – 226 б.
3. Жұбанов А.Қ., Жаңабекова А.Ә. Корпустық лингвистика: оқу құралы / А.Қ. Жұбанов, А.Ә. Жаңабекова. – Алматы: Қазақ тілі, 2017. – 336 б.
4. Сулейменова Ә.Д. Макросоциоллингвистика: монография / Ә.Д. Сулейменова. – Алматы: Қазақ университеті, 2011. – 404 б.
5. University of Colorado Boulder Libraries [Электронды ресурс] URL: <https://libguides.colorado.edu/c.php?g=456725&p=7724222>
6. Corpus of Contemporary American English [Электронды ресурс] URL: <https://www.english-corpora.org/coca/>
7. The Turkish corpus [Электронды ресурс] URL: <https://tscorpus.com/academic/>
8. English-Corpora.org [Электронды ресурс] URL: <https://www.english-corpora.org/corpora.asp>
9. Фромкин В., Родман Р., Хайамс Н. Тіл біліміне кіріспе: оқулық / В. Фромкин, Р. Родман, Н. Хайамс. – Алматы: Ұлттық аударма бюросы, 2018. – 608 б.
10. Сүлейменова Ә.Д. Қазақ тілі үшін ұлттық корпус керек пе? [Мәтін] / Ә.Д. Сүлейменова // ҚазҰУ хабаршысы, Филология сериясы. – 2011. №1 (131). 76-79 б.
11. Бияров Б. Жер-су аттарының сөзжасамдық үлгілері: монография. – Алматы: Мемлекеттік тілді дамыту институты, 2012. – 460 б.

К.К.Рысберген, Н.М.Рсалиева, З.Б.Садырбаева, Д.М.Пашан, А.Б.Сейтбатқал

Институт языкознания имени А.Байтурсынулы, Алматы, Казахстан

Создание ономастического подкорпуса и мировой опыт

Аннотация. Институт языкознания имени А. Байтурсынулы разрабатывает ономастический подкорпус охватывающий область топонимики, являющийся отдельной частью Национального корпуса. В статье обсуждается направление, принципы построения ономастического подкорпуса. Таким образом, определяются основные критерии построения ономастического подкорпуса. Рассматривается Национальный корпус казахского языка и описывается опыт, особенности крупных корпусов в мире. Основная цель работы – определение видов ономастических обозначений. Определены обозначение топонимов в корпусах английского, русского языка, способ передачи лингвистической и экстралингвистической информации, частота встречаемости ономастических терминов. В результате этих анализов делаются сравнения и высказываются предложения, необходимые для разработки топонимических обозначений. Из опыта зарубежных корпусов видно, что чем больше подкорпусов, тем совершеннее Национальный корпус. Появление ономастического подкорпуса принесло ученым-ономастам много новых возможностей для исследований. Потому что в корпусе появляется возможность быстро находить тексты, содержащие топонимы. Это может не только облегчить поисковую работу, но и способствовать комплексным исследованиям в области ономастики. Кроме того, переиздание обновленной версии словарей также будет

быстро реализована с помощью ономастического подкорпуса. Ономастический внутрикорпус незамечен не только с точки зрения утверждения и стандартизации географических названий, но и с точки зрения систематизации топонимов, их эффективного сохранения и научного изучения и для повышения познавательного интереса. Научно-теоретические рекомендации в статье способствуют совершенствованию ономастического подкорпуса, а также определяют роль ономастического подкорпуса в проведении лингвистических исследований. Ожидается, что в ближайшее время диапазон ономастического подкорпуса, начинающегося с топонимов, пополнится и другими видами онимов.

Ключевые слова: корпусная лингвистика, корпусная ономастика, ономастический подкорпус, топонимы, qazcorpus.kz, цифровизация казахского языка.

K.K.Rysbergen, N.M.Rsaliyeva, Z.B.Sadyrbayeva, D.M.Pashan, A.B.Seitbatkal

A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan

Creation of onomastic subcorpus and a few words about world experience

Annotation. The Institute of Linguistics named after A. Baitursynov is developing an onomastic subcorpus in the field of toponymy, which is an independent part of the National corpus. The paper discusses the directions and principles of construction of the onomastic subcorpus. Thus, the main criteria for the design of the onomastic inner corpus are determined. The experience and features of the National corpus of the Kazakh language and some examples of big corpora in the world are described. The main purpose of the work is to determine the types of onomastic designations. Definition of toponyms in the corpus of English and Russian, transmission of linguistic and extralinguistic information, the frequency of occurrence of onomastic terms are determined. As a result of this analysis, comparisons and recommendations are made regarding the development of toponymic designations. As we have seen from the experience of foreign corpora, the more subcorpora, the more mature the National corpus. The emergence of onomastic subcorpus brings new opportunities for research by onomastic scholars. Because in the corpus you can quickly find texts where place names are found. This not only simplifies the search, but also contributes to a comprehensive study of onomastics. In addition, the printing of updated versions of onomastic dictionaries with the help of the subcorpus will be carried out quickly. The onomastic subcorpus is necessary not only in terms of approval and standardization of geographical names, but also in terms of systematization of toponyms, access to this national store for the purpose of their preservation and research and satisfaction of cognitive interests. The scientific-theoretical recommendations of the research paper contribute to the improvement of the definition of the onomastic corpus, as well as determine the role of the onomastic corpus in linguistic research. It is expected that the scope of the onomastic subcorpus, which began with the store of toponyms, will expand its scope with other types of onyms in the future.

