

TIL KORPUSLARI MUHIM LINGVISTIK MANBA SIFATIDA

Maxammadiyeva Diyora,
SamDU 2-kurs magistranti

Annotatsiya: Ushbu maqolada o'zbek tilining milliy korpusini yaratsak, turli soha vakillari uchun til korpusidan foydalanishning ahamiyati yoritilgan. Korpusning lug'atshunos, adabiyotshunos, o'qituvchi, dasturchi, jurnalist, pragmalingvist, sud-huquq lingvistlari uchun nechog'li ahamiyatli ekanligi dalillangan.

Kalit so'zlar: *Raqamlashtirish jarayoni, til korpusi, elektron kutubxona, talimiy korpus, milliy korpus, korpusdan foydalanish, korpusning ahamiyati.*

Abstract: The article emphasizes the importance of the national corpus as a factor that has a positive impact on the development of various sectors. It has been proven how important the corpus is for lexicographers, literary critics, teachers, programmers, journalists, pragmalinguists, forensic linguists.

Keyword: *corpus, educational corpus, national corpus, corpus use, corpus importance.*

Mamlakatimizda tilga e'tibor ma'naviyatga e'tiborning ustuvor yo'nalishlaridan biri darajasiga ko'tarildi. Shu bois ona tilimizni avaylab-asrash, boyitish, undan amaliy foydalanish samaradorligini oshirish bilan birga, o'zbek tilining zamonaviy axborot-kommunikatsiya tizimida keng qo'llanishiga erishish kechiktirib bo'lmaydigan dolzarb vazifaga aylandi. Chunki ona tilimizning jahonga chiqishiga erishish milliy ma'naviyatni takomillashtirish va yuksaltirishning asosiy yo'llaridandir. Raqamli texnologiyalar rivojlangani sari ulardan foydalanish kundalik ehtiyojga aylanib bormoqda. Yangi O'zbekistonda bugun har bir soha, raqamlashtirilish jarayonida. Raqamli iqtisodiyot, elektron hukumat, elektron murojaat kabi tushunchalar hayotimizning bir bo'lagiga aylangan. Tilshunoslik ham

