

MUALLIFLIK KORPUSI KENG QAMROVLI QIDIRUV TIZIMI

Akramova Shohista Islom qizi

akramovashohista88@gmail.com

ToshDO‘TAU doktoranti

Annotatsiya. Ushbu maqolada mualliflik korpuslari muallif asarlari tilini to‘liq va aniq ko‘rasata olishi bilan boshqa axborot banklaridan ajralib turishi haqida ma‘lumotlar keltirilgan. Korpusning dunyodagi ko‘plab sohalarda ma‘lumotlarning keng va to‘liq manbayi vazifasini o‘tay olishi bilan amaliy ahamiyati oshib borayotgani va korpuslarsiz raqamli texnologiyalarga asoslangan axborot asrida ta‘lim jarayonini tasavvur qilish juda qiyinligi izohlangan.

Abstract. This article provides information on how author corpora differ from other information banks by being able to fully and clearly display the language of author's works. It is explained that the practical importance of the corpus is increasing as it can serve as a wide and complete source of information in many areas of the world, and it is very difficult to imagine the educational process in the information age based on digital technologies without the corpus

Аннотация. В данной статье представлена информация о том, чем авторские корпуса отличаются от других информационных банков возможностью полного и четкого отображения языка авторских произведений. Поясняется, что практическая значимость корпуса возрастает, поскольку он может служить широким и полным источником информации во многих регионах мира, а образовательный процесс в информационную эпоху, основанный на цифровых технологиях, очень сложно представить без корпус

Kalit so‘zlar: *korpus lingvistikasi, mualliflik korpusi, ta‘limiy ahamiyat, milliy korpus.*

Axborot asriga kelib, korpus lingvistikasi jahon tilshunosligining shiddat bilan rivojlanayotgan sohasiga, korpusning o‘zi esa tilshunoslikning ish quroliga aylandi. Tilning milliy korpusi axborot manbayi sifatida quyidagi qulayliklarni yaratdi:

1) korpuslashgan tilda yaratilgan og‘zaki, yozma yodgorliklar, milliy, madaniy meros namunalari elektron shaklda internet tarmog‘idan joy oladi;

2) korpus tabiiy tilning elektron shakldagi, qidiruv dasturiga joylashtirilgan matnlar yig‘indisi bo‘lganligidan bir marta tuzilib, mukammal razmetkalangan korpus lingvistik tadqiqotlar samaradorligini ta‘minlashda baza vazifasini bajaradi;

3) korpus elektron kutubxona, lug‘at, grammatikalar yaratish uchun asos bo‘ladi. U keng ko‘lamli bo‘lganligi uchun ma‘lumotning o‘ziga xosligini kafolatlab, til hodisalarining barcha qirralarini to‘liq namoyish etishni kafolatlaydi;

4) turli ma‘lumotlar til korpusida o‘zining tabiiy kontekstual shaklida joylashadi, bu esa ularni har tomonlama, chuqur, obyektiv o‘rganishga asos bo‘ladi;

5) korpus tilni tadqiq etish (soʻzning oʻzgarishi, istorizm, neologizm vujudga kelishi, maʼno torayishi, maʼno kengayishi; yangi frazeologizmlarning paydo boʻlishini kuzatish), til oʻrganish, lugʻat tuzishda eng zamonaviy, keng imkoniyatli dasturlashtirilgan tizimdir

Korpus lingvistikasi XX asrning 60-yillarida kompyuter lingvistikasining boʻlimi sifatida shakllangan soha boʻlib, kompyuter texnologiyalaridan foydalangan holda lingvistik korpuslarni (matnlar korpuslarini) qurish va ulardan foydalanishning umumiy tamoyillarini ishlab chiqish bilan shugʻullanadi. Lingvistik (leksik) korpus yoxud til (matnlar) korpusi muayyan lingvistik muammolarni hal qilishga moʻljallangan, maxsus saralangan va teglangan elektron shakldagi matnlar majmuasi hisoblanadi va pragmatik ahamiyatga ega yirik elektron baza sanaladi.[Abduraxmonova, 2021]

