

8-26-2020

CORPUS LINGUISTICS IS A PRIORITY AREA OF MODERN APPLIED LINGUISTICS

Shakhnoza Kakhramonovna Gulyamova
teacher of the department of uzbek linguistics, PhD, BSU

Follow this and additional works at: <https://uzjournals.edu.uz/buxdu>



Part of the [Linguistics Commons](#)

Recommended Citation

Gulyamova, Shakhnoza Kakhramonovna (2020) "CORPUS LINGUISTICS IS A PRIORITY AREA OF MODERN APPLIED LINGUISTICS," *Scientific reports of Bukhara State University*: Vol. 4 : Iss. 4 , Article 9.
Available at: <https://uzjournals.edu.uz/buxdu/vol4/iss4/9>

This Article is brought to you for free and open access by 2030 Uzbekistan Research Online. It has been accepted for inclusion in Scientific reports of Bukhara State University by an authorized editor of 2030 Uzbekistan Research Online. For more information, please contact sh.erkinov@edu.uz.

УДК: 8/ 81.33

**КОРПУС ЛИНГВИСТИКАСИ – ЗАМОНАВИЙ АМАЛИЙ
ТИЛШУНОСЛИКНИНГ УСТУВОР ЙЎНАЛИШИ**

**КОРПУСНАЯ ЛИНГВИСТИКА – ПРИОРИТЕТНОЕ НАПРАВЛЕНИЕ
СОВРЕМЕННОЙ ПРИКЛАДНОЙ ЛИНГВИСТИКИ**

CORPUS LINGUISTICS IS A PRIORITY AREA OF MODERN APPLIED LINGUISTICS

Гулямова Шахноза Кахрамоновна

БухДУ ўзбек тилшунослиги кафедраси ўқитувчиси, ф.ф.ф.д. PhD

Gulyamova Shakhnoza Kakhramonovna

teacher of the department of uzbek linguistics, PhD, BSU

Аннотация. Ушбу мақолада компьютер тилшунослигининг мустақил бўлими – корпус лингвистикаси замонавий амалий тилшуносликнинг асосий ва бугунги кун илми учун муҳим истикболли йўналиш эканлиги ҳақида баён қилинган. Корпус лингвистикасининг моҳияти, мақсад ва вазифалари, дунё тилшунослигида бу борада эришилган натижалар асосланиб, бир қатор олимларнинг илмий қарашлари умумлаштирилган. Ўзбек тили миллий корпусини яратиш фан олдида турган долзарб вазифалардан бири эканлиги таъкидланган ва бу борада мулоҳаза билдирилган.

Таянч сўзлар: компьютер лингвистикаси, корпус, матн корпуслари, тил корпуси, электрон кутубхона, сунъий интеллект, электрон луғат.

Аннотация. В статье описывается самостоятельная отрасль компьютерной лингвистики – корпусная лингвистика, которая является основным и наиболее перспективным направлением современной прикладной лингвистики. Основываясь на сущности, целях и задачах корпусной лингвистики, результатах, достигнутых в мировой лингвистике, обобщаются научные взгляды ряда ученых. Было отмечено, что создание национального корпуса на узбекском языке является одной из актуальных задач, стоящих перед наукой, и были высказаны замечания по этому поводу.

Ключевые слова: компьютерная лингвистика, корпус, текстовый корпус, корпус языка, электронная библиотека, искусственный интеллект, электронный словарь.

Abstract. This article describes an independent branch of computational linguistics - corpus linguistics, which is the main and most promising direction of modern applied linguistics. Based on the essence, goals and objectives of corpus linguistics, the results achieved in world linguistics, the scientific views of a number of scientists are summarized. It was noted that the creation of a national corpus in the Uzbek language is one of the urgent tasks facing science, and comments were made on this.

Key words: computational linguistics, corpus, text corpus, language corpus, electronic library, artificial intelligence, electronic dictionary.

Кириш қисми. Корпус лингвистикаси компьютер тилшунослиги бўлими бўлиб, компьютер технологиялари ёрдамида лингвистик корпус (матнли корпус)нинг қурилиши ва ундан фойдаланишнинг умумий тамойилларини ишлаб чиқиш билан шуғулланади. Лингвистик корпус ёки матнли корпус остида ўзига хос тил муаммоларини ҳал қилиш учун яратилган катта ҳажмдаги машинада ўқиладиган, бирлаштирилган, ёрлиқланган, форматланган филологик жиҳатдан баркамол тил маълумотлари тўплами тушунилади.

