



## KORPUS LINGVISTIKASI

*Andijon davlat chet tillari instituti*

*Sotiboldiyeva Muniraxon Abbosxon qizi*

*Ilmiy maslahatchi: Xomidova Malohat*

**Annotatsiya:** Ushbu maqolada korpus lingvistikasi tushunchasi, uning tilshunoslikdagi o‘rni va ahamiyati haqida so‘z yuritiladi. Shuningdek, korpus lingvistikasining asosiy vazifalari, tilni o‘rganishdagi imkoniyatlari hamda zamonaviy texnologiyalar bilan bog‘liqligi yoritib beriladi. Korpus lingvistikasi yordamida til birliklarini real matnlar asosida tahlil qilish, lug‘at tuzish va ilmiy tadqiqotlar olib borish masalalari ko‘rib chiqiladi.

**Kalit so‘zlar:** *korpus lingvistikasi, tilshunoslik, matn korpusi, til birliklari, kompyuter lingvistikasi, til tahlili, zamonaviy texnologiyalar.*

**Аннотация:** В данной статье рассматривается понятие корпусной лингвистики, её роль и значение в языкознании. Также освещаются основные задачи корпусной лингвистики, её возможности в изучении языка и связь с современными технологиями. Рассматриваются вопросы анализа языковых единиц на основе реальных текстов, составления словарей и проведения научных исследований с использованием корпусных данных.

**Ключевые слова:** *корпусная лингвистика, языкознание, текстовый корпус, языковые единицы, компьютерная лингвистика, анализ языка, современные технологии.*

**Annotation:** This article discusses the concept of corpus linguistics, its role and significance in linguistics. It also highlights the main tasks of corpus linguistics, its possibilities in language study, and its connection with modern



technologies. The paper examines the analysis of language units based on real texts, dictionary compilation, and conducting scientific research using corpus data.

**Keywords:** *corpus linguistics, linguistics, text corpus, language units, computational linguistics, language analysis, modern technologies.*

Korpus lingvistikasi tilshunoslikning zamonaviy va tez rivojlanayotgan yo‘nalishlaridan biridir. Bu yo‘nalish tilni o‘rganishda katta hajmdagi matnlar to‘plamidan, ya’ni korpuslardan foydalanishga asoslanadi. Korpus lingvistikasi til birliklarining real qo‘llanilish holatini aniqlash, ularni tahlil qilish hamda tilning rivojlanish jarayonlarini o‘rganishda muhim ahamiyat kasb etadi.

“Korpus” so‘zi lotincha “corpus” – “tana” yoki “majmua” ma’nosini anglatadi. Tilshunoslikda esa korpus deganda kompyuterda saqlanadigan va maxsus dasturlar yordamida tahlil qilinadigan katta hajmdagi matnlar to‘plami tushuniladi. Ushbu matnlar turli manbalardan olinishi mumkin. Masalan, badiiy adabiyotlar, ilmiy maqolalar, gazeta va jurnal materiallari, internet sahifalari, suhbat yozuvlari va boshqa matnlar korpus tarkibiga kirishi mumkin.

Korpus lingvistikasining asosiy maqsadi til birliklarining real matnlardagi qo‘llanilishini o‘rganishdir. An’anaviy tilshunoslikda ko‘pincha misollar olimlarning shaxsiy kuzatuvlari yoki kichik matnlar asosida keltirilgan bo‘lsa, korpus lingvistikasi katta hajmdagi haqiqiy matnlarga tayanadi. Bu esa tilni o‘rganishda yanada aniq, obyektiv va ishonchli natijalarga erishishga yordam beradi.

Korpus lingvistikasi yordamida so‘zlarning qanchalik tez-tez ishlatilishi, ularning boshqa so‘zlar bilan qanday birikmalarda kelishi, grammatik shakllari va ma’nodagi o‘zgarishlari tahlil qilinadi. Buning uchun maxsus kompyuter dasturlari qo‘llaniladi.

Ushbu dasturlar matnlar ichidan kerakli so‘z yoki iboralarni tezda topish, ularning ishlatilish holatini ko‘rsatish hamda statistik tahlil qilish imkonini beradi.



Bugungi kunda korpus lingvistikasi nafaqat tilshunoslikda, balki boshqa sohalarda ham keng qo'llanilmoqda. Masalan, lug'atlar tuzishda korpus ma'lumotlari juda muhim ahamiyatga ega. Chunki lug'at tuzuvchilar so'zlarning real matnlardagi qo'llanishini korpus orqali ko'rib chiqishlari mumkin. Bundan tashqari, tarjima jarayonida ham korpusdan foydalanish tarjimonlarga so'z va iboralarning to'g'ri qo'llanishini aniqlashga yordam beradi.

