

KORPUSNING LINGVISTIK AHAMİYATI

Unarova Dilafruz Abdimajit qizi

SamDCHTI akademik litseyi

ona tili va adabiyot fani o'qituvchisi

dilafruzunarova1@gmail.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.

Abstract. This article discusses corpus linguistics and its practical significance in all fields. The inextricable link and development of corpus linguistics with the science of linguistics, its place, and its practical significance are analyzed and interpreted. The article explains in detail that the corpus is not only a means of accelerating the technical process, but also an information system of a particular modern form of language, a system that can answer unexpected questions, deal with linguistic phenomena, and pose urgent problems that have not been posed before in the field.

Абстрактный. Абстрактный. В данной статье обсуждается корпусная лингвистика и практическое значение корпусов во всех областях. Анализируется и осмысливается неразрывная связь и развитие корпусной лингвистики с наукой языкознания, ее роль и практическое значение. Корпус — это не только

инструмент ускорения технического процесса, но и информационная система для конкретной современной формы языка, которая также может отвечать на неожиданные вопросы.

способен реагировать, имея дело с языковыми явлениями, до начала поля

В статье говорится, что это система, способная решать актуальные проблемы, которые до сих пор не решены.

подробно объяснено.

Kalit soʻzlar: kompyuter lingvistikasi, korpus lingvistikasi, til korpusi, lingvodidaktika va korpus, lingvistika, matnni avtomatik qayta ishlash, tarjima dasturi, korpusning lingvistik funksiyalari.

Ключевые слова: компьютерная лингвистика, корпусная лингвистика, языковой корпус, лингвистическая дидактика и корпус, лингвистика, автоматическая обработка текста, программное обеспечение для перевода, лингвистические функции корпуса.

Keywords: computational linguistics, corpus linguistics, language corpus, linguistic didactics and corpus, linguistics, automatic text processing, translation software, linguistic functions of the corpus.

Korpusning lingvistik ahamiyati. Korpusda tilni keng, toʻliq, obyektiv koʻrsatish imkoniyati bor, shu tufayli u har qanday axborot bankidan afzalligi bilan ajralib turadi. Korpusning qaysi sohadagi amaliy ahamiyati toʻgʻrisida toʻxtalishimizdan qatʻi nazar, eng avvalo, uning reprezentativlik xususiyati asosiy omil boʻlib qolaveradi. Korpus turli xildagi tadqiqot uchun asos, manba, vosita vazifasini oʻtay olganligi uchun ham dunyoda juda koʻp sohalar korpusga tayanib ish koʻradi. Yuqorida aytilganidek, korpus maxsus qoʻshimcha maʼlumotlar bilan boyitilgan, lingvistik tadqiqotlar uchun asos boʻla oladigan shunday matnlar massividan iboratki, ular oddiy elektron kutubxonadan farq qiladi. Taʼkidlash joizki, til korpuslari, eng avvalo, tilshunos uchun kerak. Zero, korpusiz 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. Aslida, korpus texnik jarayonni tezlashtiruvchi vosita bo'libgina qolmay, ma'lum til zamonaviy shaklining axborot tizimi, kutilmagan savollarga ham javob bera oladigan, tilshunoslik oldiga avval qo'yilmagan dolzarb muammoni qo'ya oladigan tizimdir⁶. Korpus asosida ish ko'radigan eng birinchi soha – leksikografiya, u katta hajmli lug'atlarni tuzish uchun asosiy, betakror manba hisoblanadi. Yaratilayotgan barcha zamonaviy, so'nggi lug'atlar korpusga asoslangan, ular misollarining haqiqiyliги, ishonarliligi bilan baholanadi. Chunki korpusda til jamiyatda qanday yashasa, shunday aks etadi, natijada, lug'atdagi misol ishonarli, asosli bo'ladi⁷. Vaqt o'tishi bilan korpuslar turli lingvistik yo'nalishlar uchun ahamiyatli bo'lishi bilan kuchli informatsion resursga aylandi. 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. Bundan tashqari, korpus tilning lug'at boyligida bo'layotgan o'zgarishlar (neologizm, istorizm, arxaizm hodisalari)ni kuzatishning eng qulay vositasidir. 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

