

## TIL TA'LIMIDA O'QUV LUG'ATCHILIKNI RIVOJLANTIRISHNING KORPUSGA ASOSLANGAN YONDASHUVI

**Nilufar Abduraxmonova Zaynobiddin qizi**

O'zbekiston milliy universiteti v.b. professori, filologiya fanlari doktori

**Xoliyorova Gulruxsor G'ayratovna**

Uchtepa tumanidagi 228-maktab o'qituvchisi

<https://doi.org/10.5281/zenodo.7440189>

**Annotatsiya.** Kompyuter lingvistikasida korpus muhim ahamiyatga ega, ayniqsa, elektron lug'atlar bazasi hamda grammatik qoidalarni yaratishda matnlar majmuasi (korpus)dan foydalaniladi. Korpus kompyuter lingvistikasining mashina tarjimasi, nutq sintezatori, matn tahlili, sentiment tahlil va boshqa qator yo'nalishlar uchun tadqiqot obyekti vazifasini o'taydi.

Korpusga asoslangan lingvistik tadqiqotlar shuni ko'rsatadiki, korpus tarkibidagi tezkor qidiruv tizim, adabiy tur va janlar bo'yicha xilma-xillik, qolaversa, statistik tahlil qilish imkoniyatini beradi. O'zbek tilining elektron korpusi o'quv lug'atchilikni rivojlantirishda muhim resurs bo'lib xizmat qiladi.

**Kalit so'zlar:** o'zbek tili korpusi, o'quv lug'atchilik, chastotali lug'atlar, kompyuter lingvistikasi.

## КОРПУСНЫЙ ПОДХОД К РАЗВИТИЮ СЛОВАРЕВЕДЕНИЯ В ЯЗЫКОВОМ ОБРАЗОВАНИИ

**Аннотация.** В компьютерной лингвистике большое значение имеет корпус, особенно корпус текстов используется при создании базы данных электронных словарей и грамматических правил. Корпус служит объектом исследования для машинного перевода, синтезатора речи, анализа текста, сентимент анализ и других областей компьютерной лингвистики.

Лингвистические исследования на основе корпусов показывают, что корпуса обеспечивают возможность быстрого поиска, разнообразие литературных жанров и жанров, а также статистический анализ. Электронный корпус узбекского языка служит важным ресурсом в разработке словарей для учащихся.

**Ключевые слова:** корпус узбекского языка, учебная словарь, частотные словари, компьютерная лингвистика.

## CORPUS APPROACH TO THE DEVELOPMENT OF VOCABULARY IN LANGUAGE EDUCATION

**Abstract.** In computer linguistics, the corpus is important, especially the corpus of texts is used in the creation of the database of electronic dictionaries and grammatical rules. The corpus serves as a research object for machine translation, speech synthesizer, text analysis, sentiment analysis and other areas of computer linguistics.

Corpus-based linguistic research shows that corpora allow for fast searchability, diversity in literary genres and genres, and statistical analysis. The electronic corpus of the Uzbek language serves as an important resource in the development of learner dictionaries.

**Keywords:** corpus of the Uzbek language, learner dictionary, frequency dictionaries, computational linguistics linguistics.

**Kirish**

Fransuz lug‘atshunosi Alan Rey “Zamonaviy sivilizatsiya bu lug‘atlar sivilizatsiyasidir [1]”, deya o‘rinli ta’kidlasa, “Lug‘at yaratish fikrlashning o‘ziga xosligini talab qilmaydi, bu kabi faoliyat turi ishga nisbatan oliy tafakkur, professionallik va mehnatsevarlikni talab qiladi. Bundan tashqari, lug‘at yaratish jarayoni muallifga mamnunlik keltirishi kerak” [2], deydi Landau. Bugungi kunda o‘zbek tili ta’limida o‘quv lug‘atchilikni yanada rivojlantirish, ayniqsa, zamonaviy yangi avlod lug‘atlarini yaratishda kompyuter texnologiyalaridan unumli foydalanish juda muhim sanaladi.

V.V.Dubichinskiy [3] leksikografiyaga izoh berar ekan lug‘at yaratish, o‘rganish va foydalanish haqidagi fan deb ta’rif beradi. Boshqa lingvistik sohalar qatorida alohida fan sifatida, shuningdek, universal metodik sohasi hamdir. Olimning nazdicha, barcha fanlarni o‘zlashtirish uchun lug‘at vositachi ekan, leksikografiyani sintetik fan sifatida baholash mumkin.

