

БУХОРО ДАВЛАТ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМИЙ ДАРАЖАЛАР БЕРУВЧИ
DSc.03/04.06.2021.Fil.72.03 РАҚАМЛИ ИЛМИЙ КЕНГАШ

БУХОРО ДАВЛАТ УНИВЕРСИТЕТИ

КАРИМОВ РУСТАМ АБДУРАСУЛОВИЧ

ЎЗБЕК-ИНГЛИЗ ПАРАЛЛЕЛ КОРПУСИНИ ТУЗИШНИНГ
ЛИНГВИСТИК ВА ДАСТУРИЙ МАСАЛАЛАРИ

10.00.06 – Қиёсий адабиётшунослик, чоғиштира тилшунослик ва
таржимашунослик

ФИЛОЛОГИЯ ФАНЛАРИ бўйича фалсафа доктори (PhD) диссертацияси
АВТОРЕФЕРАТИ

уЎК: [811.512.133:811.111]`322:004.9

Филология фанлари бўйича фалсафа доктори (PhD) диссертацияси автореферати мундарижаси
Оглавление автореферата диссертации доктора философии (PhD) по филологическим наукам
The contents of the dissertation abstract for a Doctor of Philosophy (PhD) in Philology

Каримов Рустам Абдурасулович

Ўзбек-инглиз параллел корпусини тузишнинг лингвистик ва дастурий
масалалари..... 5

Каримов Рустам Абдурасулович

Лингвистические и программные вопросы составления узбекско-
английского параллельного корпуса 25

Karimov Rustam Abdurasulovich

Linguistic and Software Issues of Compilation of Uzbek-English Parallel
Corpus..... 49

Эълон қилинган ишлар рўйхати

Список опубликованных работ
List of published works 53

БУХОРО ДАВЛАТ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМИЙ ДАРАЖАЛАР БЕРУВЧИ
DSc.03/04.06.2021.Fil.72.03 РАҚАМЛИ ИЛМИЙ КЕНГАШ

БУХОРО ДАВЛАТ УНИВЕРСИТЕТИ

КАРИМОВ РУСТАМ АБДУРАСУЛОВИЧ

ЎЗБЕК-ИНГЛИЗ ПАРАЛЛЕЛ КОРПУСИНИ ТУЗИШНИНГ
ЛИНГВИСТИК ВА ДАСТУРИЙ МАСАЛАЛАРИ

10.00.06 – Қиёсий адабиётшунослик, чоғиштира тилшунослик ва
таржимашунослик

ФИЛОЛОГИЯ ФАНЛАРИ бўйича фалсафа доктори (PhD) диссертацияси
АВТОРЕФЕРАТИ

Бухоро-2022

Фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида B2020.2.PhD/Fil587 рақам билан рўйхатга олинган.

Диссертация Бухоро давлат университетида бажарилган.

Диссертация автореферати уч тилда (ўзбек, рус, инглиз (резюме)) Илмий кенгаш веб саҳифасида (www.buxdu.uz) ҳамда «Ziynet» Ахборот таълим порталида (www.ziynet.uz) жойлаштирилган.

Илмий раҳбар:	Менглиев Бахтиёр Ражабович филология фанлари доктори, профессор
Расмий оппонентлар:	Ахмедов Ойбек Сапорбаевич филология фанлари доктори, профессор Атабоев Нозимжон Бобожон ўғли филология фанлари бўйича фалсафа доктори (PhD)
Етакчи ташкилот:	Термиз давлат университети

Диссертация ҳимояси Бухоро давлат университети ҳузуридаги DSc.03/04.06.2021.Fil.72.03 рақамли Илмий кенгашнинг 2022 йил «22» апрель соат 10:00 даги мажлисида бўлиб ўтади. (Манзил: 200118, Бухоро шаҳри, М.Иқбол кўчаси, 11-уй. Тел.: (0 365) 221-29-14; факс: (0 365) 221-57-27; e-mail: buxdu_rektor@buxdu.uz)

Диссертация билан Бухоро давлат университетининг Ахборот-ресурс марказида танишиш мумкин (125 рақами билан рўйхатга олинган). (Манзил: 200118, Бухоро шаҳри, М.Иқбол кўчаси, 11-уй. Тел.: (0 365) 221-25-87.)

Диссертация автореферати 2022 йил «06» апрель кuni тарқатилди.
(2022 йил «06» апрель даги 22-01 - рақамли реестр баённомаси.)



