O‘ZBEK TILI MILLIY KORPUSINING DAVLAT TILI TARAQQIYOTIDAGI O‘RNI

Jalilova Dilnoza Farxod qizi,
ToshDO‘TAU O‘FF 2-kurs talabasi
dilnozajalilova31@gmail.com +99899 354 34 22
Ilmiy rahbar: F.f.d.,Prof, Zulxumor Xolmanova

Annotatsiya: *Ushbu maqolada o‘zbek tili milliy korpusining qanday imkoniyatlarga ega ekanligi, morfologik analizator va lug'atlarning xususiyatlari keltirib o‘tildi. Til taraqqiyoti davlat rivoji uchun muhim omil ekanligi, sun'iy intellekt bilan milliy til korpusining chambarchas bog'liq ekanligi, o‘zbek tilining internet tili sifatida ravnaq topishi uchun korpus muhim ahamiyatga ega ekanligi dalillar yordamida yoritildi.*

Kalit so‘zlar: *korpus lingvistikasi, til korpusi, morfologik analizator, ta’limiy korpus, lug‘atlar, sun’iy intellekt, internet tili.*

Abstract: *This article outlines the capabilities of the Uzbek National Corpus and describes the features of its morphological analyzer and dictionaries. It highlights, with supporting evidence, the importance of language development for national progress, the close link between the national language corpus and artificial intelligence, and the corpus's vital role in the advancement of the Uzbek language as a language of the Internet.*

Keywords: *corpus linguistics, language corpus, morphological analyzer, educational corpus, dictionaries, artificial intelligence, Internet language.*

Аннотация: *В статье представлен обзор возможностей Национального корпуса узбекского языка, а также описаны характеристики его морфологического анализатора и словарей. На основе соответствующих данных обосновывается важность развития языка для прогресса нации, подчеркивается тесная связь между национальным языковым корпусом и искусственным интеллектом, а также отмечается ключевая роль корпуса в становлении узбекского языка как языка интернет-пространства.*

Ключевые слова: *корпусная лингвистика, языковой корпус, морфологический анализатор, учебный корпус, словари, искусственный интеллект, язык интернета.*

Til inson kishilik jamiyati taraqqiyotida erishilgan eng buyuk ma’naviy ne’mat, azaliy qadriyat hisoblanadi. Tilning ijtimoiy vazifalari insoniyatning turli ko‘rinishdagi ma’naviy ehtiyojlarini qondirishga xizmat qiladi. Antropotsentrik yo‘nalishdagi tadqiqotlarda tilning voqelikni aks ettirish, moddiy asoslar haqida tasavvur berish, ularning xususiyatlarini o‘zlashtirish, iqtisodiy va modiiy ehtiyojlarni qondirish vositasi sifatida ham ijtimoiy vazifa bajarishi ilmiy jihatdan asoslangan.

Davlat taraqqiyotida iqtisodiy, ijtimoiy-siyosiy kabi omillar bilan bir qatorda, jamiyatning o‘zagi bo‘lgan milliy tilning ham hissasi bor. O‘zbek tilining rivojida 2019-yilda qabul qilingan “2020 – 2030-yillarda o‘zbek tilini rivojlantirish va til siyosatini takomillashtirish” konsepsiyasi muhim ahamiyat kasb etdi. Bu konsepsiya doirasida til taraqqiyoti uchun turli ishlar amalga oshirilmoqda. Xususan, o‘zbek tili milliy korpusi ishlab chiqildi. Korpus matnlarni yig‘ish, tasniflash, annotatsiyalash kabi vazifalarni bajaradigan[74] korpus lingvistikasining bir yo‘nalishi bo‘lib, milliy til birliklarining xususiyatlarini aniqlash maqsadida elektron qidiruv imkoniyati mavjud tizim, tabiiy tilning raqamlashgan yozma va og‘zaki matnlari jamlanmasidir.[75] M.Abjalovaning ma’lumotlariga qaraganda, o‘zbek milliy korpusi quyidagi xususiy (ichki) korpuslarga bo‘linadi:

1. Ta’limiy korpus.
2. Ilmiy matnlar korpusi.
3. Gazetalar korpusi.
4. Badiiy matnlar korpusi.
5. Parallel matnlar kospusi.
6. Rasmiy matnlar korpusi.
7. Og‘zaki matnlar korpusi.
8. Mualliflik korpusi.
9. Poetik matnlar korpusi.

Demak, o‘zbek tili milliy korpusi jamiyatning barcha sohalariga tegishli og‘zaki va yozma resurslarni qamrab olgan. Bu lohiyada morfologik analizator, lug‘atlar, tillararo erkin tarjima mavjud. Morfologik analizator matnda keltirilgan so‘zlarning qaysi so‘z turkumiga tegishli ekanligini, asosi qaysi turkumligini, o‘zakka qanday qo‘shimchalar qo‘shilganligini, nechta so‘z turkumidan foydalanganligini statistika orqali ko‘rsatib beradi. Lug‘atlar bo‘limi orqali so‘zlarning izohini, imlo qoidalariga asosan yozilgan variantini va qaysi turkumga

[74] Korpus lingvistikasi. darslik / N.Z.Abduraxmonova. – Toshkent, 2023. – 346 b.
[75] Korpus lingvistikasi. Uslubiy qo‘llanma / M.A. Abjalova. – Toshkent: Bookmany print, 2022. – 103 b.

