

**Айнұр Аташбекқызы Сейітбекова<sup>1\*</sup>, Нұржұма Елесбай<sup>2</sup>**<sup>1\*</sup> автор-корреспондент, филология ғылымдарының кандидаты,

А. Байтұрсынұлы атындағы Тіл білімі институты, Қазақстан, Алматы қ.,

ORCID: 0000-0003-4268-1572 E-mail: ainurseit@mail.ru

<sup>2</sup> педагогика ғылымдарының магистрі, М. Әуезов атындағы Әдебиет және өнер институты,  
Қазақстан, Алматы қ., ORCID: 0000-0002-0482-1223 E-mail: n.a.sh\_januya@mail.ru**ТАРИХИ ПОЭТИКАЛЫҚ ІШКОРПУС:  
ЕСКІ ҚАЗАҚ ПОЭТИКАЛЫҚ МӘТІНДЕР БАЗАСЫ**

**Андатпа.** Мемлекеттік тілдің ақпараттық-инновациялық базасы, ғылыми-зерттеу және оқыту интернет-ресурсы ретінде Қазақ тілінің ұлттық корпусының (бұдан әрі – ҚТҰК) «Тарихи поэтикалық ішкорпусын» әзірлеудің қажеттілігі зерттеудің өзектілігін көрсетеді. «Тарихи поэтикалық ішкорпус» мәтіндер базасына VI-XIX ғғ. жазба мұралар мәтіндері қамтылуы қажет. Алайда қазіргі «Тарихи поэтикалық ішкорпус» әзірлеудің алқашқы кезеңінде корпус базасына XV-XIX ғғ. ескі қазақ поэтикалық мәтіндері енеді. Бұл кезеңдегі жазба мұралар мәтіндері араб графикасымен жазылған. Жазба мұралар қолжазбаларының екі немесе бірнеше нұсқасы болады және қолжазба мәтіндерінің басым көпшілігі қазіргі таңда кирил графикасына түскен. Осындай мәтіндерді корпус базасына енгізбестен бұрын лингвотекстологиялық өңдеулер мен түзетулер жүргізіледі. Мақаланың мақсаты – өзге тілдердің поэтикалық корпус әзірлеу тәжірибесін зерттеу, араб графикалы және кирилграфикалы мәтіндерге лингвотекстологиялық талдау жасау, ҚТҰК-ның «Тарихи поэтикалық ішкорпус» мәтіндеріне метабелгіленім қою мәселесін қарастыру. Бұдан зерттеудің мынадай міндеттері келіп шығады: *біріншіден*, поэтикалық ішкорпус әзірлеуде өзге тілдердің корпус тәжірибелеріне талдау жасап, артықшылықтарын анықтау; *екіншіден*, арабграфикалы және кирилграфикалы мәтіндерді салыстырып, лингвотекстологиялық талдау жүргізу; *үшіншіден*, тарихи поэтикалық метабелгіленімдердің параметрлерін анықтау. *Зерттеудің практикалық мәні.* Тарихи поэтикалық жазба мұралар мәтіндерін цифрландырып, корпуслық мәтіндер базасын әзірлеу – жаңа филологиялық, этнолингвистикалық білімдер жүйесін қамтамасыз етуге, ғылыми тілдік зерттеулер жасауға ықпал етеді.

**Тірек сөздер:** поэтикалық корпус; поэтикалық метабелгіленім; жазба мұралар; мәтіндер базасы; қолжазба**Қаржыландыру көзі:** Мақала BR21882249 «Мәдениетаралық коммуникация құралы ретіндегі қазақ тілінің ұлттық корпусын (ҚТҰК) жетілдіру және оның ішкорпустарын кеңейту» атты бағдарламалық-нысаналы қаржыландыру жобасы аясында дайындалды.**Сілтеме жасау үшін:** Сейітбекова А.А., Елесбай Н. Тарихи поэтикалық ішкорпус: ескі қазақ поэтикалық мәтіндер базасы. *Tiltanym*, 2024. №3 (95). 140-150-66.DOI: <https://doi.org/10.55491/2411-6076-2024-3-140-150>**Айнұр Аташбековна Сейітбекова<sup>1\*</sup>, Нұржұма Елесбай<sup>2</sup>**<sup>1\*</sup> автор-корреспондент, кандидат филологических наук,

Институт языкознания имени А. Байтұрсынұлы, Казахстан, г. Алматы,

ORCID: 0000-0003-4268-1572 E-mail: ainurseit@mail.ru

<sup>2</sup> магистр педагогических наук, Институт литературы и искусства имени М. Ауэзова,  
Казахстан, г. Алматы, ORCID: 0000-0002-0482-1223 E-mail: n.a.sh\_januya@mail.ru**ИСТОРИЧЕСКИЙ ПОЭТИЧЕСКИЙ ПОДКОРПУС:  
БАЗА СТАРОКАЗАХСКИХ ПОЭТИЧЕСКИХ ТЕКСТОВ**

**Аннотация.** Актуальность исследования обусловлена необходимостью разработки «Исторического поэтического подкорпуса» Национального корпуса казахского языка (далее – НККЯ) как информационно-инновационной базы государственного языка, научно-исследовательского и обучающего интернет-ресурса. Текстовая база «Исторического поэтического подкорпуса» должна охватить тексты письменных памятников VI-XIX вв. Однако на начальном этапе разработки «Исторического поэтического подкорпуса» в базу будут включены староказахские поэтические тексты XV-XIX вв. Письменное наследие этого периода написано арабской графикой. Рукописи письменных наследий имеют две и более версии, и подавляющее большинство рукописных текстов теперь написано кириллицей. До внесения таких текстов в базу корпуса проводится лингвотекстологическая обработка и корректировка. Цель статьи – изучить опыт создания поэтического подкорпуса других языков, провести лингвотекстологический анализ арабграфических и кириллографических текстов, рассмотреть вопросы разработки метаразметки текстов «Исторического поэтического подкорпуса» НККЯ. Из этого вытекают следующие задачи исследования: *во-первых*, проанализировать опыт других языков в разработке поэтического подкорпуса и выявить

их преимущества; *во-вторых*, провести лингвотекстологический анализ с сопоставлением арабографических и кириллографических текстов; *в-третьих*, определить параметры исторической поэтической метаразметки. *Практическая значимость исследования.* Оцифровка текстов исторического поэтического наследия и разработка базы корпусных текстов способствует обеспечению новой системы филологических, этнолингвистических знаний, проведению научных лингвистических исследований.

**Ключевые слова:** поэтический корпус; поэтическая разметка; письменные наследия; текстовая база данных; рукопись

**Источник финансирования:** Статья подготовлена в рамках проекта программно-целевого финансирования №BR21882249 «Совершенствование и расширение подкорпусов национального корпуса казахского языка (НККЯ) как средства межкультурной коммуникации».

**Для цитирования:** Сейтбекова А.А., Елесбай Н. Исторический поэтический подкорпус: база староказахских поэтических текстов. *Tiltanyum*, 2024. №3 (95). С. 140-150. (на каз. яз.)