Endilikda leksikografiya turli fanlar kesimida rivojlanib, nafaqat nazariy, balki amaliy jihatdan kompyuter texnologiyalari, axborot qidiruv tizimi, tarjimashunoslik, korpus lingvistikasi, kompyuter lingvistikasi kabi qator fanlar uchun asosiy o‘rganish obyekt sifatida integrallashayotganini kuzatish mumkin. Shuningdek, leksikografiya tilni boyitish hamda til sathlarni tizimli o‘rganishda va ularni ilmiy-nazariy jihatdan boyitishda muhim ahamiyat kasb etadi.

## II. Adabiyotlar tahlili

V.V.Dubichinskiy to‘g‘ri ta’kidlaganidek: “Lug‘at - tilning leksik tizimini ifodalashning muhim usullaridan biridir. Boshqa vositalardan farqli o‘laroq, lug‘atning leksikografik talqini, qoida tariqasida, tahlilning ko‘p qirraliligi (tilning barcha darajalari pozitsiyalarining xususiyatlari), to‘liqligi (universalligi, lug‘at tuzilishining keng qamrovliligi) va dalillar bilan tavsiflanadi. Olim leksikografiyani til tizimini modellashtirish bosqichi sifatida ham baholash mumkinligi va har bir leksikograf tilning butun tizimidan lug‘atda tilning muayyan lug‘at qatlami doirasida materialni tizimlashtirishga erisha olishini ta’kidlab o‘tadi. Demak, tilni tizim sifatida qaralsa, lug‘at mana shu tizim tarkibidagi yana bir tizim deb qarash xato bo‘lmaydi.

Ayniqsa, matn<-korpus>-statistika uchligida tekstometrika yo‘nalishi mavjud. Ushbu yo‘nalish Fransiyada 1970-yillarda Pierre Guiraud va Charlez Myuller tomonidan leksik statistika bo‘yicha ishlar birinchilardan bo‘lib amalga oshirilgan. Myuller matndagi lug‘at boyligini baholash va matn lug‘atini tavsiflash borasida o‘zining ilmiy qarashlarini ommaga e‘lon qilgan. Tekstometriya ma‘lumotlarni tahlil qilish metodi sifatida Jen Paul Benzekri tomonidan tilshunoslik sohasida qo‘llanilgan va rivojlantirilgan. Olim matnlar va so‘zlarning sintetik va vizual xaritasini ularning o‘zaro bog‘liq va qarama-qarshi tomonlarini umumlashtirishda korpus tahlilidan foydalangan.

Soha mutaxassislari leksikografiyani fan ichidagi fan deb baho berar ekan, fanning asosi bo‘lmish lug‘atning o‘ziga xos belgisi, ya’ni til hamisha dinamik ravishda o‘zgarib borishi, qolaversa, leksik qatlam jamiyat bilan hamohang shaklda o‘z aksini topishi fanning barqaror emasligini ko‘rsatuvchi omil sifatida baho beradilar. Biroq bizningcha, barcha ilmiy nazariyalar, qarashlar, hatto tasdig‘ini topgan bilimlar ham vaqt o‘tib o‘zgarishi yoki umuman o‘zining ilmiy qiymatini yo‘qotishi sir emas.

Til hamisha kishilik jamiyatning eng ajralmas aloqa vositasi bo‘lib xizmat qilgan. Uning yordamida yangi bilimga erishiladi va egallangan bilim tafakkur va tajriba asosida qayta shakllanadi. Bugun til sanoatida kompyuter lingvistikasi, mashina tarjimasini, til texnologiyasi, tabiiy tilni qayta ishlash, sun‘iy intellekt texnologiyasi kabi terminlar tez-tez qo‘llanilib

kelinmoqda. Bularning har biri insoniyat tomonidan yaratilgan lingvistik resurs orqali o'z tadrijiy takomiliga ega. Endilikda lingvistik resurslarning asosiy turlaridan biri sifatida elektron shaklidagi korpuslar nazarda tutilmoqda.

Korpus lingvistikasi matnlarni yig'ish, tasniflash, annotatsiyalash kabi vazifalarni bajaradi. Korpus mashina uchun tabiiy tilni tushunish jarayonini lingvistik jihatdan aniq va to'g'ri ko'rsata olishi bois tabiiy tilni qayta ishlash (NLP-natural language processing) sohasida uning o'rni muhim ahamiyatga ega. Shu bois ilmiy manbalarda korpus lingvistikasi tabiiy tilni qayta ishlash sohasining (NLP) bir yo'nalishi sifatida qayd etiladi. Ayrim ma'lumotlarda esa kompyuter lingvistikasi yoki amaliy lingvistikaning muayyan sohasi deb qaraladi.