lugʻatshunoslarda soʻzlarning qoʻllanilish holatlariga doir juda katta matnlar toʻplami bilan ishlash imkoniyati paydo boʻldi. Korpus vositasida soʻzning serqirraligi, bir paytning oʻzida bir necha semantik kategoriyaga mansub boʻla olishi, bu semantik farqni ajratib olish mumkinligi haqida E.V.Nedoshivina: “Leksikografiya boʻyicha tadqiqotlar semantika borasidagi izlanishlar bilan chambarchas bogʻliq. Korpusda u yoki bu lingvistik birlikning qurshovini kuzata turib, ushbu birlikni xarakterlovchi maʼlum semantik belgilarni aniqlash mumkin. Soʻz bir paytning oʻzida bir necha semantik kategoriyaga mansub boʻlishi mumkin, shuning uchun soʻzning u yoki bu kategoriyaga tegishlilik darajasiga qarab fikr yuritish lozim. Daraja esa turli kategoriya boʻyicha taqsimlanish chastotasini sanash yoʻli bilan aniqlanadi”, deb yozadi. Demak, korpus yordamida soʻzlarning qaysi semantik kategoriyalarga mansubligini aniqlash, har bir kategoriyadagi maʼnosi bilan tanishish mumkin. Til korpusi yordamida oʻtkazilgan ilk lingvistik tadqiqotlar turli birlikning tilda qoʻllanilish chastotasini statistik aniqlashni koʻzlaydi. Koʻpincha, bunday element soʻz, baʼzida grafema, morfema, soʻz birikmasi ham boʻladi. Korpus lingvistikasining keyingi taraqqiyot bosqichida statistik tadqid metodi kompyuter tarjimai, nutqni sintezlash, tanish, aniqlash, orfografik tekshiruv kabi lingvistik amallarni bajarishda qoʻl kela boshladi. Korpusdan foydalanish leksik birlikni kontekstda oʻrganish imkonini beribgina qolmay, soʻzshakl, lemma, grammatik kategoriyalar chastotasi, ularning birikish xususiyati haqida maʼlumot olish uchun ham muhim. Ibora yoki turgʻun birikmalar semantik jihatdan boʻlinmas birlikni tashkil etgani uchun leksikografiyada matnga avtomatik ishlov berish jarayonida buni inobatga olish masalasiga alohida eʼtibor qaratmoq zarur. Korpus materiali asosida statistik metod orqali qaysi soʻzlar doim birgalikda qoʻllanishi, buning natijasida turgʻun birikmaga qanchalik aloqador ekanligini aniqlash mumkin. Rus, Yevropa tilshunosligida til boʻyicha qilinadigan barcha tadqiqotlarning til korpusiga asoslanishi odatiy holga aylangan, hatto baʼzi tadqiqotlarda majburiy ham sanaladi. Rus tilshunoslari bu sohada AQSH, Evropa, Yaponiyadan orqada qolganligini afsus bilan tan olishadi, vaholanki, rus korpus lingvistikasi ancha taraqqiyot yoʻlini bosib oʻtib, 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. Sintaksisni o'rganishning ahamiyati tilning jonli qurilishini, so'zlarning o'zaro birikish imkoniyatini tahlil qila olish bilan xarakterlanishini bilamiz. An'anaviy tilshunoslikdagi misolni olishda badiiy asar tiliga tayanish tajribasidan ko'ra korpusga tayanish misolning bugungi kun uchun ham ishonarliligini ta'minlaydi. 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 sohalaridagi ahamiyati. "Korpus tor mutaxassislar uchungina kerakmi yoki yana qaysi soha vakili undan foydalanishi mumkin?" degan savollar kun tartibiga chiqadi. Korpus lingvistikasiga oid qator adabiyotlarda korpusning 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. Korpusdan tashqari yana qaysi axborot banki shunday savollarga tez, oson javob berishi mumkinligini tasavvur etib bo'lmaydi. Shuning uchun korpus lingvistikasi tadqiqotchilari muayyan korpuslar jurnalist, muxbir, muharrir, o'qituvchi hamda dasturchilar uchun maxsus yaratila boshlangan, degan xulosaga ham kelishadi. Biz bugungacha darslik, akademik grammatika, lug'atga qanday tayangan bo'lsak, zamonaviy soha vakillari korpusga shu qadar ehtiyoj sezishadi, undan foydalanishadi.

Foydalanilgan adabiyotlar:

1. Захаров В.П., Богданова С.Ю. Корпусная лингвистика.— Иркутск: ИГЛУ, 2013. – 161 с.
2. Информационные технологии в лингвистике: учеб. Пособие / Л.Ю.Щипицина. –М.: ФЛИНТА: Наука, 2013. –128 с.
3. Ҳамроева Ш.М. Корпус лингвистикаси атамаларининг қисқача изоҳли луғати. – Тошкент: Камалак, 2018. – 96 б.
4. Ҳамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Монография. – Германия: GlobeEdit, 2020. – 250 б.
5. Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.
6. Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти. Монография. – Тошкент, 2018. – 152 б.