

## TA'LIM KORPUSINI YARATISH USULLARI

### **Madatov Xabibula Axmedovich**

*Urganch davlat universiteti axborot texnologiyalari kafedrasi mudiri, fizika-matematika fanlari nomzodi, dotsent:*

+998929223272

[xabibulla3272@gmail.com](mailto:xabibulla3272@gmail.com)

### **Narimonova Muhayyo**

*Urganch Davlat Universiteti Kompyuter lingvistikasi yo'nalishi magistranti*

+998943474466

[muhayyonarimonova@gmail.com](mailto:muhayyonarimonova@gmail.com) +998943474466

**Annotatsiya.** *Bugungi kunda ta'lim va fan sohasini rivojlantirish uchun ilmiy va innovatsiya yutuqlarini amaliyotga joriy etish, uning samarali mexanizmlarini yaratish, ulug' ajdodlarimiz tomonidan asos solingan ilmiy maktablar faoliyatini chuqur o'rganish muhim poydevor bo'ldi. Bu esa yillar mobaynida o'z qadrini yo'qotmasdan millat ruhini sog'lom saqlashga xizmat qilib kelmoqda. Taraqqiyotimiz manbai bo'lgan bobomeros ilm an'analari zamonaviy shaklda davom ettirish har qachongidan ham muhim hisoblanadi. Buning uchun biz manbashunoslik materiallari hamda zamonaviy tilshunoslik sohalariga tayangan holda ish olib borishimiz zarur. Mazkur maqolada ham o'zbek tilshunosligining yangi sohasi hisoblangan korpus lingvistikasining asosiy tamoyillari korpus, uning tuzilishi, korpus turlari, korpus yaratish usullari hamda korpusni tuzish jarayoni haqida ham so'z yuritiladi. Bugungi kunda korpusning nazariy hamda amaliy ahamiyati o'zbek tilshunosligida tilning mavjud imkoniyatlarini kengroq o'rganishda, tilshunoslikdagi muammoli tomonlarini aniqlash, elektron lug'atlar yaratishda, tilni o'rganishda zamonaviy axborot texnologiyalaridan foydalanish samaradorligini oshirishda, qidiruv va kompyuter tahlili kabi masalalarni hal qilishda, tilning maxsus sohalar bo'yicha korpusini qurish zarurati mavjudligini ko'rsatadi.*

**Kalit so'zlar:** *korpus, korpus turlari, korpus tuzilishi, ta'lim korpusi,*

**Абстрактный.** *Сегодня для развития образования и науки важной основой стало внедрение научных и инновационных достижений, создание эффективных механизмов, углубленное изучение деятельности научных школ, основанных нашими великими предками. Это помогло сохранить дух нации здоровым, не потеряв своей ценности с годами. Как никогда важно продолжать древние научные традиции, являющиеся источником нашего прогресса, в современной форме. Для этого нам нужно работать на основе источников и современной лингвистики. В данной статье также рассказывается об основных принципах корпусной лингвистики, которая считается новой областью узбекского языкознания, корпусе, его структуре, типах корпусов, методах создания корпусов и процессе составления корпусов. Сегодня теоретическое и практическое значение корпуса в узбекском языкознании заключается в более широком изучении существующих возможностей языка, выявлении проблемных моментов в языкознании, создании электронных словарей,*

пovyшeнии эффeктивнoсти иcпoльзoвaния cовpeмeнных инфoрмaциoнных тeхнoлoгий в язькe. oбyчeния для рeшeния тaких зaдaч, кaк пoиск и кoмпьютeрный aнaлиз, пoкaзывaeт, чтo сyщeствyeт нeoбхoдимoсть в пoстрoении кoрпyсa язькa в cпeциaльных oблaстях.

**Ключевые слова:** *корпус, типы корпусов, конструкция корпуса, учебный корпус*

**Annotation.** *Today, for the development of the field of education and science, the implementation of scientific and innovative achievements, the creation of effective mechanisms, and the in-depth study of the activities of scientific schools founded by our great ancestors has become an important foundation. It has been serving to keep the spirit of the nation healthy without losing its value. It is more important than ever to continue the ancient scientific traditions, which are the source of our progress, in a modern form. For this, we need to work based on sources and modern linguistics. This article also talks about the main principles of corpus linguistics, which is considered a new field of Uzbek linguistics, the corpus, its structure, corpus types, methods of corpus creation, and the process of corpus compilation. Today, the theoretical and practical importance of the corpus in Uzbek linguistics is to study the existing possibilities of the language in a wider way, to identify problematic aspects in linguistics, to create electronic dictionaries, to increase the efficiency of using modern information technologies in language learning, to solve problems such as search and computer analysis, shows that there is a need to build a corpus of the language in special fields.*