Keywords: Corpus linguistics, onomastics, corpus onomastics, onomastic subcorpus, toponyms, qazcorpus.kz, digitalization of the Kazakh language.

References

1. Ultyq kartografialyq-geodezialyq qory [National Cartographic and Geodetic Fund] Available at: <https://nkgf.kz/> (Accessed: 2021)
2. Qazaqstan Respublikasynyñ geografialyq ataularynyñ memlekettik katalogy [state catalog of Geographical Names of the Republic of Kazakhstan]. (14 volumes. – Almaty: 2012. – Vol. 1. – P. 226.) [in Kazakh]
3. Jubanov A.Q., Janabekova A.A. Korpustyq lingvistika [Corpus Linguistics]: textbook / A.K. Jubanov, A.A. Janabekova. – Almaty: Qazaq tili, 2017. – P. 336.
4. Suleimenova E.D. Makrososiolingvistika [Macrosociolinguistics]: monography / E.D. Suleimenova. (– Almaty: Qazaq universiteti, 2011. – P. 404.) [in Kazakh]
5. University of Colorado Boulder Libraries. Available at: <https://libguides.colorado.edu/c.php?g=456725&p=7724222>
6. Corpus of Contemporary American English. Available at: <https://www.english-corpora.org/coca/>
7. The Turkish corpus. Available at: <https://tscorpus.com/academic/>

8. English-Corpora.org. Available at: <https://www.english-corpora.org/corpora.asp>
9. Fromkin V., Rodman R., Hyams N. Til bilimine kirispe [An Introduction to Linguistics]: textbook / V. Fromkin, R. Rodman, N. Hyams. (– Almaty: Ұлттық аударма бюросы, 2018. – P. 608.) [in Kazakh]
10. Suleimenova, E.D. Qazaq tılı үшін ұлттық корпус керек пе? [Do we need a National Corpus for the Kazakh language?] / E.D. Suleimenova // QazҰU habarşysy, Filologia seriasy. – 2011. №1 (131). – P.76-79.
11. Biarov B. Jer-su attarynyñ sözjasamdyq ülgileri [Derivational Patterns of Toponyms]: monography. (– Almaty: Memlekettik tildi damyту институты, 2012. – P. 460.) [in Kazakh]

Авторлар туралы мәліметтер

Рысберген Қыздархан Құрмашқызы – филология ғылымдарының докторы, профессор, Ономастика бөлімінің меңгерушісі, А.Байтұрсынұлы атындағы Тіл білімі институты, Алматы, Қазақстан.

Рсалиева Нұрсауле Мақсұтқызы – филология ғылымдарының кандидаты, Ономастика бөлімінің жетекші ғылыми қызметкері, А.Байтұрсынұлы атындағы Тіл білімі институты, Алматы, Қазақстан.

Садырбаева Зүбайра Боранбековна – Ономастика бөлімінің ғылыми қызметкері, А.Байтұрсынұлы атындағы Тіл білімі институты, Алматы, Қазақстан.

Пашан Дана Мұсабекқызы – филология ғылымдарының кандидаты, Ономастика бөлімінің аға ғылыми қызметкері, А.Байтұрсынұлы атындағы Тіл білімі институты, Алматы, Қазақстан.

Сейтбатқал Әрлен Берікұлы – магистр, Ономастика бөлімінің кіші ғылыми қызметкері, А.Байтұрсынұлы атындағы Тіл білімі институты, Алматы, Қазақстан.

Рысберген Қыздархан Құрмашқызы – доктор филологических наук, профессор, Заведующий отдела Ономастики, Институт языкознания имени А.Байтұрсынұлы, Алматы, Казахстан.

Рсалиева Нұрсауле Мақсұтқызы – кандидат филологических наук, Ведущий научный сотрудник отдела Ономастики, Институт языкознания имени А.Байтұрсынұлы, Алматы, Казахстан.

Садырбаева Зүбайра Боранбековна – Научный сотрудник отдела Ономастики, Институт языкознания имени А.Байтұрсынұлы, Алматы, Казахстан.

Пашан Дана Мұсабекқызы – кандидат филологических наук, Старший научный сотрудник отдела Ономастики, Институт языкознания имени А.Байтұрсынұлы, Алматы, Казахстан.

Сейтбатқал Арлен Берікұлы – магистр, Младший научный сотрудник отдела Ономастика, Институт языкознания имени А.Байтұрсынұлы, Алматы, Казахстан.

Rysbergen Kyzdarhan Kurmashkyzy – Doctor of Philology, professor, Head of the Department of Onomastics, A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan.

Rsaliyeva Nursaule Maksutkyzy – Ph.D., Leading Researcher of the Department of Onomastics, A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan.

Sadyrbayeva Zubaira Boranbekovna – Researcher of the Department of Onomastics, A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan.

Pashan Dana Musabekqyzy – Ph.D., Senior Researcher of the Department of Onomastics, A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan.

Seitbatkal Arlen Berikuly – Master's degree, Junior Researcher of the Department of Onomastics, A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan.