LINGVISTIK TADQIQOTLARDA TA’LIMIY KORPUSLARDAN FOYDALANISH

Jo‘raqulova Madina Subhonovna

O‘zMU Kompyuter lingvistikasi va

amaliy tilshunoslik kafedrasi o‘qituvchisi

Ulug‘berdiyeva Mahliyo Sunnatulla qizi

O‘zMU talabasi

Ibodullayeva Madinabonu Xusniddin qizi

O‘zMU talabasi

Annotatsiya: Ushbu maqolada ta’limiy korpus lingvistikasi sohasida olib borilgan tadqiqotlar keng tahlil qilingan. Bunda korpus lingvistikasi ta’rifi, uning dolzarbligi va zamonaviy til o‘rganishdagi samaradorligi yoritilib berilgan. Jahondagi va O‘zbekistondagi ta’limiy korpuslar, ularning imkoniyatlari, tarkibi va texnologik jihatlari batafsil o‘rganilgan. Mazkur maqola o‘quv materillarini yaratishda korpus lingvistikasining keng imkoniyatlarini namoyon etadi.

Kalit so‘zlar: korpus, ta’limiy korpus, rus tili milliy korpusi, elektron axborot bazasi, lug‘at, matn.

Аннотация: В данной статье представлен всесторонний анализ исследований в области корпусной лингвистики в образовании. Охвачено определение корпусной лингвистики, её актуальность и эффективность в современном изучении языков. Подробно изучены образовательные корпуса в мире и в Узбекистане, их возможности, состав и технологические аспекты. В

статье продемонстрирован широкий потенциал корпусной лингвистики в создании учебных материалов.

Ключевые слова: корпус, обучающий корпус, национальный корпус русского языка, электронная информационная база, словарь, текст.

Abstract: This article provides a comprehensive analysis of research in the field of educational corpus linguistics. It highlights the definition of corpus linguistics, its relevance and effectiveness in modern language learning. Educational corpora in the world and in Uzbekistan, their capabilities, composition and technological aspects are studied in detail. This article demonstrates the broad potential of corpus linguistics in creating educational materials.

Key words: corpus, educational corpus, Russian National Corpus, electronic information database, dictionary, text.

Zamonaviy tilshunoslik sohasida korpus lingvistikasi katta ahamiyatga ega. Bu katta hajmdagi matnlarni tahlil qilish orqali, tilning tuzilishi va o‘rganilishiga yordam beradi. Korpus texnologiyasida tabiiy til grammatikasini tahlil qilish, matnlardagi yangi so‘zlar va ularning ishlatilish chastotasini aniqlash, til birliklari o‘rtasidagi semantik munosabatlar, xususan, sinonim, antonim, paronim va omonimlik hodisalari tadqiqini o‘rganish mumkin bo‘ladi. Ayniqsa, korpuslardan ta’lim sohasida foydalanish o‘quv jarayonini samarali tashkil etishni ham ta’minlaydi. Ta’lim sohasida, xususan, chet tillarini o‘qitishda korpuslardan foydalanish interaktiv va moslashtirilgan metodlarni taqdim etishda muhim rol o‘ynaydi. Professor V.A.Plungyan korpus haqida so‘z yurita turib shunday deydi:.. “endi tilni o‘zlashtirish uchun odamga ikkita narsa emas, balki uchta narsa kerak: lug‘at, grammatika va til korpusi” [6, 261]. Til o‘rganuvchilarga korpus nutq, grammatika, leksikada mavjud materiallarni taqdim etadi.

Korpus atamasiga sohada tadqiqot olib borgan olimlar tomonidan turlicha ta’riflar beriladi. Jumladan, professor B.Mengliyev tomonidan korpus – ma’lum maqsadda yig‘ilgan matnlar majmuini tashkil etuvchi, til birliklari yig‘indisi, tabiiy tildagi elektron shaklda saqlanadigan yozma yoki og‘zaki, kompyuterlashtirilgan qidiruv tizimiga dasturiy ta’minot asosida joylashtirilgan onlayn yoki oflayn tizimda ishlaydigan matnlar jamlanmasi deb ta’rif beriladi [4, 3]. Korpus lingvistikasiga N.Abduraxmonova quyidagicha ta’rif beradi: “Korpus lingvistikasi – bu nutqning barcha ifodalarini kuzatish, tahlil qilish va o‘rganish imkonini beruvchi yo‘nalish bo‘lib, korpus esa tilning qisqartirilgan modeli sifatida nutqning turli diskurslardagi ifodasini tabiiy nutq holatida namoyon etuvchi tizimli matnlar majmuasidir”[1, 16].

