

XFTAP 16.31.21

DOI: [10.59102/kufil/2026/iss2pp137-149](https://doi.org/10.59102/kufil/2026/iss2pp137-149)

**Қ.Қ. Сарекенова<sup>1</sup>, Ж.Қ. Өмірбекова<sup>1</sup>**

<sup>1</sup>Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, 010000, Қазақстан

## **ФОЛЬКЛОРЛЫҚ ДИСКУРС ЖӘНЕ ОНЫҢ ҚАЗАҚ ТІЛІНІҢ ҰЛТТЫҚ ҚОРПУСЫНДАҒЫ КӨРІНІСІ**

*Бұл мақалада фольклорлық дискурс қазақ тіл білімінің заманауи бағыттары тоғысында кешенді түрде зерттеледі. Мақаланың басты нысаны - фольклорлық мәтіндердің тілдік, құрылымдық және функционалдық ерекшеліктерін қазақ тілінің ұлттық корпусының деректері негізінде айқындау. Фольклорлық дискурс әртүрлі жанрды қамтитын, тұрақты құрылымдарымен, бағалауыштық жүйесімен, ұлттық мәдени кодты жеткізу қабілетімен ерекшеленетін күрделі жүйе ретінде сипатталады. Зерттеу барысында қазақ тілінің ұлттық корпусындағы фольклор мәтіндерінің берілуі, олардың жиілік, вариативті және мәнмәтіндік ерекшеліктері қарастырылады. Фольклор мәтіндерін, оларға қатысты деректерді корпуста ендіру, таңбалау және жүйелеу мәселелері талданып, осы үдерістің ғылыми-әдіснамалық негіздері ұсынылады. Зерттеудің мақсаты – фольклорлық дискурстың корпус моделін сипаттау және қазақ тілінің ұлттық корпусындағы фольклорлық дискурс материалдары мен деректерінің ғылыми әлеуетін айқындау. Зерттеу нәтижелері қазіргідей цифрландыру тілімізге дендеп жатқан кезеңде фольклорлық мұраны сақтау және талдаудың тиімді жолдарын көрсетуге мүмкіндік береді.*

*Кілт сөздер: ұлттық корпус, фольклорлық дискурс, цифрландыру, фольклор, ішкорпус*

### **НЕГІЗГІ ЕРЕЖЕЛЕР**

Зерттеу барысында фольклорлық мәтіндердің жанрлық табиғатын, құрылымы мен поэтикасын айқындаған С.А. Қасқабасов, Ә. Қоңыратбаев сияқты ғалымдардың іргелі еңбектері мен мәтіндік дискурсты, корпустық лингвистиканың әдістерін негіздеген П. Кростуайт, Г. Флинтеруд сынды шетелдік зерттеушілердің тұжырымдамалық тәсілдері басшылыққа алынады. Фольклорлық дискурс тек қатып қалған мәтіндер жиынтығы емес, ұлттық мәдени кодты, мифологиялық танымды және тілдік бағалауыштық жүйені тасымалдаушы тірі құрылым ретінде лингвокультурологиялық және корпустық лингвистика тоғысында кешенді қарастырылады. Қазақ тілінің ұлттық корпусы негізінде фольклорлық мәтіндерді цифрландыру мен таңбалаудың (аннотациялаудың) арнайы ішкорпустық моделін енгізу ауызша дәстүрге тән варианттылық пен формулалық стильдің (мысалы, тұрақты эпитеттер, бастау-аяқтау формулалары) статистикалық заңдылықтарын дәл айқындауға мүмкіндік береді деген ғылыми болжам алға тартылады. Мақалада дәстүрлі филологиялық сипаттама әдістері қазіргі заманауи корпустық тәсілдермен ұштастырыла отырып қолданылады. Фольклорлық дискурстың корпустық көрсеткіштерін дәлелдеу арқылы болашақта жасанды интеллект пен автоматты мәтін тану жүйелеріне қазақ фольклорының лингвистикалық ерекшеліктерін дәл тануға негіз болатын цифрлық база қалыптасады.

### **КІРІСПЕ**

Қазақ халқының рухани және мәдени мұрасында фольклор ерекше орын алады. Фольклорлық мәтіндер – ұлттың тарихи жадын, дүниетанымын, әлеуметтік тәжірибесін сақтайтын маңызды дереккөз. Алайда фольклордың дәстүрлі түрде ауызша таралуы оның ғылыми айналымға ену үдерісінде қиындықтар туғызады. Осыған байланысты қазіргі кезеңде фольклорлық мәтіндерді цифрландыру және оларды ұлттық корпуста енгізу – өзекті ғылыми

және практикалық міндеттердің бірі. Соңғы жылдары корпус лингвистикасы тіл білімінің дербес әрі қарқынды дамып келе жатқан саласына айналды. Корпус зерттеулері тілдік деректерді сандық және сапалық тұрғыдан кешенді талдауға мүмкіндік береді. Бұл фольклорлық дискурсты да корпуслық тұрғыдан зерттеуге мүмкіндік береді. Өйткені фольклорлық мәтіндер лингвистикалық материал болумен қатар мәдени, прагматикалық және когнитивтік ақпараттың күрделі қабатын қамтиды. Дискурс ұғымы қазіргі лингвистикада мәтіннен кең ауқымды категория ретінде қарастырылады. Ол тілдік құрылыммен бірге экстралингвистикалық факторларды қамтиды. Осы тұрғыдан алғанда фольклорлық дискурс – белгілі бір этностың дәстүрлі мәдени кеңістігінде қалыптасқан, ұжымдық сананың көрінісін бере алатын дискурстық жүйе.

Әлемдік тіл білімінде табиғи тілді модельдеу, тілдік алгоритмдер құру және мәтінді автоматты өңдеу бағыттары ғылыми зерттеулердің жаңа парадигмасын қалыптастырды. Бұл үдерістер корпус лингвистикасының қарқынды дамуына негіз болды. Ол лингвистикалық деректерді үлкен көлемде жүйелеуге, талдауға және таныстыруға мүмкіндік береді. Аталған бағыттар фольклорлық мәтіндерді дәстүрлі сипаттамалық тұрғыдан ғана емес, дискурстық және сандық-статистикалық әдістер арқылы кешенді зерттеуге мүмкіндік береді. Отандық тіл білімінде де цифрландыру мен корпусқа негізделген зерттеулердің маңызы артып келеді. Әсіресе, ұлттық мәдени мұраның ірі қабатын құрайтын фольклорлық мәтіндерді қазақ тілінің ұлттық корпусы аясында ғылыми айналымға енгізу – өзекті мәселелердің бірі. Бұл тұрғыдан фольклорлық дискурсты тілдік, құрылымдық және функционалдық қырларынан корпус деректері негізінде талдау қазақ тіл біліміндегі жаңа ғылыми бағыттардың қалыптасуына ықпал етеді.