Д.С.Ўраева

Илмий даражалар берувчи Илмий кенгаш раиси ўринбосари, филол. ф. д., профессор

З.И. Расулов

Илмий даражалар берувчи Илмий кенгаш илмий котиби, филол. ф. н., доцент

М.М. Жўраева

Илмий даражалар берувчи Илмий кенгаш қошидаги илмий семинар раиси, филол.ф.д., профессор

КИРИШ (фалсафа доктори (PhD) диссертацияси аннотацияси)

Диссертация мавзусининг долзарблиги ва зарурати. Жаҳон замонавий тилшунослигида компьютерология ва лингвистик таълимотлар кесишувида юзага келган компьютер лингвистикаси, унинг энг юқори потенциалга эга бўлган алгоритми – корпус лингвистикасига эътибор кучайиб бормоқда. Лисоний билимларни амалда қўллашнинг рефератив усулларидан электрон форматларига ўтиш параллел корпуслар яратиш кўламини ва уларнинг лингвистик-дастурий функционалини кенгайтириш заруриятини белгилайди. Бу, ўз навбатида, вақт тежамини таъминлаган ҳолда тил тизими ва нутқий-фикрий фаолиятнинг таркибий қисми бўлган матн ва унинг турли кўринишларини лингвистик ҳақиқат ракурсида идрок этиш, таржима матнларини чоғиштириш орқали таржимон стратегиялари асосини ташкил этувчи тамойилларни белгилаш, уларнинг турли контекстларда воқеланадиган дискурсив ва кросслингвистик қиёфасини илмий баҳолашга хизмат қилади.

Дунё компьютер лингвистикасида охириги 30 йил давомида параллел корпусларни автоматик таржима платформаси, тезаурус, электрон луғат сифатида қўллаш ва уларни илмий-назарий жиҳатдан ўрганиш ҳаракати авж олди. Шу жараёнда ўзбек-инглиз, инглиз-ўзбек параллел корпусини тузишнинг дастурий ва лингвистик тамойилларини асослаш, танланган бирликларни лингвистик ва экстралингвистик теглаш, параллел корпус тузиш алгоритминини ишлаб чиқиш каби устувор вазифалар морфология, синтаксис, таржимашунослик соҳаларида узуснинг формал ва ноформал регистрини аниқлаш, фақат тилшунос интуициясига таянилган, субъективликка мойил тадқиқотларнинг ишончлилиги ва объективлигини таъминлаш, янги авлод корпус луғатлари ва корпус грамматикаларини яратишга кенг имкон яратади.

Мамлакатимизда глобал илмий-техник ислоҳотлар амалга оширилаётган бугунги кунда компьютер лингвистикаси доирасида автоматик таржима, ўзбек тилини тушуниш ва қайта ишлаш, уни бошқа тиллар билан чоғиштиришда сунъий интеллектнинг ролини белгилаш, икки ва кўп тилли параллел корпуслар яратишга кучли эҳтиёж сезилмоқда. Шу маънода “...давлат тилининг софлигини сақлаш, уни бойитиб бориш ва аҳолининг нутқ маданиятини ошириш; давлат тилининг замонавий ахборот технологиялари ва коммуникацияларига фаол интеграциялашувини таъминлаш”¹ каби устувор йўналишлар таҳлил қилинаётган тил фактларини тизимлаштириш, академик мақсадларда миллий ва хорижий тил кўникмаларини эгаллаш, тилни реал тасвирлаш, шунингдек, амалий тилшуносликда тил қўлланиши эталонларини оптималлаштиришга кенг йўл очади.

Ўзбекистон Республикаси Президентининг 2016 йил 13 майдаги ПФ-4997-сон “Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетини ташкил этиш тўғрисида”ги, 2017 йил 7 февралдаги ПФ-4947-

¹Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 20 октябрдаги «Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида»ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>

сон “Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида”ги, 2019 йил 21 октябрдаги ПФ-5850-сон “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”ги, 2020 йил 20 октябрдаги ПФ-6084-сон “Мамлакатимизда ўзбек тилини янада риволантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида”ги Фармонлари, 2020 йил 24 январдаги Ўзбекистон Республикаси Президентининг Олий Мажлисга Мурожаатномаси, Ўзбекистон Республикаси Вазирлар Маҳкамасининг 2018 йил 18 майдаги 376-сон «Жаҳон адабиётининг энг сара намуналарини ўзбек тилига ҳамда ўзбек адабиёти дурдоналарини чет тилларига таржима қилиш ва нашр этиш тизимини такомиллаштириш чора-тадбирлари тўғрисида»ги Қарори ҳамда бошқа меъриёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу тадқиқот муайян даражада хизмат қилади.

Тадқиқотнинг республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Диссертация республика фан ва технологиялари ривожланишининг I. «Ахборотлашган жамият ва демократик давлатни ижтимоий, ҳуқуқий, иқтисодий, маданий, маънавий-маърифий ривожлантиришда инновацион ғоялар тизимини шакллантириш ва уларни амалга ошириш йўллари» устувор йўналишига мувофиқ бажарилган.