Korpus lingvistikasini kompyuter lingvistikasining asosiy yo'nalishi sifatida e'tirof etish mumkin, chunki matnni qayta ishlovchi dasturlar korpusga asoslanadi. NLP tizimi uchun lingvistik bilimlar juda muhim. Ular qabul qilingan muayyan lisoniy modellar, qoidalar, lug'atlar shaklida ifoda etilsada, biroq tilning nutqiy omili bilan bog'liq bo'lgan ekstralingvistik ma'lumotlarga ham ehtiyoji katta. Shu ma'noda tilning konseptual tasviri aks etgan semantik tarmoqlar, ontologiya ko'rinishidagi bilimlar ta'minoti zarur. Demak, tilning murakkab funksional imkoniyatlarini pragmatik, neyrolingvistik, psixolingvistik yoki diskursiv parametrlarini baholashda korpus texnologiyasi oldida muhim vazifalar mavjud.

Kaliforniya universiteti mutaxassisi Stefan Gries o'zining ilmiy qarashlarida korpus uchun berilgan ta'riflar "*u metod, nazariya yoki modelmi*", degan savolga nisbatan uni metod(ologiya) deb baholaydi. Olim fikrini asoslab, til nazariyasining generativ nuqtayi nazaridan tilga aloqasi yo'q, degan to'xtamga keladi. U deskriptiv va amaliy jihatdan ayrim metodlardan foydalanib, misollar yordamida o'z fikrini isbotlashga harakat qiladi. Unga ko'ra, agar tilshunos biror leksemani korpus ichida grammatik birlik sifatida o'rganmoqchi bo'lsa, demak u grammatik nazariyaga, agar u yoki bu millatning ikkinchi til sifatida murakkab konstruksiyalardan qay darajada qo'llash imkoniyatini baholamoqchi bo'lsa, ikkinchi tilni o'rganish nazariyasiga asoslanadi, ya'ni har bir izlanuvchi o'ziga xos metodni tanlaydi, degan munosabatni bildiradi.

### III. Lingvistik tadqiqotlarda korpusga asoslangan yondashuvlar

Bugungi kunda korpusning imkoniyatlari shu darajaga yetdiki, ushbu sohada erishilgan yutuqlardan nafaqat NLP sohasi yoki kompyuter lingvistikasi, balki tilga o'qitishning metodika va pedagogika sohalari, diskurs tahlil, mashina tarjimai va lingvistikaning o'nlab sohalari (sotsiolingvistika, gendershunoslik, psixolingvistika va h.k.) unumli foydalanib, sohaga doir ijobiy natijalarga erishib kelinmoqda. Shu bilan birga Jurafsky, Martin, Manning, Schütze, Roark, Sproat kabi olimlar nuqtayi nazaricha kompyuter lingvistikasi matnlar korpusini annotatsiyalash va uni lingvistik tahlil qilishda o'zining optimal yechimlarini turli algoritmlar misolida tavsifa qiladi.

Korpus yozma va transkripsiya qilingan og'zaki nutqni lingvistik tahlil etishda asos vazifasini bajaradi. G.Kennedining fikricha, korpus lingvistikasi kompyuterlar rivojlanish davri bilan boshlanmagan bo'lsa-da, matnga asoslangan lingvistik tahlil uchun ma'lumotlar bazasi kompyuterlar keskin ravishda o'sishiga sabab bo'ldi. Korpus turli maqsadlardan kelib chiqib, uning dizayni, hajmi hamda har bir korpusning individual xarakterini shakllantirishga xizmat qiladi. Ko'pchilik lingvistik korpuslar tilning turli leksik, grammatik, diskurs va pragmatik jihatdan tahlil qilish uchun yo'naltiriladi. Jumladan, maxsus maqsadlarga yo'naltirilgan korpuslarda o'quv lug'atida qaysi so'zlar bo'lishi va ularning ma'nolari qaysi kontekstda

xususiylashuvi aks etadi. Shuningdek, korpus jamiyatning muayyan soha vakillari tomonidan qo'llaniladigan lug'at chastotasini aniqlash, tilning aktiv va passiv lug'atini nazorat qilish kabi vazifalarni bajarishga yordam beradi.