Oʻtgan asrning 40-yillarida Blumfeld, Fritz va Bonderslar tomonidan korpus lingvistikasi boʻyicha tadqiqotlar boshlangan boʻlib, Nils Frensis va Genri Kuchera birinchi boʻlib korpus yaratish tamoyillarini ishlab chiqdi. Korpus lingvistikasi termini birinchi marta 1963-yilda ingliz tilidagi matnlar asosida yaratilgan ilk korpus – Braun korpusining yaratilishi bilan qoʻllanila boshlandi. Rossiyada bu termin 1996-yilda Sidney Grinbaumning maʼruzalari orqali maʼlum boʻldi. XX asrning 90-yillariga kelib korpus lingvistikasiga oid ilmiy tushunchalar mukammalroq shakllandi, bu yangi soha haqida manbalarda bir-biriga yaqin taʼriflar berila boshlandi. V.Zaxarov til korpusiga elektron shakldagi katta hajmda tizimlashtirilgan (muayyan strukturaga ega), lingvistik izohlar – maxsus elektron belgilar qoʻyilgan, filologik jihatdan keng qamrovli va muayyan lingvistik masalalarni hal etishga moʻljallangan til maʼlumotlari majmuyidir, deya taʼrif beradi. [Ataboyev, 2019]

N.Kozlova esa korpus oddiy matnlardan raqamlashtirilganligi bilan farqlanib, yozma va ogʻzaki shakldagi matnlar yigʻindisi hisoblanishi hamda maxsus lingvistik izohlar qoʻyilgan holatda kompyuterga joylashtirilgan matnlar yigʻindisi ekanligi hamda maʼlumotlarni tartibga soladigan annotatsiyalardan iboratligini aytadi. Korpus kiritilgan matnlarning oʻziga xos jihatlari toʻgʻrisidagi muhim qoʻshimcha maʼlumotlarni annotatsiya yoki lingvistik izoh koʻrinishida saqlaydi.

Lingvistik izoh (rus tilida pazметка, ingliz tilida tag) – korpusning asosiy bosh maʼlumotnomasi boʻlib, u korpusni zamonaviy internet tarmogʻida mavjud boʻlgan oddiy matnlar toʻplami (yoki kutubxonalar)dan ajratib turadigan, masalan, Ziyonet kutubxonasi, Hilol virtual kutubxonasi, Maksim Morshkov kutubxonasi yoki rus virtual kutubxonasi kabilardan farqlanadigan matnlardir. Korpuslar umumiylikiga koʻra umumiy va yakka muallifli boʻladi. Umumiy korpuslar muayyan tilning bir yoki bir necha masalalari boʻyicha yaratiladi. Korpuslar umumiylikiga koʻra umumiy va yakka muallifli boʻladi. Umumiy korpuslar muayyan tilning bir yoki bir necha masalalari boʻyicha yaratiladi. Unda matnlarning grammatik yoxud semantik teglari berilishi, parallel tarjimai taqdim etilishi, matnlar manbasi haqida metamaʼlumotlar koʻrsatilishi mumkin. Yakka muallifli korpuslar muayyan

