



KORPUS LINGVISTIKASINI TILSHUNOSLIK BO‘LIMLARIGA MUNOSABATI

Raimnazarova Nasiba

TerDPI dotsenti

Xoliyorova To‘lg‘anoy Ravshan qizi

TerDPI 1-kurs magistranti

Annotatsiya. Ushbu maqolada korpus lingvistikasi va tilshunoslik bo‘limlari orasidagi munosabatlar haqida fikr yuritiladi.

Kalit so‘zlar. Korpus, lingvistika, emperik yondashuv, kompyuterlashtirilgan tekstlar, nutqni sintezlash, orfografik yondashuv.

Kirish.

Korpus –bu matnlar yoki matnlardan maxsus tanlangan qismlaridan ma’lum bir tarzda tashkil etilgan og‘zaki birlik. Korpus –bu lingvistik tahlil uchun muayyan tildan olingan matnlar yoki yozma matnlar to‘plami. Korpus –bu elektron shaklda saqlanadigan va kompyuterlashtirilgan qidiruvni tashkil etishga imkon beradigan har qanday tildagi og‘zaki yoki yozma matnlardan iborat bo‘lgan tabiiy matnlar to‘plami. Korpus –tilni yoki tilning o‘zgarishini to‘liq aks ettirish uchun tashqi mezonlar asosida tanlangan elektron shakldagi matnlarning qismlari to‘plami. Lingvistik tadqiqotlar uchun ma’lumotlar manbayi sifatida ishlaydi. Korpus tabiiy ravishda mavjud bo‘lgan matnlarning tizimli to‘plami sifatida belgilanishi mumkin (*ham yozma, ham og‘zaki nutq*). Korpus har qanday sistematik matn to‘plamiga murojaat qilishi mumkin bo‘lsa-da, u bugungi kunda ham tor ma’noda ishlatiladi va odatda kompyuterlashtirilgan muntazam matn to‘plamlariga murojaat qilish uchun ishlatiladi. O‘zbekiston Respublikasi Prezidentining 2019-yil 21-oktabrdagi “O‘zbek tilining davlat tili sifatidagi nufuzi va mavqeyini tubdan oshirish chora-tadbirlari to‘g‘risida”gi PF-5850-sonidan keltirilgan iqtibos til korpusini jamiyat hayotida, xususan tilshunoslik bo‘limlarida muhimligini ko‘rsatadi



Tabiiy tillarni qayta ishlash bo'yicha zarur bo'lgan birinchi vazifalardan biri korpusdir. Tilshunoslikda va tabiiy tillarni qayta ishlashda korpus matnlar to'plamiga aloqador bo'ladi. Bunday to'plamlar bir tilning yoki bir necha tillarning matnlaridan shakllantirilishi mumkin.

Asosiy qism

Korpus matnlarning katta to'plami bo'lishi bilan birga, bu lingvistik tahlilga asoslanadigan yozma yoki og'zaki materiallar asosi hamdir. Korpusda turli xildagi matnlar kiritilishi mumkin. Masalan:

- aniq bir yozuvchining yoki yozuvchilarning matnlari;
- zamonaviy matnlar;
- tarixiy matnlar.

Yuqoridagi matnlar shakllantirilgandan keyin, quyidagilarni qidirishimiz mumkin.

- eng ko'p tanlangan so'z;
- lug'atdagi o'zgarish;
- muayyan yozuvchining so'z va iboralar qo'llashdagi o'ziga xoslik;
- so'z birliklaridan foydalanadigan yashirin modellar;
- to'g'ridan-to'g'ri kontekstdagi so'zlarning barcha shakllari.

