

## MUALLIFLIK KORPUSINI LOYIHALASH (ERKIN VOHIDOV SHE'RIYATI MISOLIDA)

**Abdurahmanova Muqaddas Tursunaliyevna**

**\*O'zbekiston Milliy universiteti dotsenti, f.f.n.**

**Arislanova Muxlisa Zakir qizi**

**\*O'zbekiston Milliy teleradiokompaniyasi "Yoshlar" teleradiokanali katta muharriri**

**<https://doi.org/10.5281/zenodo.13923444>**

**Annotatsiya:** maqolada amaliy o'zbek tilshunosligining muhim sohasidan biriga aylangan, korpus lingvistikasi tamoyillari hamda mualliflik korpusini loyihalash va tuzish (Erkin Vohidov asarlari asosida) jarayoni haqida so'z boradi. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari keltiriladi. Erkin Vohidov asarlari korpusini tuzishning muhim bosqichlari sanab o'tiladi.

**Kalit so'zlar:** mualliflik korpusi, janriy-mavzuviy tuzilish, chastotali lug'at, kompyuterlashtirish, korpus interfeysi, razmetkalash.

## CRITERIA FOR CREATING A LINGUISTIC CORPUS OF ERKIN VAHIDOV'S WORKS

**Annotation:** The article discusses the principles of corpus linguistics, which has become one of the most important areas of applied Uzbek linguistics, as well as the process of designing and composing the author's corpus (based on the works of Erkin Vahidov). The linguistic bases of creation of the Uzbek language author's corpus are given. Important steps in the formation of the corpus of works of Erkin Vahidov are given.

**Keywords:** authorshipcase, genre-themed structure, frequency dictionary, computerization, case interface, marking.

Bugungi axborotlashgan asrda tabiiy tilni uning bor boyligi va imkoniyatlari bilan birga saqlab qolishga harakat qilinmoqda. Buning birgina ishonchli yo'li, tabiiy tilni raqamlashtirish va shu asosda milliy til korpusini yaratishdir.

Korpusning ahamiyati uning bir tildagi keng ko'lamli matnlarni bir joyga yig'ib, tartib berilganida emas, shu sababli uni tuzishda bir necha mezon asosida ish ko'riladi. Korpus materialini saralashda korpusning asosiy birligi nimadan iboratligi, uning hajmi qanday bo'lishi (unda qancha so'z bo'lgani ma'qul), yozma matn qaysi manbaga asoslanishi, qancha miqdorda bo'lishi, unga kiruvchi matn tilning qaysi sohasiga tegishli bo'lishi kabi masalalarga yechim topiladi.<sup>1</sup> Ushbu savolning ilk javobi 1965-80 yillar oralig'ida R.G.Piatrovskiy, uning shogirdlari tomonidan berilgan edi. Ular chastotali lug'at, lingvostatistik tadqiqot o'tkazish uchun matn tanlash tamoyillarini tuzib chiqishgan.

Milliy til korpusi o'zida bir qancha yo'nalishlardagi korpuslarni birlashtiradi. Bulardan biri, mualliflik korpusidir.

Mukammal tuzilgan, chuqur razmetkalangan mualliflik korpusi leksikologiya, leksikografiya, til tarixi, dialektologiya, sotsiolingvistika kabi qator sohalarning rivojiga hissa

<sup>1</sup>[Shayxislamov, N. (2020). Til korpusi yaratishning umumiy tamoyillari xususida *Prospects of Development of Science and Education*, 1(2), 247-249]

qo'shadi. Mualliflik korpusi yordamida ma'lum yozuvchi yoki shoir tili konkordansini tuzish, milliy korpus, ma'lumotlar bazasini yaratish mumkin.<sup>2</sup>

