

RUS TILI DIALEKTAL SUBKORPUSLARI VA KORPUS TUZISH TAMOYILLARI

Xolova Muyassar Abdulhakimovna
TerDU O'zbek tilshunosligi dotsenti,
filologiya fanlari falsafa doktori (PhD)
e-mail: xolovam@tersu.uz

Annotatsiya. Ushbu tadqiqotda rus tilida dastlabki til korpuslarning paydo bo'lishi va rus tili dialektal korpusda qidiruv sistemasida tizimida tizimlilik haqida so'z boradi. Bundan tashqari rus tili dialektal korpuslarida shevalarni fonetik metarazmetkalanishi, shevalar korpusi uchun birinchi qism sifatida shevalar bankini yaratish, ikkinchi qism sifatida audio va videofayllar uchun transkriptlar ifodasi doirasida qarashlar akslangan.

Kalit so'zlar: NRKY, rus xalqi shevalari elektron bazasi, elektron kutubxona, Starling formati, sheva janri, fonetik xususiyat, korpus taqdimoti fonetik transkripsiya, grammatik teglanishi, shevalar bankini yaratish, dala qo'lyozmalari, audiofayllar, videoyozuvlar, fotosuratlar

Rus tili shevasiga oid korpuslar (NRKY) bo'yicha dastlabki korpus Rus tili Upsal mashina fondi hisoblanib, Shvestiyaning Upsal universitetida ishlab chiqilgan¹. Vaqt o'tishi bilan NRKY bir qancha korpuslarni o'z tarkibida saqlayotgan katta korpus darajasiga ko'tarildi. Ayni kunda faoliyat ko'rsatayotgan rus dialektal korpuslar foydalanish bosqichiga ko'tarilgan hisoblanadi, bular:

rus xalqi shevalari elektron bazasi² (Starling formatida yuklash uchun ruxsat etilgan);

Arxangelsk viloyati, Ustyan tumani dialektal matnlarning multimediyali korpusi (onlayn qidiriv tizimli)³;

rus xalqi shevalari elektron kutubxonasi (onlayn ko'rish va eshitish uchun ruxsat etilgan)⁴ va hokazolar.

Rus xalqi shevalari elektron kutubxonasida hozirga qadar mavjud bo'lgan hamda 2008–2009 yillar oralig'idagi fonogrammalar lingvistik tahlil qilingan va

¹ URL: <http://www.slaviska.uu.se/korpus.htm/> https://ozlib.com/1010217/literatura/korpusy_russkogo_yazyka

² URL: http://www.ruslang.ru/krylov_dialect

³ URL: <http://www.slavist.de/Pushkino/>

⁴ URL: <http://dialekt.corpus.tatar/>

tilshunoslar ushbu yozuvlarni eshitish hamda manbadan foydalanish imkoniyatiga egadir. Rus tili elektron kutubxona orqali boshqa sheva korpuslariga o'tish imkoniyatlari mavjuddir: "Rus dialektologiyasi" axborot markazi, Rus milliy korpusining ichki dialektal korpusiga, rus tili dialektologik xarita⁵, "Rus qishloq tili" dialektologik atlas, "Regional etnolingvistika" kabi bazalar sayti qisman ochiq foydalanish imkoniyatini beradi.

Shevalar nutqi korpusi (jamlanmasi) o'z ichiga: sheva vakilining ismi, sharifi, otasining ismi; yozuv hududi (Arxangelsk, Vladimir, Voronej Tver viloyati va hk.); mavzu (subyektiv hikoya, e'tiqod bayramlari ta'rifi, nikoh to'yi ta'rifi va hk.)ni oladi.

Endi dialektal korpusni **NRKY** (Корпус национального русского языка <http://www.ruscorpora.ru/search-dialect.html>.) misolida ko'rib chiqamiz: ushbu korpus 2005 yildan buyon o'z faoliyatini davom ettirib kelmoqda.

