

KORPUSNING LINGVISTIK AHAMİYATI VA UNING TILDAGI O‘RNI

Unarova Dilafruz Abdimajit qizi
SamDCHTI akademik litseyi
ona tili va adabiyot fani o‘qituvchisi
dilafruzunarova1@g.mail.com

Annotatsiya. Ushbu maqolada korpus lingvistikasi va korpusning barcha sohadagi amaliy ahamiyati to‘g‘risida so‘z boradi. Korpus lingvistikasining tilshunoslik fani bilan uzviy bog‘liqligi va rivoji, tutgan o‘rni, amaliy ahamiyati tahlil va talqin etilgan. Korpus texnik jarayonni tezlashtiruvchi qolaversa, zamonaviy yo‘nalishning texnik jarayonni tezlashtiruvchi vositagina emas, muayyan til zamonaviy shaklining axborot tizimi bo‘lib, kutilmagan savolga ham javob bera oladigan, til hodisasi bilan shug‘ullanadigan, soha oldiga avval qo‘yilmagan dolzarb muammolarni qo‘ya oladigan tizim ekanligi maqolada batafsil yoritib berilgan.

Kalit so‘zlar: kompyuter lingvistikasi, korpus lingvistikasi, til korpusi, lingvodidaktika va korpus, lingvistika, matnni avtomatik qayta ishlash, tarjima dasturi, korpusning lingvistik funktsiyalari.

Korpusning lingvistik ahamiyati. Korpusda tilni keng, to‘liq, obyektiv ko‘rsatish imkoniyati bor, shu tufayli u har qanday axborot bankidan afzalligi bilan ajralib turadi. Korpusning qaysi sohadagi amaliy ahamiyati to‘g‘risida to‘xtalishimizdan qat’i nazar, eng avvalo, uning reprezentativlik xususiyati asosiy omil bo‘lib qolaveradi. Korpus turli xildagi tadqiqot uchun asos, manba, vosita vazifasini o‘tay olganligi uchun ham dunyoda juda ko‘p sohalar korpusga tayanib ish ko‘radi. Yuqorida aytilganidek, korpus maxsus qo‘shimcha ma’lumotlar bilan boyitilgan, lingvistik tadqiqotlar uchun asos bo‘la oladigan shunday matnlar massividan iboratki, ular oddiy elektron kutubxonadan farq qiladi. Ta’kidlash joizki,

til korpuslari, eng avvalo, tilshunos uchun kerak. Zero, korpussiz bugungi kunning nazariy hamda amaliy filologiyasini tasavvur etish qiyin. Tilshunoslikka oid tadqiqotlar hamisha dalillar asosida ish ko‘rishni, muayyan tadqiq mavzusi doirasida faktlar yig‘ishni, to‘plangan materiallarni sistemaga solishni taqozo etadi. Bunday katta hajmli ishni bajarishda korpus tilshunosning vaqtini, mehnatini tejaydigan bebaho ish qurolidir. Chunki korpus leksikografiya sohasi uchun boy manba hisoblanib, ular asosida kompyuter yordamida lug‘atlar avvalgiga nisbatan tezlik bilan tuziladi, qayta ishlanadi. Shu yo‘l bilan ishning boshlanishidan tugash jarayoni (nashr)gacha tilni aks ettirib turadi, eskirishga ulgurmaydi. Ikkinchidan, korpusning leksikologiya sohasidagi ahamiyati shundan iboratki, so‘zning qo‘llanish davri, chastotasini aniqlashda hech qanday vosita unga tenglasha olmaydi. Korpus asosida ma’lum so‘zning chastotasini aniqlash uchun berilgan qidiruv natijasida diagramma, grafiklar yordamida so‘zning tartib raqami uning chastotasiga teskari proporsional bo‘ladi, chunki ikkinchi tartib raqamida joylashgan so‘z birinchi raqamli so‘zga nisbatan kamroq, to‘rtinchisi uchinchisiga nisbatan kamroq ishlatilishi aniq8 . Birorta chastotali lug‘at korpuschalik aniq ma’lumot berolmaydi, chunki til doim o‘zgarishda, so‘zning chastotasi ham nisbiydir So‘zlarning leksik-semantik birikish imkoniyati tahlilida korpus metodini qo‘llash yangi avlod lug‘at, grammatikalari, xususan, turg‘un birikmalar lug‘atini yaratish imkonini berdi. Korpus yaratilishi, taraqqiy etishi, til korpuslari asri kirib kelishi bilan lug‘atshunoslarda so‘zlarning qo‘llanilish holatlariga doir juda katta matnlar to‘plami bilan ishlash imkoniyati paydo bo‘ldi. Korpus vositasida so‘zning serqirraligi, bir paytning o‘zida bir necha semantik kategoriyaga mansub bo‘la olishi, bu semantik farqni ajratib olish mumkinligi haqida E.V.Nedoshivina: “Leksikografiya bo‘yicha tadqiqotlar semantika borasidagi izlanishlar bilan chambarchas bog‘liq. Korpusda u yoki bu lingvistik birlikning qurshovini kuzata turib, ushbu birlikni xarakterlovchi ma’lum semantik belgilarni aniqlash mumkin. So‘z bir paytning o‘zida bir necha semantik kategoriyaga mansub bo‘lishi mumkin, shuning uchun so‘zning u yoki bu kategoriyaga tegishlilik darajasiga qarab fikr yuritish lozimKorpusdan foydalanish leksik birlikni kontekstda

