

**Sobirova Shohsanam Sobirjonovna**

*UrDU Tarjima nazariyasi va amaliyoti nemis tili yo'nalishi talabasi*

*E-Mail: shokhsanamsobirova@gmail.com*

**Annotatsiya:** *Ushbu maqola korpusga asoslangan lingvistika, tilni o'rganish va tilni tahlil qilishning muhim yondoshuvlarni va albatta korpusga asoslangan lingvistikada til bo'yicha yuzaga keladigan muammolar haqida bayonot beradi.*

**Kalit so'zlar:** *lingvistika, korpus, avtomatik tarjima, sintaktik, veb-sayt, chet tili, Google, internet.*

**Аннотация:** *В этой статье представлен обзор важных подходов к корпусной лингвистике, изучению языка и языковому анализу, а также связанных с языком проблем в корпусной лингвистике.*

**Ключевые слова:** *лингвистика, корпус, автоматический перевод, синтаксис, веб-сайт, иностранный язык, Google, Интернет.*

**Abstract:** *And here is an overview of important approaches to corpus linguistics, language study and linguistic analysis, as well as, finally, problems related to language and corpus linguistics.*

**Key words:** *linguistics, corpus, automatic translation, syntactic, website, foreign language, Google, internet.*

## **KIRISH**

Korpusga asoslangan lingvistika, korpuslar yordamida tili o'rganish va tili tahlil qilishga asoslanadi. Korpus, bir yoki bir nechta tillarning yozilgan matnlarining to'plamidir. Korpuslar, tilni o'rganish, tili tahlil qilish, tilni avtomatik tarjima qilish, tilni avtomatik imlo tekshirish va boshqa amallarni bajarish uchun foydalaniladi. Korpuslar, tilda o'zgarishlarni kuzatish, tilda so'zlarni va ularning manolarini topish, tilda grammatikasini o'rganish va tilda tarixini o'rganishga yordam beradi. Shu sababli korpusga asoslangan lingvistika, tilda yuzaga kelgan muammolarni hal qilish uchun juda muhimdir.

## **ASOSIY QISM**

Korpusga asoslangan tadqiqot odatda korpus ma'lumotlardan nazariya yoki gipoteza yaratishda foydalanadi. Korpus lingvistikasining metod sifatidagi ta'rifi korpus ma'lumotlardan lingvistikada foydalanish yondashuvini tasdiqlaydi. Korpusga yo'naltirilgan lingvistika korpus lingvistikaning metod sifatidagi xususiyatlarinin rad etadi va korpusning o'zi til haqidagi gepotezaning yagona manbasi bo'lishi kerakligini takidlaydi. Aytish mumkinki, korpusning o'zi til haqida alohida nazariyani ifodalaydi.

Korpusga yo'naltirilgan lingvistika tushunchasi lingvistika tadqiqot olib boruvchi olimlar ishi bilan yaqindan bog'langan korpusga asoslangan yoki korpusga yo'naltirilgan lingvistika o'rtasidagi xilma-xillik korpus lingvistikasi bo'yicha olib

boriladigan tadqiqotlarni yagona yoki boshqa bir guruhlarga ajratishda asosiy farqni yaratadi. Biroq, korpusning o'zi nazariy asosga ega ekanligi haqidagi tushunchani rad etgan holda aytish mumkinki, korpusga asoslangan va korpusga yo'naltirilgan lingvistika o'rtasida farq mavjud emas, va ikkala tarmoq ham bir so'z bilan korpusga asoslangan deya ta'riflanishi mumkin.

### **Ma'lumot to'plash tartiblari.**

Korpus tadqiqotlari korpus ma'lumotlarni tadqiqot savollari bilan uyg'unlashtirishi kerak. Bu mutanosiblikning qoniqarli darajada ekanligiga qanday ishonch hosil qilishimiz mumkin? Agar biz Zamonaviy ingliz tilining grammatik xususiyatlarini tadqiq qilmoqchi bo'lsak, foydalanadigan ma'lumotlarimizni tadqiqot vazifalari bilan bog'lashimiz zarur. Og'zaki ingliz tili haqida umumiy xulosalar chiqarish uchun, mavzuga doir munosib axborotlar bazasiga ega bo'lishimiz talab qilinadi. Faqatgina bir insonning nutqi bu kabi umumiy xulosalar chiqarish uchun yetarli asos bo'la olmaydi. Shuning uchun korpus yaratish va alohida bir turdagi ma'lumotlarni to'plash korpus lingvistikasida asosiy muhim masala sifatida yuzaga keladi. Axborot to'plashda ulardan qaysilarini tanlash kerakligi yuzasidan 2 xil yondashuv mavjud:

- 1) kuzatuvchi korpus tamoyili, bunda korpus doimiy kengayib, yanada ko'proq ma'lumotlarni qamrab boradi,
- 2) barqaror korpus tamoyili, tilning faqat ma'lum vaqt davomidagi xususiyatlarini, ma'lum bir shklda aks ettiradi.

### **Kuzatuvchi korpus tamoyili**

Bu tamoyil fanga Jon Sinkler tomonidan kiritilgan bo'lib, turli xil ma'lumotlarni qamrab olgan va vaqt davomida doimiy hajmi ortib boruvchi axborot bazasi yaratishda qo'llaniladi. Turli xildagi ma'lumotlarning nisbat proportsiyasi vaqt mobaynida bir biridan farq qilishi mumkin.

