

## TEXT SELECTION FOR EDUCATIONAL BUILDING

**Kholiyorov Ural Mengliyevich**

Termez State University  
Senior Lecturer, Department of Uzbek Linguistics,  
Doctor of Philosophy (PhD) in Philology  
E-mail: xoliyorov@tersu.uz

**Abstract:** Educational corpora have a special place in the system of language corpora: they differ from other corpus by the linguodidactic nature of the materials. There are such special types of corpora in English and Russian, and even different models of the educational corpora have been created. For example, the Russian National Corpus has an educational corpus: it is designed as an internal corpus of the national corpus and is an integral part of the national corpus. This material differs from the texts in the National Corpus in that it is age-appropriate. The creation of the Educational Multimodal Corpus (UMKO) at the Irkutsk State Linguistic University is underway. This corpus material is a video of unprepared educational dialogues on a specific topic by Russian, Chinese speakers, and people who use it as a second language. The corpus is marked on the basis of the ELAN program, takes the form of a parallel corpus, the dialogues are divided into spiritual groups. For example, the speech of a Russian speaker is compared with the speech of a Chinese student studying Russian on the basis of a dialogue on the same topic. The purpose of this corpus is primarily to provide a linguistic-didactic purpose, in particular, to find typical errors, to find ways to exclude these errors from speech during the lessons, and to help the student to work independently.

### TA'LIMIY KORPUS UCHUN MATNLAR TANLASH XUSUSIDA

**Xoliyorov O'ral Mengliyevich**

Termiz davlat universiteti  
o'zbek tilshunosligi kafedrasida katta o'qituvchisi,  
filologiya fanlari bo'yicha falsafa doktori (PhD)  
E-mail: xoliyorov@tersu.uz

Til korpuslari tizimida ta'limiy korpuslarning alohida o'rni bor: ular materiallarning lingvodidaktik xususiyati bilan boshqa korpuslardan farq qiladi. Ingliz va rus tillarida korpusning shunday maxsus turlari farqlanadi, hatto ta'limiy korpusning xilma-xil namunalari ham yaratilgan. Masalan, rus tili milliy korpusi tarkibida ta'limiy korpus (obuchayushiy korpus)<sup>1</sup> mavjud: u milliy korpusning ichki korpusi sifatida ishlab chiqilgan, milliy korpusning tarkibiy qismi sanaladi. Uning materiali Milliy korpusdagi matnlardan o'quvchining yoshiga mosligi bilan ajralib turadi. Shuningdek, Irkutsk davlat tilshunoslik universitetida Ta'limiy multimodal korpus (UMKO) tuzish davom etmoqda. Bu korpus materiali rus, xitoy tili tashuvchisi, ikkinchi tili sifatida ishlatuvchi shaxslarning ma'lum mavzu borasida oldindan tayyorlanmagan ta'limiy dialoglari videoyozuvidir. Korpus ELAN dasturi asosida razmetkalangan, parallel korpus shaklini olgan, dialoglar ma'noviy guruhlarga bo'lingan holda joylashtirilgan. Masalan, rus tili egasi nutqi bilan rus tilini o'rganayotgan xitoy tili vakili nutqi ayni bir mavzudagi dialog asosida qiyoslanadi. Ushbu korpus, birinchi navbatda, lingvodidaktik maqsadni, xususan, tipik xatolarni topish, mashg'ulotlar davomida

<sup>1</sup> <https://ruscorpora.ru/new/search-school.html>

shu xatolarni nutqdan chetlashtirish yo'lini izlash hamda talabning mustaqil ishiga ko'maklashishni ko'zlaydi<sup>2</sup>.

Korpusning ushbu turi ingliz tilida alohida korpus sifatida mavjud: og'zaki ingliz tilini o'rganish uchun Lankaster ingliz tili so'zlashuv korpusi (Lancaster/IBM Spoken English Corpus<sup>3</sup>, Corpus of Spoken American English<sup>4</sup>), lingvodidaktik tadqiqotlar uchun ingliz tili o'rganuvchilarining xalqaro korpusi<sup>5</sup> shular jumlasidan.

