

KORPUS TUSHUNCHASI VA KORPUS LINGVISTIKASI HAQIDA

Saydamatova Dilafruz Xamidullo qizi,
Namangan davlat universiteti tayanch doktoranti
<https://orcid.org/0009-0006-8691-7240>

Annotatsiya. Mazkur maqola korpus tushunchasi va korpus lingvistikasi haqida qisqacha ma'lumotlar berishga bag'ishlangan. Tadqiqotda dastlab, "korpus" atamasining yuzaga kelishi uning lug'aviy ma'nosiga to'xtalib o'tilgan. Korpus lingvistikasi tillarning tabiati, tuzilishi va qo'llanilishini tavsiflovchi dalil manbai sifatida og'zaki va yozma matnlar to'plamini jamlash va tahlil qilishni o'z ichiga olishi haqida ma'lumotlar keltirib o'tilgan. Bundan tashqari, maqolada asosiy e'tibor nofilologik yo'nalishlarda sohaviy leksik kompetensiyani rivojlantirishning korpus texnologiyasini takomillashtirish masalalariga ham qaratilgan.

Kalit so'zlar: korpus, korpus lingvistikasi, xorijiy tillar, tilshunoslik, lingvistik obyektlar, nofilologik yo'nalishlar, ingliz tili, korpus texnologiyasini takomillashtirish.

О ПОНЯТИИ КОРПУСА И КОРПУСНОЙ ЛИНГВИСТИКЕ

Сайдаматова Дилафруз Хамидулло кизи
Докторант Наманганского государственного университета

Аннотация. Данная статья посвящена предоставлению кратких сведений о понятии корпуса и корпусной лингвистики. Первоначально в исследовании появление термина «корпус» было ориентировано на его словарное значение. Утверждается, что корпусная лингвистика предполагает сбор и анализ коллекций устных и письменных текстов как источника данных, описывающих природу, структуру и использование языков. Кроме того, основное внимание в статье уделяется совершенствованию корпусной технологии развития лексической компетенции в нефилологических областях.

Ключевые слова: корпус, корпусная лингвистика, иностранные языки, языкознание, лингвистические объекты, нефилологические направления, английский язык, совершенствование корпусной технологии.

ABOUT THE CONCEPT OF CORPUS AND CORPUS LINGUISTICS

Saydamatova Dilafruz,
Doctoral student at Namangan State University

Abstract. This article is dedicated to providing brief information about the concept of corpus and corpus linguistics. At first, in the study, the emergence of the term «corpus» was focused on its dictionary meaning. It is stated that corpus linguistics involves the collection and analysis of collections of spoken and written texts as a source of evidence describing the nature, structure, and use of languages. In addition, the main focus of the article is on improving the corpus technology for the development of lexical competence in non-philological fields.

Key words: corpus, corpus linguistics, foreign languages, linguistics, linguistic objects, non-philological directions, English language, improvement of corpus technology.

Kirish. Jahonda mamlakatlarning ijtimoiy-iqtisodiy sohadagi integratsiyalashuvi, hususan oliy ta'lim sohasida Baloniya jarayoniga qo'shilishi natijasida chet tillarni halqaro mezonlarda belgilangan malaka talablariga muvofiq o'qitish metod va texnologiyalarini takomillashtirish muoammasi dolzarblashmoqda. Xalqaro mexnat bozorida mutaxassislarining raqobatbardoshligi nafaqat ularning mutahassislik kompetensiyasiga balki ingliz tilidan kommunikativ kompetensiyasiga egaligi bilan ham baxolanayotganligi sababli ingliz tilini maxsus maqsadlarda o'qitish, ingliz tilini muayyan maqsadlarda o'qitish va ingliz tilini kasbiy maqsadlarda o'qitish metodikalarini takomillashtirish muhim ahamiyat kasb etmoqda. Bo'lg'usi mutahassislarni taaluqli fan sohalaridagi so'ngi yutuqlar bilan o'z vaqtida tanishib borishlari va turli tillarda so'zlashuvchi ayni kasb egalarining o'zaro fikr almashishlari uchun vositachi til sifatida ingliz tilining ahamiyati oshib bormoqda.