### **Veb-sayt korpus sifatida**

Veb-sayt korpus tushunchasi (Kilgarriff va Grefenstet 2003) monitor korpusi g'oyasiga ko'p jihatdan juda o'xshash. Buning sababi bor boshlang'ich nuqtasi har doim o'sib borayotgan va yig'ilgan ma'lumotlar to'plamidir, va undan til o'rganishda foydalanadi (masalan, veb-ga asoslangan antonimlarni o'rganish). Shuningdek

vab-sahifani korpus sifatida o'rganish uchun Google kabi standart qidiruv tizimlaridan foydalanish, tadqiqotchilar, shu jumladan, uni qo'llab-quvvatlash uchun mo'ljallangan interfeyslarni ishlab chiqdilar.

Veb-sayt korpus yondashuvi muayyan muammolar mavjud. Aksariyat korporatsiyalardan farqli o'laroq, bu veb-saytning aralashmasi ehtiyotkorlik bilan tayyorlangan va tahrirlangan matnlarni va qanday qilib tasodifan "tasodifan" deb atash mumkin tayyorlangan material. Internetning tarkibi ham janrga bo'linmaydi - shuning uchun ham Veb-qidiruv orqali qaytarilgan materiallar farqsiz massa bo'lishga intiladi, bu esa muhim guruhlarga ajratish uchun juda ko'p ishlov berishni talab qilishi mumkin. Bundan tashqari, vebdagi ko'plab matnlar xatoliklarga yo'l qo'ymasligi

shubhasizdir. Misol uchun, biz ushbu kitobni yozayotganda biz receive va receive so'zlarini Googlega yozdik va receive 300 milliondan ziyod, receive

8million 670 ming ballni qo'lga kiritdi. Bu kabi ma'lumotlar ham juda qiziqarli bo'lishi mumkin imlo islohotini qo'llab-quvvatlash bo'yicha tadqiqotlar. Biroq, agar siz bunday narsaga ega bo'lmasangiz manfaatdor tomonlar, ma'lumotlardagi bunday xatolar, agar kerak bo'lsa, kiruvchi shovqinni ta'minlaydi tahlilchi veb-sahifani korpusga yaqinlashtiradi. Ushbu turdagi shovqin internetdagi tilning barcha darajalarida mavjud bo'lganligini hisobga olsak, u foydalanuvchilarning veb-sayt korpus sifatida ko'rib chiqilishi kerak. Shunga qaramay, veb-sayt shubhasiz korporativ ishlab chiqarish uchun tanlab olinadigan va tayyorlanadigan ma'lumotlarning katta miqdori turli maqsadlar uchun javob beradi.

### **XULOSA**

Korpus lingvistikasi so'nggi o'n yillikda jadal rivojlanayotgan sohaga aylandi. Chunki dunyo tili sanalgan ingliz, nemis, rus, xitoy, arab, ispan, fransuz tillarining milliy korpuslaridan tashqari, bu tillarning ko'pida gazetalarning kompyuter korpusi, multimediali korpus, maxsus matnlar korpusi, og'zaki va yozma korpus, umummilliy va mualliflik korpuslari ham yaratildi. Shuningdek, belarus, polyak, polyak-ukrain, chex, slovak milliy korpuslari, serb, xorvat, bosniya, bolgar, bolgar-rus, makedon, shotland, ivrit, niderland, niderland-fransuz, shved, dat, norveg, island, farer, o'rta asr fransuz tili, ispan, italyan, portugal, rumin, litva, latish, grek, sharqiy arman, osetin, alban, hind, sigan, xett, fin, ural tillari, eston, veps, venger, udmurt, gruzin, ingliz-gruzin, lezgin, turk, tatar, boshqird, qrim-tatar, qalmiq, buryat, mo'g'ul, arab, amxar, yapon, qadimiy yapon, baman, tojik, esperanto tillari korpuslari ham mavjud.

Yurtimizda o'zbek tili korpuslarini yaratish qoniqarli emas. Bu sohada olib borilayotgan ilmiy tadqiqot ishlari o'zbek korpus lingvistikasining nazariy asoslarini ishlab chiqish, namuna sifatida bir-ikki korpus lavhasi yaratish jarayoni bosqichida turibdi, xolos.

### **FOYDALANILGAN ADABIYOTLAR RO'YHATI:**

1. Ne'matov H. Bozorov O. Til va nutq, - Toshkent: O'qituvchi, 1993.
2. Safarov Sh. Kognitiv tilshunoslik. -Jizzax: Sangzor, 2006.
3. Лурия А.Р. Основные проблемы нейролингвистики. – МГУ, 1975
4. Фердинанд де Соссюр. Труды по языкознанию. – М.: 1977.
5. Варобев В.В Лингвакультурология. М.: Издательство РУДН, 2006
6. Zaharov V. "Korpuslingvistika" M.: 2005.
7. Анисимов А.В. Компьютерная лингвистика: мифы, алгоритмы, язык. – Киев: Наукова думка. 1991.
8. Лурия А.Р. Основные проблемы нейролингвистики. – МГУ, 1975.
9. Нурмонов А., Йўлдошев Б. Тилшунослик ва табиий фанлар. Тошкент, 2001.- 160 бет.

10. Po'latov A., Muhamedova S. Kompyuter lingvistikasi (metodik qo'llanma).-  
Toshkent: Universitet, 2009.- 104 bet.