Korpus yaratilishining boshlang'ich bosqichi dialektal matnlar korpusida barcha shevalar yozuvning qayerda qay holatda yozib olinganligi haqidagi hujjat "Yozuv pasporti", grammatik xususiyatlari, matnlarning mavzuviy-janrli xususiyatlari va alohida olingan shevaga oid so'zlarning semantik ma'nosini aniqlashga e'tibor qaratilgan. So'zlar orfografik shakllardan foydalangan holda berish maqsadga muvofiq topilib, yaxlit matn shakldagi tahlillar mavjud, so'zlarning ichki semantik ma'no noziklari esa hisobga olinmagan. 2014 yildan boshlab D.V. Sechenova va I.B. Kachinskaya boshliq guruhlar dialektal korpusning yangi standartlarini ishlab chiqishadi. Unga ko'ra olimlar dialektal korpus bir vaqtning o'zida ham matnlar bilan, ham leksemalarning lingvistik tahlilini berish bilan shug'ullanadi. Hozirgi kunda foydalanuvchiga qisqa matnlar bilan ishlash emas (boshqa ko'pgina ichki korpuslar uchun ma'lumotlar bazasi sifatida badiiy asarlar yoki ommaviy-axborot vositalari: jurnal, gazeta va internet ma'lumotlari xizmat qiladi), to'laqonli keng matnlar bilan ishlash imkoniyatlari bor.

Bundan tashqari NRKYning dialektal korpusi ushbu hududda istiqomat qiluvchi, xususan, Rossiya'ning Yevropaga yaqin qismi, ayni shu hududda ilgari yashagan aholi shevasi, shimoliy rus, keyinchalik yashash hududi sifatida tan olingan Sibir, Uzoq Sharq, Don kabilarning shevalariga oid materiallari asosiga qurilgan materiallar tahlilini o'z ichiga olgan. Korpusning tarkibiy qismiga ilmiy sayohat yozuvlari, ovoz-shifrlil yozuvlar, video-shifrlil yozuvlar, qisqa tirajga ega bo'lgan to'plam va nashrlar, majmualardan joy olgan matnlar tahlili kiradi.

⁵ URL: <http://gramota.ru/book/village/dialects.html>

Korpusda matndan foydalanayotgan tadqiqotchiga dastlab fonetik transkripsiya va urg'ular qo'yilishi saqlanishi bilan taqdim etilgan. Bu kamtirajli, transkripsiyalangan matnlarga ega shevaga oid ilmiy asarlardan foydalanayotgan tadqiqotchilar uchun juda muhim edi. I.B. Kachinskaya NRKYdagi barcha matnlar grammatika darajasini barcha xususiyatlarini to'la qamrab olmagan holda, shu bilan birgalikda korpus tuzish jarayonida sheva matnlarning mavzuviy-janr va fonetik xususiyatlariga umumiy yondashgan holda (xususan, omonimlik) tuzilganligini ta'kidlaydi⁶. Ta'kidlash joizki, korpusda qidiruv tizimi tirli mezonlar, metarazmetkalar darajasida ishlarni amalga oshiradi:

1. Yozuv hududi asosida.
2. Mavzular bo'yicha.
3. Janrlar bo'yicha.
4. Ro'yxatga olish yillari bo'yicha.
5. Tarixiy davr bo'yicha yoki keltirilgan voqeaning vaqtiga asosida.
6. Fonetik xususiyatlar (urg'u) aks etgan holda.
7. Informatorlar haqida ma'lumot (jins, yosh, ta'lim) asosida.
8. Alohida leksemani semantik maydoniga ko'ra.
9. Grammatik dialekt xususiyatlarini hisobga olgan holda.
10. Semantik moslik asosida (ba'zan chetlashlar kuzatiladi).
11. Leksemaning sinonimi va yasalish variativligi asosidagi qidiruv har doim ham emas.

Shevashunosning tadqiqot maydoni doirasida. Matnlar leksik-semantik, grammatik, mavzuviy so'rovnamalarga javob bera olishi uchun razmetkalanishi zarur. Buning uchun maxsus tizimlangan dastur – "Shevashunosning tadqiqot maydoni" yaratilishi kerak hamda dialektologning tadqiqot maydonida nafaqat bitta so'z, balki yaxlit bir matnlar ham o'z tahlilini topishi kerak (razmetkalanishi kerak)⁷.

Metarazmetka uch yo'nalishda olib borilgan:

- a) Kuzatuv manzili;
- b) Fonetik tarkib;
- c) Dialektal matnlar tahlili.