Korpus lingvistikasi - tilshunoslik bo'limlarida tez rivojlanib borayotgan soha. Ushbu soha tildagi lug'aviy birliklar, ularning ishlatilishi haqidagi ma'lumotlarni yig'ish, saralash va tahlil qilish bilan shug'ullanadi. Bunda til materiallari katta miqdordagi elektron korpuslar shaklida o'rganiladi. Korpus lingvistikasining so'nggi yillardagi o'sishi, kompyuter texnologiyalarini rivoji bilan bog'liq. Korpusda turli xildagi matnlar, tarjima matnlarini ko'rishimiz mumkin.

Korpus lingvistikasi tilshunoslikning boshqa sohalaridan farqli o'laroq, emperik yondashuvni qo'llaydi, real til foydalanish holatlariga suyanadi. Bu sohaga



oid ilk ishlar 20-asrning o'rtalarida paydo bo'lib, to'liq ravishda kompyuterlashtirilgan tekstlar bazasidan foydalanishni o'z ichiga olgan.

Korpus katta hajmli lug'atlarni tuzish uchun manba vazifasini o'taydi. Vaqt o'tishi bilan korpus turli lingvistik yo'nalishlar uchun ahamiyatli bo'lishi bilan, kuchli informatsion resursga aylandi. Korpus asosida kompyuter yordamida lug'atlar avvalgiga nisbatan tez tuziladi va qayta ishlanadi. Shu yo'l bilan ish boshlanishidan tugash jarayonigacha (*nashrgacha*) tilni aks ettirib turadi. Tilning zamonaviy holatini aks ettiruvchi korpus turli davrlarda yashab ijod etgan ijodkorlarning mualliflik korpuslari uchun asos bo'lib xizmat qiladi.

So'zning qo'llanish davri va chastotasini aniqlashda hech qanday vosita korpusga tenglasha olmaydi. Korpus lingvistikasining keyingi taraqqiyot bosqichida statistik tadqiq metodi kompyuter tarjimasi, nutqni sintezlash va tanish, aniqlash, orfografik tekshiruv kabi lingvistik amallarni ajarishda qo'l kela boshladi. Rus va Yevropa tilshunosligida til bo'yicha o'tkaziladigan barcha tadqiqotlarning korpusga asoslanishi odatiy holga aylangan va hatto ba'zi tadqiqotlarda majburiy ham sanaladi. Gap tuzilishini o'rganishdagi ahamiyati tilning jonli qurilishini, so'zlarning o'zaro birikish imkoniyatini tahlil qilishda ko'rinadi. An'anaviy tilshunoslikdagi misolni olishda badiiy asar tiliga tayanish tajribasidan ko'ra korpusga tayanish misolning bugungi kun uchun ham ishonarliligini ta'minlaydi.

Xulosa

Turli tipdagi ma'lumot til korpusida qo'llanish shakliga ko'ra joylashadi. Bu esa uni har tomonlama va xolis o'rganish uchun asos bo'ladi. Bir marta tuzilgan va tayyorlangan axborotlar massivi bir necha tadqiqotchi tomonidan ko'p marta, turli maqsadda ishlatilishi mumkin. Korpus — tilni tadqiq etish (*til birligining o'zgarishi, eskirishi, yangilarining paydo bo'lishi, ma'nosining kengayishi va torayishi; yangi iboralarning paydo bo'lishini kuzatish*), o'rganish, an'anaviy va zamonaviy lug'atlar tuzishda keng imkoniyatli dasturlashtirilgan tizim. Korpusiz bugungi kun nazariy va amaliy filologiyasini tasavvur etib bo'lmaydi.



Tilshunoslikka oid tadqiqotda fakt bilan ish ko‘riladigan hollarda material yig‘ilishi, sistemaga solinishi lozim. Bunday katta hajmli ishni bajarishda korpus vaqt va mehnatni tejaydigan ish quroli vazifasini bajaradi. U 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. Korpus asosida ish ko‘radigan eng birinchi soha leksikografiya bo‘lib, katta hajmli lug‘atlarni tuzish uchun asosiy va takrorlanmas manba sanaladi. Vaqt o‘tishi bilan korpuslar turli lingvistik yo‘nalishlar uchun ahamiyatli bo‘lishi bilan kuchli informatsion resursga ayandi. Chunki korpus leksikografiya sohasi uchun boy manba hisoblanib, ular asosida kompyuter yordamida lug‘atlar avvalgiga nisbatan tezlik bilan tuziladi va qayta ishlanadi. Shu yo‘l bilan ish boshlanish va tugash jarayonigacha (*nashrgacha*) tilni aks ettirib turadi, eskirishga ulgurmaydi.