Korpuslar turli maqsadlarda yaratiladi va tarkibida bir qancha subkorpuslarni birlashtiradi. Korpuslar tarkibida subkorpuslar sifatida mualliflik korpusi, parallel korpus, ta’limiy korpuslarni qayd etish mumkin. Jumladan, *o‘zbek tili milliy (uzbekcorpus.uz) korpusi* tarkibida mavjud ta’limiy korpus, qolaversa, Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universitetida 2021-yilda o‘zbek tilida *ta’limiy korpus* yaratishga doir dastlabki amaliy ishlar olib borish boshlandi. *(uzchoolcorpara.uz).*

Ingliz tilida ta’limiy korpus alohida korpus bo‘lib, ingliz tilini o‘rganuvchilarning xalqaro korpusi sifatida, og‘zaki ingliz tilini o‘rganish uchun *Lankaster Spoken English Corpus, Corpus of Spoken American English* korpuslarini misol qilib keltirsak bo‘ladi. N.F.Sidiropulos va N.Hatsigeorgiuning tadqiqotida ta’limda korpusdan foydalanishning afzalligi va kamchiliklari haqida fikr yuritib korpusning afzalligi uning misollar ko‘pligi, qidiruv tizimiga kirish imkoniyati hamda ko‘p sonli grammatika va til xususiyatlarini batafsil o‘rgata oladigan manba deb ko‘rsatsa, kamchiligi sifatida korpusda tildan foydalanishda kontekstning yo‘qligi va aniq tekshiruvi yo‘qligini keltiradi, chunki ta’limda tasodifiy elementlar mavjud emasligi ta’kidlab o‘tadi [9, 1-5].

Ta’limiy korpuslarning qurilishi, uning tarkibi va ishlatilishi o‘quvchilarga til o‘rgatishda qo‘llanishi ilgari o‘rganilgan korpus turlarining samaradorligi va ta’limdan foydalanishning afzalligini belgilab beradi. Sabine Braun ELISA – til o‘rganish maqsadlari uchun pedagogik jihatdan boyitilgan korpus haqidagi tadqiqotida til o‘rganish uchun mos bo‘lgan kichik va aniq maqsadli korpuslar ishlab chiqarilishi, korpusda dolzarb mavzularga oid manbalar bo‘lishi kerakligini tahlil qilib beradi [10]. John Sinclair tadqiqotining tahlil natijalari shuni ko‘rsatadiki, korpusni ta’limda qo‘llash uchun avval o‘qituvchida ko‘nikma va bilim bo‘lishi kerak. Undan keyingina o‘quvchilarga til o‘rgatishda korpusdan foydalanish kerakligi haqida va uni qanday ishlatish usullarini o‘rgata oladi. Chunki bu maqolada hozirda ko‘p talabalar, o‘quvchilar, akademik litsey o‘quvchilari til o‘qitish korpusidan bexabar ekanligi aytilgan [7, 1-3].

Rus tili milliy korpusi tarkibida ta’limiy korpus ichki korpus sifatida ishlab chiqarilgan. Uning milliy korpusdan farqi matnlarning ta’limiy korpusda o‘quvchi yoshiga mosligi bilan ajralib turadi. Rus tili milliy korpusi haqida V.V.Vinogradova ilmiy tadqiqotida Rus tili milliy korpusi uchta komponentdan iborat ekanligi haqida ma’lumot berilgan. Ularga elektron shakldagi (1) matnlar to‘plami, (2) lingvistik annotatsiya tizimi va (3) qidiruv so‘rovlarni ta’minlovchi qidiruv tizimlari kiradi. O.G.Chernyavskayaning taqdiqotida chet tilini korpus texnologiyasi bilan o‘rganishning tarixi haqida ma’lumot beriladi. Dastlab V.P.Sisoev, T.A.Chernyakova, J.Aston, P.Beyker, S.Bernardini, T.Jons va boshqalar o‘zlarining tadqiqotlarini chet tillarini o‘qitishda korpus texnologiyasidan foydalanishdagi muammolarga bag‘ishlangan edi.