bundan mustasno emas. Til ma'naviy merosimizni asrash va boyitishning asosiy vositasi hisoblanadi. Tillarni saqlab qolish milliy ma'naviyatni saqlab qolish demakdir. Zero, asl ma'naviyat faqat milliy shakldagina mavjud bo'ladi. Zamonaviy axborot texnologiyalari tilning funksional imkoniyatlaridan foydalanish borasida benihoya keng qulayliklar eshigini ochdi. Kompyuter tarjimasi, tahriri, tahlili, elektron lug'atlar va tezaurus(til xazinas)i)lar fikrimiz tasdig'idir. Ayniqsa, zamonaviy elektron lug'atlar tuzish va ulardan foydalanish madaniyatini shakllantirish til imkoniyatini kengaytirishda muhim ahamiyatga ega. Respublikamizda har bir sohada axborot-texnologiyalaridan unumli foydalanishga, dasturiy ta'minot asosida tuziladigan loyihalarga katta e'tibor qaratilmoqda. Yangi O'zbekistoningimizda bo'layotgan o'zgarishlar tilshunoslarimiz oldiga o'zbek tilining lug'at tizimini, so'zlardagi ma'no torayishi, kengayishi, ko'pma'nolilik, yangi paydo bo'lgan tushuncha va terminlarni chuqurroq tadqiq etishni kun tartibiga qo'yimoqda. XX asr texnika taraqqiyoti asri kirib kelishi bilan qator davlatlarning o'z milliy korpuslari yaratib bo'lindi. Albatta, XXI asrda internet, axborot texnologiyalarining keng quloq yoyishi tilning jahon miqyosida e'tirof etilishi, keng miqyosda targ'ib qilinishi zamonaviy kommunikatsiyalar bilan bog'liqligi bilan belgilanishi shubhasiz. Tilning, xususan, o'zbek tilining jahon hamjamiyati tomonidan o'rganilishi, e'tirof etilishi uchun , avvalo, uning asosiy leksik-grammatik bazalari shakllantirilishi va dasturiy ta'minoti yaratilishi lozim. Xo'sh, o'zbek tilini raqamlashtirsak, o'zbek tilining milliy korpusini yaratsak, qanday afzalliklarga ega bo'lamiz. Milliy korpus milliy til xazinas i demakdir. Undan lingvist, leksikograf, kompyuter lingvistlari, dasturchi, muharrir, tarjimon, jurnalist, nashrlar, olim, o'qituvchi, ta'lim oluvchilar va boshqa har qanday soha mutaxassisi keng foydalanadi. Til korpuslari — til bo'yicha tadqiqot va amaliy topshiriqlar yechimi uchun inkor etib bo'lmas 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 kutubxona matnlari til nuqtayi nazaridan ishlov berilmaganligi sababli tadqiqotlar uchun noqulaylik tug'diradi. Chunki elektron kutubxona ilmiy tadqiqot materiali bazasini tayyorlash maqsadida tuzilmaydi, balki milliy ma'naviy merosni jamlashni maqsad qilgan bo'ladi. Til korpusi esa elektron kutubxonadan farqli o'laroq, tilni o'rganish va tadqiq qilish uchun zarur, foydali va qiziqarli matnlarni to'plashni nazarda tutadi. Korpus turli ilmiy adabiyotlarda turlicha ta'riflangan. "Korpus – ma'lum maqsadda yig'ilgan matnlar majmuini tashkil etuvchi til birliklarining yig'indisi", "korpus – alohida tilning yozib olingan og'zaki hamda yozma matnlari asosidagi lingvistik ma'lumotlar majmui", "korpus – istalgan tabiiy (real) tildagi elektron shaklda saqlanadigan yozma yoki og'zaki, kompyuterlashtirilgan qidiruv tizimiga joylashtirilgan matnlar yig'indisi"[Nurillayeva, 2018]. Korpus - ma'lum maqsadda yig'ilgan matnlar majmuini tashkil etuvchi til birliklari yig'indisi, tabiiy tildagi elektron shaklda saqlanadigan yozma yoki og'zaki, kompyuterlashtirilgan qidiruv tizimiga dasturiy taminot asosida joylashtirilgan on-line yoki off-line tizimda ishlaydigan matnlar jamlanmasi[Xamroyeva, 2018] kabi. O'TILda esa korpus so'zi quyidagicha izohlangan: Korpus [lot.corpus – tana, jism, modda; bir butun narsa] 1. Mexanizm, apparat va sh.k.ning tayanchini yoki qobig'ini tashkil etuvchi qism. Soat korpusi. 2. Bir umumiy uchastkadagi, bir majmuaga kiruvchi binolardan biri, alohida bino. 3. bir necha diviziya yoki brigadadan iborat yirik harbiy qo'shilma. 4. Biror davlatdagi barcha diplomatlar, konsullar. Qonun chiqaruvchi korpus qonun chiqaruvchi ayrim muassasalarning yoki qonun chiqaruvchi organga tegishli palatalarning barcha deputatlari[O'TIL]. Ko'rinib turibdiki, O'TILda bizning tatqiq obyektimiz bo'lgan til korpusi haqida hech qanday ma'lumot keltirilmagan. O'TILning 2006-yilda nashr etilganligini hisobga olsak, korpus tilshunosligi o'zbek tilining endigina tetepoya bo'layotgan yangi yo'nalishi ekanligi ma'lum bo'ladi. Vaholanki, jahon tilshunosligi bu borada ancha-muncha yutuqlarga erishgan. Dunyo tilshunosligida korpus lingvistikasi yo'nalishi va til korpuslarini yaratish ishlari bo'yicha talay



