

UDK:811.512.133'322+821.512.133.09-1

ABDULLA ORIPOV, MUHAMMAD YUSUF, SHAVKAT RAHMON SHE'RIYATI KORPUSINI YARATISHNING AHAMIYATI

Maxammadiyeva Diyora Raxmatovna

Tayanch doktorant

diyoramaxammadiyeva@gmail.com

ToshDO'TAU

Annotatsiya. Ushbu maqolada dunyo tilshunosligida korpus lingvistikasi sohasining taraqqiyoti, rivojlangan mamlakatlar korpuslari va undan foydalanish samaradorligi haqida mulohazalar bayon etilgan. Xususan, Rus tili milliy korpusi tarkibida yaratilgan poetik asarlar korpusiga batafsil to'xtalib, o'tilgan. Xuddi shunday tajribalar asosida o'zbek tilshunosligida ham xalqimizning sevimli shoirlari Abdulla Oripov, Muhammad Yusuf, Shavkat Rahmon she'riy asarlari korpusini yaratish masalasi kun tartibiga qo'yilgan. Poetik asarlar korpusiga bo'lgan ehtiyoj, nafaqat amaliy tilshunoslik va adabiyotshunoslik uchun ahamiyatlidir, balki leksikografiya, lingvostatistika, lingvokulturologiya sohalarining rivojlanishi uchun ham foydalidir. Ushbu korpusni yaratish yo'lida amalga oshirilgan ishlar yuzasidan daslabki xulosalar va tahlillar taqdim etilmoqda.

Abstract. This article presents considerations on the development of the field of corpus linguistics in world linguistics, corpora of developed countries and the effectiveness of its use. In particular, the corpus of poetic works created as part of the Russian language national corpus is discussed in detail. Based on similar experiences, the issue of creating a corpus of poetic works of our people's favorite poets Abdulla Oripov, Muhammad Yusuf, Shavkat Rahmon has been put on the agenda in Uzbek linguistics. The need for a corpus of poetic works is not only important for applied linguistics and literary studies, but is also useful for the development of the fields of lexicography, linguostatistics, and linguoculturology. Preliminary conclusions and analyses of the work carried out to create this corpus are presented.

Абстрактный. В статье излагаются соображения о развитии направления корпусной лингвистики в мировой лингвистике, корпусах развитых стран и эффективности их использования. В частности, подробно рассматривается корпус поэтических произведений, созданных в рамках национального корпуса русского языка. На основе подобного опыта в узбекской лингвистике на повестку дня выдвинут вопрос о создании корпуса поэтических произведений любимых поэтов нашего народа Абдуллы Орипова, Мухаммада Юсуфа и Шавката Рахмона. Потребность в корпусе поэтических произведений важна не только для прикладной лингвистики и

литературоведения, но и полезна для развития таких направлений, как лексикография, лингвостатистика и лингвокультурология. Представлены предварительные выводы и анализы проделанной работы по созданию данного корпуса.

Kalit soʻzlar. *Korpus lingvistikasi, Rus tili milliy korpusi, poetik korpus, Korpus yaratishning texnologik bosqichlari, asarlarni korpusga tayyorlash jarayoni.*

XXI asrda texnologiyalar shu qadar tez surʼatlar bilan rivojlanib, takomillashib bormoqdaki, barcha tarmoqlar singari biz tilshunoslik sohasi vakillari ham lingvistik tadqiqotlar qilishning texnologik usullarini oʻrganish va yaratishga doimo tayyor turishimizni zamonning oʻzi taqozo qilmoqda. Korpus lingvistikasi oʻzbek tilshunosligida yangi paradigma hisoblanib, oxirgi yillarda ushbu yoʻnalish faol sohalardan biriga aylandi. Til korpuslarining paydo boʻlishi tilshunoslarga til faktlarini tahlil qilish uchun yangi kuchli vositani berdi. Amerika, Yevropa, Rossiya va boshqa rivojlangan mamlakatlar amaliy tilshunosligida til korpuslari oʻtgan asrning 70-80-yillarida yaratila boshlangan edi. Oradan bor-yoʻgʻi yarim asr vaqt oʻtishiga qaramay, bugun dunyo tillarining ham ogʻzaki, ham yozma matnlarini oʻzida jamlagan yirik til korpuslari paydo boʻldi. Xususan, Amerika milliy korpusi, Britaniya milliy korpusi, Lankaster Oslo/Bergen korpusi, London-Lund korpusi, Ogʻzaki amerikacha ingliz tili korpusi, Rus tili milliy korpusi, Rus tilining Upsala korpusi shular jumlasidandir. Katta hajmdagi matnlarni oʻzida jamlagan, qoʻshimcha leksik, morfologik, grammatik, semantik jihatdan ishlov berilgan til korpuslari yordamida tadqiqotlarni amalga oshirish qulay va ishonchlidir. Raqamli texnologiyalar rivojidadagi inqilobiy oʻzgarishlar har bir sohaga, xususan, tilshunoslik va unga yondosh sohalarga ham sezilarli taʼsir koʻrsatdi. Xususan, bugungi kun tilshunosligining muhim sohalaridan biri korpus lingvistikasi kompyuter tilshunosligi va amaliy tilshunoslik sohaslarining kesishgan nuqtasida paydo boʻldi.