ijodkorga mansub materiallarni qamrab olishi hamda jahon ilmida korpusshunosligida yaratilgan mualliflik korpuslari o'rganildi. Mualliflik korpusi muayyan bir muallif qalamiga mansub matnlar majmuasidan iborat tizim bo'lib, asosan, yozuvchining yozish uslubi, asarlaridagi til, matn xususiyatlarini axborot texnologiyalari yordamida semantik tasniflab, tadqiq etishga asoslangan elektron shakldagi ma'lumotlar bazasi sanaladi. Mualliflik korpusining boshqa korpuslardan farqi uning bazasida muayyan bir muallif asarlari jamlanadi, matnlar qayta ishlanadi, grammatik va semantik teglanadi, qidiruv tizimi orqali muallifga tegishli materiallardan zaruriy ma'lumotlar, lingvistik ifodalar aniqlanadi, matnlar asosida statistik ma'lumotlar taqdim etiladi, metama'lumotlar aniq beriladi. Shunday qilib mualliflik korpusi keng imkoniyatli qidiruv tizimiga ega bo'lib, muallif ijod qilgan barcha tur va janrdagi asarlar qamrab olinadigan, maxsus parametrlar asosida ham qidiruv imkoniyatiga ega, hajmi chegaralanmagan, muallif va uning ijodi bilan bog'liq manbalarni qulay va tezkor olish mumkin bo'lgan elektron ma'lumotlar bazasi hisoblanadi. Jahon tilshunosligida hozirgi kunga qadar mualliflik manbalarini yaratish borasidagi tadqiqotlar amalga oshirilmoqda. Xususan, 1956-1961-yillarda Pushkin tilining lug'ati, shuningdek, alohida asarlarning lug'atlari, masalan, 1978-yilda Lev Tolstoyning “Urush va tinchlik” romanining chastotali lug'ati nashr etildi. 1981-yilda R.Qo'ng'urov va S.Karimovlar tomonidan “Zulfiya poeziyasining lug'ati. Konkordans” tuzilgan. 1986-yilda esa S.Rizayev va N.Bo'ronovlar tomonidan “A.Qahhor “Sinchalak” povesti tilining chastotali lug'ati” yaratilgan. Bunday tadqiqotlar, shubhasiz, dolzarb hisoblanib, hozirgi kunda yozuvchilarning zamonaviy mualliflik korpuslarini yaratish borasidagi izlanishlarning shakllanishiga zamin bo'lgan.

Umuman, mualliflik korpuslarini yaratish asosi mualliflik lug'atlarining tuzilishi bilan bog'liq va MK yaratish ishlari shakllanishi o'z tendensiyasiga ega. Buni quyidagicha aks ettirish mumkin:

Rus adiblarining mualliflik korpuslari bo'yicha izlanishlar nafaqat rus tilshunosligida, balki jahon korpusshunosligida shiddat bilan rivojlanib, uning taraqqiyoti natijasida juda ko'plab ijodkorlarning elektron lug'at va korpuslari yaratildi. Mualliflik korpuslarining yaratilishi bir qancha yozuvchi, shoir tili konkordansini tuzish, milliy korpusda ularning asarlari bazasini yaratish imkonini berdi. Bu borada rus korpus lingvistikasida amalga oshirilgan ishlar va erishilgan yutuqlar diqqatga sazovor.

Mualliflik korpusshunosligining asosiy vazifasi va mualliflik korpuslaridan foydalanish sabablari quyidagicha izohlandi:

- muayyan ijodkor shaxsini o'rganish;
- adib uslubini tadqiq etish;
- muayyan shoir ijodini lingvopoetik tahlil qilish;
- ijodkorning so'z qo'llash imkoniyati va mahoratini tadqiq etish;
- mualliflik lug'atlarini yaratish;
- mualliflik iboralarini jamlash;

– MKdagi ijodkor shaxsi va uslubini namoyon etuvchi parametrlar orqali anonim asarlar muallifini topish;