Nofilologik yo'nalish talabalariga ingliz tili o'qitish metodikasi Buyuk Britaniyaning Bermingam,

Oksford, Kembrij, Norvich, Nayl universitetlari, Kanadaning Toronto universiteti, Amerika Qo‘shma Shtatlarining Arizona, Baylor, Bruklin, DePaul, Kaliforniya, Chikago universitetlari, Rossiya federatsiyasining Sankt Peterburg universiteti, Kareliya davlat pedagogika universiteti, Bunin nomidagi Yelets davlat universiteti, Orlov davlat universitetlarida olib borilgan tadqiqotlarda ilmiy asoslanmoqda.

Dunyo miqyosida talabalarni ijodkorlikka yo‘naltiruvchi produktiv va izlanuvchan yondashuvlarni takomillashtirish orqali shaxs yetukligini ta‘minlash jarayonini jadallashtirish bo‘yicha ilmiy izlanishlar olib borilmoqda. Ayniqsa, shaxsga yo‘naltirilgan va muammoli ta‘lim texnologiyalari hamda SMART texnologiyalar asosida nofiloligik ta‘lim (tarix) yo‘nalishi talabalarining tanqidiy va ijodiy fikrlashni shakllantirishning innovatsion shakllari va interfaol usullarini joriy etish masalalari hamon dolzarb bo‘lib kelmoqda. Shuningdek, bo‘lajak soha mutaxassislarining xorijiy tillarda erkin muloqot yuritishida nutq ko‘nikmalarini rivojlantirishga alohida e‘tibor qaratilmoqda. Jahon tajribasida dunyo mamlakatlaridan Britaniya ta‘lim tizimida xorijiy tillarni SMART texnologiyalar bilan integratsiyasi SLE (smart learning environment), kompyuter bilan integratsiyasi CALL (Computer Assisted Language Learning), Amerikada (Web-Based Inquiry Learning), texnologiya integratsiyasi ICT (Technology in the classroom), ingliz tilini ommaviy axborot vositalari orqali o‘rganish (English-medium instruction), sohaga oid matnlar yordamida tilni integratsiyalab o‘rganish CLIL (Content and Language Integrated Learning), tilni mashq va topshiriqlar asosida o‘rganish (Task-based language learning), loyiha asosida o‘qitish (Project-based learning) kabi ta‘lim shakllarini takomillashtirish bo‘yicha olib borilayotgan ilmiy tadqiqot ishlari ustuvor xususiyatga ega bo‘lib, mazkur tajribalarni puxta o‘rganish va amaliyotga joriy qilish maqsadga muvofiq.

Yurtimizda so‘nggi yillarda ta‘lim tizimini tubdan isloh qilish, ayniqsa, chet tillarni interfaol ta‘lim metodlari va texnologiyalari yordamida o‘qitish, ta‘lim jarayonida talabalarining ijodiy-tanqidiy fikrlashini rivojlantirishga yo‘naltirilgan innovatsion usul va vositalarni amaliyotga tadbiiq etish masalasi ustuvor vazifalardan biri etib belgilandi. “Mamlakatimizda bir necha xorijiy tillarni biluvchi zamonaviy kadrlar tayyorlash, xorijiy tillar bo‘yicha ilmiy ishlar olib borish, til o‘rgatish metodologiyasini takomillashtirish lozimligi” kabi vazifalar belgilangan.

Adabiyotlartahlili va metodologiya. Mamlakatimizda chet til o‘qitish metodikasi va korpus lingvistikasi J.J.Jalolov, G.X.Bakieva, L.T.Axmedova, G.T.Maxkamova, D.Djumanova., D.Xashimova, M.Irskulov, S.A.Misirovlar tomonidan ilmiy-nazariy jihatdan tadqiq qilingan. Nofilologik yo‘nalishlarda tillarni o‘qitish muammolari F.Sh.Alimov, T.K. Sattarov, G.A. Asilova, G.N. Irmuhamedova, I.N.Ilxomovalar tomonidan o‘rganilgan. Dars jarayoniga pedagogik texnologiyalarni joriy qilish masalalari o‘zbek pedagogikasida N.Azizxo‘jaeva, Z.K.Ismailov, R.S.Musayev, A.A.Shoyusupova, N.A.Muslimov, Z.T.Rahimov, A.A.Xojayev, H.SH.Qodirov, O‘.Q.Tolipov, M.Usmonboyeva, xorijda M.Altun, W.L.Bedwell, E.Salas, S.Smith, Y.T.Sung, CH.Kuo-En, L.Tzu-Chien, K.Uluc, rus pedagogikasida I.Baydenko, T.L.Gerasimenko, A.A.Gin, N.G.Moloshonok, V.V.Serikov, V.P.Tixomirov, A.V.Xutorskiy, A.N.Shukin, kabi olimlarning tadqiqotlarida yoritilgan.

