

KORPUS LINGVISTIKASI METODOLOGIYASI VA MATN TAHLILI**Kaxorova Nargiza Nusratovna**Buxoro davlat universiteti "Ingliz adabiyotshunosligi
va tarjimashunoslik" kafedrası assistenti

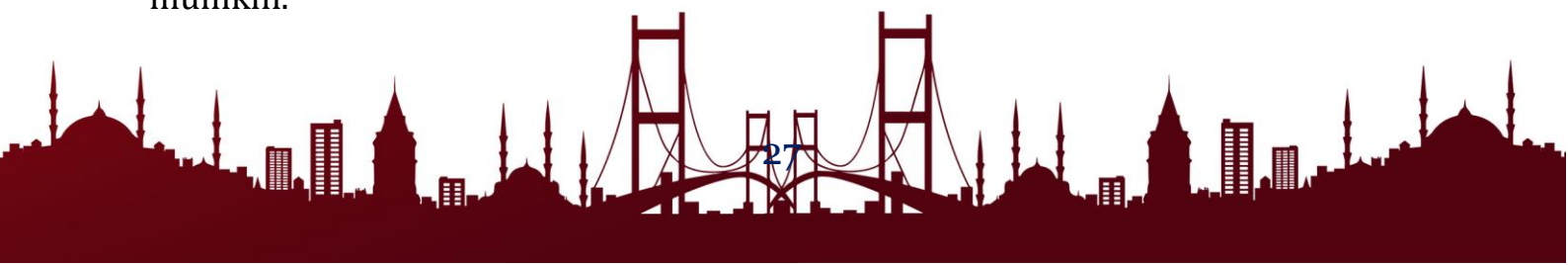
Kaxorovanargiza5@gmail.com

<https://doi.org/10.5281/zenodo.14044706>

Korpus lingvistikasi - bu tabiiy ravishda uchraydigan og'zaki va yozma matnlarning katta, elektron to'plamlaridan foydalanish orqali tildan foydalanishning kompyuterga asoslangan empirik tahlillarini (ham miqdoriy, ham sifat) o'z ichiga olgan metodologiya. Korpus lingvistikasi tillarning tabiati, tuzilishi va qo'llanilishini tavsiflovchi dalil manbai sifatida og'zaki va yozma matnlar to'plamini to'plash va tahlil qilishni o'z ichiga oladi. Hozirgi korpus lingvistikasi tilning barcha darajalarini, jumladan, fonologiya, leksika, grammatika va nutqni o'rganuvchi, lingvistik tavsif uchun empirik asos sifatida matnlardan foydalanishning uzoq an'alaridan kelib chiqdi. Graeme Kennedy lingvistik tadqiqotlarda korpusdan foydalanishning rivojlanishiga oid tadqiqotlar olib borgan. Anke Ludeling va Kyoto Merja tilshunoslikning bir nechta sohalarida korpus lingvistikaning o'rnini o'rgangan. Anatol Stefanowitsch esa korpus lingvistikasiga oid muammolar borasida tadqiqot olib borgan. Randi Reppen va Rita Simpson-Vlach tilni tahlil qilish uchun korpus lingvistikasidan foydalanish bo'yicha tadqiqot olib borgan. William Crawford va Eniko Csomay esa mavjud korpuslarda qidiruvlar o'tkazib, korpus yaratish va miqdoriy natijalarni tahlil qilish bo'yicha bir qator izlanishlar olib borishgan. O'zbek olimlaridan esa O.J Xidirov Milliy korpus uchun parsing dasturi yaratishning lingvistik asoslari bo'yicha tadqiqot olib borgan.

Tilshunoslikda korpus- bu lingvistik tahlil va tavsif uchun asos bo'lib xizmat qiluvchi yozma matn yoki transkripsiyalangan nutq to'plami sanaladi. Shuningdek, korpus elektron ma'lumotlar bazasidagi matnlar to'plami bo'lib, lingvistik tahlil uchun maxsus tuzilgan matn to'plamlaridan iborat. Bu butun matnlardan olingan uzluksiz matn namunalari iborat bo'lishi, ba'zan iqtiboslar to'plamidan ham tuzilishi mumkin. Korpus nafaqat biz tilda foydalanadigan tizimlarni tashkil etuvchi tizimli elementlar va namunalarni aniqlash uchun, balki ushbu tizimlardan foydalanishimizni xaritalash uchun ham empirik asosdir. Korpuslarni tahlil qilish va boshqa korpuslar bilan solishtirish orqali tildagi o'zgaruvchanlikni samarali o'rganishimiz mumkin. Masalan, tildagi muayyan fonologik, leksik, grammatik, nutqiy yoki pragmatik xususiyatlar qanchalik tez-tez va qayerda sodir bo'lishini korpus tahlili orqali o'rganamiz.