Korpus lingvistikasi sun'iy intellekt va kompyuter texnologiyalari bilan ham chambarchas bog'liqdir. Mashina tarjimasi, nutqni tanish tizimlari va matnni avtomatik tahlil qilish kabi texnologiyalarni yaratishda korpus ma'lumotlari muhim manba hisoblanadi. Shu sababli ham korpus lingvistikasi zamonaviy kompyuter lingvistikasining muhim tarkibiy qismi hisoblanadi.

Hozirgi kunda dunyoning ko'plab tillari uchun milliy korpuslar yaratilgan. Masalan, ingliz, rus, nemis va boshqa ko'plab tillarning katta hajmdagi korpuslari mavjud. Bu korpuslar tilshunoslar, tadqiqotchilar va talabalar uchun muhim ilmiy manba bo'lib xizmat qiladi.

O'zbek tilida ham milliy korpus yaratish ishlari olib borilmoqda. O'zbek tilining korpusi tilni ilmiy asosda o'rganish, yangi lug'atlar yaratish va tilni rivojlantirish uchun muhim imkoniyat yaratadi. Shuningdek, u o'zbek tilini chet ellarda o'rganayotgan talabalar uchun ham foydali bo'lishi mumkin.

#### Korpus lingvistikasi turlari

Korpuslar maqsad va tarkibiga qarab bir necha turga bo'linadi:

Milliy korpus – bir tilning katta matnlar bazasi, masalan British National Corpus

Parallel korpus – bir matnning bir nechta tildagi tarjimalari, tarjima sifatini tekshirishda ishlatiladi

Maxsus korpus – ma'lum bir soha (tibbiyot, huquq, iqtisod) matnlarini o'z ichiga oladi

O'quv korpusi – til o'rganuvchilar nutqi asosida tuziladi



Tarixiy korpus – tilning tarixiy rivojlanishini o‘rganish uchun ishlatiladi

Korpus turini tanlash tadqiqotning maqsadiga bog‘liq. Masalan, tarjima sifatini o‘rganish uchun parallel korpus, til o‘rganuvchilar nutqini tahlil qilish uchun o‘quv korpusi ishlatiladi.

#### Korpus tahlili metodlari

Korpus lingvistikasi bir qator ilmiy metodlardan foydalanadi:

Statistik tahlil – so‘zlarning ishlatilish chastotasini aniqlash

Konkordans tahlili – so‘zning turli matnlardagi kontekstini ko‘rish

Chastota tahlili – eng ko‘p ishlatiladigan so‘zlarni aniqlash

Kollokatsiya tahlili – so‘zlarning bir-biri bilan birga ishlatilish xususiyatlarini o‘rganish

Korpus lingvistikasi ilmiy tadqiqotlar uchun aniq va ishonchli manba hisoblanadi. Elektron matnlar bazasi yordamida olingan ma’lumotlar asosida tilning tarixiy rivojlanishi, zamonaviy ishlatilishi va turli janrlar bo‘yicha tahlil qilish mumkin. Shu sababli, bu yo‘nalish ilm-fan va amaliyotning birlashgan ko‘rinishi sifatida tilshunoslikda o‘z o‘rniga ega. Korpus lingvistikasi tilni o‘rganishda yangi, samarali va ilmiy asoslangan metodlarni taqdim etadi. Bu soha tilshunoslik tadqiqotlarini chuqurroq va realroq qiladi, til o‘qitishni yanada samarali qiladi, tarjima va kompyuter tizimlarida ishlatilishini yaxshilaydi. Shu bilan birga, korpus lingvistikasi til evolyutsiyasini kuzatishda, yangi til birliklarini aniqlashda va tilni amaliy jihatdan qo‘llashda muhim rol o‘ynaydi. Bugungi kunda korpus lingvistikasi nafaqat ilmiy, balki amaliy jihatdan ham tilshunoslikning eng asosiy va istiqbolli yo‘nalishlaridan biri sifatida e’tirof etilad

#### Foydalanilgan adabiyotlar:

1. McEnery T., Hardie A. *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press, 2012.

2. Biber D., Conrad S., Reppen R. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press, 1998.



3. Sinclair J. Corpus, Concordance, Collocation. Oxford University Press, 1991.
4. O‘zbek tilining milliy korpusi materiallari.
5. Rahmatullayev Sh. Hozirgi o‘zbek adabiy tili. Toshkent, 2006.
6. Johansson S. Corpus Linguistics and Lexical Semantics. John Benjamins Publishing, 2007.
7. Kilgarriff A., Grefenstette G. Introduction to the Special Issue on Web as Corpus. Computational Linguistics, 2003.