Foydalanilgan adabiyotlar:

1. O‘zbekiston Respublikasi Prezidentining 2003-yil 11-dekabrda - Axborotlashtirish to‘g‘risidagi qonuni.
2. O‘zbekiston Respublikasi Prezidentining 2012-yil 1-fevraldagi - Joylarda kompyuterlashtirish va axborot-kommunikatsiya texnologiyalarini yanada rivojlantirish uchun shart-sharoitlar chora-tadbirlari to‘g‘risidagi qarori.
3. Беляева Л.Н. Лексикографический потенциал параллельного корпуса текстов Труды международной конференции “Корпусная лингвистика – 2004”. – СПб., 2004.
4. Богданова С.Ю. Исследование слова и предложения компьютерными методами Слово в предложении: кол. монография Под ред. Л.М. Ковалевой (отв. ред), С.Ю. Богдановой, Т.И. Семеновой. – Иркутск: ИГЛУ, 2010.



5. Борисова Е.Г. Слово в тексте. Словарь коллокаций (устойчивых словосочетаний) русского языка с англо-русским словарем ключевых слов. – Москва, 1995.

6. Гарабик Р., Захаров В.П. Параллельный русско-словацкий корпус Труды международной конференции “Корпусная лингвистика– 2006”. – СПб., 2006.

7. Гвишиани Н.Б. Практикум по корпусной лингвистике: Учеб. пособие по английскому языку. – М.: Высшая школа, 2008.

8. Герд А.С. РНК и академическая лексикография Труды международной конференции “Корпусная лингвистика – 2006”. –СПб.: Изд-во С.-Петербур. ун-та; Изд-во РХГА, 2006.

9. Герд А.С. Несколько слов о Специальном корпусе текстов (СКТ) Труды международной конференции “Корпусная лингвистика –2006”. – СПб.: Изд-во С.-Петербур. ун-та; Изд-во РХГА, 2006.

10. Гришина Е.А., Савчук С.О. Национальный корпус русского языка как инструмент для изучения вариативности грамматических норм Труды международной конференции “Корпусная лингвистика – 2008” 6-10 октября 2008 г. – СПб., 2008.

11. Засорина Л.Н. (ред.). Частотный словарь русского языка. – М., 1977.

12. Захаров В.П. Корпусная лингвистика: Учебно-метод. пособие. –СПб., 2005.

13. Захаров В.П. Веб-пространство как языковой корпус Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции “Диалог-2005” (Звенигород, 1-6 июня 2005 г.) – М., 2005.

14. Раимназарова, Н. (2023). Формоструктурное описание диалектизмов в толковых словарях. Традиции и инновации в исследовании и преподавании языков, 1(1), 316–319.



15. Sayfullayev Najibullo, Boboniyozova Manzura. (2024). IMPORTANCE OF DIALECTS IN PLACE NAMES. International Multidisciplinary Journal for Research & Development, 11(02).

16. Недовшина Е.В. Программа для работы с корпусами текстов: обзор основных корпусных менеджеров. Работа с системой DDC Языковая инженерия: в поиске смыслов. (электронный ресурс).

17. Rahimov A. Kompyuter lingvistikasi asoslari. – T., 2011.

Internet saytlari.

1. Wikipedia.uz
2. www.natlib.uz
3. www.kutubxona.com
4. www.ziyouz.com
5. www.ziyonet.uz
6. www.kh-davron.uz
7. www.ziyouz.com kutubxonasi