mansub ekanligini, talaffuzini va imo-ishoralar tilida qanday aks etishini bilish mumkin. Milliy korpus o‘zbek tilini qiziqib o‘rganayotgan chet elliklarga tilning tabiiy kontekstida kerakli ma’lumotlarni tezda to‘plash va tahlil qilishda yengillik yaratadi. Avvallari bitta so‘zning ma’nosini bilish uchun izohli lug‘atga, to‘g‘ri yozilganligini tekshirish maqsadida imlo lug‘atiga murojaat qilingan bo‘lsa, endilikda hammasi o‘zbek tili milliy korpusining lug‘atlar bo‘limida mavjud. Ta’limiy korpusning “video”, “she’rlar”, “izlash”, “xorijliklar” kabi bo‘limlari orqali ta’limga oid resurslarda so‘z va so‘z birikmalari qanday ma’no anglatishini, manbalarda qanday shaklda kelganligini, qo‘llanish darajasini aniqlash mumkin. Korpuslar o‘quvchilarning, yosh izlanuvchilarning vaqtini tejashga xizmat qiladi, ma’lumotlarni tez va oson topishga yordam beradi. Sh.Hamroyeva korpusning imkoniyatlarini qayd etar ekan, ushbu qidiruv tizimi haqida shunday umumiy xulosaga keladi: “Korpus – millionlab so‘z kontekstlariga havola qiluvchi, maxsus qidiruv tizimi asosida ishlaydigan elektron matnlar yig‘indisi. U Internet tizimidagi elektron kutubxona, lug‘at va lingvistik grammatikadir. Korpus – tabiiy (real) tilning elektron shakli bo‘lib, uni to‘ldirish, boyitish, tuzatish, tahrir qilish imkoniyatiga ega”[76]. Shuningdek, milliy til korpusi sun’iy intellektning o‘zbek tilidagi ma’lumotlar bazasi ortishiga yordam beradi. Agar sun’iy intellekt qidiruviga o‘zbek tilida so‘rov yuborilsa, ma’lumotlar bazasida yetarlicha axborot bo‘lmaganligi uchun ishonarli, faktlarga asoslangan javob berishga qiynaladi. Yoki “google chrome”, “mozilla firefox”, “safari” kabi brauzerlarga o‘zbek tilida so‘rov yuborilsa, to‘liq ma’lumot yetkazib bera olmaydi. Aksincha, internet tili bo‘lgan ingliz yoki rus tilida so‘rov yuborilsa, dalillarga asoslangan, aniq ma’lumotlar topishga yordam beradi.

O‘zbek tilining leksik-semantik strukturasini, morfologik, sintaktik xususiyatlarini raqamlashtirish davlat tilini xalq madaniyati, qadriyatlarini dunyoga tezroq yoyish uchun zarur bo‘lgan internet tili maqomiga erishadi. Shuningdek, internet foydalanuvchilari, tadqiqotchilar uchun kerakli axborotni topishdagi qiyinchiliklarning oldi olinadi. Shu bilan bir qatorda, tadqiqodchi, olimlarning butun dunyo foydalanadigan mashhur brauzerlarga kiritilgan kitoblari, maqolalari, tadqiqodlaridan kerakli ma’lumotlarni olib foydalanish mumkin bo‘ladi.

Bundan tashqari, o‘zbek tili milliy korpusiga kiritilgan manbalarni ko‘rish, o‘rganish natijasida chet elliklarning yurtimizga bo‘lgan qiziqishi yanada ortishi, tashriflar soni ko‘payishi orqali turizm rivojlanadi. Buning natijasida davlat iqtisodi

[76] Hamroyeva Sh. O‘zbek tili mualliflik korpusini tuzishning lingvistik asoslari: filol. fan. bo‘yicha falsafa doktori (PhD) diss. avtoref. – Qarshi, 2018

boyiydi, ham yurtimiz rivojlangan davlatlar safiga qo‘shilishi uchun imkoniyat yaratiladi.

Xulosa qilib aytganda, o‘zbek tili milliy korpusi orqali barcha og‘zaki, yozma matnlarni raqamlashtirish, ommalashtirish imkoniyati mavjud. Til o‘rganuvchilar uchun lug‘atlar bilan ishlash, so‘zlarning morfem tarkibini tezda anglab olish, matnlarni tillararo tarjima qilish, o‘zbeklarning madaniyati aks etgan she’rlarni o‘qish va tinglash, qisqa videolar ko‘rish orqali tezda o‘zbek tilini o‘rganish va milliy qadriyatlarimiz haqida ham foydali ma’lumotlarga ega bo‘lishi mumkin. Shuningdek, o‘zbeklar urf-odatiga, madaniyatiga, moddiy va ma’naviy manbalariga qiziquvchilar sonining ortishi turizmning yanada rivojlanishiga yordam beradi. Sun’iy intellekt, google chrome, safari kabi dunyo brauzerlaridagi ma’lumotlar bazasining o‘zbek tilidagi raqamlashgan manbalar bilan boyishi va foydalanuvchilar sonining ortishi tilimizning internet tili safiga qo‘shilishi uchun imkoniyat yaratadi.

Foydalanilgan adabiyotlar:

1. Korpus lingvistikasi. darslik / N.Z.Abduraxmonova. – Toshkent, 2023. – 346 b.
2. Korpus lingvistikasi. Uslubiy qo‘llanma / M.A. Abjalova. – Toshkent: Bookmany print, 2022. – 103 b.
3. Hamroyeva Sh. O‘zbek tili mualliflik korpusini tuzishning lingvistik asoslari: filol. fan. b. fals. dok. (PhD) diss. avtoref. – Qarshi, 2018
4. Internet resurs havolasi: https://tafsiri-irfon.uz/static/maqola/13-maqola.pdf