Oʻzbek tilshunosligida ham keyingi yillarda Til birliklarini lingvostatistik tadqiq etishga qaratilgan ishlar[1,2,3,4,5,6], til korpuslarini tuzish metodikasi va texnikasi tadqiqiga bagʻishlangan tadqiqotlar koʻpaydi[7,8,9,10,11,12,13,14,15]. Ular asosida bir necha korpuslar ishlab chiqildi. Ammo turli oʻzbek shoirlari asarlari korpusini tuzish va unga asoslangan lingvostatistik tadqiqot olib borish dolzarbligicha qolmoqda. Abdulla Oripov, Muhammad Yusuf va Shavkat Rahmon sheʼriyati korpusini tuzish mazkur ijodkorlar madaniy merosini kelajak avlodga yetkazish, internet tizimi yoki axborot qidiruv tizimlarida sifatli saqlash, qayta ishlash, ular ustida turli tadqiqotlar olib borishda dolzarblik kasb etadi.

Hozirgi vaqtda korpus – texnik taraqqiyotning bir qismi, zarur til birliklariga oid namunalarni topish uchun qulay vosita, shuningdek, til oʻrganish va oʻrgatishning yangi shakli hamdir.

Jahondagi eng yirik korpuslardan biri Rus tili milliy korpusi ajoyib korpus interfeysi va mukammal korpus annotatsiyasi bilan ahamiyatlidir. Rus tilining Milliy korpusi 2004-yil aprelidan boshlab, Internetda www.ruscorpora.ru sayti [15]da mavjud bo'lib, hozirgi vaqtda u 133554 ta matn, 389 milliondan ortiq so'zlarni o'z ichiga oladi. Korpusga XVIII asr boshidan XXI asr boshlarigacha bo'lgan tarixiy, publisistik, ijtimoiy, og'zaki ko'rinishdagi matnlar kiritilgan. Ushbu korpusda she'riy matnlar korpusi alohida bo'limni tashkil etib, 103626 matn, 14 milliondan ortiq so'zlardan iborat. She'riy matnlarni rus tili milliy korpusi tarkibiga joylashtirish ishlari Rossiya Fanlar akademiyasining Rus tillari Instituti jamoasi tomonidan 2006 yildan boshlab amalga oshirilgan. V.A.Plungyan rahbarligidagi Rossiya Fanlar akademiyasi Falsafa instituti ko'magida “Rus tilining milliy korpusi doirasida “Rus she'riyati korpusi” axborot tizimini yaratish” loyihasi doirasida she'riy matnlarni to'plash va korpus yaratish ishlari boshlangan. She'riy matnlarni belgilash tamoyillari V. A. Plungyan tomonidan E. A. Grishina, K. M. Korchagin, N. V. Pertsov, A. E. Polyakov va D. V. Sichinava ishtirokida ishlab chiqilgan. She'riy matnlarni to'plash va belgilash bo'yicha eng asosiy ishlar T. A. Arxangelskiy, D. A. Ivanova, K. M. Korchagin, B. V. Orexov, D. V. Sichinava va K. A. Florinskiylar tomonidan amalga oshirilgan. Korpusni boyitishda L. A. Barkova, N. K. Bogomolova, A. V. Bulannikov, I. S. Bulkina, A. A. Vasyatkina, E. A. Grishina, E. N. Lovlya, D. A. Paramonova, A. V. Praxt, D. E. Smirnov, D. A. Ershler kabi tadqiqotchi-olimlar ham faol ishtirok etishgan. Ishni muvofiqlashtirish va she'riy subkorpusning belgilarini tahrirlash D. V. Sichinava ishtirokida V. A. Plungyan va K. M. Korchagin tomonidan amalga oshirilgan [16]. 2006- yildan bugungi kungacha rus tili milliy korpusi tarkibida She'riy matnlar korpusi rivojlanib, boyitilmoqda. Ushbu korpus XVIII boshlab bugungi rus zamonaviy she'riyati namunalari va ijodkorlarning she'riy tarjimalari bilan boyitilgan. Manbalar mualliflarning eng to'liq va mukammal nashrlari (shoir shaxsiy kutubxonasi, mukammal asarlari) materiallariga asoslangan. Korpusga yosh avlod shoirlarining jamoaviy antologiyalari ham kiritilgan. Odatiy semantik va morfologik belgilarga qo'shimcha, meta-belgilashning bir qismi sifatida maxsus versifikatsion belgi taqdim etiladi.