ishlar qilingan. Ko'pgina mamlakatlarda XX asrning 80-yillaridan boshlab bunday korpuslar tuzila boshlandi. Ular turli maqsad va vazifalarga xizmat qiladi. Buyuk Britaniyada Ingliz tili Banki (Bank of English) hamda Britaniya milliy korpusi (British National Corpus, BNC), Rossiyada Rus tilining milliy korpusi loyihalari ishlab chiqildi. Masalan, Rus tilining Milliy Korpusi hajmi hozirgi kunda 149 mln so'zdan iborat. Keyingi yillarda Internet tizimining rivojlanishi virtual matnlar korpusi yuzaga kelishiga olib keldi. Ya'ni Internetdagi qidiriv saytlari, elektron kutubxonalar, virtual ensiklopediyalar korpus vazifasini bajarmoqda. Korpusning janri va tematik rang-barangligi internetdan foydalanuvchining qiziqishlariga bog'liq. Korpus foydalanuvchilari turli soha, kasb vakillari bo'lishi mumkin. Korpus foydalanuvchilari doirasi biz o'ylaganimizdan ko'ra ancha keng bo'lib, quyida shularni sanab o'tamiz. Lug'at va grammatikadan faqat tilshunos foydalanmaganidek, korpus gumanitar fanlar tadqiqotchisi, adabiyotshunos hamda tarixchi uchun ham birdek zaruriy baza, deyish mumkin. O'qituvchi uchun korpus — tengsiz xazina. Masalan, u topshiriqni tayyorlash uchun faqat badiiy adabiyotdan (yoki badiiy bo'lmagan matndan: sport, avtobiografiya yoki shunchaki ma'lum ijodkor asari bilan cheklanmoqchi) foydalanmoqchi. Shu o'rinda ta'limiy va milliy korpusni farqlab olish zarurati paydo bo'ladi. Ta'limiy korpusning milliy korpusdan eng asosiy farqi shundaki, ta'limiy korpus birlik (matn)lari didaktik xarakterga ega bo'ladi. Milliy korpusda bosh maqsad jonli tilni ko'rsatish bo'lgani sababli unda turli uslub va janrdagi matnlar korpus birligi sifatida ishlatiladi, milliy korpusda reprezentativlik (matnning janriy xilma-xilligi) ustunlik qiladi. Ta'limiy korpusdagi matnlar til ta'limi, o'quvchilarni ma'naviy-axloqiy tarbiyalashga qaratilgan matnlardan iborat bo'ladi. Ta'limiy korpusga ko'proq maqol, matal, naql, ertak, rivoyat janrlari kiritiladi. Asl maqsad — ona (o'zbek) tilini o'rgatish bilan birga, tarbiyaviy maqsad ham asosiy planda bo'ladi. Milliy til korpusi - o'zbek tilidagi ilmiy, badiiy, publitsistik, so'zlashuv va rasmiy uslubga doir shuningdek, zamonaviy, tarixiy, she'riy, nasriy, dramatik va shevaga doir matnlardan iborat



milliy kontentdir. milliy til korpusi yaratilsa, ta'limiy korpus yanada takomillashadi. Milliy til korpusi tadqiqiy, tahliliy xarakterdadir. Ta'limiy korpus esa tadbiqiy, amaliy maqsadga yo'nalgan elektron resursdir[Raupova, Zaripboyeva, 2021]. Korpus turli janr va uslubdagi matnlarni qamrab olgani sababli har qanday talabni qondirish imkoniga ega. Misol uchun, jurnalistika fakulteti talabalari uchun yozilgan darslikda publitsistik matnlar ancha eskirgan. Har kuni muntazam boyitiladigan korpusdan yangi misollar olish, ularni talabaga taqdim etish yoki unga ham shunday topshiriq berish, albatta, ta'limni hayotga yaqinlashtiradi. Ko'pincha qayta-qayta nashr etilgan darslikda misollar eskiligicha qolib ketadi. Bugun filolog yoki jurnalist bo'lib yetishayotgan mutaxassisga tilimizning barcha nozik qirralarini his etishi uchun faqat eski — 20-30 yil oldingi matn namunasi yetarli emas, u bugun yozilgan barcha soha, uslub va janrdagi asarlar bilan ishlashi lozim. Shundagina ta'lim ijtimoiy buyurtmani bajargan hisoblanadi. Til korpusidan o'qituvchi, talaba va maktab o'quvchisi ham ancha unumli foydalanishi mumkin. Chunki faqat korpus orqali juda osonlik bilan kam ishlatiladigan so'z, ibora va birikmani topish, qo'llanishi va yozilishi (orfografiyasi)ni o'rganish mumkin. Ta'kidlash joizki, til korpusida til grammatika va darslik muallifi tavsiflaganidek emas, balki jamiyatda qanday yashasa va ishlasa, shunday aks etadi. Bu esa o'quvchining o'zini o'rab turgan muhit — umumxalq tili va adabiy tilni o'rganishida eng sarmahsul vosita bo'lib xizmat qiladi. Til korpusiga eng ko'p ehtiyoj sezuvchi mutaxassis bu — matnni avtomatik qayta ishlash (masalan, tarjima dasturi) va turli qidiruv tizimlari bilan ishlaydigan dasturchi. Chunki u tabiiy til bilan ish ko'radi hamda ushbu tilda yozilgan barcha matnlar strukturasini (tabiiy va jonli tilda, ilmiy grammatika va darsliklardagi misollarga tayanib emas!) mukammalroq «tushunishi», his etishi lozim. 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 televideniya xodimlari. Chunki bu mutaxassislar muayyan so'z, ibora yoki gapning qo'llanish holati, darajasi, kim, qachon ilk marta