DOI: <https://doi.org/10.55491/2411-6076-2024-3-140-150>

**Ainur Seitbekova<sup>1\*</sup>, Nurzhuma Yelesbai<sup>2</sup>**

<sup>1\*</sup>Corresponding author, Candidate of Philological Sciences, A. Baitursynuly Institute of Linguistics, Kazakhstan, Almaty, ORCID: 0000-0003-4268-1572 E-mail: ainurseit@mail.ru

<sup>2</sup>Master of Pedagogical Sciences, M. Auezov Institute of Literature and Art, Kazakhstan, Almaty, ORCID: 0000-0002-0482-1223 E-mail: n.a.sh\_januya@mail.ru

### **HISTORICAL POETIC SUBCORPUS: DATABASE OF OLD KAZAKH POETIC TEXTS**

**Annotation.** The relevance of the research is due to the need to develop a «Historical poetic subcorpus» of the National Corpus of the Kazakh Language (hereinafter – the NCKL) as an information and innovation base of the state language, a scientific research and educational Internet resource. The text base of the "Historical Poetic subcorpus" should cover the texts of written monuments of the VI – XIX centuries. However, at the initial stage of the development of the «Historical Poetic subcorpus», Old Kazakh poetic texts of the XV-XIX centuries will be included in the database. The written legacies of this period are written in Arabic graphics. Manuscripts of written legacies have two or more versions, and the vast majority of handwritten texts are now written in Cyrillic. Before entering such texts into the corpus database, linguistic and textual processing and correction are carried out. The purpose of the article is to study the experience of creating a poetic subcorpus of other languages, to conduct a linguotextological analysis of Arabic and Cyrillographic texts, to consider the issues of developing meta-labeling of texts of the "Historical poetic subcorpus" of the NCKL. This implies the following research objectives: firstly, to analyze the experiences of other languages in the development of a poetic subcorpus and to identify their advantages; secondly, to conduct a linguotextological analysis with a comparison of Arabic and Cyrillic texts; thirdly, to determine the parameters of historical poetic meta-markup. *The practical significance of the study.* The digitization of texts of historical poetic heritage and the development of a database of corpus texts contributes to the provision of a new system of philological, ethno-linguistic knowledge, the creation of scientific linguistic research.

**Keywords:** poetic corpus; poetic markup; written heritage; text database; manuscript

**Source of financing:** The article was prepared within the framework of the project of program-targeted financing of the No. BR21882249 "Improvement and expansion of the subcorpus of the national corpus of the Kazakh language as a means of intercultural communication".

**For citation:** Seitbekova A., Yelesbai N. Historical Poetic Subcorpus: Database of Old Kazakh Poetic Texts. *Tiltanyum*, 2024. №3 (95). P. 140-150. (in Kazakh)

DOI: <https://doi.org/10.55491/2411-6076-2024-3-140-150>

### **Кіріспе**

Жалпы тіл білімінде тілдік зерттеулерді іске асыруда ауқымды цифрланған мәтіндер базасы немесе терең аннотацияланған деректер қорын қамтитын тілдік корпус мүмкіндіктерін кең пайдалану тәжірибесі қарқынды дамып келеді. Тілдік корпус – толық әрі ыңғайлы метабелгіленімдерімен қамтамасыз етілген ірі лингвистикалық база ретінде әлемнің көптеген мемлекеттерінде тілдің даму үдерісіне қатысты бірқатар мәселелерді айқындайтын таптырмас көмекші құрал есебінде пайдаланылады. Бұл өз кезегінде бүгінгі пайдаланушының сұраныстарына сәйкес корпус базасын толықтырып, жетілдіру жолында тілтанушы мамандар мен бағдарламашылардың үздіксіз ізденістері мен қажырлы еңбегін талап етеді. Отандық қолданбалы лингвистика ғылымының негізін салған профессор А.Жұбанов қазақ тіліндегі мәтіндерді автоматтандыру мәселесінің маңызы мен қажеттілігі жайында былайша атап өтеді: «Қазақ мәтіндерінің стильдік, құрылымдық, мағыналық, функционалдық және т.б. сипаттарын анықтау өзекті мәселелердің бірі. Сондықтан автоматтанған қазақ тіліндегі мәтіндер

корпусының компьютерлік базасын құру да ғылыми және практикалық жағынан аса құнды мәселе» (Жұбанов, 2010: 191-197).

2016 жылы Ахмет Байтұрсынұлы атындағы Тіл білімі институтында А.Фазылжанованың жетекшілігімен «Қазақ тілінің ұлттық корпусын әзірлеу» жұмыстары басталды. Қазіргі таңда ҚТҰК ауқымды корпустық ресурс ретінде көпшілікке қолжетімді форматта үздіксіз жұмыс жасап жатыр. А.М. Фазылжанова: «Әлемдік корпустық лингвистика жетістіктері мен ұлттық корпустарды жасау мен пайдаланудың озық тәжірибелеріне (Ресей Федерациясы, Финляндия, АҚШ, Жапония және т.б.) сүйене отырып, әртүрлі лингвистикалық белгіленімдер мен метабелгіленімдер енгізілген, қазақ тілінің стильдік жанрларын толық қамтитын Қазақ тілінің ұлттық корпусының кең ауқымды, үйлестірілген әрі жан-жақты болуын қамтамасыз ету жұмыстары болашақта жүргізіле бермек», – дейді (Фазылжанова, 2023: 18). А.Байтұрсынұлы атындағы Тіл білімі институтының мамандары әзірлеген Қазақ тілінің ұлттық корпусы бірнеше ішкорпустан тұрады. Әр ішкорпустың өзіндік функционалдық, тілтанымдық ерекшелігі бар. Сонымен қатар Қазақ тілінің ұлттық корпусы қазақ тілінің байырғы графикасынан, жазу ерекшелігінен, авторлардың қолтаңбасынан, лексикасынан, сөз саптауынан, өлең құрылысынан, сөз қолдану шеберлігінен, стилінен, тұрақты сөз орамдарынан тілтанымдық ақпарат беретін *тарихи поэтикалық ішкорпусымен* де толығу қажет. Орыс тілі поэтикалық корпусының маңызы мен қажеттілігі жайында орыс тілінің корпусын құрастырушы мамандар келесідей пікір білдіреді: «Поэтические тексты интересны не только специалистам по поэтике; это – один из важнейших источников для изучения языка в целом, как в синхронии, так и в диахронической ретроспективе и перспективе» (Орехов, 2019: 41-50). Қазіргі таңда тарихи поэтикалық ішкорпус әзірлеудің алғашқы кезеңінде XV-XIX ғғ. халық ауыз әдебиеті үлгілерінің поэтикалық мәтіндер базасын жинақтап, өңдеп, қолжазба мәтіндерінің нұсқаларына лингвотекстологиялық түзетулер жүргізіп, поэтикалық метабелгіленім параметрлерін әзірлеуді мақсат етіп отырмыз.

### **Материал және әдістер**

Жұмыстың зерттеу материалдары ретінде ауыз әдебиеті үлгілерінің мәтіндері пайдаланылды. Атап айтқанда, «Қалдан Серен Абылайды тұтқынға алғаны» тарихи жыры, «Қарға батыр», «Қисса Шеризат – Күлшат» ғашықтық дастандары.

Зерттеу барысында поэтикалық корпусты әзірлеудегі әлемдік тәжірибелерді талдауда *сипаттау, салыстырмалы әдістер*, ауыз әдебиеті үлгілерінің қолжазба және кирил графикасына түскен мәтіндерін салыстыруда *лингвотекстологиялық, мәнмәтіндік әдістер*, поэтикалық мәтіндердің метамәтіндік белгіленімдері параметрлерін анықтауда *сипаттау, тарихи-салыстырмалы әдістер*, тарихи поэтикалық мәтіндердің базасын және олардың топтарын нақтылауда *топтастыру, жүйелеу* әдістері қолданылады.