Корпус тилшунослиги атамаси биринчи марта 1977 йилда ишлатилган, бу эса корпус тилшунослигини жуда ёш илмий йўналиш деб аташга имкон беради. Аммо шу қисқа вақт ичида корпус тилшунослиги ҳисоблаш тилшунослигининг ёндашувларидан бири сифатида замонавий тилшуносликнинг етакчи йўналишларидан бирига айланди. Россияда бу атама инглиз тилининг ҳалқаро корпуси асосчиларидан бири Сидней Гринбаумнинг маърузалари туфайли фақат 1996 йилга келиб маълум бўлди.

Ҳозирги кунда “корпус” тушунчасига кўплаб таърифлар мавжуд. Масалан, Э.Финеганнинг дарслигида берилган таъриф қуйидагича: корпус, одатда, маърузачи, муаллиф, адресат ёки аудитория ҳақида маълумотларни ўз ичига олувчи машинада ўқиладиган матнлар тўплами ҳисобланади.

Асосий қисм. Тил фаолиятининг замонавий шартлари нафақат нутқ намуналарини, балки алоҳида хронологик ёки фазовий ҳолатлар, нутқ амалиётини тизимлаштиришни ўрганишни ўз ичига олади. Ўрганиш ва умумлаштиришга тизимли ёндашув нутқ амалиётини ўзига хос бирлик сифатида белгилайди. Анъанага биноан фақат доимий, субъектив омиллар таъсирсиз эмпирик материални таҳлил қилиш имконини беради. Аммо бугунги кунда компьютерда тилшуносликнинг объекти учламчи восита ва алоқа воситачилик вақти яратилади, сақланади ва “ҳужжат”да ўқиб бўлмайдиган жойда амалга оширилади.

Албатта, нутқ амалиётини ёритиш бугунги кунда қўлланиладиган объектив мезонлар туфайли реалистик кўриб чиқилади. Бунда янги “лингво-информацион” формат орқали янги универсал-статистик ва махсус фанлар тил форматларини танлашда янги имкониятлар яратади. XX асрнинг 60 йилларида пайдо бўлган корпус тилшунослиги, нутқ амалиёти учун янги методологик асосни, тадқиқотнинг устувор йўналишларини яратди. Корпус лингвистикаси металингвистик восита сифатида тил методологиясини куриш, ундан фойдаланиш ва нутқ нақшларининг лингвистик талқини – компьютер воситалари ёрдамида “эски” ва “янги” тил муаммоларини ҳал қилишда имкон яратади [2].

Жаҳон тилшунослигида компьютер ва корпус лингвистикаси муаммоларини ўрганиш XX асрнинг 40 йилларида бошланиб, бу соҳада дастлабки илмий фаразлар айтилди. Хусусан, ўтган асрнинг 60 йилларида мазкур жараён жадаллашди, XXI аср бошларида ўзида миллионлаб сўзларни акс эттирувчи юзлаб тил корпуслари пайдо бўлди.

Корпус тилшунослиги замонавий амалий тилшуносликнинг устувор йўналиши ҳисобланади. Корпус тилшунослиги замонавий компьютер технологияларидан фойдаланган ҳолда лингвистик маълумотлар корпуслари (матн корпуслари) куришнинг умумий тамойилларини белгилайди, ёзма ва оғзаки матнларнинг ҳақиқий лингвистик ҳодисаларини типлаш методологиясини, шунингдек, уларни сақлаш ва таҳлил қилиш усулларини ишлаб чиқади. Матнли корпуслар билан ишлаш тадқиқотчини субъективлигидан маълум даражада мавҳумлаштиришга ва тилни объектив ўрганишга имкон беради [13].

Википедия корпусни статистик таҳлил қилиш ва гипотезани синаш, вазиятни текшириш ёки муайян соҳалар учун тил қоидаларини асослаш учун ишлатиладиган (ҳозирда электрон шаклда) катта ва тузилган матн тўпламлари сифатида белгилайди. Т. Мак Энери (Т. McEnergy) ва Э. Вилсон (А. Wilson) қуйидаги таърифни беради: корпус – тил модели сифатида фойдаланиш учун аниқ тил мезонларига мувофиқ танланган тил парчалари тўплами [9]. Корпусни феномен сифатида муҳокама қилиб, уни ҳал этишда Д.Синклер, Т.МакЭнери, Д.Байбер, А.Килгариффом, В.П.Захаров, В.А.Плунгян, В.Ш. Рубашкин, А.Н.Барановлар каби олимлар ўз хиссаларини кўшишган.

