

Islahot

O'zbek tili milliy korpusi —

Til ma'naviy merosni saqlash va boyitishning asosiy vositasi, shu bois til millatning borligi hisoblanadi. Tillarni saqlab qolish milliy ma'naviyatni, butun bir millatni saqlab qolish ekanligi ziyolilar tomonidan e'tirof etilmoqda. Zero, asl ma'naviyat faqat milliy shakldagina mavjud bo'ladi. Ma'naviyat asosi va ifodalovchisi til. Savollar esa bir talay: Bugungi axborotlashgan asrda tilni qanday saqlab qolish mumkin? Qanday qilib o'zbek tilini dunyo tillari qatoridan o'rinlatish mumkin? Javoblar esa bitta va mushtarak: tabiiy tilni raqamlashtirish va shu asosda milliy til korpusini yaratish lozim. Bu borada mamlakatimizda qanday ishlar amalga oshirilmoqda?



Baxtiyor Mengliyev — Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti "Amaliy tilshunoslik va lingvodidaktika" kafedrasini mudiri, filologiya fanlari doktori, professor

Baxtiyor Mengliyev: — O'zbekiston Respublikasi Prezidentining 2020-yil 20-oktabrdagi "Mamlakatimizda o'zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish chora-tadbirlari to'g'risida"gi farmoni bilan tasdiqlangan "2020–2030-yillarda o'zbek tilini rivojlantirish va til siyosatini takomillashtirish konsepsiyasi"da belgilangan davlat tilining zamonaviy axborot texnologiyalari va kommunikatsiyalariga

faol integratsiyalashuvini ta'minlash ustuvor yo'nalishidagi o'zbek tiliga oid barcha ilmiy, nazariy va amaliy ma'lumotlarni o'zida jamlagan elektron ko'rinishdagi o'zbek tili milliy korpusini yaratish va o'zbek tilini internet jahon axborot tarmog'ida ommalashtirish, unda munosib o'rin egallashini ta'minlash vazifalarining ko'rsatib berilishi biz mutaxassislarga katta mas'uliyat yukladi.

Zamonaviy axborot texnologiyalari tilning funksional imkoniyatlaridan foydalanish borasida benihoya imkoniyatlar eshigini ochdi. Kompyuter tarjimasini, avtomatik tahrir va tahlil, yozma matnni ovozlantiruvchi nutq sintezatorlari, og'zaki nutqni yozma matnga aylantiruvchi nutqni tanish dasturlari, elektron lug'atlar, lingvistik mobil ilovalar, tezaurus(til xazinasilari)lar va til ontologiyasi fikrimizning dalilidir. Ayniqsa, zamonaviy elektron lug'atlar tuzish va undan foydalanish madaniyatini shakllantirish til imkoniyatini egallashda samarador ekanligi o'z isbotini topgan. Xususan, tilning imkoniyatini namoyon qilish va egallash borasida dunyo miqyosida tez sur'atlarda yaratilayotgan til korpuslarining roli beqiyos.

Til korpuslari — til bo'yicha tadqiqot va amaliy topshiriqlar yechimi uchun zarur ish quroli. U oddiy elektron kutubxonadan farqlanadi. Elektron kutubxonaning maqsadi — xalqning ijtimoiy-siyosiy, ma'naviy, iqtisodiy hayotini aks ettiruvchi badiiy va publitsistik asarlarni nisbatan to'liq qamrab olish. Elektron kutubxonada matnlari til nuqtayi nazaridan ishlov berilmaganligi sababli tadqiqotlar uchun noqulaylik tug'diradi. Chunki elektron kutubxonada ilmiy tadqiqot materiali bazasini tayyorlash maqsadida tuzilmaydi, balki milliy-ma'naviy merosni to'plashni maqsad qilgan bo'ladi. Til korpusi esa elektron kutubxonadan farqli o'laroq, tilni o'rganish va tadqiq qilish uchun foydali va qiziqarli matnlarni to'plashni nazarda tutadi.