### **Әдебиетке шолу**

Қазақ тілінің ұлттық корпусын әзірлеген отандық ғалымдардың: А.Жұбанов, А.Жанабекова, А.Фазылжанова, орыс тілінің поэтикалық корпусын зерттеген: Е.А. Гришина, К.М. Корчагин, В.А. Плунгян, Д.В. Сичинава, поэтикалық метабелгіленімдерді нақтылауда: Р.Сыздық, Қ.Өміралиев, Қ.Жанабаев т.б. ғалымдардың теориялық еңбектері зерттеу жұмысымызға арқау болды.

Бүгінде ғаламтор желісінде қалың оқырманға қолжетімді бірнеше поэтикалық корпус үлгілері белгілі. Олар: башқұрт тілінің поэтикалық корпусы, орыс тілінің поэтикалық ішкорпусы, чех тілінің поэтикалық корпусы, парсы тілінің поэтикалық корпусы және т.б. Соның ішінде, көлемі жағынан ауқымды, арнайы поэтикалық белгіленімге ие 13 млн сөз қолданыстан тұратын ірі электронды ішкорпустың бірі – *орыс тілінің поэтикалық ішкорпусы*. Орыс тілінің поэтикалық ішкорпусы проф. В.А. Плунгянның басшылығымен Ресей Ғылым Академиясына қарасты Орыс тіл білімі институтында жасалып, 2006 жылдың желтоқсан айында көпшілік назарына ұсынылды. Берілген поэтикалық ішкорпус мәтіндері екі түрлі белгіленім жүйесінен тұрады. Біріншісі, жалпы корпусқа тән лингвистикалық белгіленім (грамматикалық, морфологиялық және т.б.), ал екіншісі әдебиеттану және тіл мамандарына арналған өлеңнің құрылымдық ерекшеліктерін ажыратушы поэтикалық белгіленім (метр, тармақ ұзындығы, клаузула, шумак, ұйқас және т.б.) (Плунгян, 2009).

Орыс тілінің поэтикалық ішкорпус интерфейсында іздеу жүйесі пайдаланушыға поэтикалық шығармаларда кездесетін жеке сөздер мен қолданыстарды табумен ғана шектелмей, өлеңнің құрылымдық жүйесін құрайтын *метр, строфа* (өлең жолдарының үндесуі), *клаузула* (өлең өлшемдерінің соңғы бөлігі) сынды параметрлер арқылы іздеуге мүмкіндік береді. Корпус мәтіндерінің дереккөзі – электронды академиялық кітапханалар қоры (Орыс тілінің поэтикалық корпусы, 2023). Сондай-ақ, орыс тілінің поэтикалық ішкорпусын құрастырушы мамандар корпус базасына мәтін жинақтау шарттарын анықтауда орыс тілі поэтикасын зерттеуші белгілі ғалым М.Л. Гаспаровтың зерттеулерін негізге алған (Орехов, 2015: 113-127). Нәтижесінде поэтикалық мәтіндерге қойылатын келесідей талаптарды белгілейді:

1. Корпусқа енетін автор шығармаларының «Библиотека поэта» сынды академиялық басылымдарға еніп, жинақталуы;
2. Корпусқа енетін автор шығармаларына қатысты зерттеу жұмыстарының болуы;
3. Корпусқа енетін соңғы 50 жылдағы автор шығармаларына қатысты сыни әдебиеттердің болуы (Корчагин, 2008).

Жоғарыда берілген талаптарға сәйкес құрастырушылар ішкорпус базасына енетін поэтикалық шығармаларды қатаң іріктеуден өткізеді. Бұл бірінші кезекте орыс тілінің поэтикалық ішкорпус базасына жинақталатын шығармалардың негізінде XVIII-XXI ғғ. шығармаларды толық қамтуға және осы арқылы екшелген мәтіндер базасын жасауға мүмкіндік береді. Мәтіндерді іріктеуде арнайы критерийлердің болуы – орыс тілінің поэтикалық ішкорпусының басты артықшылықтарының бірі деп есептейміз.

Белгіленім жүйесі бойынша мәтіндері лингвистикалық (морфологиялық) және поэтикалық (метрлік) белгіленімге ие келесі поэтикалық корпус – *башқұрт тілінің поэтикалық корпусы*. Берілген корпус проф. Б.В. Ореховтың басшылығымен Башқұрт мемлекеттік университетінде жасалып, 2013 жылдың қазан айында қолданысқа ұсынылды. Башқұрт тілінің поэтикалық корпусына барлығы 1,8 млн сөзқолданысты құрайтын, XX-XXI ғасыр аралығында жазылған 101 ақынның 17 мыңға жуық поэтикалық шығармалары енгізілген (Башкирский поэтический корпус, 2023). Әрбір поэтикалық шығарманың авторлық құқығы сақталған. Башқұрт тілінің поэтикалық корпусы башқұрт поэтикасын зерттеуші мамандарға ғана емес, жалпы башқұрт тілін зерттеуші кез келген қолданушыға арналған қажетті құрал болып есептеледі. Поэтикалық корпус мәтіндерінде лексика-грамматикалық және поэтикалық белгіленім арқылы іздеу жүргізуге болады. Сондай-ақ, башқұрт поэтикалық корпусының басты артықшылықтарының бірі – қолданушы корпус интерфейсіндегі іздеу жүйесінде жеке авторлар арқылы іздеу параметрін тандап, корпус базасында қамтылған әр автордың шығармашылығымен тереңірек таныса алады. Олай болса башқұрт тілінің поэтикалық корпусының оңтайлы іздеу жүйесімен жабдықталған интерфейсін әрбір жеке автордың сөзқолданыс ерекшеліктеріне статистикалық, лингвистикалық әрі поэтикалық талдаулар жүргізуге мүмкіндік береді.

*Чех тілінің поэтикалық корпусы* XIX-XX ғғ. чех тілінің поэтикалық шығармаларын қамтиды. Жалпы көлемі – 14,6 млн сөзқолданыс. Чех тілінің поэтикалық корпусын құрастыру жұмысына Чех әдебиеті институтының зерттеуші мамандары қатысып, жұмыстың алғашқы кезеңі 2013 жылы аяқталды. Поэтикалық корпус базасындағы әрбір сөзтұлғаға (словаформа) фонетикалық транскрипция беріліп, грамматикалық категориясы мен әр шығарманың құрылымдық ерекшелігіне сай поэтикалық метабелгіленімі анықталған. Чех тілінің поэтикалық корпус базасы негізінде чех өлеңдерінің дыбыстық әуезділігін анықтауға мүмкіндік беретін «Эвфонометр» қосымшасы жасалған (Чех тілінің поэтикалық корпусы, 2023). Чех тілінің поэтикалық корпусының артықшылығы: поэзия тілін зерттеуші мамандарға дыбыс әуезділігін қарастыратын поэтика саласын дамытуда таптырмас көмекші құрал бола алады.