**Keywords:** *corpus, corpus types, corpus structure, educational corpus*

Tilimiz rivojlanib borar ekan, undagi turli tushunchalar, o'zgarishlar hayotimizga katta yangiliklar olib kirmoqda. Tilning rivojlanish jarayonlari ko'plab tadqiqotchi olimlar tomonidan xarakterlanib kelmoqda. Til korpusini yaratish, uning asoslarini o'rganish va hayotga tatbiq qilish bugungi kompyuter lingvistikasi mutaxassislari uchun birlamchi vazifalardan biri bo'ladi desak, yangilishmagan bo'lamiz. Bu jarayonning rivojlanishi tilning bevosita taraqqiyoti bilan bog'liq. Bu jarayonni rivojiga, korpus tizimining taraqqiyotiga ko'pgina olimlar o'z hissalarini qo'shgan. V.P.Zaxarov va S.Y.Bogdanovalar korpusning janriy-mavzuviy tuzilishi xususida yangicha fikrlar bildirganlar. Korpusni loyihalashtirish jarayoni bugungi kunda eng asosiy vazifalardan biridir.

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 kompyuter lingvistikasi mutaxassislarga katta mas'uliyat yukladi.

Tilning dastlabki rivojlanishi bilan solishtirganda, orfografiya va savodxonlikni o'zlashtirish jarayonida til tizimidagi navbatdagi o'zgarishlar hali batafsil o'rganilmagan. Imlo, lekin grammatik xatolar uchun tuzatgan so'z shakli. Asl va maqsadli so'z shakllari xarakter jihatidan mos keladi va maqsadli so'z shakllari fonologik, bo'g'in va morfologik ma'lumotlar bilan boyitiladi. Hozir hatto qadimgi tillarning ham korpuslari ishlab chiqilgan. Axborot asrida rivojlanishga yuz tutgan har bir til sun'iy intellekt tiliga aylanishi kerak. Ona tilimizning dunyo miqyosidagi nufuzi faqat tashviqot-targ'ibot bilan rivojlanib qolmaydi. Sun'iy intellektga asoslanmagan tilning taraqqiyoti haqida gap ochish ham to'g'ri emas. Bu muammolarni hal qilish maqsadida olimlar tomonidan ko'plab ilmiy tadqiqot va amaliy ishlar bajarilishi kerak. Buning natijasi sifatida o'zbek sun'iy intellekti, o'zbek intellektual sistemasi yoki o'zbek ekspert sistemasi kabi yangi ilmiy tushunchalar zamonaviy fanlar doirasida kengyoyish choralari ko'rish maqsadga muvofiqdir. Jahon tillarining juda ko'pchiligi mukammallik darajasi va matni ilmiy qayta ishlash imkoniyati bilan farq qiluvchi o'z milliy korpuslariga ega. Milliy til va davlat tili taraqqiyotida milliy korpuslar ahamiyati va o'rnining qanchalik muhimligini davr ko'rsatmoqda.

Sun'iy intellekt metodlari asosida aniq bir predmet, til sohasiga tegishli intellektual masalalarni yechishga mo'ljallangan dasturiy yoki texnik sistemalar intellektual sistemalar deb yuritiladi. Muayyan predmet sohasidagi masalalarni amaliy yechishga mo'ljallangan intellektual sistemalar esa ekspert sistemalardir. Xullas, bularni o'zbek tili doirasida yaratish uchun milliy kompyuter lingvistikasi sohasi jadal va keng qamrovli, tizimli va aniq maqsad-rejali rivojlanmog'i lozim. Demak, zamonaviy tadqiqotlar, hatto kompyuter lingvistikasi taraqqiyoti uchun ham til korpuslarini yaratish zarur ekan.