Жаҳон тилшунослигида корпус соҳасидаги мақсадли тадқиқотлар 40 йилларда Блумфильд, Фрайс ва Бонджерслар томонидан бошланган [4]; Н.Френсис ва Г.Кучера илк марта корпус тузиш принципларини ишлаб чиққан [11]. Рус тилшунослигида В.П.Захаров, А.Б.Кутузов, Е.В.Недошивина, В.В.Риков, В.Плунгянлар корпус, унинг турлари, корпус тузиш ва теглаш тамойиллари борасида тадқиқот олиб боришган [7].

“Тилшунослик ҳар доим муҳим бўлган илмий соҳа захираси. Шуни айтиш мумкинки, тилшунослик усуллари бунга мисол бўла олади ва ҳатто муаммолари бўлган бошқа фанлар учун намуна бўлмоқда, уларнинг сони доимий равишда ўсиб бориб, жами фикрлар йўналиши инсонпарварликни қўллаб-қувватлайди”, – деб 70 йилларда Э.Бенвенист томонидан ифода этилган фикр бугунги кунда ҳам муҳим аҳамият касб этади. Компьютер воситалари “даврида”ги коммуникация (алоқа) кўпгина информацион ва техник жиҳатдан қўллаб-қувватлаш ижтимоий-маданий ривожланишнинг ютуғидир.

Замонавий тилшунослик парадигмасида равон тилшунослик услубий жиҳатдан кўп қирраллиги ва самарадорлиги билан ажралиб туради.

Бугунги кунда корпус тилшунослиги кўпинча тилшуносликда компьютер ва электрон корпусдан фойдаланган ҳолда “реал ҳаётда” тилдан фойдаланишни ўрганишга нисбатан янги ёндашув сифатида тушунилади. Корпус тилшунослиги мустақил интизом позициясини даъво қилиш учун камида иккита хусусиятга эга: 1) ишлатилган оғзаки материалнинг табиати; 2) воситаларнинг ўзига хос хусусиятлари. Агар тилшуносликнинг синтаксис, семантика ва социолингвистик бўлимлари тил тузилишини ёки лингвистик фойдаланишни тавсифлашга ёки баҳолашга қаратилган бўлса, корпус тилшунослиги кенгроқ тушунча, лингвистик тадқиқотнинг кўплаб жиҳатларида қўлланилиши мумкин бўлган методология ҳисобланади. Корпус тилшунослигини баъзида “тилшунослик тадқиқотларининг турли соҳаларидаги усуллар тўплами” деб аташади [6].

Лингвистик таҳлил усули сифатида корпус тилшунослиги уларни қиёсий ўрганиш жараёнида умумий ва тиллар, диалектлар ёки тил вариантлари ўртасидаги алоҳида фактларни аниқлашга қаратилган қарама-қарши тадқиқотлар билан ҳам боғлиқдир [5]. Лингвистик таҳлилнинг кўплаб турлари эмпирик маълумотларнинг мустаҳкам ва кенг базасида энг яхши ишлаб чиқилган. Э.Финеган корпус тилшунослигини тузиш ва ундан фойдаланиш учун зарур бўлган ва тилнинг табиий ишлатилишини ўрганишга йўналтирилган фаолият сифатида белгилайди [6]. Ушбу таъриф корпус тилшунослигининг ижодий йўналишини таъкидлайди. Корпус тилшунослигининг икки томонлама табиати (матн корпорасини яратишда ҳам, фойдаланишда ҳам диққат марказида) унинг объектининг икки томонлама табиати билан белгиланади – бир томондан корпус тилшунослиги ва бошқа томондан, бу корпус тилшунослигининг маҳсулидир.

XIX аср охири XX асрнинг бошларида корпуслар лингвистик тадқиқотлар ўтказиш учун ёки кўпинча амалий муаммоларни ечиш учун (масалан, тил бирликларининг частотасини ҳисоблаш учун) яратила бошланди. Корпус тилшунослиги учун компьютерларнинг ихтиро қилиниши ва кенг қўлланилиши билан ривожланишнинг янги босқичи бошланади – яратилган корпуслар мавжудларидан нафақат сақлаш форматида, балки ҳажмларида ҳам фарқ қилади [3].