Mustaqil Davlatlar Hamdo‘stligi mamlakatlarida ingliz tilini muloqot va mutaxassislik kompetensiyasini shakllantirish vositasi sifatida o‘qitish muammolari I. V. Zaytsyeva (Зайцева И.В., 2004), O.Yu.Ivanova (Иванова О.Ю., 2005) N.A.Gazova (Газова Н.А., 2007), O.N.Fedorova (Федорова О.Н., 2007), V.V.Tarasenko (Тарасенко В.В., 2008), T.M.Tatarina (Татарина Т.М., 2009) I. V. Nuja (Нужа И.В. 2010), T.B.Vepreva (Вепрева Т.Б., 2012), Gryaznova M.A. (Грязнова М.А., 2012) lar tomonidan tadqiq etilgan.

Mamlakatimizda chet tillarni turli yo‘nalish talabalariga o‘qitish muammolari D.M.Isroilova (Nofilologik ta‘lim yonalishlarida kasbga yo‘naltirilgan ingliz tilini fanlararo bog‘liqlikda o‘qitish metodikasi), N.H.Kushiyeva (Talabalarga ingliz tili o‘qitishda klasterli yondashuv xususiyatlari (qishloq xojaligi ta‘lim yo‘nalishlari misolida)), H.F.Maqsudova (Farmasevtika yo‘nalishida ingliz tili o‘qitish jarayonida talabalarining kasbiy layoqatlarini rivojlantirish), F.S.Artikova (Huquqshunoslik yo‘nalishi talabalariga ingliz tili sohaviy leksikasini o‘qitish metodikasi), G.O.Yaxshiboeva (Farmatsevtikaga oid terminlarni o‘rgatishda o‘qish va yozma nutq ko‘nikmalarini rivojlantirishning zamonaviy metodlari (ingliz tili misolida)) va boshqalar tomonidan tadqiqotlar olib borilgan.

Tadqiqot jarayonida muammoga doir psixologik, pedagogik va metodik adabiyotlarni o‘rganish va qiyosiy tahlil etish, ilg‘or tajribalarni o‘rganish, pedagogik kuzatuv, pedagogik eksperiment, modellashtirish, loyihalashtirilgan metodikalar, suhbat, intervyu, matematik-statistik tahlil etish va

natijalarni umumlashtirish va grafik tasvirlash kabi usullaridan foydalanildi.

Muhokama. Mavzuning dolzarbligi quyidagi omillar bilan belgilangan: birinchidan, ushbu masala yuzasidan hozirgacha nufuzli ilmiy-pedagogik tadqiqot olib borilgan emas; ikkinchidan, umumiy o'rta ta'lim maktablari boshlang'ich sinf o'quvchilariga ingliz tilida og'zaki nutqni (tinglab tushunish va gapirish) o'rgatish bilan bog'liq nimani, qachon va qanday o'qitish masalasi, shuningdek chet til ta'limining texnologiya, vosita, usul, yondashuv, va metodlari ta'rif-tavsif etilmagan; uchinchidan, gapirish nutq faoliyat turining umumiy o'rta ta'lim maqomidagi xususiyati metodik nuqtai-nazardan yoritilmagan; to'rtinchidan, umumiy o'rta ta'lim maktablari boshlang'ich sinflarida chet tilda muloqot qilishni o'rgatishda leksika, fonetika va grammatikaning mavqei aniqlanmagan; beshinchidan O'zbekiston sharoitida ya'ni ingliz tili tabiiy muhiti mavjud bo'lmagan vaziyatda, umumiy o'rta ta'lim maktab sharoitida o'quvchilarining psixofiziologik-tipologik jihatlari, aqliy faoliyat yuritish qobiliyatlari, hayot va til tajribalari va o'qishga bo'lgan munosabatlari bilan til o'rganish/o'rgatish jarayoniga o'ziga xos yondashuvni talab etilishi maxsus tadqiq etilmagan; oltinchidan boshlang'ich sinf o'quvchilarining ingliz tilida muloqot ko'nikmalarini (communicative competence) shakllantirishda jismoniy harakatli kommunikativ o'yinlarning mazmun-mohiyati va ahamiyati borasida tadqiqot ishlari bajarilmagan; yettinchidan bolalarda ona tilida nutq shakllanishining asosiy omillari va davrlari ularning chet tilni o'zlashtirishlari bilan qiyosiy tadqiq etilmagan.

