

O'ZBEK TILI MUALLIFLIK KORPUSI UCHUN YARATILGAN DASTURNING TAVSIFI

Abdiraxmanova Dilnoza Abdukarimovna,

abdurahmonovadilnoza9@gmail.com

Sh.Rashidov nomli Samarqand davlat universiteti

2-kurs magistranti

Annotatsiya: Ushbu maqolada o'zbek tili milliy mualliflik korpusi uchun yaratilgan Mualliflik korpusi dasturidan foydalanish tavsiflangan.

Kalit so'zlar: *mualliflik korpusi, konkordans, fayldan ko'rish, klasterlar/N-grammlar, so'zlar ro'yxati.*

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 oldi;

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 (so'zning o'zgarishi, istorizm, neologizm vujudga kelishi, ma'no torayishi, ma'no kengayishi; yangi frazeologizmlarning paydo bo'lishini kuzatish), til o'rganish, lug'at tuzishda eng zamonaviy, keng imkoniyatli dasturlashtirilgan tizimdir.

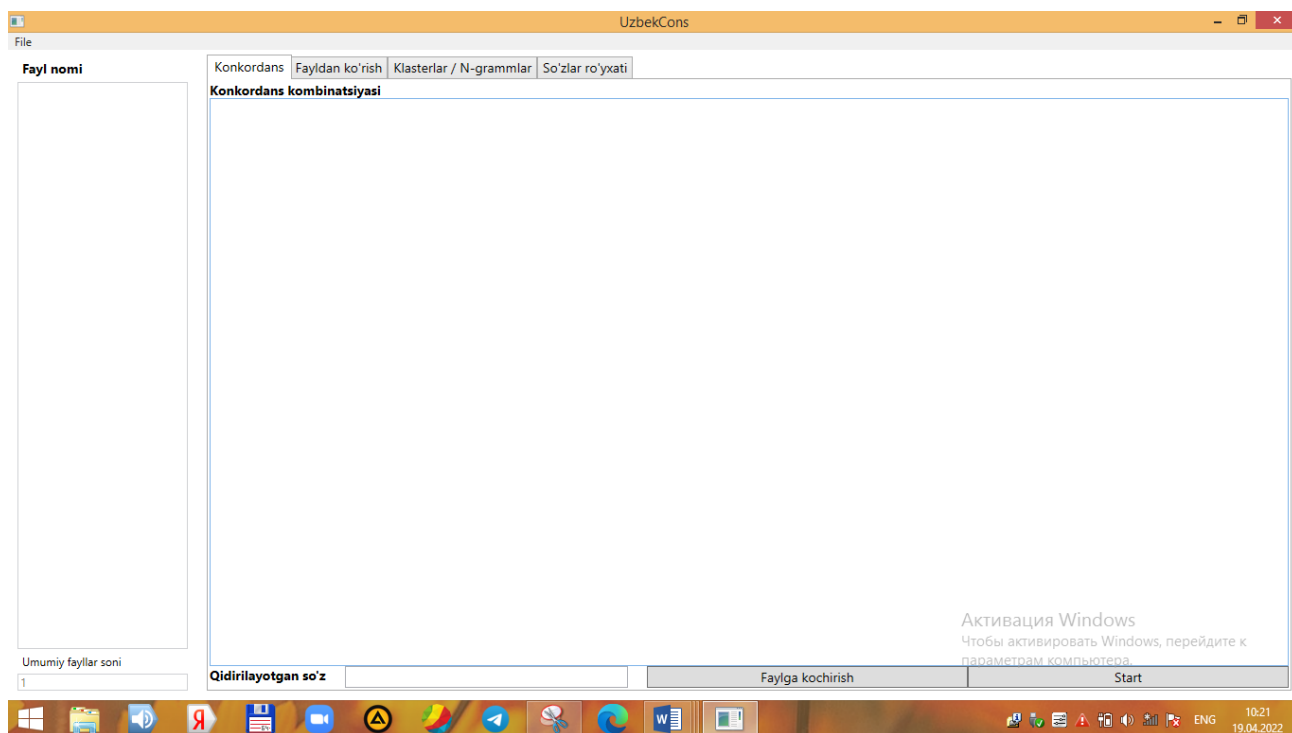
Mualliflik korpusi dasturining interfeysi quyidagi oynalardan iborat:

1.Konkordans

2.Fayldan ko‘rish

3.Klasterlar/N-grammlar

4.So‘zlar ro‘yxati

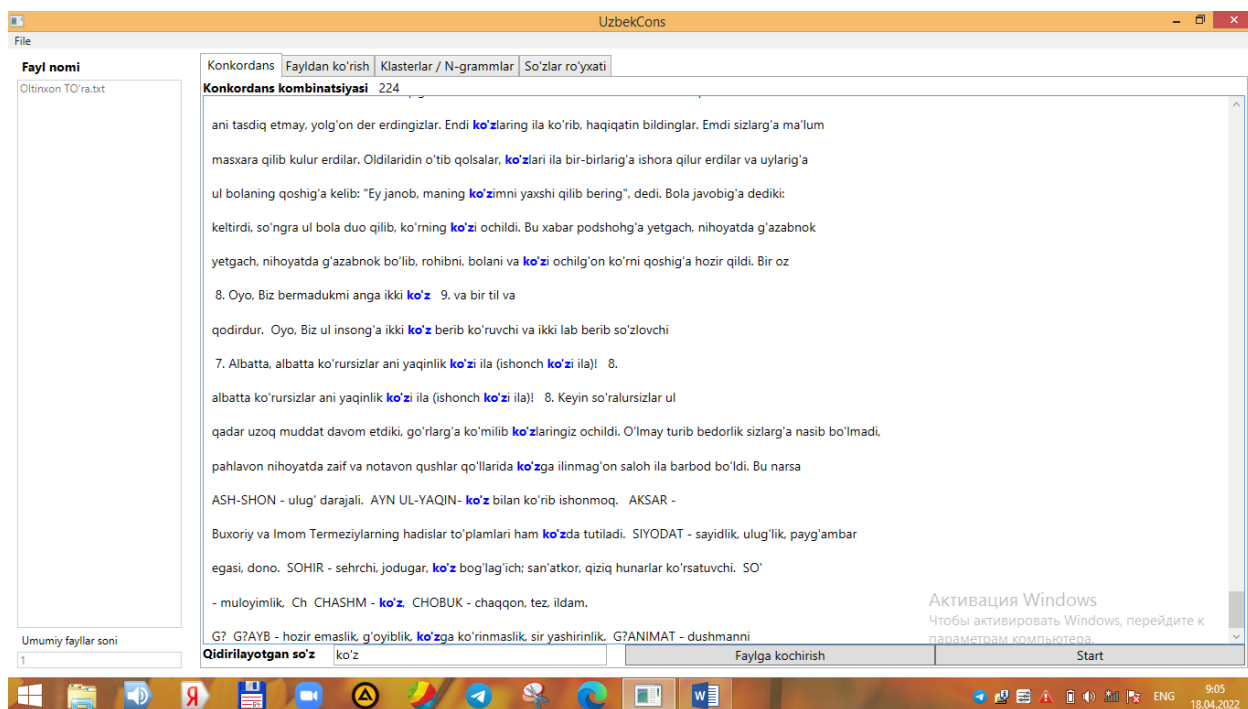


1-rasm. Mualliflik korpusi dasturining interfeysi

Fayl nomi ishchi dasturiga lingvistik tadqiqot bajariladigan matn kiritiladi. Masalan, Qur'oni karimning Oltinixon To'ra tomonidan tarjima qilingan “Al - Qur'on Al-karim” (tarjima va sharhlar).

