MILLIY TIL KORPUSI LINGVISTIK TA'MINOTI XUSUSIDA AYRIM MULOHAZALAR

Rabbimova Gulchehra Toyir qizi
SamDCHTI o'zbek tili
jurnalistika va kompyuter lingvistikasi kafedrasi
stajyor- o'qituvchisi
Tel: +998902969923
gulchehraqoshmonova@gmail.com

Qabul qilindi: 29.10.2025
Ko'rib chiqildi: 30.10.2025
Tasdiqlandi: 31.10.2025

Tayanch so'zlar:
Korpus lingvistikasi, korpus, lug'at, milliy korpus, lingvistik ta'minot, olmosh.

Annotatsiya. Mazkur tezisda so'nggi yillarda o'zbek tilshunosligida yil sayin rivojlanib borayotgan yangi yo'nalish – korpus lingvistikasining ahamiyati, lingvistik ta'minot masalasi, olib borilayotgan tadqiqotlar haqida umumiy fikr-mulohazalar beriladi.

Olmoshlarning morfologik xususiyatlari ularning til tizimidagi o'rnini belgilaydi. Ular shaxs, son, kelishik va egalik kabi grammatik kategoriyalarni ifodalaydi hamda gapda turli sintaktik vazifalarni bajaradi. Bu xususiyatlar olmoshlarni o'zbek tili grammatikasining muhim bo'g'iniga aylantiradi. Ularning tahlili nafaqat nazariy tilshunoslik, balki amaliy grammatikada ham dolzarb ahamiyatga ega. So'nggi yillarda olmoshlar korpus lingvistikasi asosida ham o'rganilmoqda. O'zbekiston milliy til korpusida (uzbekcorpora.uz) olmoshlar PRON yoki olm. belgisi bilan ajratiladi. Korpus asosida ularning matnlarda ishlatilish chastotasi, morfologik shakllari va kontekstual funksiyalari tadqiq qilinmoqda. Bu esa zamonaviy tilshunoslik uchun muhim manba bo'lib xizmat qilmoqda. Olmoshlar o'zbek tili grammatik tizimida asosiy so'z turkumlaridan biri sifatida katta ahamiyatga ega. Ularning morfologik shakllari, semantik turlari va sintaktik funksiyalarini o'rganish nafaqat nazariy, balki amaliy tilshunoslik uchun ham dolzarbdir. Ayniqsa, zamonaviy korpus lingvistikasida olmoshlarning

roli va ishlatilish xususiyatlarini aniqlash til tizimini chuqur tahlil qilishga imkon beradi.

Til korpusini shakllantirishda tilshunosning roli aniq nazariy maqsadni qo‘ya bilishida, vazifalarni aniq shakllantirishida, qanday va qancha hajmdagi materialdan foydalanish haqida qaror qabul qilish, to‘plangan materialni saralash, dastur ishining oraliq natijalarini baholash, zarurat tug‘ilganda maqsad va vazifalarni tahrir qilish, xulosa qilish hamda natijalarni tavsiflashdan iboratdir. Dasturchi esa aniq masalalarga javob topish uchun vositani tanlaydi, dastur tuzadi va tahrir qiladi, lingvistga ishonarli ko‘rsatkichlarni taqdim qiluvchi statistikani amalga oshiradi.

O‘zbek tilining milliy korpusini shakllantirish va boyitish filologlar tomonidan dasturchilar yordamida amalga oshiriladi. Shunisi ham borki, filolog ma’lumotlar bazasi, lingvistik ta’minot, annotatsiyalash, razmetkalash, morfologik, sintaktik, semantik razmetka tushunchalari haqida muayyan bilimga ega bo‘lishi lozim. O‘z navbatida, dasturchi ham morfologik, sintaktik, grammatik shakl, semantik analiz, lug‘aviy ma’no, matn, grammatik ma’no, ilmiy, rasmiy, publisistik stil xususiyatlari haqida bilishi lozim. Korpus lingvistikasida Samarqand davlat universiteti tilshunos olimlari va Toshkent axborot texnologiyalari universiteti Samarqad filiali olimlari tomonidan ishlab chiqilgan o‘zbek tili milliy korpusi – uzbekcorpora.uz dasturida morfologik razmetkalash quyidagi model va teglarga asosan amalga oshirilgan:

Morfologik razmetkalashning lingvistik teglari

Olmosh = [olm.]