Dunyo kompyuter lingvistikasida milliy til korpusining mavjudligiga tilning yashovchanlik va kompyuter tiliga aylanish mezon sifatida qaralmoqda. Til korpusining tuzilishi olimlar tomonidan bir necha bosqichlarga ajratilgan:

1. Korpusga kirishni ta'minlash;
2. Matni avtomatik qayta ishlash;
3. Matnga dastlabki ishlov berish;
4. Konversiyalash, grafematik tahlil;
5. Nostandart elementni belgilash. [1:45]

O'zbek tilini o'rganish istagida bo'lgan xorijliklar uchun sifatli qo'llanmalar yaratish hozirgi kunning dolzarb masalalaridan biri hisoblanadi. Xorijliklarga o'zbek tilini o'rgatish metodikasi hali-hanuz to'liq shakllantirilmagan. Sifatli qo'llanmalar mavjud emas. Til bo'yicha qo'llanmalarining sifatli yoki sifatsizligi, samarali yoki samarasizligi tez ko'zga tashlanadigan holat emas. Chunki, til egalari ona tilini faqat ta'lim yo'li bilangina egallamaydi. Oila, jamiyat, radio-televideniyaning ham til o'rgatuvchi bir vosita deb qarash mumkin. Ammo, xorijliklar uchun samarali natija beradigan mukammal vositalar yaratish zarur. Tilimizning murakkabligi ilmiy aniqlanib, shunga asoslangan holda turli leksik-grammatik darajadagi, ketma-ketlikka asoslangan qo'llanmalar yaratish bugungi kunning eng dolzarb vazifasi sanaladi.

Korpus – til birliklarining xususiyatlarini aniqlash maqsadida qidiruv dasturiga bo'ysundirilgan matnlar majmui, tabiiy tildagi elektron shaklda saqlanadigan yozma yoki

ogʻzaki, kompyuterlashtirilgan qidiruv tizimiga dasturiy taʼminot asosida joylashtirilgan online yoki offline tizimda ishlaydigan matnlar jamlanmasi. [2:587-593]

Korpus tilshunosligi esa kompyuter tilshunosligining tarkibiga kiruvchi tilshunoslik boʻlimi boʻlib, u korpus tuzilishining umumiy tamoyillarini ishlab chiqishda, matnlar korpusiga kompyuter texnologiyasini qoʻllash orqali loyihalar yaratishda ishtirok etadi. Tilshunoslik yoki til nuqtayi nazaridan matnlar korpusining tanasi katta, mashinadan oʻqiladigan shaklda koʻrinadigan, yagona, tizimli, belgili, filologik jihatdan malakali til majmuasi hamda tilshunoslikka oid muayyan maʼlumotlar beruvchi baza sifatida tushuniladi.

Xulosa oʻrnida aytishimiz joizki, milliy texnik baza rivojida korpus tilshunosligining, milliy korpusning, milliy korpus tamoyillarining ahamiyati va xususiyatlarini tilga olish oʻrinlidir. Chunki bu tuzum faoliyatini yoʻlga qoʻyish eng ahamiyatli muammolardan biridir. Keyingi oldimizga qoʻyilgan maqsadlardan biri milliy korpus boʻyicha aniq dasturiy tuzulmalarni yaratish va ularni hayotga tatbiq qilish. Kompyuter texnologlari tomonidan shu masala boʻyicha dastur va qoʻllanmalar tuzishdir.

### FOYDALANILGAN ADABIYOTLAR ROʻYXATI:

1. B.Sobitxonova. Milliy korpus. Toshkent. “Sharq”, 2012. 45.
2. Махмудов, Қ. Ш., Санакулов, З. И., Шайхисламов, Н. З., Қаршиева, Д. Э. (2020). Педагогик таълимда “Кластер”. Science and Education, 1(7), 587-593.
3. Internet materiallari; <https://uz.drunkentgu.com/chto-takoe-3237>
4. “Oʻzbek tilining ham milliy korpusi yaratilishi lozim”. Baxtiyor Mengliyev
5. New Perspectives on Corpus Translation Studies. Vincent X. Wang, Lily Lim, Defeng Li.

#### **Madatov Xabibulla Axmedovich**

Oʻzbekiston Respublikasi Xorazm viloyati

Urganch davlat universiteti Urganch davlat universiteti axborot texnologiyalari  
kafedrasi mudiri Fizika-matematika fanlar nomzodi, dotsent

Taʼlim korpusi yaratish usullari Manbashunoslik, arxeologiya va etnologiya masalalari  
+998929223272 [xabibulla272@gmail.com](mailto:xabibulla272@gmail.com)

#### **Narimonova Muhayyo Bahodir qizi**

Oʻzbekiston Respublikasi Xorazm viloyati

Xorazm viloyati Yangibozor tumani 14-son umumiy oʻrta taʼlim maktabi Oʻqituvchi  
Magistrant Taʼlim korpusi yaratish usullari

Manbashunoslik, arxeologiya va etnologiya masalalari  
+998943474466muhayyonarimonova@gmail.com