

INGLIZ TILI KORPUSINING FUNKSIONAL XUSUSIYATLARI

Nazarova Lola Maqsadilla qizi

Termiz Iqtisodiyot va Servis Universiteti magistranti

ФУНКЦИОНАЛЬНЫЕ ОСОБЕННОСТИ

АНГЛИЙСКОГО КОРПУСА

Назарова Лола Мақсадилла қизи

Магистрант Термезского университета экономики и сервиса

FUNCTIONAL FEATURES OF THE ENGLISH CORPUS

Nazarova Lola Maksadilla kizi

Master's student of Termez University of Economics and Service



Tel: +998906087056

Orcid: 0009-0007-4311-

5835

Nazarova.22@bk.ru

Annotatsiya: Tillar xalqlarning ma'naviy merosini boyitishning muhim omili ekan, ularni asrab-avaylash milliy ma'naviyat rivojiga hissa qo'shish demakdir. Hozirgi kompyuterlashgan asrda tillarning yo'qolib ketishining oldini olishning asosiy yechimi tillarni raqamlashtirish va shu asosda lingvistik korpusni yaratishdir. Shuningdek, korpus tilshunosligi sohasi boshqa tillarni, jumladan, ingliz tilini mukammal o'rganish va lingvistik tahlil qilish uchun katta ahamiyatga ega. Ushbu maqolada biz til korpusining tarixi va ahamiyatini bir nechta til korpuslari misolida ko'rib chiqamiz.

Kalit so'zlar: korpus lingvistikasi, lisoniy korpuslar, BNC, COCA, elektron kutubxona.

Аннотация: Поскольку языки являются основным фактором обогащения духовного наследия народов, их сохранение означает содействие развитию национальной духовности. Основным решением по предотвращению исчезновения языков в современный компьютеризированный век является оцифровка языков и создание на их основе лингвистических корпусов. Также область корпусной лингвистики имеет большое значение для идеального изучения и лингвистического анализа других языков, включая английский. В этой статье мы рассмотрим историю и значение языковых корпусов на примере нескольких языковых корпусов.

Ключевые слова: корпусная лингвистика, лингвистические корпуса, BNC, COCA, электронная библиотека.

Abstract: Since languages are the main factor of enriching the spiritual heritage of nations, preserving them means contributing to the development of national spirituality. The main solution to prevent the disappearance of languages in today's computer age is to digitize languages and create linguistic corpora based on this. Also, the field of corpus linguistics is of great importance for the perfect study and linguistic analysis of other languages, including English. In this article, we will consider the history and importance of language corpora on the example of several language corpora.

Key words: corpus linguistics, linguistic corpora, BNC, COCA, e-library.

KIRISH (ВВЕДЕНИЕ / INTRODUCTION).

Korpus bu til birliklarining xususiyatlarini aniqlash maqsadida qidiruv dasturiga birlashtirilgan katta hajmga ega bo'lgan va tizimlashtirilgan matnlarning elektron bazasidir. Korpus lingvistikasi esa kompyuter tilshunosligining tarkibiy qismiga kiruvchi tilshunoslik bo'limi bo'lib, u ma'lum til birliklarining barcha ko'rinishlarini, belgilangan so'z bilan birikish imkoniyatiga ega so'zlar

ro'yxatini, shuningdek, til foydalanuvchilarining tildagi so'zlardan foydalanish statistikasini aniqlaydi va tahlil qiladi.

Til korpusi va elektron kutubxona tushunchalari bir biriga o'xshasha-da, tubdan farq qiladi. Elektron kutubxona bu xalqning ijtimoiy, siyosiy, ma'naviy, iqtisodiy hayotini aks ettiruvchi badiiy asarlar majmui bo'lib, uning tarkibiga kiruvchi matnlarga til nuqtayi nazaridan ishlov

berilmagan. Ya'ni, elektron kutubxonalar milliy ma'naviy merosni to'plash maqsadidagina tuziladi. Til korpusi esa tilni har tomonlama o'rganish, tahlil qilish, tadqiqotlar olib borish uchun tuziladigan, foydali va muhim ma'lumotlarni o'zida qamrab oluvchi tilning elektron bazasidir. Til korpusining yana bir muhim jihati shundaki, matnlarni tahlil qilish jarayonida, korpuslar matn tahliliga ketadigan vaqtning tejalishiga imkon yaratadi. Shu bois, har bir tilda o'zining xususiyatlarini qamrab oluvchi korpus yaratish tilshunoslar oldiga qo'yilgan muhim talabdir [2].

ADABIYOTLAR TAHLILI (АНАЛИЗ ЛИТЕРАТУРЫ / LITERATURE ANALYSIS). Jahon tilshunosligida til korpuslarini yaratishga bo'lgan qiziqishlar o'tgan asrning 60-yillarida boshlangan. Til korpusining ilk namunasi esa 1961–1964-yillarda Brown Universiteti tilshunoslari Genru Kuchera va Nelson Frensis tomonidan yaratilgan va jahonga statistik ma'lumotlarni taqdim qilgan «Computational Analysis of Present-Day American English» nomli majmuadir [4]. Ushbu majmuada Kuchera va Nelson tilshunoslik, psixologiya, statistika va sotsiologiya sohalariga oid so'zlarni va ularning joylashuvini aniqlovchi matnlarni jamladilar [6].