Mazkur yo'nalishlarning har biri o'z ichki tasnifiga asoslaniladi, masalan,

- ✓ kuzatuv manzili:

⁶ И.Б. Качинская, Д.В. Сичинава., Московский государственный университет им. М. В. Ломоносова., Институт русского языка им. В. В. Виноградова РАН/ Дialeктный подкорпус сегодня.

⁷ [Архангельский Т.А., Качинская И.Б., Сичинава Д.В. Корпус диалектных текстов Национального корпуса русского языка: Грамматическая, фонетическая и метатекстовая разметка. Новый стандарт подачи.](#) 15 мая 2009 - 20 июля 2010

- ✓ Matn nomi;
- ✓ Matn taqdim etayotgan shaxs (yozib olayo'tgan shaxsning ish joyi, lavozimi, bog'lanish uchun manzili);
- ✓ Yozuv manzili;
- ✓ Yozuv yili;
- ✓ Informant haqidagi ma'lumotlar (yozib olinayotgan shaxsning ish joyi, lavozimi, kasbi);
- ✓ Matn e'lon qilingan bo'lsa, u haqidagi qaydnomalar;
- ✓ Yozuv saqlanayotgan joy.

Fonetik metarazmetka. NRKYda rus tili shevalaridagi fonetik xususiyatlarini to'la qamrab olinmaganligi ta'kidlanadi⁸:

- unlilar doirasida urg'u olish yoki olmaslik bo'yicha farqlanish⁹;
- undoshlar doirasida portlovchi, sirg'aluvchi va qorishqliklik (ts-lovchilik).

2-rasm.



<http://gramota.ru/book/village/#fon> saytidan olingan rasm

Rus tili korpus dialektologiyasi shevalarni korpusga solish masalalarining yechimini berish barobarida nafaqat katta doirada, balki kichik litsey, maktab o'quvchilari¹⁰ning mustaqil foydalanishi uchun qulay darajaga keltirganligini ko'rib o'tish mumkin.

Korpus prezintatsiyasi fonetik transkripsiyada matnni ikki variantda - T-1: dastlabki matn yozuvlariga to'g'ri keladi, ya'ni matn ilgarigi ko'rinishi asosida; T-2 mukammallashtirilgan yunikodlar asosida beriladi. Agarda kontekst boshqa shriftida

⁸ <https://ruscorpora.ru/new/>

⁹ <https://www.budde.ru/dialektologiya/udarnyy-vokalizm.html>

¹⁰ <http://gramota.ru/book/village/>

bo'lsa, avtomatik tarzda unikodlarga o'girilish ehtimoli mavjud. T-1 dan maxsus tayyorlikka tayangan holda, T-2 avtomatik ravishda detranskriptor-1 dasturi yordamida yaratilgan T-1 va T-2 muvofiqligi:

1-jadval

T-1	T-2
в'й т'ер, н'ан'к'а, н'ос'у	вй тер, нянькя, нёс'у
кон', з'ем'ёл'к'е	конь, земёльке
јес'т', пайд'у, см'ијоца, ружй'о	йёс'ть, пайд'у, смиеца, ружье

Ba'zi murakkabliklar T-1 shaklini qo'llash jarayonida yuzaga chiqish mumkin. Masalan, kirill yozuviga asosidagi Ъ yuqori tor, til oldi, lablanmagan fonemani olsak, u I tovushini bildirish bilan birgalikda yumshatish belgisi sifatida detranskriptorlar tomonidan tanilishi mumkin. T-2da aynan shu belgi, yunikod sifatida avtomatik ravishda I tovushiga o'girish jarayoni kechadi.

Aytish joizki, ushbu masala o'zbek dialektologik korpusi oldidagi muammolarning ham biridir. O'zbek shevalarida jahon transkripsiyasi¹¹ asosida tavsiya etilgan transkripsiyasidan foydalanilgan holda korpus tuzish masalalari yechimida ba'zi qiyinchiliklarga duch kelinishi mumkin. Masalan:

2-jadval

Kiril yozuvi	Kiril transkripsiyasi	Lotin yozuvi	Lotin transkripsiya
Йўл	йӨл	yo'l	Jöl
журибди	жҮрьбдь	juribdi	Žúribdi
Қил	қыл	qil	qìl