So'nggi o'n uch yillikda kompyuterlashtirilgan ma'lumotlar bazalarida saqlanadigan korpuslarni yig'ish va tahlil qilish korpus lingvistikasi deb nomlanuvchi sohani yaratilishiga olib keldi. Bu soha lingvistik tavsiflash metodologiyalarini o'zida mujassam etgan bo'lib, lingvistik elementlarning tarqalishini miqdoriy aniqlash tadqiqot faoliyatining bir qismi hisoblanadi. Korpus lingvistikasi til miqdoriy tahlilining ko'p qirrali metodologiyasi bo'lib, undan grammatika (Biber), nutq (Baker), turli ijtimoiy aktyorlar va jarayonlarni (Baker) o'rganish uchun foydalanish mumkin. Ingliz olimlaridan McEnery va Hardie korpus lingvistikasi tilga ilmiy yondashuv bo'lib, u empirik tekshirish an'alariga, ya'ni tadqiqotda asosiy e'tibor ma'lumotlarni to'plash va izohlashga qaratilganligini ta'kidlaydi. McEnery korpus- elektron formatdagi matnlar to'plami bo'lib, nafaqat sahifama-sahifa o'qish imkoniyatini, balki matnlardagi so'z yoki iboralarning miqdoriy ko'rsatkichlarini aniqlash imkonini beradi deb ta'kidlaydi. Korpus og'zaki nutqga oid asl namunalarni va asl matnlarni o'z ichiga oladi. Shuning uchun til korpuslari uchun maxsus matnlar yaratilmaydi, balki gazeta maqolalari, ilmiy maqolalar, kitobiy matnlar namunalari to'planadi. Ammo, korpus- bu tasodifan uchragan matnlar to'plami emas, balki puxta ishlab chiqilgan ma'lumotlar bazasi bo'lib, bizga ma'lum bir matn haqida ko'proq ma'lumotga ega bo'lish imkonini beradi. Korpus korpora so'zining ko'plik shakli bo'lib, biz kompyuter orqali qidirishimiz mumkin bo'lgan minglab sahifali matnlar va og'zaki til transkriptlarini o'z ichiga oluvchi lingvistik ma'lumotlarning alohida turi sanaladi. Tilshunoslar har doim tilning tabiati, elementlari, tuzulishi va funksiyalari haqidagi nazariyalar uchun dalil manbalariga muhtoj bo'lib kelgan. Turli zamonlarda bunday dalillar sezgi yoki tajribadan, shuningdek, og'zaki yoki yozma matnlarda sodir bo'lgan kuzatuvlarga asoslangan tavsiflardan kelib chiqqan. Korpusga asoslangan tadqiqotlarda esa dalillar to'g'ridan-to'g'ri matnlardan olinadi. Shu jihatdan korpus lingvistikasi sezgi yoki tajribaga asoslangan boshqa til yondashuvlardan farq qiladi. Boshqa tilshunoslik sohalari kabi korpus lingvistikasi ham birinchi navbatda, til va tillarning tabiati, tuzulishi va qo'llanilishini tavsiflash bilan birga, tildagi o'zgaruvchanlik yoki o'zgarish masalalari bilan shug'ullanadi. Shuningdek, korpus lingvistikasi tilshunoslikda o'ziga xoslikni, ya'ni faqat sintaksisga emas, balki leksika hamda leksik grammatikaga e'tibor bilan yondashuv tendensiyasini olib kirdi. Standart korpus lingvistik tadqiqotlar yordamida biz oldin hech qachon o'ylamagan tilga oid faktlarni ko'rishimiz mumkin.



Foydalanilgan adabiyotlar ro'yxati:

1. Graeme Kennedy. An Introduction to corpus linguistics - Routledge, 2014. – 6 b.
2. Anke Ludeling, Kyto Merja. Corpus Linguistics – Mouton de Gruyter, 2008. – 13 b.
3. Anatol Stefanowitsch. Corpus linguistics: A guide to the methodology, 2020. – 23 b.
4. Randi Reppen, Rita Simpson-Vlach. Corpus linguistics – An introduction to applied linguistics, 2019. – 91 b.
5. William Crawford, Eniko Csomay. Doing corpus linguistics – Routledge, 2015. – 28 b.

