



“O‘ZBEK FOLKLORI VA SHEVALARI TADQIQOTLARI: AMALIYOT, METODOLOGIYA, YANGICHA YONDASHUV”

mavzusidagi xalqaro ilmiy-nazariy
konferensiya materiallari

2026-yil 23-may

*(Professor Samixon Ashirboyev tavalludining
80 yilligi munosabati bilan tashkil etilgan)*



“O‘zbek folklori va shevalari tadqiqotlari: *amaliyot, metodologiya, yangicha yondashuv*” mavzusidagi xalqaro ilmiy-nazariy konferensiya materiallari
/ Mas’ul muharrir: Sh.Sirojiddinov. – Toshkent, 2026. bet.

Mas’ul muharrir:
Shuhrat Sirojiddinov
filologiya fanlari doktori, akademik

Tahrir hay’ati:

Normurodova N.Z. filologiya fanlari doktori, professor;
Jabborov N.A. filologiya fanlari doktori, professor;
Ashirboyev S. filologiya fanlari doktori, professor;
Tilavov A.X. filologiya fanlari nomzodi, dotsent;
Xidraliyeva Z.R. filol. fanlari bo‘yicha falsafa doktori (PhD), dotsent v.b.;
Aliqulova H.Q. filol. fanlari bo‘yicha falsafa doktori (PhD), katta o‘qituvchi;
Turdialiyev A.X. filol. fanlari bo‘yicha falsafa doktori (PhD), dotsent v.b.;
Alimardanov E.I. filol.fanlari bo‘yicha falsafa doktori (PhD), kattao‘qituvchi;
Qodirov U.A., katta o‘qituvchi.

Ushbu to‘plamda Alisher Navoiy nomidagi Toshkent davlat O‘zbek tili va adabiyoti universitetida 2026-yil 23-mayda o‘tkazilgan **“O‘zbek folklori va shevalari tadqiqotlari: *amaliyot, metodologiya, yangicha yondashuv*”** mavzusidagi xalqaro ilmiy-nazariy konferensiya materiallari hamda filologiya fanlari doktori, professor Samixon Ashirboyevning tabarruk yoshi munosabati bilan yo‘llangan tabriklar o‘rin olgan.

To‘plam materiallaridan filolog mutaxassislar, ilmiy tadqiqotchilar, magistratura va bakalavriat talabalari, shuningdek, keng o‘quvchilar ommasi foydalanishlari mumkin. Maqolalar sifati, mazmuni va haqiqiyliги uchun muallifning o‘zi javobgar hisoblanadi.



60	Saida NORMAMATOVA, Surxondaryo qipchoq shevalarida chorvachilik oid leksika	703
61	Feruza QOSIMOVA, O‘zbek dialektologiyasida qo‘llanilgan transkripsion belgilar tizimi	707
62	Dilnoza RAIMBERDIYEVA, “Oychinor” dostonining fonetik dialektizmlari tahlili	711
63	Shohida MAMARAJABOVA, Uzun tumani areali bo‘ylab qo‘ng‘irot shevalarining tarqalishi	713
64	Shahzoda MURODOVA, “Go‘ro‘g‘li” dostonlarida boshlanma	717
65	Mohidil RAJAMATOVA, Parkent tumani folklori namunalari tahlili	723
66	Holida HUSANOVA, “Zarbulmasal”dagi maqollar tahlili	726
67	Gulsevar TURSUNOVA, Anvar Obidjonning bolalar adabiyotiga qo‘shgan hissasi	730
68	Malika AMINJONOVA, Namangan o‘zbek shevalarining o‘zbek adabiy tiliga munosabati	735
69	Behruz BOBOYOROV, O‘zbek tilida blingvizm hodisasi va uning o‘rganilish talqinlari	739
70	Nazokat QODIROVA, Sayfi Saroyining "Guliston bit-turkiy" asari leksikasining tematik guruhlar	745
71	Shodiya QOSIMOVA, Muyassar XOLOVA, O‘zbek milliy korpusiga yozma matnlarni kiritishning metodologiyasi: tematik va metama’lumotli (Mingbuloq tumani misolida)	748
72	Muborak XAYDAROVA, Muyassar XOLOVA, Dialekt va nutq madaniyati	751
73	Muxlisa KUCHIMOVA, Mifologiyada hayvon obrazlarining badiiy ifodasi	748
74	Marjona TOG‘AY, “Qoraqayun” romanida qo‘llangan leksik dialektizmlar tahlili	752

**O'ZBEK MILLIY KORPUSIGA YOZMA MATNLARNI KIRITISHNING
METODOLOGIYASI: TEMATIK VA METAMA'LUMOTLI**

(Mingbuloq tumani misolida)