O'zbek tilshunosligida esa she'riy matnlar korpusini tuzish ishlari kun tartibidagi birinchi masalalardan biridir. Bunda korpus tuzuvchisi oldidagi dastlabki qadam – davri, maqsadi, izohlanishi, matn hajmi, izoh xususiyatiga ko'ra yaratilajak korpus turini va korpus materialini aniqlashtirib olishdan iborat. Professor Sh.Hamroyevaga ko'ra[17] Korpus tuzishning texnologik jarayoni bir necha bosqichlarni qamrab oladi. 1. Belgilangan manbaga muvofiq holda matnning korpusga kirishini ta'minlash. 2. Matnni avtomatik o'qilish shaklida qayta ishlash, ya'ni korpusga kiritiladigan elektron shakldagi matn turli usul bilan olingan bo'lishi mumkin: qo'lda terilgan, skanerlangan, mualliflik nusxasi, hadya, ayirboshlash, internet, nashriyotlar tomonidan korpus tuzuvchisiga beriladigan orginal maketlar.

3. Tahlil, matnga dastlabki ishlov berish bosqichida turli manbalardan qabul qilingan matnlar filologik tekshiruv, tahrirdan o'tadi. 4. Konversiyalash, grafematik tahlil, ba'zi matnlar qayta kodlashtirish jarayoni amalga oshadigan ilk mashina ishlovidan qayta-qayta o'tadi, nomatniy qismlar, jadvallar, rasmlar o'chiriladi yoki o'zgartiriladi. Matndagi bo'g'in ko'chirish, chegaralar bekor qilinadi, tire va boshqa belgilar bir xilligiga erishiladi. 5. Grafematik tahlil korpusga kiruvchi matnni qismga (so'z, bog'lovchi) ajratish, nomatniy elementni o'chirish kabi amallarni bajaridan iborat. 6. Nostandart (noleksik) elementni belgilash, rasmiylashtirish, maxsus matniy elementni (qisqartma asosida yozilgan nom (ism, familiya), boshqa alifboda yozilgan o'zlashma leksema, rasmga berilgan izoh, nom, zarvaraq, adabiyotlar ro'yxati va b.) bir xil me'zon asosida qayta ko'rib chiqish. Ammo she'riy asarlar uchun korpus yaratish jarayinida yuqorida sanab o'tilgan bosqichlardan tashqari yana ayrim qo'shimcha ishlar amalga oshirilishi maqsadga muvofiq. Hozirda bizning tadqiqotimiz korpus tuzishning 1-2-3-texnologik jarayonlari yakunlangan bo'lib, dastlab O'zbek xalqining buyuk shoirlari Abdulla Oripov, Muhammad Yusuf, Shavkat Rahmon ijodiga mansub asarlar Ziyo.net kutubxonasidan[18] elektron shaklda pdf formatda kompyuter xotirasiga saqlab olindi. Abdulla Oripovning 26 ta (shoirning dastlabki to'plamlari tarkibidagi she'rlar mukammal asarlari to'plamida takrorlanganligi sababli tadqiqot obyekti sifatida 8 jildli mukammal asarlari to'plamini oldik). Muhammad Yusufning 13 ta, Shavkat Rahmonning 5 ta she'riy to'plami mavjud ekanligi aniqlandi. Pdf formatidagi barcha kitoblar PDF to Word Converter dasturi orqali docx format holatiga keltirildi, Bexato telegram boti orqali kitoblar kril alifbosidan lotin alifbosiga o'girildi, ushbu jarayonda yuzaga kelgan texnik va imloviy kamchiliklar filologik tekshiruv va tahrirdan o'tkazilib, she'riy matnlar nashr holati bilan bir xillikka erishildi. Asarlardagi nashriyot nomi, kirish so'zlar, mundarija kabi asasiy matn tarkibiga kirmaydigan qismlar olib tashlandi. Abdulla Oripov, Muhammad Yusuf, Shavkat Rahmon asarlari va ularning umumiy hajmini (sahifada) quyidagi jadvalda keltirdik.