bu gapni qo'llagani, 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. Biz bugungacha darslik, ilmiy grammatika va lug'atga qanday tayangan bo'lsak, zamonaviy soha vakillari korpusga shu qadar ehtiyoj sezishadi, undan foydalanishadi[Nurillayeva, 2018]. Korpusga katta ehtiyoji bor soha lingvodidaktika bo'lib, u ona tilini o'qitishda ham, xorijiy tillarni o'qitishda ham birdek muhim. Tilni o'rgatishda lug'at boyligining kattaligini ko'rsata olish, so'zning qo'llanish imkoniyatini u yoki bu grammatik qurilish orqali tushuntirish uchun misollar massivini ko'rsatishda korpus juda qo'l keladi. Til ta'limi uchun muhim bo'lgan misolning doimiy yangilanib borishi, buni aks ettirib turish xususiyati hamda imkoniyati faqat korpusga xos. O'qituvchi yangi, ishonarli, cheksiz hamda xilma-xil misollarni shu yerdan topa oladi, topshiriq va mashqlarni belgilashda qiynalmaydi, bir necha daqiqada mavzu bo'yicha yangi misollardan iborat mashq tayyorlay oladi[Baxriddinova, Bahodirova, 2021]. Prezidentimiz Sh.M.Mirziyoyev ta'kidlaganidek, “ona tilimizni jozibali va qiziqarli tarzda taqdim etish o'zbek tilining kompyuter dasturlarini, onlayn darsliklar, elektron lug'atlarni yaratishni taqozo etadi”[Mirziyoyev, 2019, <http://uza.uz/oz/politics/>]. Hatto tilshunoslikning nisbatan yangi sohalari bilan shug'ullanuvchi psixolingvist, pragmalolingvist, sud-huquq lingvistlari ham korpusdan unumli foydalanishlari, kutilgan natijaga soniyalarda erishishlari mumkin. Buning uchun avvalo ochiq milliy korpus dasturini yaratish lozim va soha fidoyilari uning tarkibini turli janrdagi og'zaki va yozma matnlar bilan doimiy to'ldirib borishi talab etiladi. Xususan pragmalolingvist korpus yordamida ma'lum lisoniy birliklarni kommunikatsiya jarayonidagi pragmatik bahosini aniqlashda turli xildagi yuzlab matnlarni topishi va qisqa fursatlarda saralab olishi mumkin. Mavzuga chuqurroq yondashadigan bo'lsak, ayrim so'zlar semantikasini, so'zlashuvning pragmatik maqsadini turli yosh, jins, ijtimoiy tabaqa va guruhdagi insonlar turlicha talqin qiladilar. Masalan,

o'smir yoshlar nutqida “yomon”, “dahshat” ”jinni” sifatlari bugungi kunda ijobiy boholash vositasi bo'lishi mumkin. Ammo keksa avlod vakillari yoki ota-onalar uchun esa buning aksi bo'lishi ehtimol. Hozirda jahon tilshunosligida sud lingvistlari tez-tez til korpuslaridan foydalanadilar. Chunki mukammal ishlangan til korpuslarida o'z joniga qasd qilish to'g'risidagi xatlardan tortib, uyali telefon xabarlari, politsiya bayonotlari, politsiya bilan suhbat yozuvlari, guvohlarning bayonotlarigacha qamrab olingan. Lingvistik ekspertiza uchun korpusning quyidagi ahamiyatlari mavjud:

- Turli hudud vakillari nutqining fonetik, leksik grammatik, xususiyatlari;
- Plagiatlik holatlarida aynan va o'xshashlik jihatlarini aniqlash;
- Haqorat, qarg'ish va obro'sizlantirish kabi holatlarni taftish qilish;
- Murojaat. Chaqiruv ko'rinishidagi ekstermistik materiallarning milliy etnik