Олим В.А.Плунгян ўз вақтида тўғри баҳо берганидек, энди тилни ўрганиш учун одамга иккита эмас, учта нарса керак бўлади: луғат, грамматика ва маълум тилдаги матнларнинг корпуси. У луғат ҳам, грамматика ҳам, умуман олганда, ушбу яшаш майдонидан ташқарида фойдасиздир, чунки тил аслида ишлайди, деб таъкидлайди.

Лингвистик корпус – маълум принципларга мувофиқ, белгиланган (изоҳланган), махсус кидирув тизими томонидан тақдим этилган матнлар тўпламининг вакили. Корпус тилшунослиги тор маънода фақат корпусларни қандай яратиш ва улардан қандай фойдаланиш ҳақидаги илмдир, аммо бу тил илмининг янги мафкураси ролини кўпроқ талаб қилади [13].

Корпус бўйича билдирилган фикрлар, илмий назариялар хусусида куйидагиларни умумлаштириб, қайд этиш мумкин:

- корпус – “тил танаси” деб аталадиган матнлар тўпламининг бир тури. Аммо оддий матнлар тўплamidан фарқ қилади;
- корпус – маълум бир тарзда ташкил этилган оғзаки бирлик;
- корпус – маълум бир тилдан олинган шаклдаги лингвистик маълумотлар тўплами таҳлили учун мавжуд бўлган ёзма сўзлар ёки матнлар;
- корпус – исталган тилда оғзаки ёзи ёзма равишда ёзилган табиий матнлар тўплами.

Шундай қилиб, корпуснинг методологик муҳимлиги В.А.Плунгян томонидан нафақат “мантик”, балки эмпирик амалиётнинг умумлашмалари “замонавий тадқиқотчилар тилшуносликда тобора кўпроқ конструкцияларни эмас, далилларни ўрганишга, кузатиладиган ҳодисаларнинг моделлари эмас, балки хусусиятлари ҳақида гапиришга чақирди. Тилшунослик билан боғлиқ муаммо шундаки, корпус идеология сифатида маълум маънода мафкура, назарий мавҳумликни ташкил этади”.

А.Д.Шмелёвнинг фикри ҳам шундан далолат беради. У тадқиқотчи корпусдан танқидий фойдаланганда, корпусдан олинган маълумотлар самарали бўлади, – деб

таъкидлайди. Демак, корпус – амалга оширишингиз мумкин бўлган тўққизта матнлар тўплами: сўзлар, иборалар, грамматик шакллар ва сўз маънолари тизимлари.

Бугунги глобаллашув даврида ҳар бир халқ, ҳар қайси мустақил давлат ўз миллий манфаатларини таъминлаш, бу борада, аввало, ўз маданиятини, азалий қадриятларини, она тилини асраб-авайлаш ва ривожлантириш масаласига устувор аҳамият қаратиши табиийдир. Ўзбек тилининг халқимиз ижтимоий ҳаётида ва ҳалқаро миқёсдаги обрў-эътиборини тубдан ошириш, униб-ўсиб келаётган ёшларимизни ватанпарварлик, миллий анъана ва қадриятларга садоқат, улуғ аجدодларимизнинг бой меросига ворислик руҳида тарбиялаш, мамлакатимизда давлат тилини тўлақонли жорий этишни таъминлаш [1] мақсадида корпус лингвистикаси истиқболли илмий йўналиши сифатида ўзбек тили миллий корпусини яратиш фанамиз олдида турган долзарб вазифалардан биридир. Бу ўта муҳим ишни амалга ошириш ўз ўрнида корпус лингвистикасининг назарий муаммоларини ечишни тақозо этади.