— "Milliy korpus nima?" degan savol hali ham ko'pchilikni qiziqtiradi. Mamlakatimizda endi shakllanayotgan til korpusi aksariyat kishilar uchun yangilik.

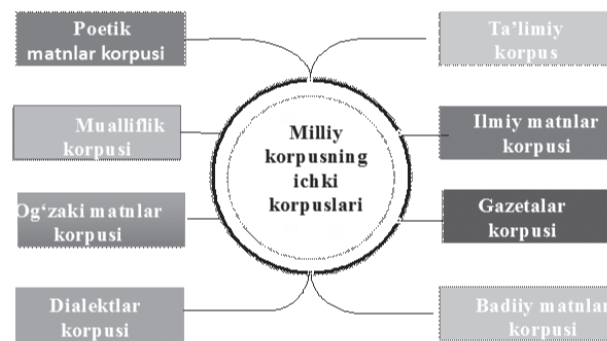
Shahlo Hamroyeva:

— Milliy til korpusi — til birliklarining xususiyatlarini aniqlash maqsadida elektron qidiruv imkoniyati mavjud tizim, tabiiy tilning raqamlashgan yozma va og'zaki matnlar jamlanmasi. Milliy korpus — til birligining o'zgarishi, eskirishi, yangilarining paydo bo'lishi, ma'nosining kengayishi va torayishi, yangi iboralarning paydo bo'lishini kuzatish, an'anaviy va zamonaviy lug'atlar tuzishda keng imkoniyatli dastur-



Shahlo Hamroyeva — Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti doktoranti, filologiya fanlari bo'yicha falsafa doktori (PhD)

lashtirilgan tizim. Korpusdagi milliy so'zi nafaqat tilning, balki korpus tuzilishi va tarkibining ham o'ziga xosligini anglatadi. Milliy korpusning boshqa tur til korpuslaridan farqi ta'limiy korpus, mualliflik korpusi, poetik matnlar korpusi, ilmiy va rasmiy matnlar korpusi, og'zaki matnlar, badiiy matnlar va gazetalar hamda dialektlar korpusi kabi ichki korpuslarni o'z tarkibiga qamrab olganligi bilan belgilanadi.



Undan lingvistlar, leksikograflar, kompyuter lingvistlari, dasturchilar, muharrirlar, tarjimonlar, jurnalistlar, noshirlar, olimlar, o'qituvchilar, ta'lim oluvchilar va boshqa har qanday soha mutaxassisi keng foydalanish imkoniyatiga ega bo'ladi. Korpusiz na nazariy, na amaliy filologiya taraqqiy etadi.

Foydalanishda deyarli kasbiy tabaqalanishga yo'l qo'ymaydigan til korpuslari barcha fan-soha vakillarini birday qiziqtirishi tabiiy. Ayniqsa, ona tili va chet tillarini o'qitish va o'rganish borasida korpuslarning ahamiyati beqiyos. Bugungi kunda dunyo miqyosida til o'rgatish tizimi korpuslarga yo'naltirilayotganligi ham fikrimizning dalili. Shuning uchun ham ta'lim korpusi, sheva matnlari korpusi, mualliflik korpusi, poetik matnlar korpusi, og'zaki, ilmiy, rasmiy matnlar korpusi, gazeta matnlari korpusi kabi qator mikrokorpuslarning tuzilayotganligi ahamiyatli.

— **Bu borada ancha ishlar qilinayotganligidan xabarimiz bor. Ommaviy axborot vositalari, internet saytlari, ijtimoiy tarmoqlarda ham xabarlar berib borilmoqda. Botir aka, korpusning dasturiy tomoni mas'uliyati Sizning jamoangiz zimmasida. Albatta, bu katta kuchni talab qiladi. Shu kungacha bu borada qancha ko'lamdagi ishlar amalga oshirildi?**

Botir Elov:

— Ko'p yillik talab va ehtiyoj asosida ona tilimizni avaylab-asrash, boyitish, undan amaliy foydalanish samaradorligini oshirish maqsadida Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universitetida fidoyi olimlar tashabbusi bilan o'zbek tili milliy korpusining yaratilishiga erishildi. Bu borada universitetimiz rektori Shuhrat Sirojiddinovning qo'llab-quvvatlashlari, ijodiy sharoit va texnik imkoniyatlar eshigini ochib bergani e'tirofga loyiq. Shu o'rinda, "O'zbek tilining ta'limiy korpusini yaratish" amaliy loyihasi zamirida ta'lim korpusining yaratilishi ham ta'lim



Botir Elov — Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti "Axborot texnologiyalari" kafedrasini mudiri, texnika fanlari bo'yicha falsafa doktori (PhD)

fidoyilari uchun benazir imkoniyat eshigini ochadi. O'zbek milliy korpusini yaratish maqsadida shu kunga qadar 10 ga yaqin ilmiy ish himoyasi amalga oshirildi. Bugungi kunda bu borada kompyuter va korpus tilshunosligi va amaliy tilshunosligi sohalarini rivojlantirish, milliy korpus va boshqa tur til korpuslarini yaratish maqsadida magistrlik dissertatsiyalari, fan nomzodi va fan doktorligi bo'yicha ilmiy tadqiqot ishlari amalga oshirilmoqda. Toma-toma ko'l bo'lur deganlaridek, bugun ushbu tadqiqotlarning amaliy natijasi sifatida milliy korpus yuzaga keldi. Ushbu mas'uliyatli va sermashaqqat jarayonda ToshDO'TAU, BuxDU, QarDU, TerDU, NavDPI kabi oliy ta'lim muassasalaridagi 100 ga yaqin mutaxassislarning birdam mehnat qilganligini alohida ta'kidlash joiz. Bugun ToshDO'TAU mutaxassislarni birlashtira oldi, ularni bir maqsad yo'lida jamladi.

— **Milliy korpusning tarkibiy qismlaridan biri ta'limiy korpus. Uning ta'lim uchun ahamiyati qanday?**



O'ral Xoliyorov — Termiz davlat universiteti "O'zbek tilshunosligi" kafedrasini katta o'qituvchisi

xususiyati hamda imkoniyati faqat korpusda mavjud. O'qituvchi yangi, ishonarli, cheksiz hamda xilma-xil misollarni shu yerdan topa oladi, topshiriq, mashqlarni belgilashda qiynalmaydi, bir necha daqiqada mavzu bo'yicha yangi-yangi misollardan iborat topshiriqlarni tayyorlay oladi.

Til birligini qidirish kerak bo'lsa, bunday dasturiy ta'minot, ya'ni korpus tadqiqotchi yoki foydalanuvchiga juda katta yordam beradi. Ilgari tadqiqotchi o'z ishi uchun misolni topish, ularni kartotekaga ko'chirish (kompyuter texnologiyalari rivojlanishidan oldingi davrda)ga oylab, ba'zan yillab vaqt sarflagan bo'lsa, bugun dunyo til korpuslari yordamida sanoqli daqiqada yuzlab misol topishga, tahliliy natijalarni olishga va ular ustida ishlash imkoniga ega bo'ldi. Maxsus qidiruv tizimi korpusdan ma'lumot olishga mo'ljallangan bir qancha dasturdan iborat, u statistik axborot va qidiruv natijasini foydalanuvchiga qulay shaklda taqdim eta oladi. Tilda qanday jarayon kechayotganligini aniq tasavvur qilish uchun korpus qamrovini yanada kengaytirish, nafaqat yozma nutq, balki og'zaki nutq materialidan ham foydalanish maqsadga muvofiq. Bunday korpus yordamida taraqqiyot natijasida tilda sodir bo'lgan va kutilayotgan o'zgarish haqida aniq xulosa chiqarish mumkin.