Ta'limda korpusdan foydalanish tilni o'rganuvchi, uning tashuvchisi bo'lgan kishi nutqini qiyoslash; adabiy og'zaki/yozma tilni chog'ishtirish; til uslublarini kitob, darslik, nashr materiali bo'yicha qiyoslash, farqini aniqlash; gaplarda so'z tartibini aniqlash kabi jarayonlar uchun zamonaviy axborot-qidiruv vositasi vazifasini bajaradi. Til korpusida foydalanuvchiga taqdim etilgan ma'lumot xaotik tarzda bo'lganligi sababli material tanlashda o'qituvchi mas'uliyatni o'z zimmasiga olishi talab etiladi. Shunda ta'limiy korpus til ta'limida foydali elektron ta'limiy resurs bo'lib xizmat qiladi. Ushbu vositani qo'llash orqali ta'lim oluvchida til materialini (korpus asosida) mustaqil tahlil qilish, tanqidiy va analitik fikrlash ko'nikmasini hosil qiladi. Murakkabligiga qaramasdan, o'quv materialini ishlab chiqishda ta'limiy korpus samarali manba bo'la oladi.

O'zbek tili ta'limiy korpusi materiali sifatida hozirgi o'zbek adabiy tilidagi rasmiy, ilmiy-ommabop, badiiy hamda publitsistik uslubdagi matnlar tanlangan. Korpus reprezentativlik xususiyatini to'la namoyon etadi: uning tarkibidagi materiallar xilma-xil; tarkibidagi matnlar janr jihatdan quyidagi proporsiyaga ega:

1) o'zbek tili ta'limiy korpusining hozirgacha to'plangan matnlarining badiiy uslubga tegishli qismi Abdulla Qahhor, Said Ahmad, Tog'ay Murod, Asqad Muxtor qalamiga mansub qissa, hikoya, roman, felyeton, esdalik kabi janrlarga oid asarlarni qamrab olingan;

2) publitsistik uslubdagi matnlar tarkibiga kun.uz, daryo.uz saytlaridagi publitsistik maqolalar kiritilgan;

3) korpusning rasmiy uslubdagi materiallarini lex.uz saytida mavjud O'zbekiston Respublikasi Konstitutsiyasi, O'zbekiston Respublikasi Prezidentining Qonun, Qaror va Farmonlari, turli kodekslarning matnlari tashkil etadi;

4) ilmiy-ommabop uslubdagi materiallar, asosan, maktab darsliklari: adabiyot, fizika, matematika, kimyo, geografiya, botanika, biologiya, chizmachilik, odobnoma, ma'naviyat asoslari kabi 100 dan ortiq darslik matnlaridan iborat.

Mantlarni fragmentga bo'lish, nomatniy belgi (jadval, rasm, chizma)larni olib tashlash sublimetext dasturi vositasida amalga oshirilgan.

Bugungi kunda o'zbek tili ta'limiy korpusini ishlab chiqish bo'yicha quyidagi amaliy natijalarga erishilgan:

- o'zbek tili ta'limiy korpusi uchun axloqiy-tarbiyaviy mazmundagi matnlar tanlangan, korpusga joylashtirilgan;

- o'zbek tilida mavjud o'quv lug'atlari bir platformaga to'plangan; qulay qidiruv tizimiga bo'ysundirilgan;

- o'zbek tili ta'limiy korpusi lingvistik ta'minoti morfoleksikon shakllantirilgan.

Yuqoridagi sanab o'tilgan manbalar ta'limiy korpus uchun boshlang'ich materiallar sifatida kiritilgan. Keyinchalik badiiy uslubga tegishli yozuvchi va shoirlarning sara, tabiiyaviy ahamiyatga ega, tili soddaroq 200 dan ortiq asarlari kiritiladi. Bunda asosan maktab

<sup>2</sup> Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD) ...дисс. – Қарши, 2018. – 168 б.

<sup>3</sup><https://www.routledge.com/A-Corpus-of-Formal-British-English-Speech-The-LancasterIBM-Spoken-English/Knowles-Assistant-Taylor-Research-Williams/p/book/9780582056398>

<sup>4</sup> <https://www.linguistics.ucsb.edu/research/santa-barbara-corpus>

<sup>5</sup> Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD) ...дисс. – Қарши, 2018. – 168 б.

adabiyot darsliklarida tilga olingan asarlarga urg'u berish maqsadga muvofiq. Publitsistik uslubdagi eng ommabop saytlarda e'lon qilingan maqolalar saralangan holda, ommabop gazeta va jurnallardagi materiallar esa to'liq kiritilishi lozim. Keyinchalik baza audio, video materiallar bilan ham boyitilishi mumkin.