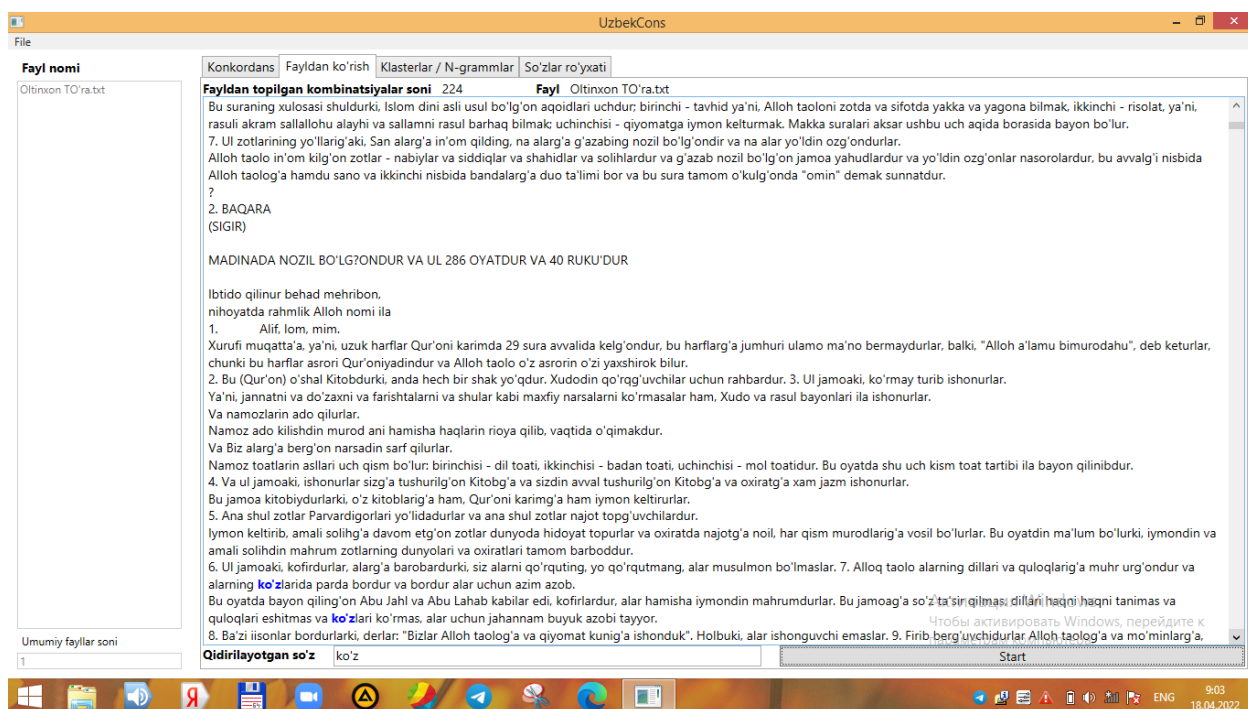
Konkordans - bu matndagi yoki korpusdagi so‘zni va uni o‘rab turgan so‘zlar bilan birga ro‘yxati keltirilgan. Odamlar korpus tilshunosligidagi kelishuvlar to‘g‘risida gaplashganda, odatda "Kontekstdagi kalit so‘zlar" (KWIC) ning oddiy muvofiqligi tushuniladi. Konkordansiyalar asosan ma'lumotlarni vizuallashtirish usuli hisoblanadi. Qidiruv atamasi va uning qo‘shma matni, matn muhitini baholash va qidiruv atamasi atrofidagi belgilarning kelishuvlarini (va so‘z ro‘yxatlarini) matnning o‘zgarishi deb ta’riflaydi va tahlilchiga matnga nisbatan turli xil qarashlarni ko‘rish imkoniyatini beradi. Misol uchun, Qur'oni karimning Oltinixon To'ra tomonidan tarjima qilingan “Al -Qur'on Al-karim” (tarjima va sharhlar) matnidagi “ko‘z” leksemasini qidirilayotgan so‘z degan joyiga kiritilsa va Start

tugmasi bosilsa, asarda *ko'z* leksemasining konkordans kombinatsiyasi soni kontekstlar bilan ishchi dasturda ko'rinadi. Dasturning ish natijasiga ko'ra *ko'z* leksemasi bilan bog'liq 224 ta kombinatsiya ishchi oynasida ko'rinadi.