T.Jons birinchi bo‘lib, 1991-yilda chet tilini o‘qitishda korpusdan foydalanishni taklif qilgan. Ushbu yondashuv *Data Driven Learning* deb ataldi. Ingliz tili fanida grammatika va lug‘at darslarida birinchi bo‘lib korpusdan foydalanildi. Tadqiqotda korpusdagi matnlar, manbalar va topshiriqlar o‘quvchi darajasiga mos bo‘lishi kerak, degan fikrlar berilgan va bir nechta ingliz tilidagi

korpuslar tahlil qilingan. Misol uchun, *The Oxford English Corpus*da hozirgacha ikki milliarddan ortiq so‘zlardan foydalanilgan. *The Corpus of Contemporary American English* korpusi esa 1990-yildan hozirgi kungacha bo‘lgan matnlarni va 445 million so‘zni o‘z ichiga oladi. *The Bank of English* korpusi 4 yarim milliard so‘zdan iborat ingliz tilidagi matnlarni o‘z ichiga oladi [5, 226].

O‘.Xoliyorovning tadqiqotida korpusda giponomik, graduonimik va partonimik munosabatlar bildiruvchi birliklar bazasini shakllantirish yuzasidan takliflar beriladi. Til korpusi – millionlab matnlarni qamrab olgan til korpuslari so‘zni izohlash, uni tildagi munosabatlari bilan birga ko‘rsata olish imkoniyatining mavjudligi bilan ajralib turadi [3, 30]. O‘zbek tili ta’limiy korpusi tuzilishining lingvistik asoslari bo‘yicha O‘.Xoliyorovning tadqiqotlari quyidagi vazifalarni o‘z ichiga oladi: o‘zbek tili ta’limiy korpusi tuzilishining nazariy masalalarini belgilash, o‘zbek tili ta’limiy korpusining lingvistik asoslarini ochib berishdan. Tadqiqotda ta’limiy korpusning ona tili ta’limi, o‘zbek tilini xorijiy til sifatida o‘qitishdagi afzalliklarini aniqlash, o‘zbek tili ta’limiy korpusiga material tanlash tamoyillarini ishlab chiqish, o‘quv lug‘atlarini ta’limiy korpusga to‘plash jarayonini aniqlash kabi ishlar amalga oshirilgan.

“Aksariyat ilmiy manbalarda dastlabki yozma korpuslar sifatida Brown (1964), LOB (1978), FROWN (1999), FLOB (1998), Kolhapur (1978), ACE (1986) manbalar qayd etilsa, nisbatan keyin yaratilgan og‘zaki korpuslar sifatida SEU, LLC, SEC (ingliz tili og‘zaki matn korpusi), Map task korpus hamda HKCSE (inglizcha og‘zaki matnlarning Gong Konk korpusi), aralash korpus (og‘zaki va yozma matnlardan tashkil topgan) sifatida ingliz tilining xalqaro ICE korpuslari e’tirof etilgan. Yuqorida nomi keltirilgan korpuslarning hajmi 1 mln. so‘zni o‘z ichiga oladi” [1, 12].

Brown korpusi – ingliz tilidagi ilk katta korpuslardan biri. U 1961-yilda AQSHning Braun universitetida Nelson Fransis va Genri Kuchere tomonidan