Tilshunoslikka oid tadqiqotda fakt bilan ish ko'rildigan hollarda material yig'ilishi, sistemaga solinishi lozim. Bunday katta hajmli ishni bajarishda korpus vaqt va mehnatni tejaydigan ish quroli vazifasini bajaradi. U texnik jarayonni tezlashtiruvchi vosita bo'libgina qolmay, muayyan tilning zamonaviy shakliga xos axborot tizimi ham bo'lib, kutilmagan savolga javob bera oladigan, til hodisasi bilan shug'ullanadigan soha oldiga avval ko'rilmagan dolzarb muammolarni qo'ya oladigan tizim.

muhim madaniy voqelik

Korpuslar turli maqsadlarda turli sohalar uchun tuzilishi mumkin. Jumladan, og'zaki yoki yozma matnlar, bir tilli yoki ko'p tilli, uslubiga ko'ra badiiy, publitsistik, folklor, bazasi o'zgaruvchan yoki o'zgarmas, matn hajmiga ko'ra to'liq matnli yoki fragmentli korpuslar bo'ladi.

— **Milliy korpus — murakkab tuzilishga ega virtual borliq. Uning tarkibiy turlari bilan birga bazasi va tuzilishi ham anchayin murakkab. Bu murakkablik korpus imkoniyatlarining kengligini ham ko'rsatadi. Shu haqda ham gapirsangiz.**

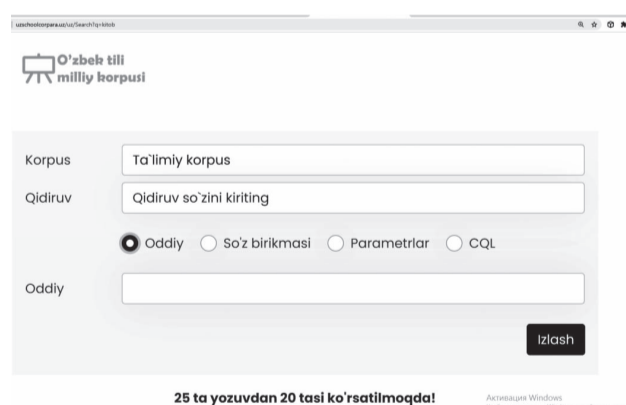
Botir Elov:

— Milliy korpuslar mushtarak va xususiy belgilarga ega. O'zbek tilining korpusi ham milliy shart-sharoitlardan kelib chiqadigan xususiyatlarga ega bo'lishi kerak edi. Korpusni yaratish konsepsiyasida bu puxta ishlab chiqildi va rejalashtirildi.

O'zbek tili milliy korpusining tuzilishi va tarkibi

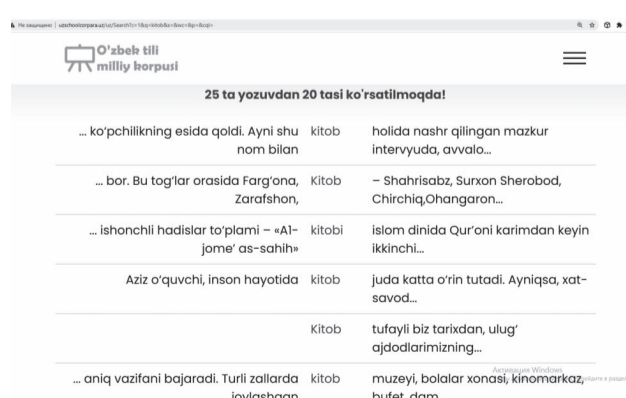


O'zbek tili milliy korpusi korpus interfeysi, qidiruv tizimi, 75 mln.ga yaqin so'z-shakl, 1 mlndan ortiq matnlar bazasi va 15 elektron lug'at manbasidan iborat. Shuningdek, o'zbek tili milliy korpusi saytining birinchi sahifasida korpus va uning tuzuvchilari haqida asosiy ma'lumot, o'ng tomondagi menyuda istalgan sahifaga o'tish imkoniyati mavjud. Korpus menyusi to'rt qismdan iborat. Bosh sahifa, saytning qidiruv oynasi, matn haqida unga birlashtirilgan qo'shimcha ma'lumot ilovasi, qidiruv uchun til birliklariga izoh yozish parametrlari, elektron leksikografik manbalar tavsifi, oxirgi blok korpus tuzuvchilari jamoasi, foydalanilgan dastur, matnlarning mualliflik huquqi haqida to'liq ma'lumotlar bazasidan tashkil topdi.