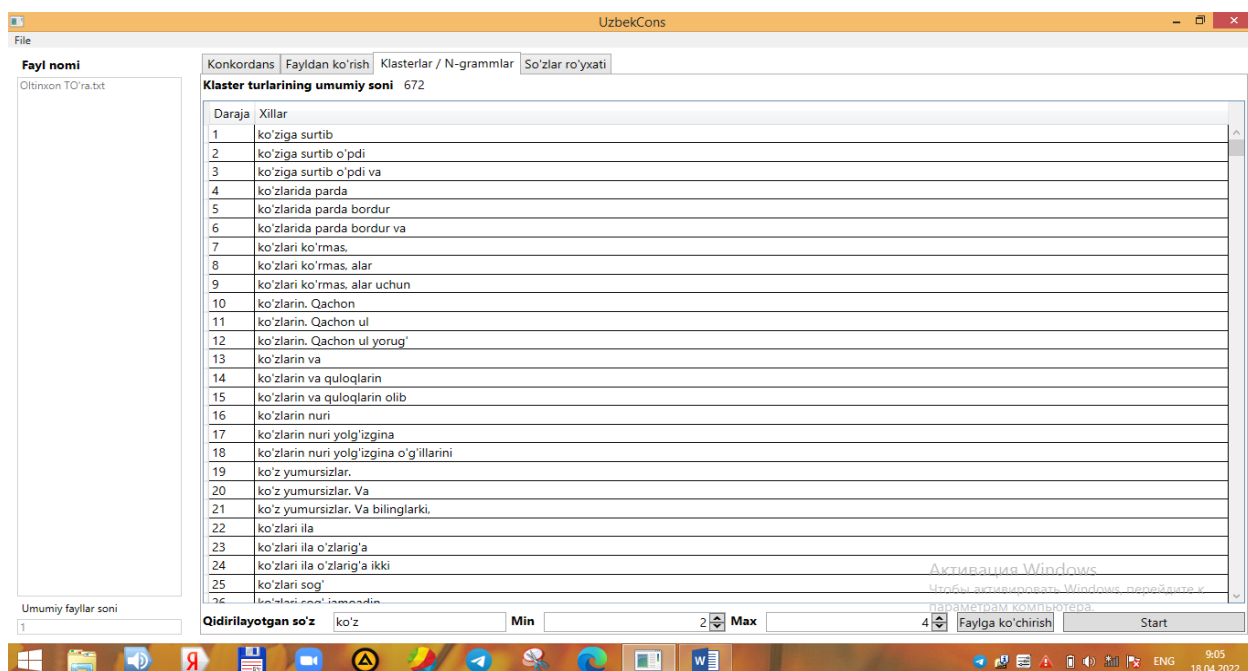
2-rasm. Konkordans ishchi dasturining ko'rinishi

Fayldan ko'rish oynasida qidirilayotgan so'zga *ko'z* leksemasi kiritilsa, butun matndan shu so'z ajratib olinadi va shu so'z o'rganilayotgan so'z bilan taqqoslanadi, so'zning o'zini ko'k rang bilan ajratib konkordans ro'yxatiga qo'shadi. Misol uchun, Qur'oni karimning Oltinxon To'ra tomonidan tarjima qilingan "Al -Qur'on Al-karim" (tarjima va sharhlar) matnidagi "ko'z" leksemasini qidirilayotgan so'z degan joyiga kiritilsa va Start tugmasi bosilsa, asarda *ko'z* leksemasi butun matn ichida ajratib ko'rsatiladi.