Amerikalik filolog Georg Zif matnda qo'llangan so'zlar chastotasi bilan matn uzunligi o'rtasidagi munosabatlarga asoslangan Zif qonunini yaratdi. Shuningdek, lingvistik tadqiqotlarda matnga asoslangan yondashuv Jorjtaun universitetida bo'lib o'tgan muhokamada o'z aksini topdi. Unga ko'ra, korpusning quyidagi yo'nalishlari bo'yicha rivojlantirishga e'tibor qaratildi: 1) og'zaki matn korpusini loyihalashtirish va takomillashtirish; 2) onlayn korpuslarning qidiruv va qayta ishlash jarayoni uchun instrumentlar yaratish; 3) onlayn korpuslar va korpusga asoslangan instrumentlarni tanqidiy baholash; 4) korpusga asoslangan tahlilning metodologik masalalarini muhokama qilish; 5) lingvistika va kompyuter lingvistikasi, diaxron lingvistika, diskurs analiz, uslubiy analiz, tilga o'qitish kabi yondosh sohalaridagi korpusga yo'naltirilgan dasturiy ta'minotlar yaratish va natijalarni o'rganish.

#### **IV. Metodologiya va metodlar**

M.Kurdi e'tiroficha, korpus texnologiyasida lingvistlar uchun empirik tadqiqot olib borish imkoniyatining mavjudligi so'z turkumlarga ajratish teggeri, morfologik analizator, sintaktik parser kabi NLP sohasida yaratilgan ko'plab uskunalar yordamida amalga oshirish imkoniyati mavjud. Kompyuter lingvistikasi va korpus texnologiyasi o'zaro to'ldiruvchi va takomillashtiruvchi sohalar sifatida statistik bilimlarga ega bo'lish orqali yuqori natijalarga erishishda bir-birini taqazo etadi.

Tekstometriya matnlar majmuasini metodologik jihatdan qayta tahlil qilish uchun lingvistik va matematik jihatdan sintagmatik hamda paradigmatic bog'lanishlar, qarama-qarshi xarakterlarni baholashda korpusdan olingan statistik ma'lumotlar natijasiga asoslanadi. NLP sohasida matnlarni qayta ishlash va ma'lumotlarni turli jihatdan tahlil qiluvchi korpusni kompleks tahlil qilish (moslik darajasini tahlil qilish, klasterlash, leksik jadvallar yaratish, murakkab leksik qurilmalarni qidirish, turli o'lchovlarga ko'ra ichki korpuslarni ajratish) uchun samarali vositalardan biri TXM platformasi hisoblanadi.

Mashina tarjimasida so'zlardagi polisemiya va omonimiya kabi muammolarni yechishda ham matnlarni tasniflash va uning statistic ma'lumotlarga tayanish muhim ahamiyat kasb etadi. Olimlar shunday masalalarni hal etishda ikki omilni taklif etadilar. Birinchi omil – kontekstologik lug'at yaratish. Bu tamoyil dastlab AMPAR mashina tarjima tizimi ilmiy guruhi tomonidan tuzilgan. Unga ko'ra, asos tildagi matn konkordansi va unga munosib tarjima (parallel matnlar) asosiga quriladi. Mazkur lug'at aniqlangan matn bilan bog'liq ravishda undagi har bir ko'p ma'noli yoki omonimlarning tarjima qilinayotgan matndagi lingvistik qurshovi aks etadi. Matn turlari mashina tarjima uchun boshqa lingvistik tarkibiy qismlar: frazeologiya, sintaktik strukturalar, morfologik xususiyatini aniqlashga yordam beradi. Ikkinchi omil – tizimini qurish asnosida parallel matnlar konkordansining xususiyatlari bilan tizim imkoniyatlari asos hisoblanadi”.

Shuningdek, Kemal Oflazer ishlarida ingliz tilidan turk tiliga tarjima qilish bo'yicha statistik mashina tarjima tizimi yaratilgan. Unda tillarning morfologik va sintaktik qoidalarigagina emas, parallel matnlardagi ma'lumotlarga tayanilgan. Jumladan, 22500 jumla namunasida so'zlarning matndagi holatiga ko'ra ingliz va turk tillarining grammatik holati aniqlangan.