Shu o'rinda korpus tushunchasi haqida, korpus katta hajmli va tizimga solingan matnlar to'plamidan iborat til manbasi hisoblanadi. Korpus tilshunosligida ular ma'lum bir til doirasida yoki tilning ma'lum bo'limida statistik tahlillarni amalga oshirish, qarashlarni, tildagi hodisalar yoki nazariy qoidalarni tekshirish uchun foydalaniladi. Korpus bir til yoki bir necha tildagi matnli ma'lumotlardan iborat bo'lishi mumkin. Korpus deganda, odatda, matnli korpus tushunchasi anglanadi, lekin hozirgi kunda korpuslar faqatgina matnlardan iborat bo'lmay qoldi. Shuning uchun korpus so'zi o'rniga matnli korpus tushunchasini ishlatamiz. Tilga oid tadqiqotlarni olib borishni yanada samaraliroq qilish uchun korpuslar annotatsiyalanadi [1, 2]. Masalan, korpusni annotatsiyalashning bir turi bu so'zlarni teglash hisoblanadi (POS-tagging). Bunda so'zning turkumi va shu turkum kategoriyalari asosida teglab chiqish tushuniladi. Ya'ni kitoblarga so'zi quyidagi ma'lumotlarni tashiydi: ot, ko'plik, jo'nalish kelishigi. Ayni mana shu ma'lumotlar teglar orqali so'zga biriktirib chiqiladi. Annotatsiyalashning yana bir ko'rinishi o'zaklash (lemmatizatsiya) bo'lib, u so'zning tayanch shaklini ko'rsatib berish hisoblanadi. Masalan, kitoblar, kitobning, kitobga so'zi uchta shaklda turibdi lekin ularning asosi bir xil – kitob. Ana shu lemmatizatsiya (o'zaklash) deb ataladi. Bu yerda asos hamda o'zak tushunchalarini adashtirib yubormaslik kerak. Masalan, bostirma so'zi bostir+ma shaklida yasalgan, lekin uni o'zaklashda bostir so'zini lemma deb qaray olmaymiz, bostirma yagona so'z hisoblanadi. Agar, bostirmada, bostirmaga, bostirmaning so'zlarini o'zaklash kerak bo'lsa unda bostirma so'zini olishimiz to'g'ri bo'ladi. Sodda qilib aytganda, lemma so'zning shakl yasovchi qo'shimchalari tushirib qoldirildigan qismi hisoblanadi. Sodda qilib aytganda, korpus tilshunoslikning istalgan sohasida tadqiqotlarni olib borishning sifatli va samarali bo'lishiga yordam beradi.

Korpus lingvistikasi tillarning tabiati, tuzilishi va qo'llanilishini tavsiflovchi dalil manbai sifatida og'zaki va yozma matnlar to'plamini jamlash va tahlil qilishni o'z ichiga oladi. Bu ish, odatda, muayyan kontekstlarda lingvistik obyektlar yoki jarayonlar sodir bo'lish ehtimoli haqidagi ma'lumotlarni o'z ichiga olgan holda tillarni tavsiflash uchun miqdoriy o'lchovni olib keladi. Hozirgi korpus lingvistikasi tilning barcha darajalarini, jumladan, fonologiya, leksika, grammatika va nutqni o'rganuvchi, lingvistik tavsiflash uchun empirik asos sifatida matnlardan foydalanishning uzoq an'alaridan kelib chiqdi. Korpus lingvistikasi ma'lum bir so'zlarni saqlashga moyil bo'lgan kompaniyaga va til foydalanuvchilari o'zlarini odatda ifodalash usullariga e'tibor qaratgan holda, tillarning turli tarixiy, mintaqaviy va sotsiolingvistik kontekstlarda, janrlarda va registrlarda tizimli ravishda o'zgarishiga yangi yoritib berdi. Tillarning ehtimollik tavsiflari tilshunoslar tomonidan qo'llaniladigan boshqa metodologiyalarni to'ldirishi mumkin va lingvistik tavsifdan tashqari, bir qator sohalarda ishlash uchun ta'sir ko'rsatishi mumkin. Bularga tabiiy tilni qayta ishlash, til ta'limi va kognitiv lingvistika kiradi [3].