Қазақ фольклорының ғылыми тұрғыдан зерттелуі Ш. Уәлиханов, Ә. Марғұлан, Ә. Диваев, М. Әуезов, С. Қасқабасов, Р. Бердібаев, Ә. Қоңыратбаев, Ж. Жақыпов және басқа да ғалымдардың еңбектерінде әдебиеттану, этнография және мәтінтаным бағыттары тұрғысынан кеңінен қарастырылды. Аталған зерттеулерде фольклор үлгілерінің жанрлық табиғаты, поэтикалық жүйесі, тарихи-мәдени мазмұны жан-жақты сипатталғанымен, оларды тілдік дерек ретінде үлкен көлемде жүйелі түрде талдау мүмкіндігі шектеулі болды. Қазақ тілінің ұлттық корпусындағы фольклорлық мәтіндер фольклорлық дискурстың құрылымы мен әртүрлі ерекшеліктерін айқындауға мүмкіндік береді.

Зерттеудің мақсаты – фольклорлық дискурсты қазақ тілінің ұлттық корпусы материалдары негізінде талдау арқылы оның тілдік, құрылымдық, семантикалық және стилистикалық ерекшеліктерін айқындау.

Зерттеудің мақсатына сәйкес мынадай міндеттер қойылды:

- әлемдік тіл біліміндегі фольклорлық мәтіндер корпустары мен оларды зерттеудің дискурстық-танымдық тәжірибесіне шолу жасау;
- қазақ тілінің ұлттық корпусындағы фольклорлық мәтіндердің құрамын, құрылымын және жүйелену ерекшеліктерін сипаттау;
- фольклорлық дискурсты корпус лингвистикасы әдістері арқылы зерттеудің ғылыми-теориялық маңызын негіздеу;
- фольклорлық дискурстағы тілдік бірліктердің (тұрақты тіркестер, бағалауыштық бірліктер, әртүрлі тұрақты құрылымдар және т.б.) жиілік, таралу және қолданыс ерекшеліктерін корпуслық-статистикалық талдау арқылы көрсету.

## МАТЕРИАЛДАР МЕН ӘДІСТЕР

Фольклорлық мәтіндерді цифрландыру, оларды қазақ тілінің ұлттық корпусының дерекқорына енгізу, ғылыми қауымдастыққа қолжетімді ету және білім беру кеңістігіне бейімдеу – кешенді әрі көпсатылы үдеріс. Аталған жұмыс тек лингвистика саласының шеңберінде ғана жүзеге аспай, фольклортану, мәтінтану және ақпараттық технологиялар салаларымен өзара ықпалдастықта жүргізіледі. Осы тұрғыдан алғанда фольклорлық дискурсты корпуслық негізде сипаттау – пәнаралық зерттеудің айқын көрінісі.

Қазақ тілінің ұлттық корпусындағы фольклорлық мәтіндер фольклорлық дискурстың тілдік, құрылымдық және функционалдық ерекшеліктерін жаңа ғылыми деңгейде зерделеуге мүмкіндік береді. Корпус деректері фольклорлық дискурстың әртүрлі қырларын айқындауға жағдай жасайды, сонымен қатар фольклорлық дискурсты сандық және сапалық тұрғыдан талдауға жол ашады. Зерттеу барысында төмендегідей әдістер қолданылды:

– сапалық және сандық талдау әдістері (фольклорлық дискурстағы тілдік бірліктерді жиілік, коллокация, құрылымдық және функционалдық ерекшеліктері тұрғысынан қарастыру);

– ақпаратты өңдеу әдістері (фольклорлық мәтіндерді жинақтау, талдау);

– жинақтау және жүйелеу әдісі (фольклорлық мәтіндерді жанрлық және дискурстық белгілеріне қарай біріздендіріп, қазақ тілінің ұлттық корпусы аясында құрылымдалған дерекқор қалыптастыру).

Фольклорлық дискурсты корпусық негізде жүйелеу оның тілдік табиғатын, дискурстық модельдері мен функционалдық сипаттарын жаңа ғылыми деңгейде тануға мүмкіндік береді. Осыған сәйкес фольклорлық дискурсты корпус деректері арқылы талдау дәстүрлі фольклортанулық зерттеулерді заманауи лингвистикалық және ақпараттық технологиялық құралдармен ұштастырып, тың ғылыми нәтижелерге қол жеткізуге жағдай жасайды. Зерттеу негізінде қазақ фольклорлық дискурсын цифрландыру арқылы ұлттық мәдени мұраны қазақ тілінің ұлттық корпусы шеңберінде жаңа ғылыми айналымға енгізуді және корпус лингвистикасы мен лингвофольклортану тоғысында әдіснамалық негіз қалыптастыруды көздейміз. Жұмыс нәтижелері қазақ тіл білімінің теориялық базасын толықтырып қана қоймай, фольклортану ғылымына да әдіснамалық және қолданбалы тұрғыдан үлес қосады. Сонымен қатар, зерттеу нәтижелері қазақ тілінің ақпараттық технологиялар саласындағы қолданысын (табиғи тілді өңдеу, жасанды интеллект, машиналық аударма) кеңейтуге, білім беру жүйесінде фольклорлық дискурсты меңгертудің тиімді құралы ретінде пайдалануға және ұлттық мұраны мәдени-танымдық әрі цифрлық түрде сақтауға мүмкіндік береді.

## НӘТИЖЕЛЕР

Қазіргі гуманитарлық ғылымда фольклор тек мәтіндер жиынтығы ретінде емес, дискурстық құбылыс ретінде де қарастырылады. Сондықтан фольклор мәтінін пайда болу жағдаятымен, әлеуметтік-мәдени мәнмәтінмен бірлікте талдау маңызды. Қазақ фольклорының ғылыми зерттелуі Ш. Уәлиханов, Ә. Диваев, Ә. Марғұлан, М. Әуезов еңбектерімен байланысты. Онда фольклор тарихи-мәдени дерек ретінде қарастырылса, кейінгі кезеңде теориялық тұрғыдан жүйеленді. С. Қасқабасов фольклорды көркемдік жүйе, жанрлық құрылым және поэтика тұрғысынан талдап, фольклорлық мәтіннің көпқабатты табиғатын көрсетті. Ғалымның еңбектері фольклорлық дискурстың жанрлық және стилистикалық негізін анықтауға тірек болады [1, 43].

Ә. Қоңыратбаев қазақ фольклорының тарихи дамуын жүйелеп, жанрлық классификация ұсынды. Бұл жіктеу ұлттық корпус аясында фольклорлық мәтіндерді құрылымдауға теориялық бағдар береді [2, 87].

Қазіргі корпус лингвистикасында цифрлық технологиялар мәтінді құрылымдық және семантикалық тұрғыдан кешенді зерттеуге мүмкіндік береді. П. Кростуайт және әріптестері корпус лингвистикасының заманауи гуманитарлық ғылымдардағы маңызды зерттеу бағытына айналғанын атап көрсетеді. Олардың пікірінше, корпус тілдік деректерді объективті түрде талдауға және дискурстық ерекшеліктерді анықтауға көмектеседі [3, 16]. Соңғы жылдары тіл білімінде корпус лингвистикасының дамуы дәстүрлі мәтінтану мен фольклортануға жаңа әдіснамалық мүмкіндіктер ұсынды. Корпустар үлкен көлемдегі мәтіндерді жүйелі түрде жинақтап, оларды сандық және сапалық тұрғыдан талдауға мүмкіндік береді. Бұл үрдіс фольклорлық мәтіндерді де қамтып, «фольклорлық корпус» ұғымының қалыптасуына әкелді. Г. Флинтеруд фольклордың цифрлық платформаларда жаңа сипатқа ие болатынын айтады. Оның пікірінше, әлеуметтік желілердегі фольклорлық коммуникация алгоритмдік

жүйелермен тығыз байланысты болып, дәстүрлі ауызша мәдениеттің трансформацияланған үлгісін қалыптастырады [4, 450].