*Парсы тілінің поэтикалық корпусы* – жасалу уақыты және көлемі бойынша әзірге ең жас поэтикалық корпусардың бірі. Парсы тілінің поэтикалық корпусы пайдалануға 2020 жылы берілді. Бүгінгі таңда корпус базасына IX-XVII ғғ. аралығын қамтитын 16 842 поэтикалық шығарма енген, бұл үздіксіз толығып отыратын үдеріс. Жалпы поэтикалық корпус базасы 4,3 млн сөзқолданыстан тұрады. Парсы тілінің поэтикалық корпус мәтіндері әзірге морфологиялық және жартылай поэтикалық метрлік белгіленімге ие (Парсы тілінің поэтикалық корпусы, 2023).



Берілген корпус алдағы уақытта сан және сапа жағынан толығып, парсы поэзиясы мен парсы тілін зерттеуші мамандар, шығыстанушылар мен түркітанушылар үшін пайдалы көмекші құрал болмақ.

Жоғарыда көрсетілген әлем тілдерінің поэтикалық корпус тәжірибесін саралай отырып, кез келген тілдің поэтикалық корпусы сол халықтың әдебиетінен, мәдениетінен ақпарат беретін және тіл байлығын танытатын маңызға ие екендігін атап өткеніміз келеді. Осыған орай поэтикалық корпус жобасын қолға алу, дамыту – өзекті мәселе.

### **Нәтижелер және талқылау**

Жалпы поэзия ұғымы ерте кезден қалыптасқаны белгілі. Поэзия тілінде жазылған жазба мұралар көне дәуірден бастау алады. Осыған орай тарихи поэтикалық ішкорпусқа енетін мәтіндер базасын хронологиялық тұрғыдан бірнеше кезеңге бөлуге болады:

1. V – IX ғғ. көне түркі поэзиясы
2. X – XIV ғғ. орта ғасыр поэзиясы
3. XV – XIX ғғ. ескі қазақ поэзиясы
4. XX – XXI ғғ. қазақ поэзиясы

Алайда тарихи поэтикалық мәтіндер базасын әзірлеудің алғашқы кезеңінде XV-XIX ғғ. *ескі қазақ поэтикалық* мәтіндер қамтылады. Солардың бірі қазақ халқының ертеден келе жатқан әдеби, мәдени мұрасының бірі – *ауыз әдебиеті поэтикалық үлгілері* (хикаялық дастандар, діни дастандар, батырлық жырлар, ғашықтық жырлар, қаһармандық дастандар, айтыстар, нақыл-өсиет өлеңдер т.б.). Олардың басым көпшілігі араб графикасында хатқа түскен, еліміздегі сирек қолжазбалар қорында сақталған. XIX ғасырдың 20-30-жылдарынан бастап қағазға түсіріліп, кітап болып жариялана бастады. Тәуелсіздік алған жылдардан бастап фольклорлық мұраларды М.О. Әуезов атындағы институт мамандары мемлекеттік «Мәдени мұра» бағдарламасы аясында «Бабалар сөзі» атты сериямен 100 том етіп дайындап, қалың жұртшылыққа ұсынды. Бүгінгі таңда көптомдық сериямен жарияланған бұл басылымдар тілтанушылардың да зерттеу материалдарына айналып отыр (Бабалар сөзі, 2005).

Қолымызға түскен қолжазба күйінде сақталған ауыз әдебиетін, поэтикалық үлгілерінің табылуы, жариялануы мен танылуына қарай үш топқа бөлуге болады:

*Бірінші топқа* баспа бетін көрмеген жыр-дастандар жатады. Олардың кейбіреулері шығыстың классикалық әдебиеті сюжетіне құрылған. Мәселен, Асылбек пен Гүлжиһан қиссасы, Жетім Мариям, Қарға батыр, Қисса Шеризат Күлшат Мырқы батыр, Нұғыман – Нағиманның қиссасы, Сырым батыр, Үшбу қисса Зеңгір жүһүд, Хикая Әбушаһыма т.б.

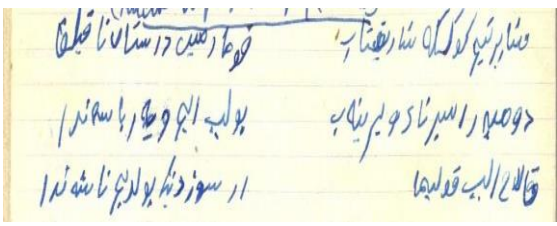
*Екінші топқа* халық арасына кең тарап баспа бетінде жарияланып үлгерген «Қозы Көрпеш – Баян сұлу», «Қыз Жібек», «Бөгенбай батыр» т.б. шығармалар. Жыр-дастандардың мәтіндері бірнеше ғасырды кешіргендіктен, олардың екі немесе одан да көп нұсқалары болады. Кейде шығарманың негізгі желісіне басқаша бір оқиғалар қосылса, екінші нұсқада ол оқиғалар болмайды, ал үшінші нұсқада оқиғалардың шиеленісі, шарықтау шегі мүлдем өзгеше болады.

*Үшінші топқа* халық арасына кең тарамаған, әрі өзге нұсқалары бар шығармалар жатады. Мәселен, Мырзаш батыр (Ниязбек нұсқасы, 1-2-3 бөлім), Мырзаш батыр (Сағидолла нұсқасы), Мырзаш батыр (Ержан Ахметов нұсқасы) т.б.

«Бабалар сөзі» көп томдық сериясын әзірлеушілер шығармалардың баспа бетін көрмей, жарияланбай қалу себептерін былайша атап өтеді: 1. Қоғамдық-идеологиялық мүдделерге сай келмеген; 2. Фольклорлық туындылардың табиғаты ескерілмей, оларға жазба әдебиет шығармаларына қолданылатын талаптар қойылған. Ал жарияланған мәтіндердің көбісі саяси-идеологиялық, қатаң цензуралық бақылау салдарынан мәтіндік қысқарту мен редакциялық түзетуге ұшыраған (Бабалар сөзі, 2005).

Араб графикалы қолжазбаны кирилге түсіргенде мынадай лингвистикалық ерекшеліктер байқалды: бұл кезеңдердегі барлық шығармаларда кірме араб, парсы сөздері цитаталық принциппен түпнұсқа тіліндегі нормамен жазылған. Түпнұсқадағыдай емес, қазақ тілінің фонетикалық ерекшелігіне сәйкес игеріліп жазылған. Мәселен, *Алла* сөзі түпнұсқада *Allah* деп берілсе де, кирил графикасында *Алла* формасы сақталған. Сонымен қатар Алланың барлық есімдері бас әріппен жазылған.

Кесте 1 – «Қисса Шеризат – Күлшат» атты қолжазбаның 2-ші бетінен үзінді  
 Таблица 1 – Отрывок из 2-й страницы рукописи «Кисса Шеризат – Кулшат»  
 Table 1 – Excerpt from page 2 of the manuscript «Qissa Sherizat – Kulshat»

	<p>Шабытым көкке шарықтап,        Құмармын дастан, нақылға.        Домбыра, сырнай үйреніп,        Болып ем үйір басында.        Қалам алып қолыма,        Әр сөздің болдым <i>нәшінде</i>.</p>
---	---

Игеріліп жазылған сөздердің бірі – *нәшін*. Түпнұсқада *нашында* деп жазылған (1-кесте). Өлең мазмұнына қарай «сөздің ырғағына» деген мағына беріп тұр. Қазіргі тілімізде *нашына/нәшіне/машына келтірді* варианттары әлі күнге дейін кездеседі. Кирилге түсіру барысында автор жіңішке вариантымен *нәшінде* деп транскрипциялаған. Бұл Р.Сыздықтың: «Қазақ тіліндегі сөздер сингармониялық тембріне байланысты жуан-жіңішке болып үндеседі. Сөздің қалай таңбаланғанына қарамастан, ішінде жуан-жіңішке дауысты дыбыс әріптері жазылғанмен, тұтасымен не жуан, не жіңішке болып таңбалануы қажет», – деген принципіне сәйкес келеді (Сыздықова, 2000: 187).