T/R	Abdulla Oripov		Muhammad Yusuf		Shavkat Rahmon	
	Asar	Sahifalari	Asar	Sahifalari	Asar	Sahifalari
1.	Mukammal asarlari to'plami 1-jild	428	Ko'nglimda bir yor	32	Gullayotgan tosh	62
2.	Mukammal asarlari to'plami 2-jild	492	Bulbulga bir gapim bor	54	Hulvo	63
3.	Mukammal asarlari to'plami 3-jild	382	Erka kiyik	70	Saylanma	215
4.	Mukammal asarlari to'plami 5-jild	399	Halima enam allalari	36	Uyg'oq tog'lar	190

5.	Mukammal asarlari to'plami 6-jild	382	Iltijo	60	Yurak qirralari	107
6.	Mukammal asarlari to'plami 7-jild	299	Ishq kemasi	220		
7.	Mukammal asarlari to'plami 8-jild	322	Iymon qo'shig'i	75		
8.			O'zingdan qo'ymasin xalqim	102		
9.			Saylanma	366		
10.			Uyqudagi qiz	206		
11.			Tanish teraklar	66		
12.			Xalq bo'l elim	189		
13.			Yolg'onchi yor	151		
jami	7 ta to'plam	2704 sahifa	13 ta to'plam	1627 sahifa	5 ta to'plam	637 sahifa

Korpus yaratishda amalga oshirilishi kerak bo'lgan dastlabki matn to'plash va tahrir bilan bog'liq texnik ishlar yuqoridagi ketma-ketlikda amalga oshiriladi.

Xulosa o'rinda shuni aytish mumkinki, Dunyo tajribasida korpus tildan foydalanishda uning statistik analizi, tabiiy tilni qayta ishlash dasturiy ta'minoti, leksik resurslarni yaratish, til o'qitish va boshqa qator maqsadlarda qo'l kelmoqda. O'zbek adabiyotining buyuk shoirlari Abdulla Oripov, Muhammad Yusuf, Shavkat Rahmon o'zining betakror she'riyati va uslubiga ega, shuning uchun ularning asarlari asosida yaratilgan she'riy korpus nafaqat o'zbek adabiyoti, she'riyatini chuqurroq tushunishga, balki zamonaviy poetikaning rivojlanishini o'rganishga ham katta imkoniyat yaratadi va o'zbek tilshunosligida korpus lingvistikasi yo'nalishida amalga oshirilayotgan tadqiqotlarning davomi sifatida ularni to'ldiradi.

Foydalanilgan adabiyotlar:

1. Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. – Ленинград, 1981.;
2. Айымбетов Н.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук.– Ташкент, 1987.;

3. Каримов С., Қаршиев А., Исроилова Г. Абдулла Қаҳҳор асарлари тилининг луғати. Алфавитли луғат. Частотали луғат. Терс луғат. – Тошкент, 2007. – 420 б.;
4. Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Фил.фан.док.дис...автореф. – Тошкент, 2008.;
5. Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: Фан, 2010.
6. Муҳаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. – Тошкент, 2006. – 45 б.
7. Хамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Фил.фан.бўйича фалсафа доктори (PhD)...дис. – Бухоро, 2018. – 250 б.;
8. Тоирова Г.И. Ўзбек тили миллий корпусини яратишнинг назарий ва амалий масалалари. Филология фанлари доктори (DSc) диссертацияси автореферати. – Бухоро, 2021.
9. Ахмедова Д. Атов бирликларини ўзбек тили корпуслари учун лексик-семантик теглашнинг лингвистик асос ва моделлари Фил.фан.бўйича фалсафа доктори (PhD) дис. – Бухоро, 2020. – 40 б.;
10. Эшмуминов А.А. Ўзбек тили миллий корпуси синоним сўзлар базасини яратиш асослари: Фил.фан.бўйича фалсафа доктори (PhD) дис. – Термиз, 2019. – 140 б.
11. Атабоев Н. Инглиз тили корпусининг функционал хусусиятлари (СОСА мисолида) . Дис. Автореферат. 15- бет.
12. Begmatova G. O'zbek milliy korpusida idiomalar bazasini yaratish. PhD dissertatsiya. 2021
13. Xidirov O. Milliy korpus uchun parsing dasturi yaratishning lingvistik asoslari. .PhD dissertatsiya 2021.
14. Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида). Филол.фан.бўйича фалсафа доктори (PhD)...дис. афтореф. – Тошкент, 2018.
15. www.ruscorpora.ru sayti
16. <https://ruscorpora.ru/corpus/poetic>
17. Хамроева Ш. Корпус лингвистикаси атамаларининг қисқача изоҳли луғати”. –Тошкент: Камалак, 2018.
Ziyo.net kutubxonasi.