Қазақ тіл білімінде корпус мәселесі – салыстырмалы түрде жаңа бағыт. Қазақ тілінің ұлттық корпусы жасалып, ғылыми айналымға ене бастағанымен, фольклорлық мәтіндердің корпустың орны, олардың дискурстық сипаттамасы және арнайы фольклорлық ішкорпус құру мәселелері әлі де жүйелі түрде зерттелген жоқ. Фольклорлық дискурс – ауызша дәстүрге негізделген, ұжымдық сана мен мәдени тәжірибені бейнелейтін дискурс түрі. Ол белгілі бір жанрлық шеңберде (ертегі, аңыз, эпос, мақал-мәтел) жүзеге асады және тұрақты тілдік модельдермен сипатталады. Ғылыми әдебиеттерде фольклорлық дискурстың келесі негізгі белгілері көрсетіледі:

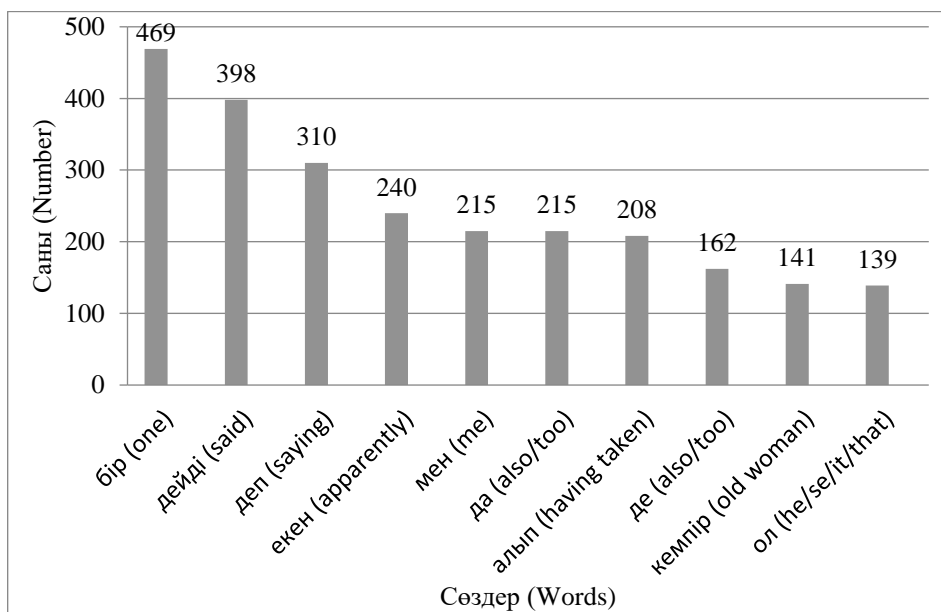
- ✓ құрылымдық – фольклорлық мәтіндерде тұрақты тіркестер мен қайталанатын құрылымдардың кең қолданылуы.
- ✓ қайталау және вариативтілік – бір сюжет немесе мотивтің әртүрлі нұсқада берілуі.
- ✓ бағалаушытық және модалдылық – кейіпкерлер мен әрекеттерге берілетін айқын аксиологиялық баға.
- ✓ мифологиялық уақыт пен кеңістік – нақты тарихи координаттардан тыс, шартты модельдер.
- ✓ ауызша табиғаты – яғни мәтіннің бастапқыда жазбаша емес, ауызша формада өмір сүруі;
- ✓ ұжымдық авторлық – фольклорлық мәтіндердің жеке авторға емес, халыққа тиесілі болуы.

Фольклорлық дискурс жанрлық тұрғыдан да күрделі құрылымға ие. Эпостық жырлар, ертегілер, аңыздар, мақал-мәтелдер, жұмбақтар мен тұрмыс-салт жырлары дискурстық құрылымы, прагматикалық мақсаты және тілдік құралдары жағынан бір-бірінен ерекшеленеді. Сондықтан фольклорлық дискурсты зерттеу кешенді, пәнаралық тәсілді талап етеді.

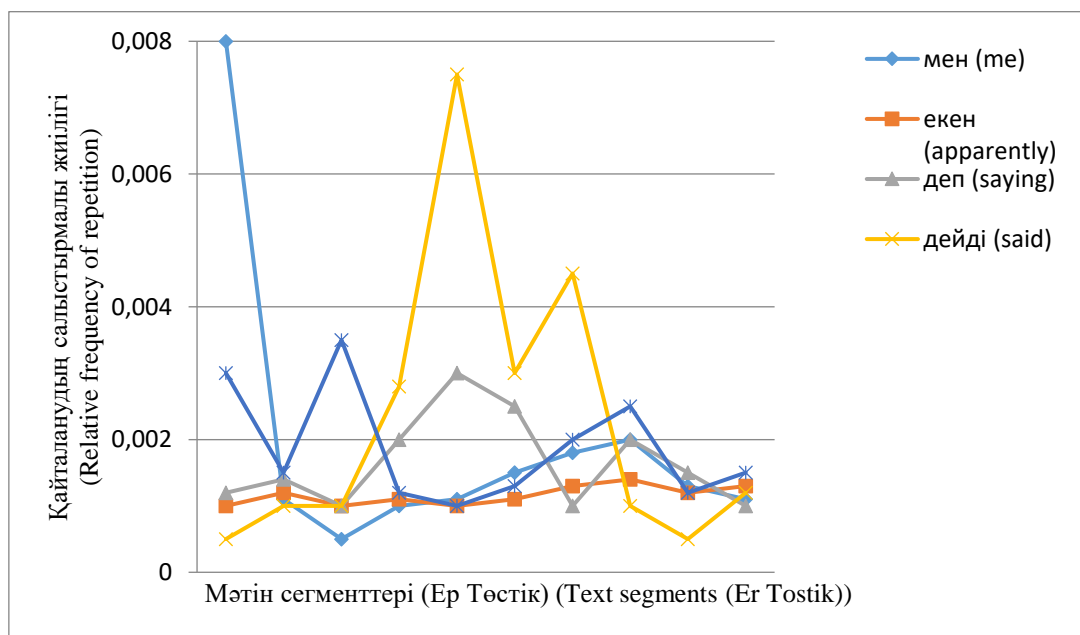
Фольклорлық дискурстың негізгі белгілерінің бірі – құрылымдық пен варианттылық. Бұл мәселе ауызша дәстүрді зерттеудің классикалық теорияларында кеңінен қарастырылған. Фольклорлық мәтіндерді цифрландыру және корпустық форматта ұсыну – әлемдік тіл білімінде кең таралған тәжірибе. Мәселен, Финляндиядағы SKVR (Suomen Kansan Vanhat Runot) жобасы – халық поэзиясының ауқымды цифрлық корпусы. Ж. Хагедорн мен С. Дарани фольклорлық мәтіндерді цифрлық форматта жүйелеу мен ашық ғылыми айналымға енгізудің маңызын атап көрсетеді. Ғалымдардың пікірінше, аннотацияланған фольклорлық корпус зерттеушілерге мәтіндерді құрылымдық, семантикалық және сюжеттік тұрғыдан салыстырмалы түрде талдауға мүмкіндік береді. Авторлар ашық корпус моделінің ғылыми зерттеулердің қайталанбалылығын (reproducible research) қамтамасыз етудегі рөлін ерекше атап өтеді. Сонымен қатар, фольклорлық мәтіндерді цифрлық таңбалау олардың жанрлық ерекшеліктерін, мотивтік құрылымын және мәдени контекстін тереңірек зерттеуге жағдай жасайды [5, 109]. Сол сияқты, Орыс ұлттық корпусында халық сөйлеуі мен ауызша мәтіндерге арналған ішкорпустар бар. И.Б. Качинская мен А.В. Малышева фольклорға жақын ауызша мәтіндерді корпустық енгізудің ғылыми-әдістемелік қағидаларын көрсетеді. Олардың пікірінше, ауызша дискурсты корпустық енгізу дискурстық белгілерді сақтай отырып таңбалауды талап етеді [6, 81]. М. Нұртазина қазақ тілінің ұлттық корпусын қалыптастыру барысында мәтіндерді жанрлық, стильдік және функционалдық ерекшеліктеріне қарай жүйелеудің маңызын атап өтеді. Автордың пікірінше, ұлттық корпус тілдің қолданыс аясын, мәдени ерекшелігін және тарихи даму үрдісін көрсетуі тиіс [7, 46]. Осы бағыттағы Ә. Қайдардың этнолингвистикалық зерттеулерінде фольклорлық құрылымдар этномәдени ақпараттың жинақталған көрінісі ретінде қарастырылады [8, 93]. Корпустық деректер бұл құрылымдардың фольклорлық дискурста дискурстық маркер қызметін атқаратынын нақтылайды.