**METHODOLOGY FOR INCLUDING WRITTEN TEXTS IN THE UZBEK
NATIONAL CORPUS: THEMATIC AND METADATA**

(on the example of Mingbuloq district)

Muyassar XOLOVA,

TerDU O'zbek tilshunosligi kafedrasida dotsenti,

Shodiya QOSIMOVA,

TerDU Kompyuter lingvistikasi magistranti

10.52773/tsuull.conf.2026/UYUZ6510

Annotation. *This thesis examines the methodological principles of integrating written texts into the Uzbek National Corpus through thematic and metadata-based approaches using the Mingbuloq district dialect as a case study. The research highlights the processes of selecting, digitizing, classifying, and linguistically annotating dialectal written texts. Furthermore, the study discusses the importance of metadata in corpus organization, including issues of standardization and systematization.*

Key words: *Uzbek National Corpus, dialect corpus, Mingbuloq dialect, written text, metadata, linguistic annotation, digitization, computational linguistics.*

O'zbek milliy korpusi — bu o'zbek tilining yozma va og'zaki namunalari asosida shakllantirilgan, tilshunoslik, kompyuter lingvistika, sun'iy intellekt va boshqa sohalarda tadqiqotlar uchun muhim bo'lgan katta hajmdagi ma'lumotlar to'plami hisoblanadi. O'zbek milliy korpusi ning asosiy maqsadi — tilning barcha aspektlarini, jumladan, shevalar, dialektlar, leksik va grammatik xususiyatlarni o'rganish uchun zarur bo'lgan ma'lumotlarni taqdim etishdir. O'zbek milliy korpusini boyitish jarayonida yozma matnlarni tanlash, tartiblash, kiritish muhim metodologik asoslarni talab qiladi. Ayniqsa, shevaga doir materiallar bilan ishlaganda bu yondashuvlar yanada ahamiyatli bo'ladi. Tematik va metama'lumotli yondashuvlar korpusni tizimli, izchil, shuningdek, foydalanuvchiga qulay qilishda muhim rol o'ynaydi.

O'zbek tilining elektron korpusi bir necha yillik ilmiy tadqiqotlar natijasida yaratilgan bo'lib, unda o'zbek tilidagi lug'atlar, veb-sahifalar, morfologik ma'lumotlar bazasi, o'quv adabiyotlar hamda turli janrdagi — ilmiy, rasmiy va badiiy — matnlar majmuasi jamlangan. Shuningdek, O'zbekiston olimlari milliy korpus yaratishni ona

tilimizni ma'naviy meros sifatida saqlab qolish va tilni keng qamrovda o'rganishning muhim vositasi deb hisoblashadi. Masalan, Botir Elov va boshqalar tilni saqlash uchun uni raqamlashtirish shu bilan birga, milliy til korpusini yaratish muhimligini ta'kidlaydi.

O'zbekiston Respublikasi Prezidenti farmoni asosida o'zbek tili milliy korpusini yaratish hamda til ma'naviy merosini jamlagan elektron bazani shakllantirishni ustuvor vazifalardan biri etib belgilangan [Mengliyev:2021]. Milliy korpus – qidiruv imkoni mavjud, raqamlashtirilgan yozma, shuningdek, og'zaki matnlar majmuasi hisoblanib, unda til birliklarining strukturasi va ularning o'zaro bog'liqligi o'rganiladi. Sh.Hamroyeva ta'kidlashicha, milliy korpus til birligining o'zgarishlari, eskirishi, yangi atamalarining paydo bo'lishi kabi jarayonlarni kuzatishda muhim vositalardan biri hisoblanadi.

Milliy korpusning boshqa tilda tuzilgan elektron kutubxonalardan farqi shundaki, korpus til o'rganilishi uchun qiziqarli matnlarni jamlaydi. Korpus lingvistlar, leksikograflar, tarjimonlar, kompyuter lingvistlari, o'qituvchilar va boshqa mutaxassislar uchun zarur bo'lgan ilmiy manba sifatida xizmat qiladi [Hamroyeva:2021].

Korpus lingvistikasi tilni statistika hamda kompyuter vositalari yordamida tahlil qilish bo'lib, bu sohada katta hajmdagi matn to'plamlari – til korpuslari talab qilinadi[Elov:2024].