o'rganish imkonini beribgina qolmay, so'zshakl, lemma, grammatik kategoriyalar chastotasi, ularning birikish xususiyati haqida ma'lumot olish uchun ham muhim. Ibora yoki turg'un birikmalar semantik jihatdan bo'linmas birlikni tashkil etgani uchun leksikografiyada matnga avtomatik ishlov berish jarayonida buni inobatga olish masalasiga alohida e'tibor qaratmoq zarur. Korpus materiali asosida statistik metod orqali qaysi so'zlar doim birgalikda qo'llanishi, buning natijasida turg'un birikmaga qanchalik aloqador ekanligini aniqlash mumkin. Rus, Yevropa tilshunosligida til bo'yicha qilinadigan barcha tadqiqotlarning til korpusiga asoslanishi odatiy holga aylangan, hatto ba'zi tadqiqotlarda majburiy ham sanaladi. Rus tilshunoslari bu sohada AQSH, Evropa, Yaponiyadan orqada qolganligini afsus bilan tan olishadi, vaholanki, rus korpus lingvistikasi ancha taraqqiyot yo'lini bosib o'tib, bu borada rivojlangan mamlakatlarga tenglashyapti. Mutaxassislarning fikricha, yaqin orada til o'rganayotgan o'quvchi yoki tilning biror jihatini tadqiq etayotgan tadqiqotchi, izlanuvchining lug'atga bo'lgan ehtiyoji korpusga ko'chishi shubhasiz. Uslubiyatni o'rganishda, jumladan, matnning uslubiy xususiyatini tahlil etishda ham korpusga tayanish mumkin. Bu amaliyotni bajarishda matnlarning statistik holati tahlili(matndagi jumlaning uzunligi, bir so'zning boshqa so'z bilan birikish holatining odatiy yoki noodatiyligi)ni aniqlashtiradi. Mazkur usulda korpus yordamida yozma nutq bilan birga og'zaki nutqni ham o'rganish mumkin bo'ladi. Masalan, rasmiy ish qog'ozlarida hujjat matnida shartnoma bo'yicha talablar shakli ma'qulmi yoki shartnoma talablari konstruksiyasi to'g'rimi? Korpusga bu tarzda berilgan savolga sanoqli daqiqalarda qaysi shakl ko'proq qo'llanilishini aniqlab beradi. Shunday qidiruv natijasida rus tili milliy korpusida shartnoma bo'yicha talablar konstruksiyasi 40 marta, shartnoma talablari shakli 3 marta qo'llanilganligi to'g'risida statistik ma'lumot beradi. Bu konstruksiyaning qo'llanilishi, albatta, rus tiliga xos, ehtimol, u o'zbek tilida boshqacha natija bersa kerak.

Korpusning boshqa sohalardagi ahamiyati. "Korpus tor mutaxassislar uchungina kerakmi yoki yana qaysi soha vakili undan foydalanishi mumkin?" degan savollar kun tartibiga chiqadi. Korpus lingvistikasiga oid qator adabiyotlarda

корпусning ahamiyati to'g'risida so'z borar ekan, uning qo'llanilish doirasining ancha kengligi ta'kidlanadi. Korpus foydalanuvchilari doirasi biz o'ylaganimizdan ko'ra ancha keng. Lug'at, grammatikadan faqat tilshunos foydalanmaganidek, korpus gumanitar fanlar tadqiqotchisi, adabiyotshunos hamda tarixchi uchun ham birdek zaruriy baza bo'la oladi. Til korpusiga eng ko'p ehtiyoj sezuvchi mutaxassis matnni avtomatik qayta ishlash (masalan, tarjima dasturi), turli qidiruv tizimlari bilan ishlaydigan dasturchi. Chunki bu mutaxassis tabiiy til bilan ish ko'radi hamda ushbu tilda yozilgan barcha matnlar strukturasini (tabiiy, jonli tilda, akademik grammatika, darslik yoki o'quv qo'llanmalardagi misollarga tayanib emas!) mukammalroq "tushunishi", his etishi lozim. Til korpusiga ehtiyoj sezadiganlardan yana biri kundalik turmush tarzida yozuv bilan shug'ullanuvchi mutaxassislar: gazeta, jurnal muharriri, jurnalist, radio, televideniye xodimi hisoblanadi. Chunki bunday xodim ma'lum bir so'z, ibora yoki konstruksiyaning qo'llanilish holati, darajasini, kim, qachon ilk marta shu konstruksiyani qo'llaganligi, qanday uslub uchun xoslanishini bilishga grammatika bilan shug'ullanuvchi olimdan ko'ra ko'proq ehtiyoj sezadi. Biz bugungacha darslik, akademik grammatika, lug'atga qanday tayangan bo'lsak, zamonaviy soha vakillari korpusga shu qadar ehtiyoj sezishadi, undan foydalanishadi.

Foydalanilgan adabiyotlar:

1. Хамроева Ш.М. Корпус лингвистикаси атамаларининг қисқача изоҳли луғати. – Тошкент: Камалак, 2018. – 96 б.
2. Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Монография. – Германия: GlobeEdit, 2020. – 250 б.
3. Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.
4. Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти. Монография. – Тошкент, 2018. – 152 б.