Korpus bo'yicha qidiruv foydalanuvchiga quyidagilarni aniqlash imkonini beradi:

1) muayyan so'zning barcha shakllari misollar massivi bilan;



2) misollarning qaysi manbadan olinganligini ko'rish;



3) so'zning filologik lug'atlardagi mohiyati va izohini;

4) qidiruvga berilgan so'z bilan uning o'ng va chap tomonidan birlashtirish imkoniyatiga ega so'zlar ro'yxatini;

5) u yoki bu muallifning ayni so'zdan foydalanish chastotasi yoki statistikasini;

6) so'zning o'z va ko'chma ma'nolarini;

7) so'z qo'llanishining yashirin modeli (imkoniyati)ni;

8) til taraqqiyotining turli davrida qo'llanish holatini.

— **O'zbek tili milliy korpusining jamiyatimizdagi roli katta. Milliy korpus madaniyatimiz taraqqiyotida muhim voqelik bo'ldi. Bu borada Manzura Abjalovanning fikrlari bilan tanishsak.**

Manzura Abjalova:



Manzura Abjalova — Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti "Axborot texnologiyalari" kafedrasida katta o'qituvchisi, filologiya fanlari bo'yicha falsafa doktori (PhD)

Lug'atlarni tuzish uchun korpus katta hajmli manba vazifasini o'taydi. Vaqt o'tishi bilan korpus kuchli informatsion resursga aylanadi. Korpus asosida kompyuter yordamida lug'atlar avvaligiga nisbatan tezlik bilan tuziladi va qayta ishlanadi. Shu yo'l bilan ish boshlanishidan tugash jarayonigacha (nashrgacha) tilni aks ettirib turadi, lug'at maqolasi "eskirish"ga ulgurmaydi. Tilning zamonaviy holatini aks ettiruvchi korpus turli davrlarda yashab ijod etgan ijodkorlarning mualliflik korpuslari uchun asos bo'lib xizmat qiladi.

So'zning qo'llanish davri va chastotasini aniqlab berishi bilan korpusning so'zshunoslik sohasidagi ahamiyati namoyon bo'ladi.

Korpus lingvistikasining keyingi taraqqiyot bosqichida statistik tadqiq metodi kompyuter tarjimasini, nutqni sintezlash va tanish, orfografik tekshiruv kabi lingvistik amallarni bajarishda qo'l kela boshladi. Rus va Yevropa tilshunosligida til bo'yicha qilinadigan barcha tadqiqotlarning korpusga asoslanishi odatiy holga aylangan va hatto ba'zi tadqiqotlarda majburiy ham sanaladi. Til o'rganayotgan o'quvchi yoki uning biror jihatini tadqiq etayotgan kishining bugungi kunda lug'atga bo'lgan ehtiyoji korpusga ko'chishi shubhasiz.

Terminologiya sohasini rivojlantirishda, bazadagi millionlab turli sohalariga oid matnlardan muayyan soha terminologiyasi shakllantiriladi. Masalan, ingliz < lotin tillaridan o'zlashgan *injiniring* atamasining bank-moliya, qurilish, axborot texnologiyalari, xalqaro munosabatlardagi tushunchalarini birgina korpusga qidiruv berish natijasida aniqlash mumkin bo'ladi.

Gap qurilishini o'rganishdagi ahamiyati tilning jonli qurilishini, so'zlarning o'zaro birlashtirish imkoniyatini tahlil qila olishida ko'rinadi. An'anaviy tilshunoslikdagi misolni olishda badiiy asar tili-ga tayanish tajribasidan ko'ra korpusga tayanish misolning bugungi kun uchun ham ishonarliligini ta'minlaydi.

Tarjima dasturlarini yaratishda parallel korpuslar katta ahamiyatga ega. Ikki yoki ko'p tillardagi matnlarning parallel tarjima bazasi yordamida tarjimon lug'atlari, avtomatik tarjima dasturlarining lingvistik bazalarini shakllantirishda yirik resurs vazifasini bajaradi.