– mualliflik parafraza, parema, hikmatlarini jamlash; tasviriy ifodalarning qo'llanilish ko'lamini ijodkor kontekstidan aniqlash mumkin bo'ladi. Ma'lum bir mualliflik korpusining tarixiy, lingvistik, ta'limiy, ijtimoiy, tarbiyaviy ahamiyatlari yoritiladi. Mualliflik korpusida muallif tilini to'liq, batafsil, obyektiv ko'rsatish imkoniyati bor, shuning uchun bunday korpuslar boshqa axborot bankidan afzalligi bilan ajralib turadi. Korpus turli tadqiqotlar uchun asos, manba, vosita vazifasini o'tay oladi. Bunday korpuslarning afzalligi yana shundaki, ular yordamida nafaqat biron bir so'z yoki jumlaning, balki butun boshli asar tillarini bilish imkoniyati mavjud. Mualliflik korpuslaridagi ma'lumotlarning ilmiy manbalar asosida tahrir qilinganligi, unda berilgan ma'lumotlarning aniqliligi va ishonchliligini kafolatlaydi hamda lingvistik hodisalarning butun spektrini to'liq aks ettirishi bilan ularni har tomonlama va obyektiv o'rganishga imkon beradi. MK ta'lim jarayonida muallif asarlarining ta'limiy, ijtimoiy, tarixiy, lingvodidaktik, ahamiyatini yoritib berishda, muallif asarlarida qo'llanilgan so'zlarning dialektal xususiyatlarini tasniflashda, Milliy tilning semantik xususiyatlarini o'rganishda muhim sanaladi. XV asr til uslubini bugungi tilga qiyosan o'quvchilarga o'rgatish mumkin. Badiiy matnlarning statistik tahlili natijasida matnda tez-tez qo'llaniluvchi til birliklari (otlar, sifatlar, kalit so'zlar, fe'llar, grammatik shakllar, jumla qurilishi, bir so'z bilan aytganda, yozuvchining uslubini ko'rsatuvchi vositalar) lingvostatistik tahlil yordamida aniqlanadi. Turli matnlardan olingan dalillarning qiyosiy tahlili bizga o'sha matnning mazmun-mundarijasini, matn tuzilgan davrni, dalillarni argumentlash xarakterini va hatto mualliflikni aniqlashga imkon beradi. Mualliflik korpuslari o'quvchilarning izlanuvchanlik salohiyatini o'stirib, kichik tadqiqotlarni ishonchli dalillar asosida bajarishga imkon beradi. Ta'kidlash joizki, mualliflik korpuslarining ijtimoiy ahamiyati – keng qamrovli hisoblanib, undan lingvistik, etnopsixolingvistik tadqiqotlarda, ona tili, adabiyot, xorijiy til ta'limida, matnga avtomatik ishlov berish, tarjima dasturlarini tuzishda foydalanish mumkin. Mualliflik korpusi orqali Ma'naviy merosimizni asrab-avaylash, ajdodlar ruhini ulug'lash ko'nikmalarini shakllantirish, mutolaa madaniyatini oshirish natijasida teran fikrga ega bo'lish, tafakkurni kengaytirishga erishiladi.

Mualliflik korpuslari ta'lim jarayonida muallif asarlarining ta'limiy, ijtimoiy, tarixiy, lingvodidaktik, ahamiyatini yoritib berishda, muallif asarlarida qo'llanilgan so'zlarning dialektal xususiyatlarini tasniflashda, milliy tilning semantik xususiyatlarini o'rganishda muhim sanaladi. Mualliflik korpusi materialidan ta'limiy maqsadda foydalanish o'ta ahamiyatli masala hisoblanadi. Darhaqiqat, til korpuslari ta'lim jarayonida eng muhim o'quv quroli. Sababi mavzulashtirilgan matn ta'lim jarayonida o'quv materialini tayyorlashda qulay axborot manbayi vazifasini bajaradi. Lingvodidaktikada mualliflik korpusining ahamiyati shundan iboratki, o'qituvchi uchun korpus tengsiz xazina, undan o'quv mashg'uloti uchun serqirra, mazmunli o'quv materialini tayyorlashi mumkin.



Mualliflik korpuslari o'quvchilarning izlanuvchanlik salohiyatini o'stirib, kichik tadqiqotlarni ishonchli dalillar asosida bajarishga imkoniyat beradi. Ta'kidlash joizki, mualliflik korpuslarining ijtimoiy ahamiyati – keng qamrovli hisoblanib, undan lingvistik, etnopsixolingvistik tadqiqotlarda, ona tili, adabiyot, xorijiy til ta'limida, matnga avtomatik ishlov berish, tarjima dasturlarini tuzishda foydalanish mumkin. Mualliflik korpuslaridan – "ish quroli – so'z" bo'lgan barcha soha vakillari: tilshunos, tarjimon, o'qituvchi, dasturchi, jurnalist, muharrir kundalik faoliyati jarayonida keng foydalanishlari mumkin[Hamroyeva, 2020].

Foydalanilgan adabiyotlar:

1. Abduraxmonova N. O'zbek tili elektron korpusining kompyuter modellari. Monografiya/ Abduraxmonova N. –Toshkent, 2021.
2. Ataboyev N.B. Korpus lingvistikasining asosiy xususiyatlari // Ўзбекистонда хорижий тиллар илмий-методик электрон журнал journal.fledu.uz №2/2019
3. Hamroyeva Sh.M. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari. Monografiya. – Toshkent, 2020.