Y fonemasining J shalkda: yo'l (standart til) ~ jöl (transkripsion): J fonemasining y fonemasi shaklida Žúribdi (transkripsion) ~ yuribdi (standart til) berilishini yoki standart til tarkibidagi o tovushi transkripsion belgi sifatida orqa qator, o'rta keng, lablangan ot, qol (adabiy tilda o't, qo'l)¹² qo'llanishi sun'iy intellekt tomonidan barcha shu kabi fonemalarning yaxlit bir belgi sifatida qabul qilinishi orqali korpusning tarkibiy qismida anglashilmovchiliklarning ko'payishiga olib kelishi mumkin va ushbu

¹¹ Kholova M. A. Transcription in the corpus of the uzbek national boysun dialect (On the example of Baysun District "J" Dialects), Jurnal of Critical Reviews 2020/5 №7/2394-5125

¹² Ashirboyev S. O'zbek dialektologiyasi. –T.:2016. – B. 21.

masalalarning yechimini topish uchun barcha fonemalar uchun ichki yunikoddan tashqari nisbatan farqlilik ega bo'lgan transkripsion belgilar yunikodlarini qo'llash doirasi ishlab chiqish darkor bo'ladi.

Asarlarda va badiiy matnlarda shevaga oid so'zlarning ko'proq orfografik me'yoriy berilishi, lotin grafikasiga asoslangan dialektal materiallarning ham yetishmasligi va ushbu korpusdan foydalanuvchilarga qulaylik tug'dirish va hozirgi texnik imkoniyatlarimizning biroz cheklanganligini, hisobga bazadagi materiallarning ko'p qismlari orfografik qoidalarga bo'ysungan holda berilishi dialektologiyaga oid tadqiqotlarda¹³ o'z nazariy asoslarini topgan.

O'zbek tili shevalar korpusi o'z ichiga aholi istiqomat qiluvchi hudud shevalariga oid bo'lgan matnlarni, shu hudud va uning atrofidan yetishib chiqqan yozuvchilarning asarlar tarkibidagi shevalar mavjud bo'lgan matnlarni qamrab olishi darkor. O'zbek milliy shevalar subkorpusining tadrijiy kengaytirish ayni shu jarayon hisobiga amalga oshiriladi.

O'zbek milliy korpusi til birliklarni bir korpusga jamlash bosqichini boshidan kechirmoqda va shevaga oid matnlarning jamlanmasi lingvistik korpus oldida turgan to'siq bo'lib emas, balki tilshunoslik korpusining to'ldirib boruvchi manba bo'lib xizmat qilmog'i lozim. Bir tomondan shevaning standart holati matnlarining me'yoriy korpusga solingan taqdimoti bilan chambarchas bog'langan. Dialektal matnlar haqida gap borar ekan D.Sichinovaning asosiy e'tiborni ko'proq shevalarning korpusda shifrlanganliga qaratish lozimligi haqidagi fikriga tayanmoq kerak¹⁴. Ammo lingvistik korpus uchun ba'zi shevalar matni me'yorlardan tashqarida shifrlangan holatida uzatilishi oddiy, sheva mutaxassisi bo'lmagan foydalanuvchi qatlam tadqiqotchilarida qiyinchiliklar tug'diradi. Shevalar korpusining boshqa korpuslardan asosiy farqlaridan biri uning leksik-grammatik teglanishi va bunda asosiy e'tibor leksik merazmetkaga berilishi kerak, sababi shevalar uchun grammatik farqlilik mavjud bo'lsa-da, yagona me'yorlashtirilgan grammatik birliklarning ishtiroki boshqa shevalarning qisman tushunilishiga olib keladi

Korpusni tuzish nisbatan murakkab ishlardan biridir. Bu ish o'z-o'zidan ikki qismni tashkil etadi.

Birinchi qism - shevalar bankini yaratish, ya'ni shevashunos olimlar tomonidan hozirga qadar olib borilgan dala, qishloq va aholi yashash hududlarda olib borilgan ekspeditsiya materiallari. Bular:

¹³ <https://ruscorpora.ru/new/corpora-other.html>

¹⁴ И.Б. Качинская, Д.В. Сичинава., Корпус диалектных тестов в национальном корпусе русского языка: сегодняшнее состояние и перспективы

✓ Oliy ta'lim muassasasida o'qiyotgan talabalar tomonidan dialektal amaliyotlar natijasida yig'ilgan dala qo'lyozmalari, audio va videoyozuvlar, fotoapparatlarda tushirilgan suratlar banki;