xoslanishi korpus bazalaridan foydalangan holda aniqlash[Musulmonova, 2021]. O'zbek tilida milliy korpus yaratish ishlari hanuz orqaga surilmoqda. O'zbek tilida ko'plab lug'atlar yaratilgan, lekin til korpusi va uning nazariy asoslarini ishlab chiqish bo'yicha hali jiddiy tadqiqot qilingani yo'q. Buxorolik mustaqil tadqiqotchi Shahlo Xamrayeva bugungi kunda mualliflik korpusiga oid tadqiqotni amalga oshirdi. Uning tadqiqoti bizga milliy mualliflik korpusi yaratish me'zonlari uchun o'zbek tilida yaratilgan muhim manba sifatida munosib o'ringa egadir. Shuningdek, bugun G.Begmatova[Begmatova,2021], A. Eshmo'minov[Eshmo'minov 2019], Xidirov O.[Xidirov, 2021], S. Karimov, A.Qarshiyev kabi tadqiqotchi- olimlarimiz, dasturchilarimiz korpus lingvistikasi va uning dasturiy ta'minoti borasida birmuncha ishlarni amalga oshirishdi . Biz ham kelgusi ishlarimizda korpus lingvistikasining mohiyati, bu sohaning xorijdagi mualliflik korpusi amaliy tajribalarga asoslanib Usmon Azim dramalarining ham til korpusida lingvistik ta'minotini yaratishni maqsadimiz etib belgiladik.

O'zbek tilining milliy mualliflik korpusi uchun Usmon Azim dramalari asosida namuna yaratish maqsadida Usmon Azimning “Adibning umri”, “Kunduzsiz kechalar”, “Abdulla

Qahhor”, “Alpomishning qaytishi”, “O‘tgan zamon hangomalari”, “Bor ekan-u yo‘q ekan”, “Bir qadam yo‘l”, “Bozor” kabi dramalari tanlandi.

Xulosa qilib aytganda, tilga oid ilgari qilingan tadqiqotlar ko‘lami va natijalari kompleks holda o‘rganilib, korpus lingvistikasiga oid nazariy bilimlar hali to‘ldirilishi, tadqiqi va amaliy ishlanishi zaruriyatdir. Yangi va keng ko‘lamda lingvistik tadqiqotlar olib borish uchun zamonaviy korpus lingvistikasining asosiy yo‘nalishlaridan biri mualliflik korpuslariga asoslanishini nazarda tutib mualliflik korpuslari yaratish oldimizda turgan vazifalardan biridir. Yaratilajak milliy korpus negizida turli sohalarga oid materiallarni joylashtirish g‘oyatda ahamiyatli. Turli tarmoq korpuslarining yaratilishi esa ilm-fan rivoji va muayyan soha taraqqiyoti uchun omil bo‘lib xizmat qiladi.

Foydalanilgan adabiyotlar:

1. Xamroyeva Sh. Korpus lingvistikasi atamalarining qisqacha izohli lug‘ati. Toshkent – 2018
2. O‘TIL. 2-tom.
3. Raupova L., Zaripboyeva N. Ta’limiy korpus: tatbiqiy, amaliy elektron resurs vositasi sifatida. O‘zbek milliy va ta’limiy korpuslarini yaratishning nazariy hamda amaliy masalalari. Vol.1.2021
4. Nurillayeva O. O‘zbek tilining milliy korpusini yaratish tamoyillari. Bitiruv malakaviy ish. Toshkent - 2018
5. Baxriddinova B., Bahodirova G. O‘zbekistonda milliy til renessansi va lug‘atchilik taraqqiyoti. O‘zbek milliy va ta’limiy korpuslarini yaratishning nazariy hamda amaliy masalalari. Vol.1.2021
6. Mirziyoyev Sh. O‘zbek tili xalqimiz uchun milliy o‘zligimizni va mustaqil davlatchilik timsoli, bebaho ma’naviy boylik, buyuk qadriyatdir. Prezident Sh.Mirziyoyevning o‘zbek tiliga davlat tili maqomi berilganining o‘ttiz



yilligiga bag'ishlangan tantanali marosimdagi nutqi. 2019-yil, 21-oktabr.

<http://uza.uz/oz/politics/>

7. Musulmonova K. milliy korpus yaratishning kriminolingvistika taraqqiyotidagi roli va istiqbollari. O'zbek milliy va ta'limiy korpuslarini yaratishning nazariy hamda amaliy masalalari. Vol.1.2021
8. Mengliyev B va boshqalar. O'zbek tilining milliy korpusi//Ma'rifat, - Toshkent, 2018.
9. Xamroyeva Sh. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari. PhD dissertatsiya avtoreferati. 2018.
10. Begmatova G. O'zbek milliy korpusida idiomalar bazasini yaratish. PhD dissertatsiya. 2021.
11. Eshmo'minov A. O'zbek tili milliy korpusining sinonim so'zlar bazasi. PhD dissertatsiya. 2019.
12. Xidirov O. Milliy korpus uchun parsing dasturi yaratishning lingvistik asoslari. .PhD dissertatsiya avtoreferati. 2021.