#### **V. Muhokama**

Til pedagogikasida o'qituvchi til o'rganuvchilarga qaysi so'zlar nechog'lik kerakligi va vaziyatlarda ularni qo'llash mumkinligi haqida ichki sezgi va tajribaga tayanib, ta'lim jarayonini tashkil etadilar. Chastotali lug'atlar o'qituvchi uchun qo'shimcha manba sifatida saralab olingan leksik minimumning nechog'lik to'g'ri tanlangani yoki metodik jihatdan tilga oid

Valent so'zlar	“tortmoq”	Sinonim so'zlar va ma'nodosh so'zlar
<i>na'ra</i>		<i>qilmoq</i>
<i>Azob</i>		<i>chekmoq</i>
<i>e'tibor</i>		<i>qaratmoq</i>
<i>Siyoh</i>		<i>yo'zmoq</i>
<i>jilovidan</i>		<i>ushlamoq</i>
<i>Fotiha</i>		<i>qilmoq</i>
<i>tarozi</i>		<i>o'lchamoq</i>
<i>otasiga</i>	<b>TORTMOQ</b>	<i>o'xshamoq</i>
<i>dorga</i>		<i>osmoq</i>
<i>uh</i>		<i>qilmoq</i>
<i>xomush</i>		<i>bo'lmoq</i>
<i>Bosh</i>		<i>(rad qilmoq)</i>
<i>qo'shin</i>		<i>surmoq</i>
<i>Saf</i>		<i>bo'lmoq (saflanmoq)</i>
<i>Jabrini</i>		<i>ko'rmoq</i>
<i>Domiga</i>		<i>yutmoq</i>
<i>Pichoq</i>		<i>urmoq</i>
<i>o'ziga</i>		<i>qaratmoq</i>
<i>Osh</i>		<i>tarqatmoq</i>
<i>ko'ngil</i>		<i>(xohlamoq)</i>
<i>Loy</i>		<i>surmoq</i>
<i>Jiddiy</i>		<i>bo'lmoq</i>

kompetensiyalarni rivojlantirish usullari uchun yo'riqnoma bo'lib hizmat qiladi.

Yuqorida qayd etilganidek, chastotali lug'at yaratishda korpus manba bo'lib xizmat qiladi. Kilgarrif so'zlarning chastotasini hisoblashning uch usulini ko'rsatadi: nusxa ko'chirish orqali, bashorat qilish, sanash orqali (korpusdan).

Rus tilining chastotali lug'atida 5000 lemma (so'z) bo'lib, quyidagi strukturaga ega edi:

chastotalarning tartib tartibi

- normallashtirilgan chastota (million so'zga)
- bir nechta bo'g'inli urg'uli so'zlar (lemma)
- so'z turkumlari (ot so'z turkumlarini bog' rod kategoriyasi haqida ma'lumot)
- korpusdan olingan so'zlarning ingliz tiliga tarjima qilingan so'zlarning namunasi

## VI. Natijalar

Ilmiy izlanishlarimiz natijasida <http://uzbekcorpus.uz> saytini yaratdik. Ushbu saytda o'zbek tili korpusi yaratilgan bo'lib, ta'limiy korpus ham ichida o'rin olgan. Yuqorida olingan *tortmoq* fe'li orqali o'zbek tili o'quv lug'atini yaratishning namunasini tahlil qildik. Ushbu

holatda *tortmoq* fe'lining ham omonim va ko'p ma'noli jihatlarini korpusdagi matnlardan olingan namunalar asosida kuzatar ekanmiz, ushbu fe'ning o'zi bir vaqtning o'zida boshqa so'zlar bilan kontekstologik sinonimik qatorni ham hosil qiladi.

### VII. Xulosa

Demak, korpusga asoslangan o'quv lug'atlarning kontentini yaratish o'quvchilarda u yoki bu so'zni og'zaki va yozma muloqtda ko'nikma hosil qilishga yordam beradi. bu kabi so'zlikni yaratish o'quvchida qaysi birliklar til birliklari va qaysi birlari nutq birliklari ekanligini farqlashga ham ko'maklashadi. Ushbu fe'ning o'rniga erkin so'z qo'llash mumkin bo'lgan holatlarni o'quvchi belgilab boradi. Natijada o'ziga kerakli birliklarni nutq jarayonida faol qo'llashga harakat qiladi. Toki o'quvchida matnda qo'llanadigan har bir so'zning yosh va auditoriyasiga mos so'zligi shakllantirilmas ekan u matnda anglashilgan mazmunga e'tibor qaratmaydi, qaratgan taqdirda ham tushuninh va uni muloqotda qo'llashga qiyinchilikka duch keladi.