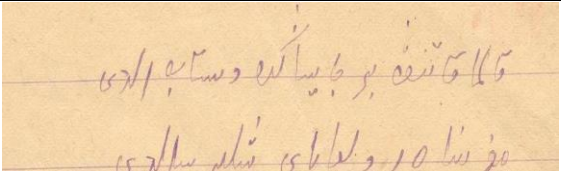
Қолжазбада кездесетін орта түркі тіліннің кітаби лексикалық және грамматикалық қолданыстары түпнұсқадағыдай өзгеріссіз түсірілген: *түрлік, тірік, дүр, бірлән, илә, сәна, үшбу, анығ, андын, әйла, соңыра*, т.б.

Әдетте текстологиялық талдаулар екі немесе бірнеше нұсқасы бар жазба мұралар мәтіндеріне жүргізіледі. Тарихи ішкорпус базасына енетін мәтіндерді өңдеуде нұсқалар арасындағы қолжазбаларды талдау мақсат етілмейді. Бұндай текстологиялық жұмыстар көп уақыт пен қажырлы еңбекті қажет етеді. Сондықтан лингвотекстологиялық түзетулер нұсқалар арасындағы мәтіндерге емес, қолжазба мәтіндері мен кирил графикасына түскен мәтіндерге жасалады. «Бабалар сөзі» 100 томдығының авторлары арабграфикалы қолжазба мәтінін кирил графикасына түсіргенде мынадай қиындықтар болғанын атап өтеді: «Араб жазуында дауысты дыбыстардың жасырын тұратыны, бас әріптің болмайтындығы, тыныс белгілердің қойылмағандығы мәтінді дұрыс оқуға қолбайлау болды. Қазақ және араб тілдері әртүрлі тілдік жүйеге жататындықтан, араб әліпбиі тіліміздегі төл дыбыстарды таңбалауға да, оның өзіндік ерекшеліктерін байқатуға да дәрменсіз екені белгілі. Сонымен қатар XIX ғасырда араб жазуын пайдаланудың жалпыға ортақ ережелерінің болмауы өз әсерін тигізбей қалмады» («Бабалар сөзі», 2005).

«Бабалар сөзі» көптомдық сериясын әзірлеген мамандар арабграфикалық қолжазбадан кирилге түсіргенде автор тарапынан кеткен бірқатар кемшіліктерді атап өтеді. Мәселен, *жеткізер бізді димешке* дегенді *жеткізер бізді бейішке* мәнмәтінге қарай түзетіп отырған. Біз де арабграфикалы қолжазба мен кирилге түскен мәтіндерді беттеу, сәйкестендіру барысында әлі де редакторлық, корректорлық түзетулер жасау қажеттігін байқадық. Кейбір мәтіндерге жасалған түзетулерден үзінді келтірейік:

«Қалдан Серен Абылайды тұтқынға алғаны» деген жырда автор «Қалмақтың *бар* жайсаңын ұстап алды» деп кирилге түсірген. «Қалмақтың *бар* жайсаңын ұстап алды» дегендегі *бір жайсаң* емес, *бар жайсаң* деп берілуінің себебі мәнмәтінге қарай анықталды. Сол сияқты өлең жолдарында Олжабай батыр қалмақтың бір батырын мойнына шылбыр салып ұстап, Абылайхан туралы нақты ақпаратты алғандығы айтылады. Сол дәуірде батырларда шынжырға қарағанда өзімен бірге шылбырды алып жүретіндігін ескерсек, онда мәтінді *шынжыр* емес, *шылбыр* деп транскрипциялаған жөн.

Кесте 2 – «Қалдан Серен Абылайды тұтқынға алғаны» атты қолжазбаның 5-ші бетінен үзінді  
 Таблица 2 – Отрывок из 5 стр. рукописи «Как Калдан Серен взял в плен Абылая»  
 Table 2 – Excerpt from page 5 of the manuscript «How Qaldan Seren Captured Abylai»

	<p>Қалмақтың бар жайсаңын ұстап алды.          Мойнына ер Олжабай шылбыр салды.          Сұрайды Абылайды білдің бе деп,          Абылай қолда екенін біліп қанды.</p>
---	--

Аталмыш қолжазба мәтінде кездесетін төмендегідей сөйлемнің аударылмай қалып кеткені байқалды: «Қазақ халқы басқа халыққа бағынады, күндердің күнінде менің бұл сөзім тарихта қалмақ, қош Олжабай, – деп қоштасты. Олжабай қайтып кетеді. Хан Абылай алған Тобыш сұлудан Қасым хан туғанда қан ұстап жалғыз туады, онан кейін Сары Наурызбай туады».

Осындай өлең жолдарының жеке сөздер мен сөз тіркестерінің өзгертілуі немесе кирил графикасына түспей қалуы жырдың көркемдігіне едәуір нұқсан келтіреді. Аталған кемшіліктер поэтикалық ішкорпусқа дайындау барысында қалпына келтіріліп отырады.

*Тарихи поэтикалық метамәтіндік белгіленімдер.* Мәтіннің метабелгіленім параметрлері ішкорпустың мазмұнына сай арнайы метабелгіленімдердің параметрлерімен ерекшеленеді. Мәселен, орыс тілінің метабелгіленімдерінде 17 параметр көрсетілген, оның 6-уы жалпы барлық мәтіндерге ортақ параметрлер, 11-і орыс тілінің поэтикалық өлең құрылысының ерекшеліктеріне сәйкес алынған параметрлер: число строк (тармақ саны), строфика (шумақ), графическая строфика (графикалық шумақ), метр (өлең өлшемдері), число стоп/иктов, слогов (буын саны), клаузула (өлең өлшемдерінің соңғы бөлігі), рифма (буын мен дыбыстардың үйлесімділігі), дополнительные параметры (қосымша параметрлер), метрическая формула (метрикалық формула), предложений (сөйлем), словоформ (сөзтұлға).