✓ tadqiqotchi olimlar tomonidan ularning dialektal sayohatlari natijasida yig'ilgan dala qo'lyozmalari, audio va videoyozuvlar, fotoapparatlarda tushirilgan suratlar banki;

✓ shevashunos olimlarning yig'ilgan dala qo'lyozmalari, audio va videoyozuvlar, fotoapparatlarda tushirilgan suratlar orqali shevalar xususida kompyuterda tayyorlangan shaxsiy xulosalari yoritilgan ma'lumotlar banki;

✓ dialektologiya yo'nalishi doirasida nashrdan chiqqan o'quv darsliklari, o'quv qo'llanmalari, majmualari, dialektologik atlaslarga oid olib borilgan hududiy fonetik, leksik, grammatik so'rovnomalar hamda shevashunos olimlarlar tomonidan olib borilgan tadqiqotlar yoritilgan ilmiy to'plamlar banklari yig'ilgan yagona bazani ko'rib chiqish.

Buning uchun birinchi navbatda barcha ma'lumotlar tadrijiy, shevaga oid matnlarni diakritikadan foydalangan holda, transkripsiyada kompyuterda bir xillik topilgan dialektal matnlarni avtomatik tanib, ajratib olish tizimi yordamida qo'lda kiritib chiqiladi. Vaqt o'tishi o'zbek milliy korpusida ichki shevalar korpusini tuzish uchun faqat qo'lyozma shakldagi yozuvlarni yagona korpusining yaratilishi audio, videofayllarning shifrlangan va me'yorlashgan fonetik transkripsiyalari ustida tadqiqotlar olib borilishi uchun zamin bo'lib xizmat qiladi.

Ikkinchi qism. Yuqorida qayd etilgan audio va videofayllar uchun yagona yunikodlarga ega bo'lgan transkriptlarni ishlab chiqish. Buning uchun **shevashunos tadqiqotchining ishlash maydoni** (SHTIM)ni [1-matn yoki 2-matn shaklida] tayyorlash kerak. SHTIMda matnlar yig'uvchi va axborot beruvchi matnlarga ajratilishi, ma'lum belgilarni ko'rsatish uchun probel qo'yish, chunonchi, so'z va so'zshakllarni ikkita alohida so'z sifatida qabul qilish: *uyuni, uyga, uyda, uydi yoki buvi, moma, sheshe* va *hokazo*. Ishning aynan shu qismi anchayin murakkab bo'lib, tadqiqotchidan ko'p kuch va vaqt talab etadi.

Yetkazilayotgan matn axboroti yagona yunikodli mustaqil, tayyorlangan bir yoki bir necha shriftlar orqali UTF-8 (Unicode-da ruxsat etilgan, elektron aloqa uchun ishlatiladigan va kod nuqtalarining barcha 1,112,064 [nb 1] kodlarini bir tizimga solishga qodir bo'lgan o'zgaruvchan kenglikdagi belgilarning tizimi.)¹⁵ txt formatida kodlanadi.

¹⁵ <https://en.wikipedia.org/wiki/UTF-8>

“Ishlov”dan o‘tgan matnlar mavjud fayllarga murojaat qilinganda maxsus tayyorlangan detranskriptorlar- 1-matnni, 2-matnga moslashtirishni amalga oshiradi, ya’ni detranskriptor-I 1-matn va 2-matn orasida oson, mukammal transkripsion tarjimani amalga oshirsa, detranskriptor-II esa 2-matn bilan 3-matn orasida muqobil orfografik variantdagi grammatik analizatorlarni ishga tushiradi. Bunda barcha analitik tahlil jarayonlarini avtomatik amalga oshirish xatoliklarning yuzaga kelishiga sabab bo‘lib qolishi mumkin, shu sababli 3-matnda to 1-matngacha bo‘lgan oraliqda qo‘lda bajarish kerak bo‘lgan amallar ham mavjuddir. 3-matn yozma shaklda amalga oshirishning sababi kompyuterda o‘zbek tiliga oid bo‘lgan imloning dasturining hali ishlab chiqilmaganligi dialektologga matnni belgilangan me’yorga keltirishni qo‘lda amalga oshirishiga to‘g‘ri keladi. 3-matnni grammatik metarazmetkasini qayta ishlash jarayonida ma’lum grammatik belgilar amalga oshiriladi, XML tugmasi va markaziy gorizontall joylashgan RM paneli matn darajasiga qarab grammatik ishni bajaradi. Matndagi barcha analizator (3-matn bo‘yicha ishlaydi), undan so‘ng grammatik shakldoshliklarni ajratib o‘tadi [asos ~ agglyutinatsiya ~ qo‘shimcha (keying sufiks) ~ shakl ~ so‘zshakl ~ o‘tish ~ oraliq ~ qaytish ~ vaqt].