Korpus lingvistikasining paydo bo'lishi va taraqqiyoti haqida birinchilardan bo'lib ilmiy izlanish olib borgan tilshunos bu G.Lichdir. S.Xyuston, D.Biber, C.Grays, M.Xalliday esa zamonaviy korpus lingvistikasi va uning shakllanishiga oid tadqiqotlar olib borishdi. Tojikistonlik olimi D.Jafarova esa korpus lingvistikasini komputer lingvistikasi va boshqa o'zaro o'xshash sohalar bilan qiyoslab o'rgandi. O'zbekistonda ham bugungi kunda korpora, unga doir metodlar va ularni amaliyotda foydalanushga doir bir qator ishlar olib borilmoqda. Jumladan, J.Jumaboyev ingliz tili korpuslari, xususan, BNC korpusining semantik gradatsiyalarini topishga doir tadqiqotlar olib borgan bo'lsa, G.Ergasheva BNC korpusi asosida genderga oid konseptual metaforalarni tadqiq qildi, G.Sobirova esa BNC korpusini tillarni o'qitishda qo'llashga e'tibor qaratdi [4].

MUHOKAMA (ОБСУЖДЕНИЕ / DISCUSSION) Tilshunoslikning turli sohalarida korpus lingvistikasi metodlaridan foydalanish asosan korpusning qulayligi va qamrovi kabi amaliy jihatlari bog'liq. Shunga ko'ra mavjud korpuslarning turlariga oid malumotlar quyidagilardir:

1. Yozma korpus – ushbu korpusning keng qo'llanilishi hamda komputer va internet tarmog'inining jadal ommalashuvi tufayli yozma korpuslarni tuzish ancha osonlashdi. Natijada, adabiy va publisistik matnlarni o'z ichiga olgan, hozirda barcha uchun onlayn foydalanish imkoniyati mavjud bo'lgan rasmiy yozma nutqning katta elektron bazasi yaratildi. Misol uchun, hatto kichik milliy tillar uchun ham bir necha million so'zlarni jamlagan korpuslar mavjud. (alban, bask tillari). Yuqorida keltirilgan korpuslar so'zlashuv suhbatlarining transkriptlari va yozuvlarini taqdim etadi.

2. Yo'qolib ketish xavfi ostidagi tillardagi matnlarni o'zida jamlagan korpus. Ushbu korpusda yo'qolib ketish xavfi ostidagi tillarga oid matnlar to'planib, onlayn tarzda taqdim etilmoqda.

3. Parallel korpus. «Parallel korpus» atamasi «ko'p tili korpus» ham deb ataladi, turli tillarga tarjima qilingan matnlar to'plamidan o'z ichiga oladi va ular bo'ylab tavsiflanadi. (Masalan, Yevropa Parlamentining Parallel Korpusi). Parallel korpuslar asosan bir-biriga zid bo'lgan tilshunoslik va tarjimashunoslik sohasidagi tadqiqotlar uchun foydalaniladi [3].

NATIJARLAR (РЕЗУЛЬТАТЫ / RESULTS). Korpusga asoslangan tadqiqotlar, tilshunoslikning boshqa kichik sohalarida bo'lgani kabi, komputer tilshunosligida ham juda dinamik tushunchadir. Ushbu dinamizm komputerlashtirilgan korpuslarning o'sishini taminlaydi, shuningdek, elektron korpuslarga onlayn kirish imkoniyatlariga ega platformalar yaratadi. Yaqin kelajakda Korpus lingvistikasini ijtimoiy va siyosiy muammolarni hal qiladigan kuchga ega mustaqil fan sifatida rivojlanishi kutilmoqda.

Korpusdan foydalanish jarayonida korpusning hajmi va undagi ma'lumotlar kun sayin oshgani sababli, korpus foydalanuvchilari korpusda qayd etilgan ma'lumotlardan doimiy xabardor bo'lishlari muhimdir. Izohlarning uzluksiz ko'payishi foydalanuvchilarning korpusdagi izohlarni to'liq tushuna olmasligiga olib kelishi mumkin. Shuningdek, u o'qituvchilar va talabalar uchun korpusning pedagogik ahamiyatiga ta'sir qiladi. Ko'rib turganimizdek, izlanish muammosi ham o'rganish muammosidir [1].

Buyuk amerikalik tilshunos Chomskiyning fikriga ko'ra, korpus tilshunoslar uchun foydali vosita bo'la olmaydi, chunki tilshunos til

kompeyensiyalarini modellashturishga harakat qilishi kerak. Haquqatdan ham korpus tilshunosligi biror til elementining nima uchun qaysidir shaklda ekanligini tushuntirib bermaydi, balki qaysi manbada qay tarzda keltirilganligi haqida ma'lumot beradi. Bundan tashqari korpora bir vaqtda faqat bir tilga oid malumotlar bilan ta'minlaydi. Unday bo'lsa korpusni yaratishdan nima foyda?