Н. Байтов. Нескончаемые сетования : «И я такой же, еще хуже...» (2000)	
Автор	Н. Байтов
Пол	муж
Год рождения	1951
Название	Нескончаемые сетования : «И я такой же, еще хуже...»
Дата создания	2000
Жанр текста	поэма
Число строк	898
Строфика	0
Графическая строфика	вольная
Метр	я
Число стоп/иктов/слогов	4
Клаузула	вольная : ж, м
Рифма	вольная
Дополнительные параметры	неравносложная рифма, нарушения строфики
Метрическая формула	Я4ж,м; Я4д
Предложений	378
Словоформ	3964

Сурет 1 – Орыс тілінің корпусындағы поэтикалық метабелгіленімдер

Рисунок 1 – Поэтические метазаметки в корпусе русского языка

Figure 1 – Poetic metadata in the Russian language corpus

ҚТҮК аясында әзірленіп жатқан тарихи поэтикалық ішкорпусының XV-XIX ғғ. халық ауыз әдебиеті үлгілері қолжазбалары мәтіндерінің поэтикалық метамәтіндік белгіленіміне 28 параметр алынды. Мәтін туралы берілетін ақпараттың түріне қарай метамәтіндік белгіленім параметрлерін үш топқа бөлуге болады:

1. Барлық корпустарға ортақ базалық параметрлер: *поэтикалық шығарманың атауы, мәтін тақырыбы, мәтін стилі, мәтін жанры, аудитория жасы, ішкорпус түрі, сөзқолданғыс*

саны, *ғасыры мен жылы, мәтіннің бет саны, мәтінді корпусқа енгізген (тұлға), мәтінді енгізген мерзімі.*

2. Ерте дәуірде жазылған жазба мұралар мәтіндері болғандықтан, олардың табылуы мен жариялануына, авторларының белгілі, белгісіздігіне қарай алынған параметрлер: *мәтін туралы қысқаша аңдатпа, мәтін жырлаушысы, қолжазбаны тапсырған, мәтін графикасы, мәтіннің алғашқы баспада жариялануы, мәтін нұсқалары, қолжазбаның сақталған орны, мәтінді аударған адам, мәтін формасы, мәтінді беттеген (тұлға), ескерту.*

3. Поэтикалық мәтіннің өлеңнің құрылысы мен өлшемдеріне қарай міндетті параметрлері: *шумақ саны, тармақ саны, бунақ саны, буын саны, ұйқас түрі.*

Тарихи поэтикалық ішкорпустың мәтіндер базасына енетін мәтіндердің поэтикалық белгіленімдерінің ішінде қажетті деп саналатын кейбір параметрлерінің алыну принциптеріне тоқталайық:

*Мәтін туралы қысқаша аңдатпа.* Шығарманың негізгі мазмұны туралы қысқаша берілетін ақпарат.

*Мәтін жырлаушысы.* Корпусқа алынып отырған кейбір шығармалардың жырлаушысы белгісіз болып жатады. Осыған орай табылған жағдайда жырлаушының аты-жөні көрсетіледі.

*Қолжазбаны тапсырған (тұлға).* Кеңес Одағы кезеңінде саяси-идеологиялық қатаң цензура салдарынан жарияланбай қалған, жоғалып кеткен немесе әлі күнге дейін үй-үйлерде сақталған жазба мұралар сирек қолжазбалар орталықтарына тапсырылып жатыр. Сондықтан қолжазбаны тапсырған тұлғаның аты-жөні көрсетілуі тиіс.

*Ғасыры мен жылы/мерзімі «анық емес» – «анық».* Кейбір қолжазбалардың жазылған жылы көрсетілмей жатады немесе көрсетілген мерзімі көмескі жазылады. Сондықтан «анық» немесе «анық емес» белгісі қойылады.

*Мәтін графикасы.* XV-XIX ғғ. ауыз әдебиеті үлгілері әртүрлі графикада (араб-қадим, араб-жадид, араб-төте) жазылғандықтан, мәтіннің қай графикада жазылғаны туралы берілетін ақпарат.

*Мәтіннің алғашқы жариялануы.* Ертеде жазылған дастандардың, жырлардың алғашқы жариялануы сақталмай қалып жатады. Сондықтан алғашқы жарияланымның көрсетілуі оқырман үшін маңызды.

*Қолжазбаның сақталған орны.* Араб, латын графикаларында жазылған ауыз әдебиеті үлгілері түпнұсқасының қайда сақталғаны туралы берілетін құнды ақпарат.

*Мәтіннің нұсқалары.* Ертеректе ауызша таралған жыр-дастандардың екі немесе бірнеше нұсқасы бар екендігі туралы берілетін құнды мәлімет.

*Ескерту.* Сирек қолжазба мәтіндерінің араб графикасындағы қолжазба мәтіндерінің көбі шифрланған, нөмірленген бума-бума (папки) дәптерлерде жазылған, өлең шумақтары дәптер бетіне әрқалай түскен. Кейбір авторлардың ірі қаріппен жазылған өлең жолдары бір бетті қамтыса, кейбір авторлардың майда қаріппен жазылған бірнеше қатар өлең жолдары дәптердің бір бетіне сыйдырылған. Осыған орай қағаз бетіне өлең жолдарының орналасу ретіне қарай «Ескерту» параметрі алынды.

*Поэтикалық параметрлер.* Қазақ поэзиясының тілін зерттеген Қ.Өміралиев: «XVIII-XIX ғасырдағы поэзия өлеңдерінің негізгі белгілерін атап өтеді: 1) бірінші кісі айтқан сөзді екінші кісінің қайталап отыруы (қайымдап айту); 2) ұйқас жағынан а/ә/а/ә, а/ә/б/ә түрінде келіп тұрақты болмауы; 3) ырғақ-интонациясының риторикалық қалыпта болуы; 4) буын саны жағынан тұрақты болмай, әрі буынды болып келуі; 5) өлең жолының бірінші бунағына екпін түсіріп отыруы (Өміралиев, 1976: 82). Біз қарастырып отырған мәтіндерде бірқатар ерекшеліктерді анықтадық және олардың әрқайсысына тоқталуды жөн санадық:

*Шумақ.* Арабграфикалы қолжазбалардың бір ерекшелігі – өлең жолдары қағаз бетіне әрқалай түскен: оннан солға қарай екі тармақты да, үш тармақты да немесе төрт тармақты да болып бір сызық бойына жазыла берген.

Қолжазбада ешбір шумаққа бөлінбей, біріне бірі жалғасып кете беретін астрофикалық (шумақсыз) өлең түрі кирилмен түскенде біртұтас ойды білдіретін, өзара ұйқасқа құрылған көптармақты өлең жолдары шумақтарға бөлінген, яғни ондағы тармақтар қатаң тәртіппен (төрт-төрттен, алты-алтыдан, сегіз-сегізден т.т.) тізбектелмей, әр жерде әр санды (4-5-тен 20-ға дейін

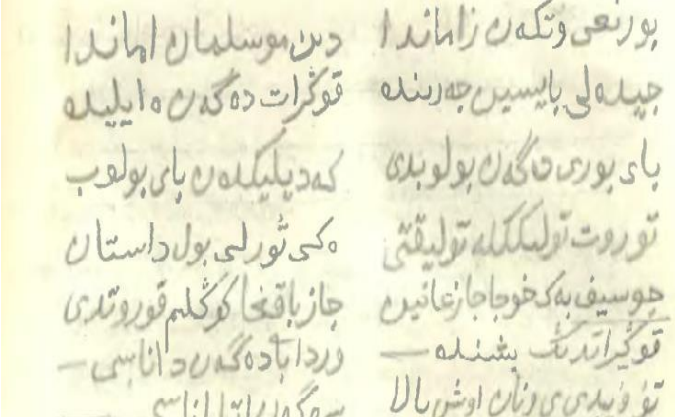


не одан астам) болып топтасқан.