Oralatib boryapti, Necha tovdning toshiga.

Bismillo deb chiqdi, Baland tovdning boshiga. Ushbu matnning birinchi qatorining grammatik metarazmetkadan o‘tkazamiz: <noun> Ora|| <verb.suf>-la|| <verb.suf>-t|| <verb.suf>-ib|| <verb> bor|| <verb.suf>-yap|| <verb.suf>-ti. Aytish joizki, matnninng ikkinchi va to‘rtinchi qatorlaridagi adabiy tilda omonimligi kuzatilmayotgan ba’zi bir grammatik birliklarning sheva korpusini tuzish holatida omonimlikning yuzaga kelishi kompyuter programmmalashtirilishi nazariy yondashuvda bir xil amallar orqali xatoliklarni keltirib chiqarishi mumkin. Masalan, matndagi *tovding* so‘zidagi *-ding* qo‘shimchasi bilan aslida qaratqich kelishigi –ning qo‘shimchasining fonetik varianti hisoblanadi, ammo metarazmetkalash jarayonidagi bu avtomatik farqlanishning imkoni yo‘q. Shu sababli ushbu qo‘shimchaning fe’lning o‘tgan zamon -di suffiksi va fe’lning shaxs-son qo‘shimchasi -ng qo‘shimchasi bilan farqlanishi razmetkalanish jarayonini qo‘lda bajarilib chiqilishi xatoliklarning oldi olinishiga olib keladi.

Shevalar korpusini yaratish oldidan shevaning o‘zbek adabiy tilda yoki o‘zbek tilining izohli lug‘atidagi tutgan o‘rnini belgilab olish ish sur‘atini tezlashtiradi. O‘TILda shevalar shakli, semantik ma’nosi va kelib chiqishiga ko‘ra shaklan bir xil va mazmunan turli xil bo‘lgan shevaga xos so‘zlar, shaklan har xil, ammo mazmunan bir xil xususiyatli so‘zlar, izohli lug‘atga qisman kiritilgan, adabiy tilda dominantasiga ega bo‘lmagan sheva so‘zlari, qadimgi turkiy hamda eski o‘zbek tilidan fonetik va

turli ma'noviy o'zgarishlar bilan yetib kelgan shevaga oid so'z kabi turlarga bo'lib o'rganiluvchi leksemalar badiiy asar tilini, ruhiyatini, mentalligini, hududiy mansublik belgilarini ochib berishga xizmat qiladi va shu bilan birga o'zbek adabiy tilining cheksiz manbasi ekanligiga yana bitta dalildir. O'zbek dialektologiyasida shevalarni bir korpusga yig'ish jarayonida jahon tajribalariga suyanmog'imiz darkor. Shevalar korpusi ustida ish olib borish uchun aniq bir tizimlilik asosida ishlarni yo'lga qo'yimoq, ham shakliy (Portugal, Polsha, Ispan, Bulg'or, Gurjiston, Xitoy-Mandarin, Moldova-Rumin), ham hududiy (Saratov, Kuban, Arxangelsk, Ustyia) korpuslarni ko'zdan kechirgan holda o'zbek shevashunosligidagi o'z yo'nalishimizni aniqlab olishimiz kerak.