Mualliflik korpusi tuzishning umumiy tamoyillari asosida Erkin Vohidov asarlari korpusi tuzish o'zbek tili hamda E.Vohidov ijodining o'ziga xos jihatlariga alohida e'tibor qilishni talab etadi. Mualliflik korpusi tarkibiy qismlarining to'g'ri joylashtirilishi, korpus mazmuni, materiali, razmetkasi hal qiluvchi omillar bo'lganligi uchun Erkin Vohidov korpusini tuzishning muhim bosqichlari:

- korpus interfeysini loyihalash;
- korpusga material tanlash, uni qayta ishlash;
- korpus razmetkasi tamoyillarini ishlab chiqish;
- morfologik, semantik, sintaktik razmetka teglarini modellashtirish;
- dasturiy ta'minot yaratishdan iborat bo'ladi.

Erkin Vohidov asarlari korpusini tuzish uchun morfologik razmetkalash tamoyillarini ishlab chiqish lozim, chunki o'zbek tilini tushunadigan – avtomatik razmetkalay oladigan dastur mavjud emas. Internetdagi oddiy elektron fayldan korpusning farqi razmetkaning mavjudligidadir. Bu muallif asari yoki uning parchalariga lingvistik, ekstralingvistik sharh qo'shish bilan amalga oshiriladi. Agar muallif asari to'liq razmetkalanagan bo'lsa, keyinchalik bu korpusdan faqat elektron asarlar yig'indisi sifatida emas, balki turli lingvistik tadqiqotlar olib borishda foydalanish mumkin bo'ladi. korpus tarkibidagi matnga quyidagi qo'shimcha ma'lumot(teg)lar yozish mumkin. asar yozilish vaqti, janri xususiyati to'liq asarga nisbatan yoziladi, bu ekstralingvistik razmetkadir. Asar tarkibidagi har bir gapga uning sintaktik strukturasi haqida ma'lumot beriladi, matndagi leksemalarga esa alohida morfologik belgilar qo'yiladi.

Korpus interfeysi dizayni, tuzilishi, qidiruv oynalari, ularning tarkibi, korpus materiali korpus lingvistikasi mutaxassisi tomonidan, dasturiy ta'minot esa dasturchi tomonidan tayyorlanadi. Erkin Vohidov korpusiga material tanlash ham muhim masala. Korpus matnlarining razmetkalanishi E.Vohidov asarlari matni ustida turli lingvistik amallarni bajarish, adabiyotshunoslik, tarix, etnografiya, lingvomadaniyatshunoslik, lingvoma'naviyatshunoslikka oid tadqiqot olib borishda ko'p funksiyali axborot manbai sifatida xizmat qila oladi. Buning uchun korpusning asosiy bazasi tuzilayotgan vaqtda to'g'ri rejalashtirilgan bo'lishi lozim.

Mualliflik korpusini yaratish bir muncha qiyin bo'lganligi sabab, jahon tilshunosligida yaratilgan milliy korpuslar tarkibidan bunday korpusning o'rin olishi ko'p vaqt talab etdi. She'riy korpusni yaratishdagi birinchi qadam she'rning rasmiy parametrlariga asoslangan batafsil belgilash tizimini ishlab chiqishdir.

Mualliflik korpusi tuzishda ikki manbaga asoslanish mumkin:

1. Har bir muallifning nashr etilgan mukammal asarlari to'plami.
2. Internet tarmog'idagi elektron fayllar.

Birinchi manbani elektron shakga aylantirish (skanerlash, matnni sun'iy intellekt tushunadigan formatga keltirish) orqali material sifatida ishlatsak, ikkinchi manbadan nisbatan tayyor holda foydalanishimiz mumkin. Erkin Vohidov she'riyatining lingvistik korpusini yaratishda shoirning Saylanma kitobi 1-jildini obyekt sifatida tanlab oldik. Korpus yaratish

<sup>2</sup>[Хамроева Ш.М. “Корпус лингвистикаси атамаларининг қисқача изоҳли луғати” Бухоро 2018]

jarayonida mazkur kitob matnlari elektron shaklga keltiriladi. To'plangan matn ma'lum tizimga solinadi.