Кесте 3 – Қисса “Алпамыс” жырынан үзінді

Таблица 3 – Отрывок из поэмы “Алпамыс”

Table 3 – Excerpt from the poem “Alpamys”

<p>Бұрынғы өткен заманда, Дін мұсылман аманда, Жиделі Байсын жерінде, Қоңырат деген елінде Байбөрі деген болыпты. Кедейліктен бай болып, Төрт түлікке толықты...</p> <p>Екі түрлі бұл дастан, Жүсіпбек қожа жазғанын Жазбаққа көңілім қорытты. Қоңыраттың ішінде Ордабай деген данасы.</p>	
--	--

Әсіресе жыр түрінде көптеген тармақтарды бір ұйқаспен қамту – дәстүрлі кең тараған құбылыс екенін көптеген зерттеушілер айтып өткен болатын (Сыздыкова, 1970: 22). Башқұрт тілінің поэтикалық шығармаларында изосиллабикалық және гетеросиллабикалық өлең формаларының ішінде 7-8 тармақтыны «қыска-кюй» және 9-10 тармақтыны «узун-кюй» деп атайды (Плунгян, 2009: 41). Осыған орай жыр жолдарында кездесетін біртұтас ойды білдіретін көптармақты өлең түрлері бір шумақ ретінде белгіленді.

*Буын саны.* Буын дауыстылар санына негізделеді. Қолжазбада 7-8, 11 буынды өлшемдер көбірек кездеседі. Ұйғыр поэзиясын зерттеген М.Хамраев: «Жеті-сегіз, сондай-ақ он бір буынды өлшемдер түркі тілдес халықтар өлеңінің көне өрнектері болып саналады», – дейді (Ғамраев, 2002: 78).

*Бунақ саны.* Өлең өлшемі жалғыз тармақтың буын санына қарай емес, тармақ ішіндегі бунақтарға да байланысты. Бунақ – тармақ ішіндегі ырғақ саны. Бунақ корпусқа алынатын мәтіндерге қиғаш сызықпен беріледі. Көбіне 2 немесе 3 бунақты болып келеді.

*Ұйқас түрі.* Поэзиядағы басты талаптардың бірі – ұйқас. Бірнеше ұйқас түрлері белгілі. Біз қарастырып отырған поэтикалық шығармаларда мынадай ұйқас түрлері көбірек кездеседі: қара өлең ұйқасы (а а б а шұбыртпалы ұйқас (а а а а ); ерікті ұйқас (а б в г б); кезекті ұйқас (а б в б ); шалыс ұйқас (а б а б ); егіз ұйқас ( а а б б); аралас ұйқас ( а а б в б в б б б); осы күнгі ерікті ұйқастар (а б а б в г д).

Кесте 4 – Өлең жолдарына жасалған поэтикалық талдау белгіленімі

Таблица 4 – Разметки поэтического анализа стихотворных строк

Table 4 – Markup of poetic analysis of poetic lines

<p>Бөгенбай батыр жырынан: Қазаққа// үлгі болған// Бөген қарты, Ерлердің// бола бермес// мұндай арты. Әкесі //Маян өлген //күндерінде, Бар екен// еркек кіндік// тоқсан алты.</p>	<p>«Қисса шеризат – Күлшат» киссасынан: Бір сұмдықтың// болғанын, Інісі де// біледі. Қайғысы деп// баланың, Зар жылап іші// күйеді. Аң-таң болып// Пархизат, Патшаның басын// сүйеді. «Қайғысы кетіп// көкемнің, Болғай да қабыл //тілегі».</p>
<p>11буынды, 3 бунақты, 4 тармақты, қара өлең ұйқасы</p>	<p>7 буынды, 2 бунақты, 8 тармақты, шалыс ұйқас</p>

Жалпы корпусқа поэтикалық белгіленімдердің берілуі – өлең өлшемдерінің табиғатын, дәстүр сабақтастығын, жұмсалыу ерекшеліктерін, эволюциясын тануға мүмкіндік береді.

### Қорытынды

Сонымен, *тарихи поэтикалық ішкорпус* – көне дәуірден келе жатқан тарихи поэтикалық мәтіндер базасы мен ауызша тілдің поэтикалық түрінің ерекшеліктерін сипаттайтын арнайы поэтикалық метабелгіленімдер мен лингвистикалық белгіленімдерден тұратын ақпараттық құрал қызметін атқаратын электрондық қор.

ҚТҰК-ның аясында «Тарихи поэтикалық ішкорпус» әзірлеудің пайдаланушылар үшін мынадай артықшылықтары бар:

- ауқымды поэтикалық жазба мұралар мәтіндерімен танысып, сол кезеңдегі поэзия тілін тереңірек зерттеуге, тілдік құралдардың қолданысына статистикалық мониторинг жүргізуге, лингвистика және әдебиеттану саласына қатысты зерттеулердің эмпирикалық базасын кеңейтуге;

- тарихи поэтикалық шығармалардың өлең құрылысынан, поэзия өнерінің даму тарихынан ақпарат беріп қана қоймай, жеке тілдік лексемалардың семасиологиялық, прагматикалық, стилистикалық өзгерістері мен ерекшеліктерін, поэтикалық өлең өлшемдерін анықтауға септігін тигізеді.

ҚТҰК-ның аясында XV-XIX ғғ. тарихи поэтикалық мәтіндер базасын әзірлеуде мынадай ғылыми нәтижелерге қол жеткізілді:

*біріншіден*, өзге елдердің тәжірибелері талданып, атап айтқанда, башқұрт, орыс, чех тілдерінің поэтикалық корпустарының артықшылықтары анықталды;

*екіншіден*, корпус базасына енетін XV-XIX ғғ. ауыз әдебиеті үлгілері қолжазбаларының жариялану және жарияланбау себептеріне сипаттама берілді;

*үшіншіден*, қолжазба мәтіні және кирил графикасына түскен аударма мәтіндері қатар көрсетілді. Корпус интерфейсіне екі мәтіннің қатар берілуі – қолжазбаның шынайы қалпын көруге, салыстыруға мүмкіндік береді.

*төртіншіден*, лингвотекстологиялық талдау нәтижесінде бірқатар жиі кездесетін кателердің түрлері анықталды: а) араб графикасындағы қолжазбада автор тарапынан кеткен кателер; ә) арабтан (төтеден) кирил графикасына түспеген кейбір мәтін үзінділері (шумақтар, тармақтар, әріптер); б) қолжазба парақтарының ескіруінен болған кейбір көмескі көрінетін әріптер;

*бесіншіден*, «Тарихи поэтикалық ішкорпустың» спецификалық ерекшелігіне қарай поэтикалық метамәтіндік белгіленімдер параметрлерінің алыну принциптері түсіндірілді.

### Әдебиеттер

Бабалар сөзі: Жүз томдық. – Астана: Фолиант. – Т. 46: Ғашықтық дастандар, 2005. – 434 б.

Башкирский поэтический корпус [Электронный ресурс]. – URL: <http://web-corpora.net/bashcorpus/search/> (дата обращения: 21.12.2023).

Жұбанов А. Қазақ әдеби тілінің электрондық корпусын түзудің теориялық бастаулары / Тіл және мәдениет: тілдің антропоэктік парадигмасы. Профессор Ж.А. Манкееваның 60 жылдығына арналған республикалық ғылыми-теориялық конференцияның материалдары. – Алматы: А.Байтұрсынұлы атындағы Тіл білімі институты, 2010. – 385 б.

Корчагин К. Поэтический подкорпус Национального корпуса русского языка как акцентологический источник, 2008 / Материалы международной конференции «Диалог 2008» [Электронный ресурс]. – URL:

Қазақ тілінің ұлттық корпусын әзірлеу тәжірибесі / Жоба жетекшісі: Фазылжан А. – Алматы: ЖК Асыл, 2023. – 446 б.