Xulosa shuki, dialektal korpusning texnik tomoniga ishlov berilishi uchun katta mablag' va ilmiy salohiyatga ega bo'lgan guruh shakllantirish lozimligini esdan chiqarmaslik lozim, korpus uchun baza vazifasini o'tovchi shevalar jamlanmasini tuzib chiqish hamda respublika shevashunoslari va soha tilshunoslari bilan yaxlit uzviylik asosida ishni yo'lga qo'yish kerak. Shevalar korpusini yaratish jarayoni uzoq va murakkab mehnatni talab etishini hisobga olgan holda korpusning dasturiy ta'minotini ishlab chiquvchi dasturchi texniklar guruhini shakllantirish lozim. Mazkur dialektal korpus tuzishni tadrijiy, ketma-ketlikda amalga oshirish maqsadida lingvistlar, shevashunolar va dastur yaratuvchi texniklar bilan o'zaro kelishgan holda tuzilayotgan korpuslarning mundarijalarini tuzib chiqishi darkor. Shevalar korpusi jahon standartiga mos, foydalanuvchi (tadqiqotchilar)ning tadqiqot doirasini to'laqonli qamrab olish imkoniyatiga ega bo'lgan korpus darajasiga etkazilsagina maqsadga erishilgan hisoblanadi. Shevalar korpusi uchun yagona shevalar bazasini shakllantirgan holda undan foydalanishning qulayligi va ommaviyligini yo'lga qo'yimoq darkor va shevalar korpusida foydalanish yo'riqnomasini ishlab chiqish kerak. Aynan shu korpus o'zbek tili (qipchoq, qarluq-uyg'ur, o'g'uz) shevalarining barcha xususiyatlarini hisobga olgan holda tuzilmog'i kerak. Hozirga qadar shevalarga oid ma'lumotlarning dialektolog olimlar shaxsiy kutubxonalarida saqlanayotganligini hisobga olib qog'oz shaklidagi shevaga oid materiallarni yagona bazaga yig'ish va ulardan ko'chirmalar olish kerak. Dialektolog olimlar va shevashunos tadqiqotchilar tomonidan tayyorlangan video va audioyozuvlar yig'ilgan yagona shevalar bazasini yaratish. Metama'lumotlar qidiruvida informantlar bergan ma'lumotlar bo'yicha yosh, jins, yashash joyi, vaqt oralig'i, mavzu, janr kesimida amalga oshirish imkoniyati bo'lishi kerak. Shevalar korpusiga yig'ilgan mahalliy aholining urf-odatlarini, yashash tarzini olgan shevalar keyinchalik faqatgina shevashunoslik emas, balki psixolingvistika, nutq madaniyati, etnopsixolingvistika, lingvokulturologiya, sostiolingvistika kabi

sohalarga; lemmalangan, ya'ni teglangan so'z, so'zshakl va birliklar esa leksikologiya, grammatika, fonetika va fonologiya sohalariga boy material bera oladi.

FOYDALANILGAN ADABIYOTLAR:

1. Ashirboyev S. O'zbek dialektologiyasi. –Toshkent, Navro'z, 2016.
2. Качинская И.Б. Сичинава Д.В., Московский государственный университет им. М. В. Ломоносова., Институт русского языка им. В. В. Виноградова РАН/Диалектный подкорпус сегодня.
3. [Архангельский Т.А., Качинская И.Б., Сичинава Д.В. Корпус диалектных текстов Национального корпуса русского языка: Грамматическая, фонетическая и метатекстовая разметка. Новый стандарт подачи.](#) 15 мая 2009 - 20 июля 2010.
4. Сичинава Д.В., Качинская И.Б. Корпус диалектных тестов в национальном корпусе русского языка: сегодняшнее состояние и перспективы.
5. Kholova M. A. Transcription in the corpus of the uzbek national boysun dialect (On the example of Baysun District “ž” Dialects), Jurnal of Critical Reviews, 2020. Vol.5 Is.7. – P. 2394
6. Летучий А. Б. [Диалектный корпус: состав и особенности разметки](#) // Национальный корпус русского языка: 2006-2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009, С. 114.

INTERNET MANBALAR:

1. URL: <https://ruscorpora.ru/new/corpora-other.html>
2. URL: [http://www.slaviska.uu.se/korpus.htm./](http://www.slaviska.uu.se/korpus.htm/)
https://ozlib.com/1010217/literatura/korpusy_russkogo_yazyka
3. URL: http://www.ruslang.ru/krylov_dialect
4. URL: <http://www.slavist.de/Pushkino/>
5. URL: <http://dialekt.corpus.tatar/>
6. URL: <http://gramota.ru/book/village/dialects.html>
7. URL: <https://ruscorpora.ru/new/>
8. URL: <https://www.budde.ru/dialektologiya/udarnyy-vokalizm.html>
9. URL: <http://gramota.ru/book/village/>