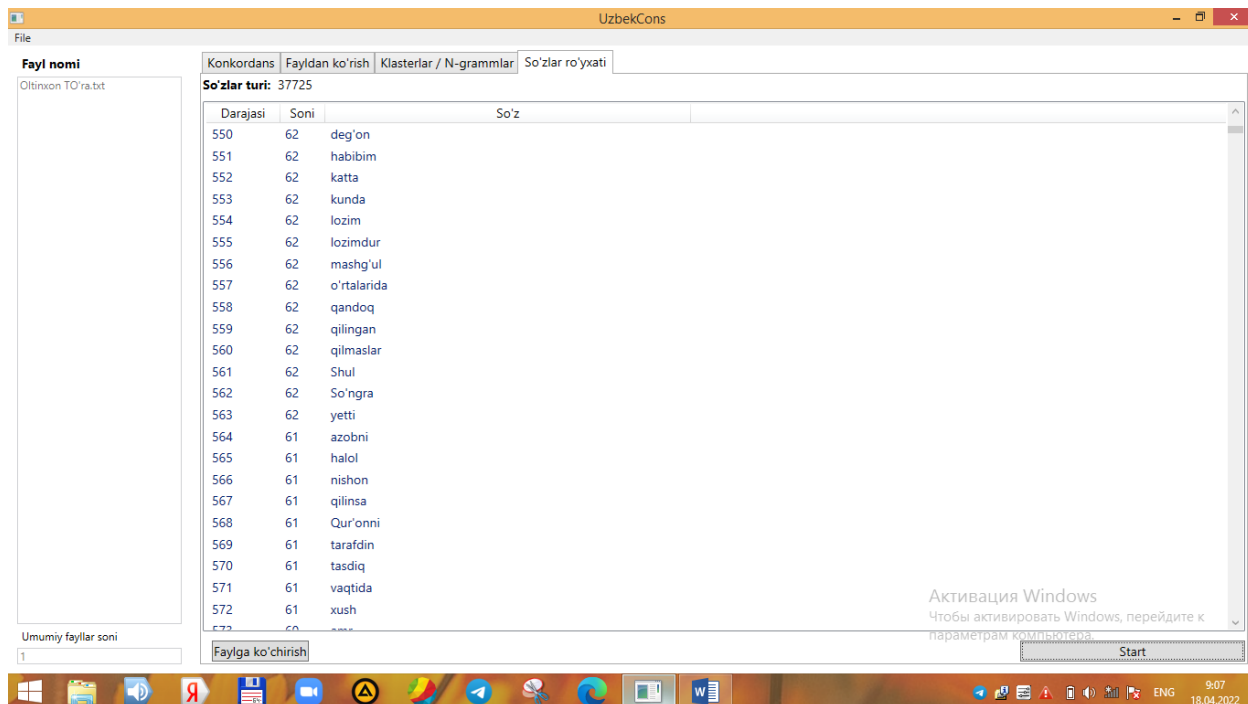
3-rasm. Fayldan ko'rish ishchi dasturining tasviri

Klasterlar/N-grammlar oynasida o'rganilayotgan so'zni chap tomondan n-grammaning minimum va maksimum qiymatining oralig'ida kontekst tuzadi va so'zning o'zini ko'k rang bilan ajratib ro'yxatiga qo'shadi. Matndan ajratilib olingan so'z, o'rganilayotgan so'zga teng bo'lmasa matnning keyingi so'zini tekshirishga o'tadi. Bu jarayon matnning oxiriga borgunga qadar davom etadi. Masalan, qidirilayotgan so'z qismiga *ko'z* so'zi kiritiladi, min qismiga 2 soni, max qismiga 4 soni kiritilsa, matn ichida *ko'z* leksemasining, ikkita, uchta, to'rttagacha bo'lgan leksemalar bilan birgalikda birikish holati ko'rsatiladi. Masalan, Qur'oni karimning Oltinxon To'ra tomonidan tarjima qilingan “Al -Qur'on Al-karim” (tarjima va sharhlar) matnidagi “*ko'z*” leksemasini qidirilayotgan so'z degan joyiga kiritilsa va Start tugmasi bosilsa, matnda *ko'z* leksemasining eng kichik 2 ta va eng katta 4 tagacha bo'lgan leksemalar bilan yonma-yon kelish holatini ko'rsatib beradi.



4-rasm. Klasterlar/N-grammlar ishchi dasturining ko'rinishi

So'zlar ro'yxati oynasida chastotali lug'at - so'zlarning (so'z shakllari, so'z birikmalarining) nutqda qo'llanish miqdori haqida ma'lumot beruvchi lug'at yaratiladi. O'rganilayotgan matnning lemmalar miqdori haqida ma'lumot beradi. Masalan, Qur'oni karimning Oltinxon To'ra tomonidan tarjima qilingan “Al - Qur'on Al-karim” (tarjima va sharhlar) matnining chastotali lug'atini yaratish mumkin bo'ladi.



5-rasm. So'zlar ro'yxati ishchi oynasining ko'rinishi



Xulosa o'rnida shuni aytish mumkinki, o'zbek tili milliy korpusini yaratish bugungi kun tilshunosligining eng muhim vazifalaridan biri hisoblanadi. Mualliflik dasturi har qanday matn ustida lingvistik tadqiqot o'tkazish uchun mo'ljallangan.

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" maqolasi ["Ўзбекистонда хорижий тиллар" илмий-методик электрон журнал journal.fledu.uz №2/2019]
3. Xamrayeva Sh.M. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari. monografiya / Xamrayeva Sh.M. – Toshkent. 2020.