Қазақ тілінің ұлттық корпусы – фольклорлық дискурсты сандық-мәтіндік тұрғыдан зерттеуге мүмкіндік беретін негізгі дереккөз. Корпус құрамында фольклорлық ішкорпус болмаса да, мәдени-репрезентативті ішкорпуста фольклор мәтіндері қамталған. Мұндағы

мәтіндер төрт түрлі бағытта (фольклорлық, авторлық ауыз әдебиеті мұралары, этнографиялық еңбектер, ғылыми еңбектер, мақалалар) жинақталған (Қазақ тілінің ұлттық корпусы). Ал, Н.Н. Аитова, М. Бақытқызы фольклорлық мәтіндерді корпуста енгізу ұлттық тілдің цифрлық репрезентациясын кеңейтетінін және лингвофольклортанудың жаңа бағытын қалыптастыратынын атап көрсетіп, Ертегілер параллель корпусының үлгісін ұсынады [9, 14]. Онда авторлар ертегі мәтіндерінің тілдік бірліктерінің сандық, сапалық талдамасын автоматты көрсету үлгісін және жиілік, параллель теңестірілген нұсқа сияқты бірнеше үлгі берген. Оларды төмендегі сызбалардан көре аламыз:



1-сызба. Н. Аитова, М. Бақытқызы ұсынған «Ер Төстік» ертегісіндегі сөздердің жиілік үлесі



2-сызба. Н. Аитова, М. Бақытқызы ұсынған «Ер Төстік» ертегісі мәтінінің сегменттерінде сөздердің қайталану жиілігі

Бұл суреттерге қарап фольклорлық мәтіндерді ұлттық корпусқа енгізу бірнеше кезеңнен тұратын күрделі үдеріс екендігін байқаймыз. Алдымен мәтіндер әртүрлі дереккөздерден (қолжазбалар, архив материалдары, баспа жинақтар, аудио жазбалар) жинақталады. Кейін олар цифрлық форматқа көшіріліп, редакцияланады. Бұл кезеңде мәтіннің түпнұсқалық сипатын сақтай отырып, орфографиялық және графикалық біріздендіру жүргізіледі. Келесі маңызды кезең – метадеректерді белгілеу. Фольклорлық мәтіндер үшін жанры, айтушысы немесе орындаушысы, жазылып алынған орны мен уақыты, жинаушысы сияқты мәліметтер аса маңызды. Бұл ақпарат фольклорлық дискурсты кеңістік-уақыттық тұрғыдан зерттеуге мүмкіндік береді. Фольклорлық мәтіндерді корпусқа енгізуде лингвистикалық белгілеу ерекше рөл атқарады. Морфологиялық, лексикалық және синтаксистік белгілеумен қатар фольклорға тән қосымша белгілерді енгізу қажеттілігі туындайды. Мысалы, құрылымдық тіркестерді, архаизмдерді, диалектілік элементтерді немесе мифологиялық атауларды арнайы белгілеулер арқылы көрсету фольклорлық дискурсты тереңірек талдауға мүмкіндік береді. Алайда фольклорлық дискурстың ерекшеліктері автоматты өңдеу барысында бірқатар қиындықтар туғызады. Варианттылық, еркін синтаксистік құрылымдар, нормадан тыс тілдік бірліктер автоматты талдау жүйелерінің дәлдігін төмендетуі мүмкін. Сондықтан фольклорлық корпус жасау барысында автоматтандырылған және қолмен белгілеу әдістерін үйлестіре қолдану тиімді.

## ТАЛҚЫЛАУ

Фольклорлық дискурстың корпустағы көрінісі ең алдымен эпостық жырлар, ертегілер, мақал-мәтелдер мен тұрмыс-салт жырлары арқылы байқалады. Бұл жанрлар дискурстық құрылымы жағынан әртүрлі болғанымен, барлығына ортақ мәдени кодтар мен тілдік модельдер тән. Фольклорлық дискурстың басты белгілерінің бірі – құрылымдық тіркестердің тұрақты қолданылуы. Фольклорлық мәтіндерде қайталанатын құрылымдардың көпшілігі ертегілерде кездеседі. Мысалы:

*Ерте, ерте, ертеде,  
Ешкі жүні бөртеде...*

Бұл құрылым ертегілік дискурстың басталуын білдіретін дискурстық маркер қызметін атқарады.

Мұндай құрылым тыңдаушыны мифологиялық кеңістікке енгізеді, мәтіннің фольклорлық табиғатын айқындайды, уақыт пен кеңістіктің шартты сипатын білдіреді.

Сондай-ақ, фольклорлық дискурста қазіргі әдеби тілде сирек қолданылатын сөздердің жиі кездеседі. Мысалы:

*Алдаспаным жарқ етті,  
Ақ сауытын кигізіп...*

Мұндағы *алдаспан*, *сауыт* сөздері – тек қару-жарақ атауы емес, батырлық жырларда жиі ұшырасатын атаулар. Бұл лексемалар көбіне батырлық эпос мәнмәтінде, ерлік, жаугершілік семантикасымен қатар қолданылады. Батырлық эпостардағы архаикалық лексемалардың (*алдаспан*, *сауыт*, *найза*) жоғары жиілігі А.А. Потебняның тілдің кумулятивтік қызметі туралы теорияларымен үйлеседі. Аталған ғалымдардың пікірінше, фольклор тілінде сақталған архаизмдер – өткен тарихи-мәдени кезеңдердің тілдік ізі [10, 217]. Қазақ тіл білімінде бұл тұжырымды Е.Жанпейісов жалғастырып, эпостық лексиканың тарихи-танымдық мазмұнын арнайы зерттеген [11, 15].