Uslubiyatni o'rganishda matnning uslubiy o'ziga xosligini tahlil etishda ham korpusga tayanish mumkin. Bu amaliyotni bajarishda matnlarning statistik holati tahlili(matndagi jumlaning uzunligi, bir so'zning boshqa so'z bilan birlashtirish holatining doimiy yoki noodatiyligi)ni aniqlashtiradi. Bunday usulda korpus yordamida yozma nutq bilan birga og'zaki nutqni ham o'rganish mumkin bo'ladi. Masalan, rasmiy ish qog'ozlari matnida *shartnoma bo'yicha talablar* shakli ma'qulmi yoki *shartnoma talablari* konstruksiyasi to'g'rimi? Korpus bu tarzda berilgan savolga javob sifatida sanoqli daqiqalarda qaysi shakl ko'proq qo'llanishini aniqlab beradi va biz rasmiy matnda mazkur birlashtirish to'g'ri ifodasini yozaveramiz. Ta'kidlash joizki, aynan hozirgi adabiy til uslubini shakllantirish uchun ham mutaxassisga zamonaviy matnlar majmuasi zarur bo'ladi.

Korpus bu — *lingvodidaktika*, u ona tili va xorijiy tilni o'rganishda birdek ahamiyatli. Bazaning jonli nutq, zamonaviy OAV materiali, badiiy adabiyot so'zligi, mumtoz nasr namunasi, turli mavzu va janrdagi matnlar bilan doimiy boyitilishi natijasida korpus ta'limda yo'naltirib o'qitish imkoniyatini ham ochadi. Ko'pincha qayta-qayta nashr etilgan darslikda misollar eskiligidan qolib ketadi va axborot bilan shiddatli yangilanib borayotgan bugungi kun uchun mavzu eskiradi, o'quvchi bilimni davr bilan hamohang o'zlashtirmaydi. O'quvchida korpusdan unumli foydalanish ko'nikmasi shakllantirilsa, u o'z nutqida yo'l qo'yilayotgan xatolar va g'aliz ifodalarni bartaraf eta oladi.

Til korpusiga eng ko'p ehtiyoj sezuvchi mutaxassis — bu filologik dasturlar uchun lingvistik ta'minotni yaratish uchun matnni avtomatik qayta ishlaydigan dasturchidir. Chunki u tabiiy til bilan ish ko'radi hamda ushbu tilda yozilgan barcha matnlar strukturasi mukammalroq "tushunishi", his etishi lozim bo'ladi.

Til korpusiga ehtiyoj sezadigan mutaxassislardan yana biri — kundalik ish jarayonida yozma va og'zaki nutq jozibasiga tez-tez murojaat etuvchilar: gazeta va jurnal muharriri, jurnalist, radio va televideniye xodimlari. Chunki bu mutaxassislar ma'lum bir so'z, ibora yoki gapning qo'llanish holati, darajasi, kim, qachon ilk marta bu gapni qo'llaganligi, qanday uslub uchun xoslanishini bilishga grammatika bilan shug'ullanuvchi olimdan ko'ra ko'proq ehtiyoj sezadi. Korpusdan tashqari hech qanday axborot banki bu kabi savollarga zudlik bilan javob berishi mumkin emas. Shuning uchun korpuslar jurnalist, muxbir, muharrir, o'qituvchi hamda dasturchi uchun maxsus yaratilgan, deysish mumkin.

— **Til ta'limi hayotdan uzilib qolmasligi uchun muntazam ravishda matnlar bazasi yangilanib boriladigan korpusga tayanish maqsadga muvofiq bo'ladi. Milliy til korpusi — Ona tilimizning jahonga yuz tutishiga erishish, milliy ma'naviyatni takomillashtirish va yuksaltirishning ham asosiy vositalaridan biri. Bu borada O'zbek tili milliy korpusi jamoasiga muvaffaqiyatlar tilaymiz.**

Davra suhbatini
Husan NISHONOV yozib oldi.