Korpus tilshunosligi atamasi umuman korpusga asoslangan lingvistik tadqiqotlarni anglatadi. Arxetip korpus ishi zamonaviy raqamli davrdan ancha oldin mavjud bo'lgan. Hozirgi vaqtda korpus tilshunosligida hisoblash va statistik vositalardan foydalanish hal qiluvchi rol o'ynasa-da, korpus lingvistikasida ikkita muhim element – tilshunoslik masalalarini o'rganish va korpuslardan foydalanish ekanligini ta'kidlash muhimdir.

Birinchi zamonaviy korpus 1960-yillarda Braun universitetida tuzilgan hozirgi Amerika ingliz tilining Braun universiteti standart korpusi (“Jigarrang korpus”; Frensis va Kučera, 1979; Kučera va Frensis, 1967). Muvozanatli, turli xil manbalardan to‘plangan taxminan bir million so‘zli matn bilan Jigarrang korpus korpus hajmi va korpus dizayni nuqtai nazaridan yutuq bo‘ldi. Ammo keyingi yillarda korpus hajmining standarti tez sur‘atlar bilan o‘tdi. Korpus tilshunosligi nisbatan qisqaroq tarixga ega, chunki u mavjud deb hisoblanadi kompyuter texnologiyalarining rivojlanishi bilan paydo bo‘ldi.

Amerika Milliy Korpusi (AMK) – 1990 yildan beri ishlab chiqarilgan 22 million so‘z yozma va og‘zaki ma‘lumotlarni o‘z ichiga olgan amerikacha ingliz tilidagi matn korpusidir. Hozirda AMK bir qator janrlarni o‘z ichiga oladi, jumladan elektron pochta, tvitlar va veb-ma‘lumotlar. Britaniya Milliy Korpusi kabi oldingi korpuslarga kiritilmagan. U nutq va lemma, sayoz tahlil va nomli shaxslar uchun izohlanadi.

Britaniya Milliy Korpusi (BMK) – bu XX asrning keyingi qismidagi Britaniya ingliz tilining yozma va og‘zaki tilining 100 million so‘zdan iborat to‘plami. Britaniya Milliy Korpusi kattaroq yozma qismdan (90%, masalan, gazetalar, akademik kitoblar, xatlar, insholar va boshqalar) va kichikroq og‘zaki qismdan (qolgan 10%, masalan, norasmiy suhbatlar, radio ko‘rsatuvlar va boshqalar) iborat. Og‘zaki qism audio formatda ham mavjud.

Xulosa. Oliy ta‘lim muassasalarida chet tillarni kasbiy yo‘naltirilgan holda o‘qitishga masul kafedralar doimiy duch keladigan bir qator muammolar to‘sqinlik qilmoqda. Ulardan eng muhimlari quyidagilardan iborat: OTMlarning fakultetlararo chet til kafedralari va mutaxassis tayyorlovchi kafedralari o‘rtasida fanlararo muvofiqlashtirish yo‘qligi sababli tilga oid bo‘lmagan oliy ta‘lim muassasalari talabalari uchun ular egallagan mutaxassislik malaka talablariga muvofiq chet tilini o‘rgatish masalalari ishlab chiqilmagan; Talabalarining chet til bo‘yicha egallagan kommunikativ kompetensiyalarini mutahassislik sohasidagi amaliy muammolarni hal qilish uchun ishlata olmasliklari.

FOYDALANILGAN ADABIYOTLAR:

1. Abjalova M. Tahlil va tahrir qiluvchi dasturning lingvistik modullari: Monografiya. – Toshkent: Nodirabegim, 2020. – 176 b.
2. Xamroyeva Sh. O‘zbek tili mualliflik korpusini tuzishning lingvistik asoslari: Monografiya. – Globe edit, 2020. – 260 b.
3. Козлова Н.В. Лингвистические корпуса: определение основных понятий и типология / Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2013. Том 11, Выпуск 1. – С. 79-88.
4. Захаров В.П., Богданова С.Ю. Корпусная лингвистика. – Иркутск: ИГЛУ, 2013. – 161 с.
5. http://en.wikipedia.org/wiki/Korpusnaya_lingvistika
6. <https://myfilology.ru>