Мақал-мәтелдер – фольклорлық дискурстың ең жинақы әрі концептуалды жанры. Корпустық деректер олардың мәдени концептілерді репрезентациялаудағы рөлін айқын көрсетеді. Мысалы: *Қонақ келсе – құт*. Бұл мәтінде қонақ концептісі тек номинативтік емес, бағалауыштық қызмет атқарады. Корпус материалдары бұл сөздің *құт*, *береке*, *ырыс* сияқты лексемалармен жиі тіркесетінін көрсетеді. Осылайша фольклорлық дискурс арқылы қазақ мәдениетінің басты құндылықтарының бірі – қонақжайлық тілдік деңгейде бекітіледі.

Фольклорлық дискурстағы мәдени концептілердің тілдік көрінісі цифрлық гуманитаристика мен корпус лингвистикасының маңызды зерттеу нысандарының бірі болып табылады. Ж. Толберт пен Е. Джонсон фольклорлық мәтіндердегі қайталанатын тілдік бірліктер мен тұрақты модельдер мәдени жадтың көрінісі ретінде қызмет ететінін атап көрсетеді. Ғалымдардың пікірінше, корпус негізіндегі талдау фольклорлық мәтіндердегі мәдени кодтар мен концептілердің дискурстық қызметін анықтауға мүмкіндік береді [12, 330]. Мысалы:

1-нұсқа:

*Жеті күн, жеті түн жол жүріпті.*

2-нұсқа:

*Қырық күн, қырық түн жол кешіпті.*

Бұл айырмашылық фольклорлық дискурстың нақты уақытты емес, шартты уақытты білдіретінін, символдық сан атауларының (жеті, қырық) мәдени мәнге ие екенін көрсетеді. Корпустық салыстыру бұл нұсқалардың географиялық және жинақталған кезеңіне байланысты өзгеретінін айқындауға мүмкіндік береді. Символдық сан атауларының (үш, жеті, қырық) жиілігі К. Леви-Стросстың мифологиялық ойлау жүйесі туралы тұжырымдарымен үндеседі. Ғалым мифтік санада сан атаулары символдық мәнге ие болатынын көрсетеді [13, 152].

Фольклорлық мәтіндерде бағалаушылық құрылымдар айқын көрінеді. Бұл құрылымдар ұлттық дүниетанымның тілдік бейнесін көрсетеді. Мысалы:

*Жаманнан жирен,*

*Жақсыдан үйрен.*

Жақсы және жаман оппозициясының фольклорлық дискурста өте жиі кездеседі. Бұл бинарлық модель моральдық нормаларды бекітеді, әлеуметтік мінез-құлықты реттейді, тәрбиелік функция атқарады. Бағалаушылық оппозициялар (жақсы-жаман, адал-арам) фольклорлық дискурстың нормативтік сипатын көрсетеді.

Жоғарыда келтірілген мысалдар фольклорлық дискурстың қазақ тілінің ұлттық корпусында толыққанды лингвистикалық дереккөз ретінде қызмет ете алатынын дәлелдейді. Корпус фольклорлық мәтіндерді сандық (жиілік, таралу), сапалық (мағына, функция), салыстырмалы (жанр, аймақ, кезең) тұрғыдан зерттеуге мүмкіндік береді. Корпус лингвистикасы фольклорлық дискурсты зерттеуде жаңа ғылыми мүмкіндіктер ұсынады. Үлкен көлемдегі мәтіндерге сүйене отырып, белгілі бір тілдік құбылыстың жиілігін, таралу аймағын және қолданыс ерекшелігін анықтауға болады. Мысалы, эпостық жырлардағы тұрақты эпитеттердің жиілігі немесе ертегілердегі белгілі бір синтаксистік модельдердің қолданысы нақты сандық деректер арқылы дәлелденеді. Сонымен қатар, фольклорлық дискурсты қазіргі әдеби тілмен салыстыру арқылы қазақ тілінің тарихи дамуын, тілдік сабақтастық пен өзгерістерді зерттеуге болады. Бұл қазақ тілінің эволюциясын түсінуде маңызды ғылыми нәтижелерге қол жеткізуге мүмкіндік береді. Сонымен, фольклорлық дискурстың негізгі тілдік белгілерін төмендегі кестеден көре аламыз:

1-кесте. Фольклорлық дискурстың негізгі тілдік белгілері

Белгі	Мысал
Құрылымдар	Ертеде, жеті күн, жеті түн, бай болыпты, төрт-түлігі сай болыпты, т.б.
Варианттылық	<b>Ұрықсат</b> алып Танаштан, Шоң батыр сонда аттанды, Жаудан кекті алмаққа Қаруын алып, саптанды  Алтайға әкеп жабады. Қырғызға бірге барам, – деп, Сұрағанда <b>рұқсат</b> , Барақ батыр бермеген
Бағалаушылық	батыр – оң, жау – теріс

Бұл белгілердің әрқайсы фольклорлық дискурс корпусын құруда қарастырылуы тиіс. Себебі корпус құруда негізгі корпус талаптарынан бөлек мәтіндерге сай және олардың барлық белгілерін қамтитын талаптар ескерілуі маңызды.

Фольклор мәтіндерінде варианттылық жиі кездеседі. Бұл – бір сөздің не қосымшаның әртүрлі формада берілуі. Варианттылықтың болуы фольклорлық мәтіндерді зерттеу барысында сол кездегі лексикалық не морфологиялық ерекшеліктерді зерттеуге мүмкіндік береді. Ендеше,

2-кесте. Фольклорлық дискурс мәтіндеріндегі варианттылық мысалдары

Варианттылық	Мысалдар
Бір қосымшаның екі түрлі формада берілуі. -менен формасы	<p>Асыл Алтай <b>сәніменен</b>                      Тастан суын қайнатып.                      Алтайда бейғам ел жатыр,                      Балығы шоршып көл жатыр,                      Жер қайысқан шерікпен                      Келе жатыр он батыр</p> <p>Алтынды-күміс <b>сәнменен</b>                      Сылдыраған <b>әнменен</b>                      Көз жанарын алады</p>
<b>Сан</b> есімдердің <b>варианттылығы</b>	<p><b>Жеті күн, жеті түн</b> той жасалды.                      Елдің басы қосылды...</p> <p><b>Қырық күн, қырық түн</b> той жасалды.                      Елдің басы қосылды...</p>
<b>Әріптерді түсіріп қолдану</b>	<p><b>Ұлықпан</b> хакім пайғамбар,                      Жәрдем бер, – деп сұрады.                      Мерттіккендер орнынан                      Төрт-бес күнде тұрады.</p> <p><b>Лұқпан</b> хакім аузында                      Шеріктердің жадында                      Жәрдем ет, хакім, – деген ед</p> <p><b>Дұшмандар</b> естен танады.                      Ер Танашты іздеген                      Абылай, Барақ елшісі                      Неше күндей таба алмай</p> <p><b>Дұшпандар</b> естен танады.                      Ер Танашты іздеген                      Абылай, Барақ елшісі                      Неше күндей таба алмай</p>

Бұл варианттылық үлгілері фольклорлық дискурс мәтіндерінің ауызша таралуымен тікелей байланысты. Дегенмен олардың өзі белгілі бір кезеңдегі тілдің құрылымдық ерекшеліктерін анықтауға және соның негізінде зерттеулер жасауға негіз бола алады.