Korpuslar tildagi elementlarni real kontekstlarda keltirilgan namunalar yordamida o'rganadi. Korpuslar quyidagi sifatlarni o'zida jamlashi muhimdir.

Miqdor – korpusda saqlangan til ma'lumotlari miqdori;

Sifat–korpusda ishlatiladigan matnlarning haqiqiyliigi;

Tenglik – tildan foydalanishda barcha sohalaridagi matnlarning bir tekis taqsimlanishi;

Qaytalanish – qayta foydalanish imkoniyati;

Asllik – matnlarning haqiqiyliigi

Ko'payuvchanlik – yangi matn namunalarini qo'shish bilan korpus hajmini muntazam ravishda oshirib borilishi [2].

Bilamizki, ingliz tili judayam shevalarga boy til, shu sababli ingliz tilining o'zida bir necha o'nlab korpuslar mavjud. BNC–British National Corpus va COCA–The Corpus of Contemporary American English bunga yaqqol misollardir.

English Corpora saytidagi ma'lumotlarga ko'ra, Zamonaviy Amerika ingliz tili korpusi (COCA, 560+million so'z) Britaniya Milliy korpusidan (BNC,100 million so'z) 5–6 baravar katta. Natijada, u ko'pincha BNCda mavjud bo'lmagan aktiv foydalanilmaydigan so'zlarni ham taqdim qiladi. BNC (British National Corpus) – bu ko'p foydalaniladigan ingliz tili korpusi. BNC britan tilida yozilgan 1980-1990-yillarda yozilgan matnlardan iborat bo'lib, matnlar ko'p turdagi manbalardan yig'ilgan. BNC turli matnlar, jurnallar, kitoblar, darsliklar, maqolalar va boshqa manbalardan yig'ilgan so'zlarni o'zida jamlagan korpusdir. Bu korpus stilistik tahlillar, tilning rivojlanishi va foydalanishining o'zgarishi haqida ma'lumotlar olishda ham ahamiyatlidir. BNC til o'qitish va o'rganish, lingvistik tadqiqotlar va tarixiy izlanishlarda foydalaniladi.

BNC 10% og'zaki / 90% yozma manbalarni o'z ichiga olsa, COCAda korpus og'zaki, badiiy asarlar, mashhur jurnallar, gazeta va akademik ishlar

o'rtasida deyarli teng taqsimlangan (har bir janrda 20%).

XULOSA (ЗАКЛЮЧЕНИЕ / CONCLUSION).

Tilning elektron bazasini shakllantirish har bir tilning tilshunoslari oldiga qo'yilgan muhim vazifa ekan, Korpus lingvistikasi kun sayin rivojlanadi. Sodda qilib aytganda, korpuslar tilshunoslikning istalgan sohasida tadqiqotlarni olib borishning sifatlil va samarali bo'lishiga yordam beradi. BNC til foydalanuvchilarini so'zlarning izohi va u bilan birikib keladigan so'zlar ro'yxati, janrlariga ko'ra matnlar tahlili, grafik shaklida jamlangan ma'lumotlar, tildagi neologizmlar va ularning paydo bo'lish statistikasi, ma'lum bir so'zning qo'llanilishiga oid bir qancha misollar va shu kabi bir qator foydali ma'lumotlar bilan ta'minlaydi. Shuningdek, COCA korpusi bilan taqqoslaganda so'zlar soni kam bo'lishiga qaramasdan, BNC korpusi ingliz tilidagi matnlar bazasini saqlashda muhim ahamiyat kasb etadi, chunki, bu korpusdagi so'zlarning ko'p qismi qadimgi ingliz tiliga oidligi tilshunos olimlar va tadqiqotchilarning izlanishlarida o'z isbotini topmoqda.

ADABIYOTLAR RO'YXATI

(ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА / REFERENCES)

1. Abdurakhmonova. N, Urdishev. K. «Corpus based teaching Uzbek as a foreign language». Journal of Foreign Language Teaching and Applied Linguistics. 2019.
2. Ataboyev. N. Ingliz tili korpusining funksional xususiyatlari (COCA misolida). Avtoreferat. – Toshkent. 2020.
3. Chomsky N. Current issues in linguistic theory // J.A. Fodor and J.J. Katz. The Structure Of Language. Engelwood Cliffs, New Jersey: 1964.
4. Genri Kuchera and W. Nelson Fransis. «Computational Analysis of Present–Day American English». Providence, RI:Brown University Press. 1967.
5. Gries S.Th. Introduction to S. Gries and A. Stefanowitsch (eds.). Corpora in Cognitive Linguistics: Corpus–based Approaches to Syntax and Lexis. Berlin: Heidelberg, New York: Mouton de Gruyter, 2006.
6. Halliday M.A.Teubert W., Yallop C. and Cermakova A. Lexicology and Corpus Linguistics, – London, 2004.