Орехов Б.В. Еще раз об исследовательском потенциале поэтического корпуса: метр, лексика, формула / Труды Института русского языка им. В.В. Виноградова, 2015, Вып. 6. – С. 449-463.

Орехов Б.В. Метр отрезков длиннее строки в башкирском силлабическом стихе. Известия РАН. Серия Литературы и языка, 2019, 78(2). – С. 41-50.

Өміралиев Қ. XV-XIX ғасырлардағы қазақ поэзисының тілі. – Алматы: Ғылым, 1976. – 82 б.

Парсы тілінің поэтикалық корпусы [Электронный ресурс]. – URL: [https://linghub.ru/persian\\_poet\\_corpus/#](https://linghub.ru/persian_poet_corpus/#) (дата обращения: 21.12.2023).

Плунгян В.А., Корчагин К., Гришина Е., Ситчинава Д. Национальный корпус русского языка: 2006-2008. Новые результаты и перспективы. СПб., 2009. – С. 71-113.

Поэтический корпус русского языка [Электронный ресурс]. – URL: <https://ruscorpora.ru/corpus/poetic?search=CgQyAggJMAE=> (дата обращения: 19.12.2023).  
<http://www.dialog-21.ru/dialog2008/materials/pdf/Korchagin.pdf>. (дата обращения: 19.12.2023).  
 Сыздық Р. Қазақ тілінің анықтағышы (емле, тыныс белгілері, сөз сазы). – Астана: Елорда, 2000. – 532 б.  
 Сыздықова Р. Абай өлеңдерінің синтаксистік құрылысы. – Алматы: Ғылым, 1970. – 173 б.  
 Һәмраев А. Нәзәрийвий нұқта вәбәдий тәсәвур. – Бишкек, 2002. – 130 б.  
 Чех тілінің поэтикалық корпусы, [Электронный ресурс]. – URL: <https://sysblok.ru/philology/resursy-dlja-cifrovyh-stihovedov-pojeticheskie-korpora/> (дата обращения: 21.12.2023).

### References

- Babalar sozi: Zhuz tomduq (2005) Astana, 434 b. [Words of ancestors: One hundred volumes (2005) Astana, 434 p.] (in Kazakh)
- Bashkirskij pojeticheskij korpus [Bashkir poetic corpus]. [Electronic resource]. Available at: <http://web-corpora.net/bashcorpus/search/> (Accessed: 21.12.2023). (in Russian)
- Cheh tilining pojetikalıyq korpusy, [Czech poetical corpus]. [Electronic resource]. Available at: <https://sysblok.ru/philology/resursy-dlja-cifrovyh-stihovedov-pojeticheskie-korpora/>. (Accessed: 21.12.2023). (in Russian)
- Һәмраев, А. (2002) Nazariivi nukta vabadii tasavur. Bishkek, 136 b. [Khamrayev, A. (2002) A theoretical point is a universal idea. Bishkek, 136 p.] (in Kyrgyz)
- Korchagin, K. (2008) Pojeticheskij podkorpus Nacional'nogo korpusa russkogo jazyka kak akcentologicheskij istochnik. Materialy mezhdunarodnoj konferencii «Dialog 2008» [Korchagin, K. (2008) Poetic subcorpus of the National Corpus of the Russian Language as an accentological source. Proceedings of the international conference “Dialogue 2008”] [Electronic resource]. Available at: <http://www.dialog-21.ru/dialog2008/materials/pdf/Korchagin.pdf> (Accessed: 19.12.2023). (in Russian)
- Omiraliev, Q. (1976) XV-XIX gasyrlardagy qazaq poezijasyning tili. Almaty: Gylym, 82 b. [Omiraliev, Q. (1976) The language of Kazakh poetry in the XV-XIX centuries. Almaty: Gylym, 82 p.] (in Kazakh)
- Orehov, B.V. (2015) Eshhe raz ob issledovatel'skom potentsiale poeticheskogo korpusa: metr, leksika, formula. Trudy Instituta russkogo jazyka im. V.V. Vinogradova, S. 449-463. [Orehov, B.V. (2015) Once again about the research potential of poetic corpus: meter, vocabulary, formula. Proceedings of the Russian Language Institute. V.V. Vinogradov, P. 449-463 p.] (in Russian)
- Orehov, B.V. (2019) Metr otrezkov dlinnee stroki v bashkirskom sillabicheskom stihe. Izvestija RAN, S. 41-50. [Orehov, B.V. (2019) Meter of segments longer than a line in a bashkir syllabic verse. Izvestiya RAN, P. 41-50] (in Russian)
- Parsy tilining pojetikalıyq korpusy [Persian poetical corpus]. [Electronic resource]. Available at: [https://linghub.ru/persian\\_poet\\_corpus/#](https://linghub.ru/persian_poet_corpus/#) (Accessed: 21.12.2023). (in Russian)
- Plungjan, V.A., Korchagin, K., Grishina, E., Sitchinava, D. (2009) Nacional'nyj korpus russkogo jazyka: 2006-2008. Novye rezul'taty i perspektivy, S. 71-113. [Plungyan, V.A., Korchagin, K., Grishina, E., Sitchinava, D. (2009) National corpus of the Russian language: 2006-2008. New results and perspectives, P. 71-113 p.] (in Russian)
- Pojeticheskij korpus russkogo jazyka [Poetic corpus of the Russian language]. [Electronic resource]. Available at: <https://ruscorpora.ru/corpus/poetic?search=CgQyAggJMAE=> (Accessed: 19.12.2023). (in Russian)
- Qazaq tilining ulttyq korpusyn azirleu tazhiribesi (2023) Zhoba zhetekshisi: Fazylzhan A. Almaty, 446 b. [Contemporary interpretation of Kazakh language national corpus. Project manager: Fazylzhan A. Almaty, 446 p.] (in Kazakh)
- Syzdyq, R. (2000) Qazaq tilining anyqtagyshy (emle, tynys belgileri, soz sazy). Astana: Elorda, 532 b. [Syzdyq, R. (2000) Determinant of the Kazakh language (spelling, punctuation, word music). Astana: Elorda. 532 p.] (in Kazakh)
- Syzdyqova, R. (1970) Abai olengderining sintaksistik qurylysy. Almaty: Gylym, 173 b. [Syzdyqova, R. (1970) Syntactic structure of Abai's poems. Almaty: Gylym, 173 p.] (in Kazakh)
- Zhubanov, A. (2010) Qazaq adebi tilining jelektronıyq korpusyn tuzuding teorijalyq bastaulary. Til zhane madeniet: tilding antropoozektik paradigmasy. Professor Zh.A. Mankeevanyng 60 zhyldygyna arналган respublikalyq gylimi-teorijalyq konferencijanyng materialdary. Almaty, 385 b. [Zhubanov, A. (2010) Theoretical beginnings of creating an electronic corpus of the Kazakh literary language. Language and culture: the anthropocentric paradigm of language. Materials of the republican scientific-theoretical conference dedicated to the 60th anniversary of Professor Zh.A. Mankeyeva. Almaty, 385 p.] (in Kazakh)

*Мақала туралы ақпарат / Информация о статье / Information about the article.*

Редакцияға түсті / Поступила в редакцию / Entered the editorial office: 14.03.2024.

Жариялауға қабылданды / Принята к публикации / Accepted for publication: 03.09.2024.

© Сейітбекова А.А., Елесбай Н., 2024

© А. Байтұрсынұлы атындағы Тіл білімі институты, 2024