Осы аталған ерекшеліктерді ескере отырып, фольклорлық дискурстың корпусық моделінің үлгісін ұсынамыз. Онда 3-сызбадағыдай белгілер болуы керек деп есептейміз.



### 3-сызба. Фольклорлық дискурстың корпуслық моделі

Фольклорлық мәтіндер корпусқа енгізілген жағдайда оларды жай көркем немесе этнографиялық дерек ретінде емес, белгілі бір дискурстық жүйе ретінде қарастыру қажет. Бұл тұрғыда фольклорлық дискурс әлеуметтік жадты, ұжымдық дүниетанымды және мәдени нормаларды тіл арқылы бекітетін тұрақты коммуникативтік модель ретінде танылады. Қазақ тілінің ұлттық корпусындағы фольклорлық мәтіндерге жүргізілген талдау олардың бірқатар дискурстық ерекшеліктерін анықтады. Атап айтқанда, құрылымдар мен қайталаулар фольклорлық дискурстың негізгі ұйымдастырушы механизмдері ретінде көрінеді. Мысалы, ертегілердегі дәстүрлі бастау формулалары («Ерте, ерте, ертеде...») мен уақытты шартты білдіретін құрылымдар («жеті күн, жеті түн») мәтіннің жанрлық сипатын белгілеп қана қоймай, тыңдаушыны дискурстық кеңістікке енгізу қызметін атқарады. Мұндай формулалар корпуслық деңгейде жиілік арқылы айқындалып, фольклорлық мәтіндерді автоматты түрде тануға мүмкіндік беретін маркерлер ретінде қарастырылуы мүмкін. Сонымен қатар, корпус деректері фольклорлық дискурстың бағалауыштық сипатын нақты көрсетуге мүмкіндік береді. Батырлық жырлар мен аңыздарда жағымды және жағымсыз кейіпкерлердің тілдік репрезентациясы айқын бинарлық оппозицияға негізделген. Бұл құбылыс фольклорлық дискурстың аксиологиялық жүйесін айқындап, оның тәрбиелік және идеологиялық қызметін көрсетеді. Корпусқа негізделген талдау мұндай бағалауыштық бірліктердің жиілігін, контекстік қолданысын және жанрлар арасындағы айырмашылықтарын салыстырмалы түрде зерттеуге жол ашады.

Фольклорлық дискурстың маңызды қырларының бірі ретінде интертекстуалдық байланыстарды атап өтуге болады. Корпус деректері негізінде әртүрлі жанрлар арасындағы ортақ мотивтер мен сюжеттік үлгілердің тұрақты қайталанып отыратыны байқалады. Бұл құбылыс фольклорлық мәтіндердің тек жеке шығармашылық акт ретінде емес, тұтас мәдени код жүйесінің элементі ретінде қызмет ететінін көрсетеді. Мысалы, батырлық жырлар мен ертегілер арасында кездесетін «жолға шығу», «сынақтан өту», «көмекші кейіпкердің пайда болуы» сияқты мотивтер дискурстық деңгейде ортақ семантикалық өрісті қалыптастырады. Осыған байланысты фольклорлық мәтіндердің вариативтілігі де ерекше назар аударуды қажет етеді. Бір сюжеттік желінің әртүрлі нұсқалары корпус ішінде сақтала отырып, олардың аймақтық, орындаушылық және тарихи факторларға байланысты өзгеріп отыратыны анықталады. Мұндай вариативтілік фольклорлық дискурстың динамикалық табиғатын айқындап, оның тұрақты құрылым мен өзгермелі мазмұн арасындағы тепе-теңдігін көрсетеді. Корпуслық талдау нәтижелері фольклорлық дискурстың прагматикалық қызметін де ашуға мүмкіндік береді. Мәтіндерде жиі кездесетін үндеу, бата беру, қарғыс айту, немесе ақыл-кеңес беру формалары тыңдаушыға тікелей ықпал етуге бағытталған коммуникативтік стратегиялар ретінде қызмет етеді. Бұл элементтер фольклорлық мәтіннің тек ақпарат жеткізу құралы емес, сонымен қатар әлеуметтік мінез-құлықты реттейтін құрал екенін дәлелдейді. Осылайша, фольклорлық мәтіндерді корпус аясында дискурстық жүйе ретінде қарастыру олардың құрылымдық, семантикалық және прагматикалық қабаттарын кешенді түрде сипаттауға мүмкіндік береді.

Фин, орыс және ағылшын тілдеріндегі фольклорлық корпусарда мәтіндер жанрлық, мотивтік және прагматикалық деңгейде арнайы таңбаланған. Ал, қазақ тілінің ұлттық корпусында фольклорлық мәтіндер бар болғанымен, олар әлі де толыққанды фольклорлық

ішкорпус ретінде жүйеленбеген. Осыған байланысты қазақ фольклорлық дискурсын цифрлық форматта тиімді зерттеу үшін бірқатар ғылыми-әдіснамалық мәселелерді шешу қажет. Біріншіден, фольклорлық мәтіндер ұлттық корпус құрамында жеке ішкорпус ретінде бөлініп, жанрлық және дискурстық белгілермен таңбалануы тиіс. Екіншіден, формулалық құрылымдар мен тұрақты мотивтерді белгілеуге мүмкіндік беретін аннотациялық модельдер әзірлеу қажет. Үшіншіден, фольклорлық мәтіндердің ауызша табиғатын ескеретін метадеректер (орындаушы, жазып алынған уақыты мен өңірі) корпус құрылымында міндетті компонент ретінде қарастырылуы тиіс. Бұл бағыт дәстүрлі сапалық талдауды сандық деректермен ұштастырып, фольклорды қазіргі цифрлық гуманитаристика аясында қайта пайымдауға мүмкіндік береді.

## ҚОРЫТЫНДЫ

Фольклорлық дискурс қазақ тілінің ұлттық корпусында ерекше ғылыми және мәдени маңызға ие. Ол қазақ халқының тілдік және мәдени мұрасын цифрлық кеңістікте сақтауға және жүйелі түрде зерттеуге мүмкіндік береді. Фольклорлық мәтіндерді ұлттық корпусқа енгізу – қазақ тілін цифрландыру үдерісінің маңызды құрамдас бөлігі. Зерттеу нәтижелері фольклорлық дискурсты корпус лингвистикасы аясында қарастырудың ғылыми әлеуеті жоғары екенін көрсетеді. Қазақ тілінің ұлттық корпусындағы фольклорлық материалдар ұлттық мәдени мұраны цифрлық форматта сақтаудың маңызды құралы бола алады. Алайда фольклорлық дискурстың ғылыми әлеуетін толық ашу үшін фольклорлық мәтіндерді арнайы ішкорпус ретінде жүйелеу, жанрлық және дискурстық аннотацияны тереңдету қажет екені анықталды. Шетелдік фольклорлық корпус тәжірибесі бұл бағытта нақты үлгі бола алады, алайда оларды механикалық түрде көшіру емес, қазақ фольклорының өзіндік ерекшеліктерін ескере отырып бейімдеу маңызды.