### REFERENCES

1. Ступин Л.П. Лексикография английского языка М.: Высшая школа, 1985. –С. 2.
2. Landau S.I. Dictionaries: The Art and Craft of Lexicography. – N. Y., 1989. –P. 4.
3. Дубичинский В.В. Лексикография русского языка (Учебное пособие) Москва Издательство «Флинта» Издательство «Наука» 2008. –С.8.
4. Дубичинский В.В. Лексикография русского языка (Учебное пособие) Москва Издательство «Флинта» Издательство «Наука» 2008. –С.13-14.
5. <https://txm.gitpages.huma-num.fr/textometrie/en/Introduction/>
6. Mohamed Zakaria Kurdi. Natural Language Processing and Computational Linguistics: Speech, Morphology and Syntax. –Great Britain, USA: Wiley-ISTE 2016. –P. 12.
7. Jurafskiy D., James H. Martin Speech and language processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 2007
8. Gries, Stefan Th. How to use statistics in quantitative corpus analysis. In Michael McCarthy & Anne O'Keeffe (eds.), The Routledge Handbook of Corpus Linguistics. New York & London: Routledge., 2021.
9. Jurafsky, D., and J. H. Martin. 2008. Speech and language processing, 2nd ed. Upper Saddle River: Prentice Hall.; Manning, C., and H. Schütze. 1999. Foundations of statistical natural language processing. Cambridge: The MIT Press.; Roark, B., and R. Sproat. 2007. Computational approaches to morphology and syntax. Oxford: Oxford University Press.; The Oxford handbook of computational linguistics (edited by Mitkov R.) – Oxford, 2003. – P. 63
10. Graeme Kennedy An introduction to corpus linguistics London: Longman, 1998. – P. 2
11. Mohamed Zakaria Kurdi. Natural Language Processing and Computational Linguistics: Speech, Morphology and Syntax. –Great Britain, USA: Wiley-ISTE 2016. – P. 12
12. <https://txm.gitpages.huma-num.fr/textometrie/en/Introduction/>
13. А. М. Лаврентьев, Ф. Н. Соловьев, А. М. Чеповский Внедрение в ТХМ дополнительных инструментов автоматической обработки текста / PROCEEDINGS OF THE INTERNATIONAL CONFERENCE «CORPUS LINGUISTICS–2019», - С. 55.
14. <http://www.nop-dipo.ru/node> Марчук Ю.Н. Типология текстов и машинный перевод.

15. I Iknur Durgar El-Kahlout, Kemal Oflazer Initial Explorations in English to Turkish Statistical Machine Translation / Proceedings of the Workshop on Statistical Machine Translation, pages 7–14, New York City, June 2006. (Association for Computational Linguistics)
16. Serge Sharoff, Elena Umanskaya, James Wilson A frequency dictionary of Russian. Routledge, London and New York, 2013. – P. 5.
17. N. Abdurakhmonova, I. Alisher and R. Sayfulleyeva, "MorphUz: Morphological Analyzer for the Uzbek Language," 2022 7th International Conference on Computer Science and Engineering (UBMK), 2022, pp. 61-66, doi: 10.1109/UBMK55850.2022.9919579.
18. Abdurakhmonova N. The bases of automatic morphological analysis for machine translation. *Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta*. 2016;2 (38):12-7.
19. Abdurakhmonova N, Tuliyeu U. Morphological analysis by finite state transducer for Uzbek-English machine translation/*Foreign Philology: Language. Literature, Education*. 2018(3):68.
20. Mengliev D, Barakhnin V, Abdurakhmonova N. Development of Intellectual Web System for Morph Analyzing of Uzbek Words. *Applied Sciences*. 2021; 11(19):9117. <https://doi.org/10.3390/app11199117>
21. Abdurakhmonova N., Tuliyeu U. and Gatiatullin A., "Linguistic functionality of Uzbek Electron Corpus: [uzbekcorpus.uz](http://uzbekcorpus.uz)," 2021 International Conference on Information Science and Communications Technologies (ICISCT), 2021, pp. 1-4, doi: 10.1109/ICISCT52966.2021.9670043.
22. Kubedinova L. Khusainov A., Suleymanov D., Gilmullin R., Abdurakhmonova N. First Results of the TurkLang-7 Project: Creating Russian-Turkic Parallel Corpora and MT Systems. Proceedings of the Computational Models in Language and Speech Workshop (CMLS 2020) co-located with 16th International Conference on Computational and Cognitive Linguistics (TEL 2020) .2020/11: 90-101