Rasmiy uslubga tegishli matnlar lex.uz saytida yetarlicha mavjud. Lekin ta'limiy korpus uchun uning barchasini korpus bazasiga kiritish to'g'ri emas. Chunki ushbu saytda 50 ortiq turdagi 100 mingga yaqin rasmiy hujjatlar mavjud. Ularning hammasi ta'limiy korpusga kiritilsa, uslublararo proporsionallik buziladi, qolaversa, o'ta tor mutaxassislikka doir hujjatlar juda ko'pligi sabab omma uchun tushunarsiz atamalar ko'payib ketadi. Shuning uchun omma uchun xizmat qiluvchi, ta'lim-tarbiyaga, omma uchun kerakli bo'lgan 500 atrofida rasmiy hujjatlar matni kiritilsa, korpusning ta'limiy qimmatini oshadi.

Ilmiy-ommabop matnlar sifatida maktab darsliklariga qo'shimcha material sifatida oliy ta'limda o'qitiladigan va birinchi blok fanlariga kiritiladigan tarix, falsafa, mantiq kabi fanlar matnlari ham bazadan o'rin olsa foydadan xoli bo'lmaydi.

Ta'limiy korpusda, albatta, matnlar hajmi chegaralangan bo'lishi kerak. Chunki uning foydalanuvchisi asosan 14-24 yoshdagi o'quvchi/talabalar bo'lganligi sababli har xil matnlarni bazaga to'plash ta'limiy-tarbiyaviy qiymatni kamaytiradi.

Ta'limiy korpusda matnlarni lingvistik tegrashda korpus kimlar uchun mo'ljallanayotganligiga qarash kerak. Masalan, maktab o'quvchilari uchun maxsus ta'limiy korpus tuzilayotgan bo'lsa, maktab darsliklaridagi ona tili bo'yicha lingvistik ma'lumotlarga suyanilgani ma'qul. Xuddi shunday litsey, kollejlarda o'quvchilari va oliy ta'lim talabalari uchun tuzilayotgan bo'lsa, o'sha bosqich darsliklarida berilgan lingvistik qoidalarga suyanilgani maqsadga muvofiq. Umumiy ta'limiy korpus tuzishda esa, hamma bosqichlar uchun umumiy lingvistik meyorlarga amal qilish kerak bo'ladi. Hozirgi vaqtda o'zbek tili uchun tuzilayotgan ta'limiy korpus umumiy bo'lganligi uchun, asosan, maktab ona tili darsliklariga, qo'shimcha tarzda litsey, oliy ta'lim darsliklaridagi tilshunoslik ma'lumotlariga, o'quv lug'atlariga asoslangan teglar kiritiladi. Milliy korpus uchun lingvistik teglar tizimi akadem grammatikalar, akadem nashrlarga asoslaniladi.

Ta'limiy korpusdan leksik birlik izlanganda quyidagi lingvistik ma'lumotlar foydalanuvchiga taqdim etish imkoniyati mavjud bo'lishi maqsadga muvofiq:

1) yozilishi; 2) bo'g'inlarga bo'linishi; 3) qaysi imlo qoidasi asosida yozilganligi; 4) imloviy variantdoshi; 5) talaffuzi va urg'usi; 6) tovush o'zgarishi; 7) ma'no(izoh)lari, ko'p ma'nolilikda ma'no ko'chish usullari; 8) sinonim(lar)i; 9) antonim(lar)i; 10) toksonimik uyasi, uyadosh(lar)i; 11) meronimik uyasi, uyadosh(lar)i; 12) darajalanish qatori; 13) shakldosh(omonimlar)i; 14) talaffuzdosh(paronim)i; 15) etimologiyasi; 16) o'z / o'zlashma qatlama mansubligi; 17) qo'llanishiga ko'ra: chegaralanmagan/chegaralangan: shevaga xos/atama/argo va jargonlar; 18) davriga ko'ra: zamonaviy/eskirgan/yangi leksik birlik; 19) dialektal (sheva) varianti; 20) tub/yasamaligi; 21) tarkib (morfemalar)i; 22) tuzilish turi: sodda/murakkab: qo'shma, juft, takror; 23) turkumiy belgisi turlari bilan; 24) old/orqa birikuvchilari; 25) uslubi; 26) chastotasi; 27) misollar.

Ta'limiy korpusda yuqoridagi lingvistik ma'lumotlar foydalanuvchi tanloviga ko'ra turlicha saralash imkoni yaratilishi lozim bo'ladi.