tuzilgan. Bu korpusning yaratilishidan dastlabki maqsad ingliz tilidagi grammatik va leksik xususiyatlarni statistik jihatdan o‘rganish bo‘lib, ilmiy tadqiqotlar va tilni o‘rganish jarayonida real hayotda qo‘llaniladigan matnlar asosida tilning sturkturasini tahlil qilish uchun yordamchi vazifasini bajarish edi. Korpus 500 ta matndan tashkil topgan bo‘lib, har biri taxminan 2000 ta so‘zdan iborat. Matnlar esa turli janrlardan olingan, misol uchun, ilmiy, badiiy, ommaviy axborot vositalari. Hozirda bu korpus zamonaviy korpuslar uchun model sifatida xizmat qiladi va tilshunoslikda qimmatli manba sifatida saqlanib qolmoqda. *LOB* korpusi – *(Lancaster-Oslo/Bergen Corpus)* Britaniya ingliz tilidagi ilk kompyuterlashtirilgan korpus bo‘lib, 1970-yilda yaratila boshlangan. Bu Brown korpusining Britaniya versiyasi. LOB korpusi Brown korpusiga o‘xshash tarzda tuzilgan bo‘lib, Britaniya ingliz tilini tadqiq qilish va Amerika ingliz tili bilan solishtirish uchun yaratilgan. *FROWN (Freiburg-Brown)* va *FLOB (Freiburg-LOB)* korpuslari o‘tgan asrning 90-yillar oxirida yaratilgan bo‘lib, zamonaviy ingliz tilini tahlil qilish va ilgari yaratilgan korpuslar bilan solishtirish uchun mo‘ljallangan. Germaniyaning Freiburg universitetida Christian Mair va uning jamoadoshlari tomonidan yaratilgan. *FROWN* Amerika ingliz tiliga asoslangan bo‘lib, uning ildizi Brown korpusiga borib taqaladi, FLOB Britaniya ingliz tiliga asoslangan bo‘lib, u esa LOB korpusiga bog‘liqdir [11, 16]. Ingliz tilida ta’limiy korpus alohida korpus bo‘lib, Gonkong Study Corpus ingliz tilini o‘rganuvchilarning yozma ishlariga alohida e’tibor qaratadi. Ushbu korpus til o‘rganuvchilarning yozish jarayonidagi xatolarini, uslublarini va til qoidalariga rioya qilish darajasini o‘rganish imkonini beradi. Bu esa ta’lim jarayonida samarali til o‘rgatish usullarini ishlab chiqishga xizmat qiladi.

Xulosa qilib shuni aytish mumkinki, maqolada ta’limiy korpus lingvistikasi sohasida olib borilgan tadqiqotlar tahlil qilinib, uning dolzarbligi va amaliy ahamiyati, ta’limiy korpusning tilni o‘qitishdagi roli, uning texnologik yondashuvlari, o‘quv jarayoniga ijobiy ta’siri ko‘rsatib berilgan, dunyodagi ilg‘or korpuslar, jumladan, ingliz, rus va o‘zbek tilidagi korpuslarning o‘ziga xos jihatlari

hamda til o‘rganishda foydalanish usullari tahlil qilingan. Tadqiqot natijalari ta’limiy korpuslarni takomillashtirish, o‘quv jarayonlarini sifatli tashkil etish va milliy til korpuslarini rivojlantirishda muhim ilmiy asos yaratadi. Shu sababli, ta’limiy korpuslar o‘quvchilarning lingvistik ko‘nikmalarini rivojlantirishda samarali vosita sifatida qo‘llanishi mumkinligi aniqlangan.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI:

1. Abduraxmonova N. Korpus linvistikasi darslik. – Toshkent, 2024. ─B.12.

2. Elov B, Abjalova M, Raupova L. O‘zbek tilining ta’limiy korpusi va uning imkoniyatlari // O‘zbekiston til va madaniyat. 2021. № 4. – 69-71 b.

3. Xoliyorov O‘. Ta’limiy korpuslar uchun giponimik, graduonimik va partonimlik munosabatlar bildiruvchi birliklar bazasini shakllantirish // Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti axboroti, 2021. № 01. – 30-b.

4. Mengliyev B, Xamroyeva Sh. Til korpusi va uning ahamiyati.– Buxoro, 2018. - 3-b. 5. Чернявская О. Дидактический потенциал корпусных технологий в преподавании иностранных языков // Ученые записки Крымского федерального университета имени В. И. Вернадского. Филологические науки. Том 2. № 2. 2016. – 226 с.

6. Марданшина Р. Hatsіональный корпус в практике лингвистических исследований и преподавании языка // Hatsіональный корпус русского языка / Н.Р. Добрушина, А.И. Левинзон // Вопросы образования, 2006. № 4. – 1-5 с.

7. Синклер Д. Как использовать корпуса в преподавании иностранного языка // How to use corpora in language teaching. – Amsterdam, 2004. № 6. – 1-3 b.

8. Холиёров Ў. Ўзбек тили таълимий корпусини тузишнинг лингвистик асослари. Филол. фан. б. фалс. док. дисс. – Термиз, 2021.

9. Sidiropulos N, Hatzigeorgiu N. Construction of a web-based corpus for educational uses // Institute for Language and Speech Processing, 2011.

10. Sabine Braun. ELISA—til o‘rganish maqsadlari uchun pedogogik jihatdan boyitilgan korpus, 2020.

11. Martin W. Practical Corpus Linguistics An Introduction to Corpus-Based Language Analysis. Oxford, 2016. – P. 16.