Зерттеу нәтижелері қазақ тілін цифрландыру үдерісінде фольклорлық мәтіндер корпусын қалыптастырудың ғылыми және практикалық маңызын негіздейді. Фольклорлық дискурсты корпус негізінде зерттеу қазақ тіл білімінде жаңа бағыттардың қалыптасуына, сондай-ақ ұлттық мәдени мұраны заманауи цифрлық кеңістікте тиімді игеруге мүмкіндік береді. Бұл мақалада ұсынылған тұжырымдар мен әдіснамалық ұсыныстар фольклортану, корпус лингвистикасы және дискурс зерттеулері саласындағы болашақ зерттеулерге теориялық негіз бола алады.

## ҚАРЖЫЛАНДЫРУ ТУРАЛЫ АҚПАРАТ

Мақала Қазақстан Республикасы Ғылым және жоғары білім министрлігі Ғылым комитеті тарапынан қаржыландырылатын «AP26103909 Фольклорлық мәтіндер корпусы: қазақ тілін цифрландырудағы лингвистикалық және IT шешімдер» атты гранттық жоба негізінде жазылды.

*Мақаланың аннотациясын аудару, ГОСТ-қа сәйкес әдебиеттер тізімін түзу, латын әрпіне негізделген нұсқасын әзірлеу кезінде жасанды интеллект пайдаланылды.*

## ӘДЕБИЕТТЕР ТІЗІМІ

- 1 Қасқабасов С.А. Қазақ фольклортануының заманауи парадигмалары және цифрландыру мәселелері // Керуен. – 2022. – Т. 75, № 2. – Б. 18-29.
- 2 Қоңыратбаев Ә. Әдеби жәдігерлер және фольклорлық мәтіндерді цифрлық форматта зерттеу трендтері // Тіл және әдебиет. – 2021. – № 3. – Б. 54-65.
- 3 Crosthwaite P., Ningrum S., Schweinberger M. Research trends in corpus linguistics: A bibliometric analysis of two decades of Scopus-indexed corpus linguistics research in arts and humanities // International Journal of Corpus Linguistics. – 2022. – Vol. 28(3). – P. 15-33.

- 4 Flinterud G. Folk in the Age of Algorithms: Theorizing Folklore on Social Media Platforms // Folklore. – 2023. – Vol. 134(4). – P. 439-461.
- 5 Hagedorn J., Daranyi S. Bearing a Bag-of-Tales: An Open Corpus of Annotated Folktales for Reproducible Research // Journal of Open Humanities Data. – 2022. – Vol. 8. – P. 105-116.
- 6 Качинская И.Б., Малышева А.В. Корпусные методы в исследовании устной народной речи: принципы лингвистической разметки // Вопросы языкознания. – 2022. – № 3. – С. 78-95.
- 7 Нұртазина М. Қазақ тілінің ұлттық корпусын құру мәселелері және фольклорлық мәтіндерді аннотациялау // Тілтаным. – 2023. – № 4. – Б. 45-53.
- 8 Қайдар Ә. Этнолингвистикалық атаулардың корпусық модельдері мен семантикасы // Тіл әлемі. – 2022. – № 2. – Б. 89-98.
- 9 Аитова Н.Н., Бақытқызы М. Қазақ фольклорлық мәтіндерінің корпусы: ғылыми талдау және зерттеу мүмкіндіктері // Л.Н. Гумилев атындағы Еуразия ұлттық университетінің хабаршысы. Серия: Филология. – 2025. – № 3 (152). – Б. 11-27.
- 10 Потебня А.А. Мысль и язык (Академическое переиздание) // Вестник славянских культур. – 2022. – Т. 64. – С. 124-135.
- 11 Жанпейісов Е. Эпос тілінің этнолингвистикалық және корпусық сипаты // Тіл және әдебиет. – 2021. – № 1. – Б. 12-23.
- 12 J.A. Tolbert, E.D.M. Johnson. Digital folkloristics: Text, ethnography, and interdisciplinarity // Western Folklore. – 2023. – Vol. 78(4). – P. 327-356.
- 13 Леви-Стросс К. Мифологическое мышление (Антология структурной антропологии) // Вопросы философии. – 2021. – № 8. – С. 142-154.

Материал 29.03.2026 баспаға түсті

### **Фольклорный дискурс и его репрезентация в национальном корпусе казахского языка**

К.К. Сарекенова<sup>1</sup>, Ж.К. Омирбекова<sup>1</sup>

<sup>1</sup>Евразийский национальный университет имени Л.Н. Гумилева, Астана, 010000, Казахстан

*В данной статье фольклорный дискурс рассматривается комплексно на стыке современных направлений казахского языкознания. Основным объектом исследования являются языковые, структурные и функциональные особенности фольклорных текстов, выявляемые на основе данных национального корпуса казахского языка. Фольклорный дискурс характеризуется как сложная система, охватывающая различные жанры и отличающаяся устойчивыми структурами, оценочной системой и способностью передавать национальный культурный код. В ходе исследования рассматривается представление фольклорных текстов в национальном корпусе казахского языка, а также их частотные, вариативные и контекстуальные особенности. Анализируются вопросы включения фольклорных текстов и связанных с ними материалов в корпус, их разметки и систематизации, а также предлагаются научно-методологические основы данного процесса. Цель исследования – описать корпусную модель фольклорного дискурса и определить научный потенциал материалов и данных фольклорного дискурса в национальном корпусе казахского языка. Результаты исследования позволяют обозначить эффективные пути сохранения и анализа фольклорного наследия в условиях современной цифровизации, когда цифровые технологии всё глубже проникают в языковую и культурную среду.*

*Ключевые слова: фольклорный дискурс, национальный корпус, цифровизация, фольклор, подкорпус.*

### **СПИСОК ЛИТЕРАТУРЫ**

- 1 Каскабасов С.А. Современные парадигмы казахского фольклороведения и вопросы цифровизации // Керуен. – 2022. – Т. 75, № 2. – С. 18-29.