I. Ma’nosiga ko‘ra:

a. kishilik = [kish.olm.];
b. ko‘rsatish = [ko‘r.olm.];
c. o‘zlik = [o‘z.olm.];
d. belgilash = [belg.olm.];

e. gumon = [gum.olm.];
f. so‘roq = [so‘r.olm.];
g. bo‘lishsizlik = [b-siz.olm.].

II. Tuzilishiga ko‘ra turi:

a. sodda olmosh = [sod.olm.];
b. qo‘shma olmosh = [qo‘sh.olm.];
c. juft olmosh = [juft_olm.];
d. takror olmosh = [tak.olm.].

Olmoshlar o‘zbek tilining muhim va funksional jihatdan faol qatlamlaridan biri. Ular nafaqat semantik yuklamasi, balki morfologik va sintaktik holatlari bilan ham tilshunoslikda chuqur o‘rganilayotgan so‘z turkumidir. Zamonaviy tilshunoslik, xususan korpus lingvistikasi asosida olmoshlarning tahlili fan uchun dolzarb masalalardan biri sanaladi. Shuningdek, ilmiy izlanish o‘zbek tilida olmosh so‘z turkumining milliy korpus uchun morfologik razmetkalangan (teglangan) tahlilini o‘z ichiga oladi. Tadqiqotda olmoshlarning turlari, ularning sintaktik vazifalari, korpusga mos belgilash mezonlari, morfologik atributlar va ularni avtomatik teglashdagi muammolar tahlil qilinadi. Morfologik jihatdan olmoshlar ot yoki sifatga o‘xshash bo‘lib, egalik va kelishik qo‘shimchalarini oladi ya’ni turlana olish imkoniyatiga ega. Milliy korpusda olmoshlarning o‘rni salmoqli. Milliy korpus -janridan qatiy nazar tilning milliy boyliklari jamlangan elektron qidiruv tizimi. [1. 39]. Olmoshlar milliy korpusda gapning strukturasi, shaxslararo munosabatlar va kontekstual bog‘lanishni aniqlashda muhim ahamiyatga ega. Ular matnlarning sintaktik va semantik tahlili uchun tayanch birliklardan biri hisoblanadi. Korpusda ularni aniqlash, morfologik atributlar bilan teglash va sintaktik munosabatlarda joylashtirish zarur. Quyida tekshirish vositasi sifatida Abdulla Qahhorning “Daxshat” hikoyasining bir qismi tanlandi. Natija dasturda quyidagicha aks etdi:

Dodxo darchani zichlab yopib, joyiga o'tirib tasbeh o'girishga kirishdi. Uning barmoqlari tasbeh donalarini tez-tez o'tkazayotgan bo'lsa ham, qulog'i g'uvillayotgan shamolda, xayoli go'ristonda edi: "Hozir go'riston qanaqa vahimali bo'lsa ekan..."

O'zbek go'ristoni o'zi xunuk, buning ustiga, go'riston haqida aytilmagan xunuk gap, to'qilmagan vahimali mish-mish qolgan emas. Haqiqatan, bunday kechalarda go'riston esiga tushgan har qanday odam, ayniqsa, dodxo singari payg'ambar yoshidan oshib, kafanligini sandiqqa solib qo'ygan kishi o'lishdan ham ko'ra go'ristonda yotishini o'ylaganida tiligacha sovuq ter chiqaradi.

Dodxo boshidan go'riston xayolini chiqarib tashlash uchun tasbehini qo'yib undan-bundan gapirgan bo'ldi, lekin hech kim bu gaplarga gap ulamadi. [2. 48].