Dialekt yoki mahalliy sheva matnlarini o'rganishda korpusdagi ma'lumotlar sun'iy bo'lmaganligiga e'tibor qaratiladi: ta'sirchan gapiruvchilar shevalarga ko'ra guruhlab olinadi va talaffuz farqlari maxsus kodlash tizimi orqali saqlanadi. Korpus mazmuniga kiritiladigan matnlar namunasi tasodifiy yoki tanlab olingan asosda bo'lishi, har bir tanlovning reprezentativligi (mahalliy aholi nutqining to'laqonli aks etishi) reja qilinishi lozim. Korpus tuzishda quyidagi bosqichlar tavsiya etiladi:

Ma'lumotlar yig'ish: Matn manbalari sifatida turli janr va mavzulardagi yozma asarlar, badiiy hikoya, maqola, hujjatlarni transkripsiyalari shu bilan birga, og'zaki nutq yozuvlari tanlanadi;

Matnlarni tayyorlash: Yig'ilgan matnlarni kiritishdan oldin raqamlashtiriladi, imlo xatolari tuzatiladi so'ngra aniq bir yozuv tizimiga (lotin yoki kirill) standartlashtiriladi;

Teglash va indekslash: Matnlardagi har bir so'z lemmazatsiya va morfologik-sintaktik teglashdan o'tkaziladi; har bir segment qidiruv imkoniyatlari uchun indekslanadi;

Korpus tizimini yaratish: Qidiruv uchun foydalanuvchi qulay interfeys ishlab chiqiladi so'ngra matnlar tegishli formati XML, TEI da joylanadi [Abdurahmonova:2023].

Mingbuloq tumani (Namangan viloyati) shevasining korpus obyekti sifatida tanlanishiga uning lingvistik o'ziga xos shevaga ega ekanligi sabab bo'ladi.

Mingbuloq tumani shevasining yozma matn sifatida korpusga kiritishda Mingbuloq tumani hududiga tegishli bo'lgan 51 ta MFY hamda aholi istiqomat qiluvchi hudud shevalariga oid bo'lgan matnlarni, audio hamda interviyu ko'rinishida shu hududdagi shevalar mavjud bo'lgan matnlarni qamrab oladi.

Mingbuloq tuman shevasiga doir yozma matnlarni O'zbek milliy korpusiga kiritishning metodologik asoslari yoritiladi. Matnlarni tanlashda mahalliy aholidan yig'ib olingan audiomatnlar, videoyozuvlar rivoyat va folklor namunalaridan foydalanildi.

Tanlash mezoni sifatida shevaning o‘ziga xos fonetik belgilari bo‘lgan (masalan, [q] tovushining [x] ga aylanishi) leksik boyligi qabul qilindi. Matnlarda tematik jihatdan qishloq xo‘jaligi (20%), kundalik hayot (30%), an‘anaviy marosimlar (25%) va boshqa mavzular (25%) sifatida tasniflangan.

Tadqiqotlar ko‘rsatishicha, Mingbuloq shevasiga faqat shu hududda qo‘llaniladigan ko‘plab leksik birliklarga kiradi. Masalan, “voyim” so‘zi (“siqilmoq, vahima qilmoq” ma’nosida) faqat shu hudud shevasidagina uchraydi; “chippang” [Qosimova:2023] esa (“suv to‘lib oqishi” holatini ifodalovchi) atamasi ham bevosita shu tumanning aholisi tomonidan qo‘llanilishini ko‘rish mumkin. Bunday noyob shevalar ma’nolari adabiy lug‘atlarda yoki boshqa hudud shevalarida mavjud emas, shu sababdan ularni korpusga kiritish o‘zbek tilining boy leksikasini mustahkamlashda muhim ahamiyatga ega [Elov: 2024]. Mingbuloq shevasining o‘rganilishi til tarixiy-jadidiy o‘zgarishlari, dialektologiya va leksikografiya uchun ham qimmatli materiallar beradi.

Foydalanilgan adabiyotlar

1. Mengliyev B. Til korpuslari - til bo‘yicha tadqiqot va amaliy topshiriqlar yechimi uchun zarur ish quroli. Ma’rifat, 2021.
2. Hamroyeva Sh. Milliy til korpusi (intervyu). Ma’rifat, 2021.
3. Elov B.B., U.S. Yodgorov, A.U. Yuldashev. Lingvistik korpuslar tipologiyasi va ularga qo‘yilgan talablar. Xorazm Ma’mun akademiyasi axborotnomasi, 2024.
4. Elov B.B., A.U. Yuldashev, U.S. Yodgorov. Til korpusi turlari va umumiy xususiyatlari. Xorazm Ma’mun akademiyasi axborotnomasi, 2024.
5. Qosimova Sh.U. Mingbuloq areali leksikasining boshqa shevalar bilan o‘xshash va farqli jihatlari. “Ommaviy axborot vositalarida til va lug‘at masalalari” (ilmiy konferensiya materiallari), – Namangan, 2023.
6. Abdurahmonova N. Korpus lingvistikasi. – Toshkent: Globe, 2023.
7. Davies M. British National Corpus (BNC) [Internet-resurs]. – Oxford University Press, 1996.
8. <http://www.natcorp.ox.ac.uk/>