Тил корпуси ва у билан боғлиқ бўлган фан, яъни корпус лингвистикаси XX асрнинг охири XXI аср бошида тилшунослик ва тил ўқитиш методикасига шиддат билан кириб келган янги йўналишдир. Тил корпусини яратиш ҳамда корпус лингвистикасини ривожлантириш тилшунослик, тил ўқитиш методикаси соҳаларидаги долзарб масалалардан бири ҳисобланади. Корпус лингвистикаси тилшуносликнинг бир бўлими бўлиб, матн корпусларини ишлаб чиқиш билан шуғулланади. Корпус лингвистикаси ривожлантирилса, сунъий интеллектнинг ўзбек тилини “тушуниш, ўқий олиш, қайта таҳрирлаш, унга мос жавоб қайтариш”ини таъминлайдиган дастурлар яратилади (бу тилшунос ва дастурчилар ҳамкорлигидаги амалга ошириладиган долзарб вазифа). Агар ўзбек тили корпуси яратилса, “Google”да ўзбек тилида қидирув тизими пайдо бўлади. Онлайн таржима (интернет таржима) ёки автоматик таржима дастурлари ривожлантирилади. Бу орқали эса ўзбек тилининг турли электрон луғатларини яратиш имконияти кенгайди, тилни дунё сахнасига олиб чиқиш имконияти кенгайди [10].

Ўзбек тилшунослигида корпус, унинг ўзига хос хусусияти, ижтимоий, лексикологик, таълимий аҳамияти, корпус турлари, муаллифлик корпусининг лингвистик қиммати, ўзбек тили муаллифлик корпусини яратишнинг лингвистик асослари масалалари Ш.Хамроева томонидан тадқиқ этилган. Олима ахборот асрига келиб, корпус лингвистикаси жаҳон тилшунослигининг шиддат билан ривожланаётган соҳасига, корпуснинг ўзи эса тилшуносларнинг бекиёс иш қуролига айланганини таъкидлаб, тилнинг миллий корпуси ахборот манбаи сифатида қуйидаги қулайликларни яратганини кўрсатади:

1) корпуслашган тилда яратилган оғзаки, ёзма ёдгорликлар, миллий, маданий мерос намуналари электрон кўринишда Интернет тармоғидан жой олди;

2) корпус табиий (реал) тилнинг электрон шаклдаги, қидирув дастурига жойлаштирилган матнлар йиғиндиси бўлганлигидан бир марта тузилиб, мукамал разметкаланган корпус лингвистик тадқиқотлар самарадорлигини таъминлашда барқарор лингвистик база вазифасини бажаради;

3) корпус электрон кутубхона, луғат, грамматикалар яратишга асос бўлади. У кенг қўламли бўлганлиги учун маълумотнинг ўзига хослигини кафолатлаб, тил ҳодисаларининг барча қирраларини тўлиқ намоиш этишни таъминлайди;

4) турли маълумотлар тил корпусида ўзининг табиий контекстуал шаклида жойлашади, бу эса уларни ҳар томонлама, чуқур, объектив ўрганишга асос бўлади;

5) корпус тилни тадқиқ этиш (сўзнинг ўзгариши, историзм, неологизм вужудга келиши, маъно кенгайиши, торайиши; янги фразеологизмларнинг пайдо бўлишини кузатиш), тил ўрганиш, луғат тузишда энг замонавий, кенг имкониятли дастурлаштирилган тизимдир [12].

Миллий корпус миллий тил хазинаси демакдир. Ундан лингвист, лексикограф, компьютер лингвистлари, дастурчи, муҳаррир, таржимон, журналист, ноширлар, олим, ўқитувчи, таълим олувчилар ва бошқа ҳар қандай соҳа мутахассиси кенг фойдаланади.

Корпус бўйича қидирув фойдаланувчига қуйидагиларни аниқлаш имконини беради:

- 1) белгиланган тил бирлигининг турли қўлланишлардаги барча шаклли кўринишини;
- 2) тилнинг луғат таркибидаги ўрни ва вариантларини;

TILSHUNOSLIK

- 3) белгиланган сўз билан бириктириш имкониятига эга сўзлар рўйхатини;
- 4) у ёки бу муаллифнинг айна сўздан фойдаланиш частотаси ёки статистикасини;
- 5) сўзнинг ўз ва кўчма маъноларини;
- 6) сўз қўлланишининг яширин модели (имконияти)ни;
- 7) тил таракқиётининг турли даврида қўлланиш ҳолатини [8].

Тилшунослик корпуси муайян хусусиятларига кўра яхлит ҳолда (тил, жанр, матн ва ш.к.) ва қидирув тизими билан таъминланган матнлар тўпламидир. Лингвистик корпус иккала ёзма матнлар (газета матни, журналлар, адабий асарлар) ва радио ва телевидение транскрипциясини ҳам ўз ичига олиши мумкин. Корпусни ташкил қилиш жуда хилма-хил бўлиши мумкин. Яратиш мақсадига қараб, корпус таркибидagi матнларни ўз ичига олиши мумкин (маълум бир тил, маълум бир тарихий даврда ёзилган адабий жанрлар).