Erkin Vohidov she'riyatining tilshunoslik korpusi illyustrativ shaklda bo'ladi. Bunday korpusdan quyidagi maqsadlarda foydalanish mumkin:

- ta'limiy topshiriqlar tuzish;
- korpus asosida tajriba o'tkazish;
- tarjima;
- matn tuzish.

She'riy korpus nafaqat kerakli so'zlarni va boshqa tillarni qidirib topishga imkon beradi, balki she'riy matnning o'ziga xos xususiyatlarini ham aks ettiradi. Ayni paytda dunyodagi ikkinchi faol she'riyat korpusi bu - boshqird tili she'riy korpusi (BPK). U boshqird she'riy matnlari keng to'plamiga asoslangan. Mazkur korpus 2013-yil oktabr oyida foydalanish uchun taqdim etilgan. U 101 shoirning 17 mingga yaqin she'rlaridan iborat bo'lib, ular jami 1,7 milliondan ortiq tokenlarni va taxminan 450 ming she'riy satrlarni o'z ichiga oladi. Mazkur korpusdan erkin va bepul foydalanishimiz mumkin. Boshqird she'riyat korpusidan foydalanish uchun kompyuter internetga ulangan bo'lishi kerak. Ushbu korpus dasturi umumiy brauzerlar: Google Chrome, Windows Internet Explorer, Mozilla Firefox, Opera kabilarda ortiqcha so'rovlarsiz ishlaydi. Boshqird tili poetik korpusida tadqiqot maqsadlari uchun mos bo'lgan turli xil ma'lumotlar mavjud. Hozirda korpusda amalga oshirilishi mumkin bo'lgan qidiruv so'rovlari 2 ga ajratiladi:

1. Leksik izlash.
2. Grammatik izlash

Adabiy til rivojiga salmoqli hissa qo'shgan buyuk yozuvchilar asarlarining leksikasini asrab qolish, nafaqat asrash, balki butun ko'lami bilan tavsiflash, til egalari uchun namuna vazifasini o'taydigan model shakliga keltirish adabiy tilni asrash, rivojlantirish omili bo'lib xizmat qiladi. Bu borada o'z millati milliy, adabiy tili asoschisi hisoblangan butun dunyoga mashhur adiblar *Shakespeare, Dante, Gyote, Pushkin*larning adabiy merosi asosida qilingan ishlar (tuzilgan korpuslar) namuna sifatida xizmat qiladi.

### References:

1. Грудева Е. В. Корпусная лингвистика: учеб. пос. 3-е изд., стереотип. М.:ФЛИНТА, 2017.
2. Джаъфарова Д.Ф. Модели лингвистического анализа текстов таджикского языка: на материале газелей Хафиза. Автореферат. -, 2
3. Захаров В. П. Корпусная лингвистика: учеб.-метод. пос. СПб.: Изд-во С.-Пе-терб. ун-та, 2005.
4. Захаров В.П., Богданова С.Ю. Корпусная лингвистика.– Иркутск: ИГЛУ, 2011. –161 с.
5. Копотев М. В. Введение в корпусную лингвистику. Прага: Animedia Company, 2014.
6. Плунгян В. Корпусная лингвистика // ПостНаука [Электронный ресурс]. 2013.
7. Фицджеральд Ф.С. Лингвостатистический анализ художественного текста: На материале коротких рассказов. Автореферат. -, 2002.
8. Ҳамроева Ш.М. “Корпус лингвистикаси атамаларининг қисқача изоҳли луғати” Бухоро 2018
9. Ҳамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари. Монография. – , 2020

10. Ҳамроева Ш.М. “Корпус лингвистикаси атамаларининг қисқача изоҳли луғати”  
Бухоро 2018.
11. <http://ruscorpora.ru>
12. <http://sara.natcorp.ox.ac.uk/>



INNOVATIVE  
ACADEMY