Avtomatik teglashdagi murakkabliklar

Olmoshlar kontekstga bog'liq holda o'zgaruvchan bo'lishi mumkinligi sababli ularni avtomatik aniqlash va teglashda qiyinchiliklar yuzaga keladi. Masalan, "*o'zi*" so'zi har doim o'zlik olmoshi bo'lmasligi mumkin, u sintaktik konstruktsiyada o'zgara oladi. Shunday bo'lishiga qaramasdan korpus bazasiga faqatgina bitta o'z so'zi kiritiladi va bu so'z olmosh turkumiga mansub lemma sifatida razmetkalanadi. U matnda qanday ma'noni ifodalashidan qat'iy nazar olmosh sifatida namoyon bo'laveradi. Sababi hozircha korpus faqat shakl jihatidan matnlarni tahlil qilish imkoniyatiga ega, bizningcha. Yuqoridagi kontekstda keltirilgan o'zi so'zi ham aslida ma'noni kuchaytirishga xizmat qilayotgan bo'lishiga qaramasdan o'z shaklida qidiruv berilganda uni oddiygina olmoshning bir turi sifatida ta'riflaydi.

Qanaqa –

so'roq = [so'r.olm.];

sodda olmosh = [sod.olm.].

Bunday –

ko‘rsatish = [ko‘r.olm.];

sodda olmosh = [sod.olm.].

Undan-bundan –

ko‘rsatish = [ko‘r.olm.];

takror olmosh = [tak.olm.].

Tilshunosligimizda har + so‘roq olmoshlari yoki hech + so‘roq olmoshlari belgilash va bo‘lishsizlik olmoshi sifatida belgilab kelinadi. Lekin dastur matndagi faqatgina har va hech so‘zlarini o‘zinigina taniy oladi. Birikma holatdagi olmoshlarni umumiy holatda olib, olmosh deb aks ettirmaydi. Masalan, yuqoridagi matn parchasida keltirilgan har qanday va hech kim olmoshlarini alohida teglaydi va izoh beradi.

Olmoshlar morfologik jihatdan murakkab bo‘lmasa-da, ularning kontekstual funksiyasi va ko‘p ma’noliligi sababli milliy korpusda alohida e’tibor bilan teglanishi zarur. Zamonaviy korpus texnologiyalarida ularni aniq identifikatsiya qilish uchun sun’iy intellekt va mashinaviy avtomatik o‘rganish yondashuvlari joriy etilmoqda. Bu, ayniqsa, stilistik va sintaktik munosabatlar tahlilida juda muhim qism sanaladi. Til korpuslari doirasida olmoshlarning avtomatlashtirilgan tahlili ularning shakliy va ma’noviy jihatlarini aniq belgilashni talab etadi. Bu esa avtomatik morfologik tahlilda noaniqliklarga olib keladi. Shu bois, olmoshlarning korpus lingvistikasi doirasida chuqur tahlili, ularni aniq belgilash va klassifikatsiya qilish korpus sifatini belgilovchi muhim omillardandir. Hozirgi zamon lingvistikasi, ayniqsa, hisoblash tilshunosligi va tabiiy tilni qayta ishlash (NLP) yo‘nalishlarida olmoshlarning aniqligi va funksional yuklamasi muhim ahamiyat kasb etmoqda. Shu sababli, korpuslar tuzishda olmoshlar uchun maxsus belgilar tizimini ishlab chiqish, ularning kontekstual rolini aniqlash va avtomatik tahlil vositalari orqali modelga kiritish bugungi kunda dolzarb masalalardan biridir.

FOYDALANILGAN ADABIYOTLAR RO'YXATI

1. Ҳамроева Ш.М. Корпус лингвистикаси атамаларининг қисқача изоҳли луғати. – Тошкент: Камалак, 2018.
2. Qahhor A. Anor. – Toshkent, 2021.