Миллий тилшунослик корпуси – турли жанрдаги, услубдаги, минтақавий ва ёзма матнлар тилида тақдим этилган ва ўрганиш учун қизиқарли бўлган ижтимоий имкониятлар тили. Миллий тилшунослик корпуси ўқиш учун хизмат қилади. Миллий корпусда бош мақсад жонли тилни кўрсатиш бўлгани сабабли унда турли услуб ва жанрдаги матнлар корпус бирлиги сифатида ишлатилади, миллий корпусда репрезентативлик (матннинг жанрий хилма-хиллиги) устуңлик қилади.

Корпус ҳар бир тилнинг моҳиятини акс эттирадиган кўзгу, имконият экан, ўзбек тилининг миллий ва таълимий корпусини яратиш ва ундан амалий фойдаланиш бугунги куннинг кечиктириб бўлмас вазифасидир.

АДАБИЁТЛАР

1. *O'zbekiston Respublikasi Prezidenti Shavkat Mirziyoevning «O'zbek tilining davlat tili sifatidagi nufuzi va mavqei tubdan oshirish chora-tadbirlari to'g'risida»gi farmoni.* // www.xabar.uz
2. Barkovich A.A. *Korpusnaya lingvistika: spetsifika sovremennix metaopisaniy yazika.* // *Vestnik Tomskogo gosudarstvennogo universiteta.* 2016. № 406. S. 5-13.
3. Barsegyan A.A. *i dr. --Analiz dannix i protsessov: ucheb. posobie /.* 3-e izd., pererab. i dop. SPb.: BXV-Peterburg, 2009. 512 s.: il. + CD-ROM (Uchebnaya literatura dlya vuzov).
4. Blumfeld L. *Yazik.* - M.: Progress, 1968. - 608 s.; Fries Ch.C. *The structure of English. An introduction to the construction of English sentences.* - L., 1969; Bongers H. *The history and principles of Vocabulary control.* - Woerden: WOCOPI, 1947.
5. Gvishiani N.B. *Praktikum po korpusnoy lingvistike. Ucheb. posobie po angliyskomu yaziku.* - M.: Visshaya shkola, 2008.
6. Zaxarov V.P., Bogdanova S.Yu. *Korpusnaya lingvistika.* - SPb, 2013. - 148 str.
7. Kutuzov A.B. *Korpusnaya lingvistika. – (Elektron resurs): Litsenziya Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс)-* // lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf; Nedoshivina Ye.V. *Programmi dlya raboti s korpusami tekstov: obzor osnovnix korpusnix menedjerov. Uchebno-metodicheskoe posobie.* - SPb, 2006. 26 s.; Rikov V.V. *Kurs leksiy po korpusnoy lingvistike.* URL: <http://rykov-cl.narod.ru/c.html>; Plungyan V. *Zachem mi delaem Natsionalniy korpus russkogo yazika? "Otechestvennie zapiski"* 2005, №2. http://magazines.russ.ru/oz/2005/2/2005_2_20-pr.html
8. Mengliev B. va b. *O'zbek tilining milliy korpusi.* // *Ma'rifat.* - 26/04/2018.
9. McEnery T., Wilson A. *Corpus Linguistics.* – Edinburgh: Edinburgh University Press, 2001. – 154 p.
10. Sobirova E. *Korpus lingvistikasi - tilshunoslik sohasidagi yangilik.* // <https://uza.uz/>
11. Frensis N., Kuchera G. *Vichislitelniy analiz sovremennogo amerikanskogo varianta angliyskogo yazika.* - M., 1967; Sinkler D. *Predislovie k knige "Kak ispolzovat korpusa v prepodavanii inostrannogo yazika". / D. Sinkler [Elektronniy resurs].* - Rejim dostupa: <http://www.ruscorpora.ru/corpora-infro.html>, свободный.
12. Xamroeva Sh. *O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari. Filol.fan.bo'yicha falsafa doktori (PhD)...dis.* - Buxoro, 2018. - 250 b.
13. <http://www.tuvancorpus.ru/?q=content/o-proekte>
14. <http://lab314.brsu.by/kmp-lite/CL/CL-Lect/LecCL51.htm>