- 2 Коныратбаев А. Литературные памятники и тренды исследования фольклорных текстов в цифровом формате // Тіл және әдебиет. – 2021. – № 3. – С. 54-65.
- 3 Crosthwaite P., Ningrum S., Schweinberger M. Research trends in corpus linguistics: A bibliometric analysis of two decades of Scopus-indexed corpus linguistics research in arts and humanities // International Journal of Corpus Linguistics. – 2022. – Vol. 28(3). – P. 15-33.
- 4 Flinterud G. Folk in the Age of Algorithms: Theorizing Folklore on Social Media Platforms // Folklore. – 2023. – Vol. 134(4). – P. 439-461.
- 5 Hagedorn J., Daranyi S. Bearing a Bag-of-Tales: An Open Corpus of Annotated Folktales for Reproducible Research // Journal of Open Humanities Data. – 2022. – Vol. 8. – P. 105-116.
- 6 Качинская И.Б., Малышева А.В. Корпусные методы в исследовании устной народной речи: принципы лингвистической разметки // Вопросы языкознания. – 2022. – № 3. – С. 78-95.
- 7 Нуртазина М. Проблемы создания национального корпуса казахского языка и аннотирование фольклорных текстов // Тілтаным. – 2023. – № 4. – С. 45-53.
- 8 Кайдар А. Корпусные модели и семантика этнолингвистических наименований // Тіл әлемі. – 2022. – № 2. – С. 89-98.
- 9 Аитова Н.Н., Бакытқызы М. Корпус казахских фольклорных текстов: научный анализ и исследовательские возможности // Вестник Евразийского национального университета им. Л.Н. Гумилёва. Серия: Филология. – 2025. – № 3 (152). – С. 11-27.
- 10 Потебня А.А. Мысль и язык (Академическое переиздание) // Вестник славянских культур. – 2022. – Т. 64. – С. 124-135.
- 11 Жанпейсов Е. Этнолингвистический и корпусный характер языка эпоса // Тіл және әдебиет. – 2021. – № 1. – С. 12-23.
- 12 J.A. Tolbert, E.D.M. Johnson. Digital folkloristics: Text, ethnography, and interdisciplinarity // Western Folklore. – 2023. – Vol. 78(4). – P. 327-356.
- 13 Леви-Стросс К. Мифологическое мышление (Антология структурной антропологии) // Вопросы философии. – 2021. – № 8. – С. 142-154.

Материал поступил в редакцию журнала 29.03.2026

### **Folklore discourse and its representation in the national corpus of the Kazakh language**

К. Sarekenova<sup>1</sup>, ЗН. Omirbekova<sup>1</sup>

<sup>1</sup>Eurasian National University named after L. Gumilyov, Astana, 020000, Kazakhstan

*This article examines folkloric discourse comprehensively at the intersection of contemporary trends in Kazakh linguistics. The main focus of the study is to identify the linguistic, structural, and functional features of folkloric texts based on the data of the national corpus of the Kazakh language. Folkloric discourse is described as a complex system encompassing various genres and characterized by stable structures, an evaluative framework, and the ability to convey the national cultural code. The study explores the representation of folkloric texts in the national corpus of the Kazakh Language, as well as their frequency, variability, and contextual features. Particular attention is given to the issues of incorporating folkloric texts and related materials into the corpus, their annotation and systematization, while also proposing the scientific and methodological foundations of this process. The aim of the research is to describe a corpus model of folkloric discourse and to determine the scientific potential of folkloric discourse materials and data within the national corpus of the Kazakh Language. The findings of the study provide effective approaches for preserving and analyzing folkloric heritage in the current era of rapid digitalization, when digital technologies are increasingly influencing language and culture.*

*Key words: folklore discourse, national corpus, digitalization, folklore, subcorpus.*

## REFERENCES

- 1 Qasqabasov S.A. Qazaq folklorytanuynyn zamandau paradigmatary jane cifrandyru maseleleri [Modern paradigms of Kazakh folklore studies and issues of digitalization]. Keruen, 2022, vol. 75, no. 2, pp. 18-29. (In Kazakh)
- 2 Qonyratbaev A. Adebı jadigerler jane folklorlyq matinderdi tsifrolyq formatta zertteu trendteri [Literary monuments and trends in the study of folklore texts in digital format]. Til jane adebiet, 2021, no. 3, pp. 54-65. (In Kazakh)
- 3 Crosthwaite P., Ningrum S., Schweinberger M. Research trends in corpus linguistics: A bibliometric analysis of two decades of Scopus-indexed corpus linguistics research in arts and humanities. International Journal of Corpus Linguistics, 2022, vol. 28(3), pp. 15-33.
- 4 Flinterud G. Folk in the Age of Algorithms: Theorizing Folklore on Social Media Platforms. Folklore, 2023, vol. 134(4), pp. 439-461.
- 5 Hagedorn J., Daranyi S. Bearing a Bag-of-Tales: An Open Corpus of Annotated Folktales for Reproducible Research. Journal of Open Humanities Data, 2022, vol. 8, pp. 105-116.
- 6 Kachinskaya I.B., Malysheva A.V. Korpusnye metody v issledovanii ustnoy narodnoy rechi: printsipy lingvisticheskoy razmetki [Corpus methods in the study of oral folk speech: principles of linguistic annotation]. Voprosy Jazykoznanija, 2022, no. 3, pp. 78-95. (In Russian)
- 7 Nurtazina M. Qazaq tilinin ulttyq korpusyn quru maseleleri jane folklorlyq matinderdi annotatsiyala [Issues of creating the national corpus of the Kazakh language and annotation of folklore texts]. Tiltanym, 2023, no. 4, pp. 45-53. (In Kazakh)
- 8 Qaidar A. Etnolingvistikalıq ataulardyn korpustyq modelleri men semantikasy [Corpus models and semantics of ethnolinguistic names]. Til alemi, 2022, no. 2, pp. 89-98. (In Kazakh)
- 9 Aitova N.N., Baqytqyzy M. Qazaq folklorlyq matinderinin korpusy: qylımi taldaıu jane zertteu mumkindikteri [Corpus of Kazakh folklore texts: scientific analysis and research opportunities]. L.N. Gumilev atyndagy Euraziya ulttyq universitetinin habarshysy. Seriya: Filologiya, 2025, no. 3 (152), pp. 11-27. (In Kazakh)
- 10 Potebnya A.A. Mysl i yazyk (Akademicheskoe pereizdanie) [Thought and language (Academic edition)]. Vestnik Slavianskikh Kul'tur, 2022, vol. 64, pp. 124-135. (In Russian)
- 11 Zhanpeisov E. Epos tilinin etnolingvistikalıq jane korpustyq sipaty [Ethnolinguistic and corpus character of the epic language]. Til jane adebiet, 2021, no. 1, pp. 12-23. (In Kazakh)
- 12 J.A. Tolbert, E.D.M. Johnson. Digital folkloristics: Text, ethnography, and interdisciplinarity. Western Folklore, 2023, vol. 78(4), pp. 327–356.
- 13 Levi-Strauss C. Mifologicheskoe myshlenie (Antologiya strukturnoy antropologii) [Mythological thinking (Anthology of structural anthropology)]. Voprosy Filosofii, 2021, no. 8, pp. 142-154. (In Russian)

Received: 29.03.2026

IRSTI 16.21.51

DOI: [10.59102/kufil/2026/iss2pp149-164](https://doi.org/10.59102/kufil/2026/iss2pp149-164)

**R. Saurbayev<sup>1</sup>, \*A. Zhetpisbay<sup>2</sup>, B. Kulbayeva<sup>2</sup>, N. Omarov<sup>2</sup>**

<sup>1</sup>Toraighyrov University, Pavlodar, 140008, Kazakhstan

<sup>2</sup>Pavlodar Pedagogical University named after Alkey Marghulan, Pavlodar, 140008, Kazakhstan

## STYLISTIC DIFFERENTIATION OF NOUNS IN THE ENGLISH LANGUAGE

*The aim of this study is to investigate whether tracing a word back to its historical roots offers a workable, unbiased benchmark for sorting English nouns into stylistic tiers within today's lexicon. Rather than treating etymology as a mere curiosity, the paper asks if pedigree can reliably signal whether a term belongs to high registers or colloquial strata.*