Муаммонинг ўрганилганлик даражаси. Жаҳон тилшунослигида корпус лингвистикаси янги соҳа бўлишига қарамай, корпус соҳасидаги мақсадли тадқиқотлар талайгина. Браун корпуси тузувчилари Н.Френсис, Г.Кучераларнинг корпус тузиш принциплари ҳақидаги фикрлари, Инглиз тили банки лойиҳаси муаллифи Ж.Синклернинг назарий ва амалий таклифларини корпус борасидаги дастлабки тадқиқотлар² сифатида баҳолаш мумкин.

Рус тилшунослигида В.П.Захаров, А.Б.Кутузов, Е.В.Недошивина, В.В.Риков, В.Плунгянлар³ корпус, унинг турлари, ўзига хос хусусияти, корпуснинг ижтимоий аҳамияти, корпус тузиш тамойиллари борасида тадқиқот олиб боришган. Параллел корпуслар В.В.Мамонтова, Д.В.Сичинава томонидан тадқиқ этилган⁴.

Ўзбек тилшунослигида компьютер ва корпус лингвистикаси йўналишида бирмунча тадқиқотлар амалга оширилган. Ўзбек-инглиз

² Френсис Н., Кучера Г. Вычислительный анализ современного американского варианта английского языка. – М., 1967.; Синклер Д. Предисловие к книге “Как использовать корпуса в преподавании иностранного языка”/ Д. Синклер [Электронный ресурс]. – Режим доступа: <http://www.ruscorpora.ru/corpora-info.html>, свободный.

³ Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс)- //lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf; Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. – Санкт-Петербург. – 2006. – 26 с.; Рыков В.В. Курс лекций по корпусной лингвистике. URL: <http://rykov-el.narod.ru/c.html>; Плунгян В. Зачем мы делаем Национальный корпус русского языка? // “Отечественные записки”. – 2005. – №2. http://magazines.russ.ru/oz/2005/2/2005_2_20-pr.html

⁴ Мамонтова В.В. корпус параллельных текстов и база данных для исследования переводческих соответствий: проблемы и процедуры формирования. – Альманах современной науки и образования Тамбов: Грамота. – 2008. – № 8 (15): в 2-хч. Ч.II. – С.117-118. – Адрес статьи: www.gramota.net/materials/1/2008/8-2/50.html; Сичинава Д.В. Европейский перфект сквозь призму параллельного корпуса. – Acta Linguistica Petropolitana. Труды института лингвистических исследований. – 2016. – № 2. – Т.12. – С. 85-114.

параллел корпусини тузиш махсус тадқиқ объекти бўлмаган эса-да, компьютер лингвистикасининг айрим муаммолари, чунончи, ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш⁵, ўзбек-инглиз тили машина таржимасининг лингвистик таъминоти⁶, ўзбек тили муаллифлик корпусини тузиш тамойиллари⁷, ўзбек тили бирликларини графематик таҳлил қилиш муаммолари⁸, тил корпуси лингвистик базасини тузиш тамойиллари⁹, инглиз тили корпусининг функционал хусусиятлари¹⁰ монографик планда ўрганилган. Юқорида кўрсатилган тадқиқотларда параллел корпусларни яратиш масаласи кун тартибига қўйилмаган. Диссертацияни тайёрлаш жараёнида санаб ўтилган илмий изланишлар атрофлича ўрганилди, зарур ўринларда уларга муносабат билдирилди ва улардан тадқиқотда фойдаланилди.

Диссертация тадқиқотининг диссертация бажарилган олий таълим муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Тадқиқот Бухоро давлат университетининг “Хорижий тил таълимини оптималлаштириш муаммоси” мавзусидаги илмий-тадқиқот режаси доирасида бажарилган.

Тадқиқотнинг мақсади ўзбек-инглиз параллел корпусини яратишнинг назарий асосларини ёритиш, унинг лингвистик базасини шакллантириш тамойилларини аниқлашдан иборат.

Тадқиқотнинг вазифалари қуйидагилардан иборат:

параллел корпуснинг умумий хусусиятларини тавсифлаш, корпуснинг бошқа турларидан фарқ қилувчи томонларини очиқ бериш;

параллел корпуснинг хорижий тил лингводидактикаси, лингвистик ва таржима соҳасидаги афзалликларини кўрсатиб бериш;

ўзбек-инглиз параллел корпусига материал танлаш тамойилларини аниқлаш ҳамда ўзбек-инглиз параллел корпуси фрагменти учун материал тўплаш;

ўзбек-инглиз параллел корпуси матнларини лингвистик ва экстралингвистик теглаш йўллариини белгилаш ҳамда матнни параллел теглаш дастурлари учун таклифлар ишлаб чиқиш.

Тадқиқотнинг объекти сифатида ўзбек-инглиз параллел корпуси танланган.

Тадқиқотнинг предметини ўзбек-инглиз параллел корпуси лингвистик базаси ва дастурлаш масалалари ташкил этади.

⁵ Муҳаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. – Тошкент, 2006.;

⁶ Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида). Фил. фан. бўйича фал. доктори (PhD)...дис. афтореф. – Тошкент, 2018. – 52 б.

⁷ Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2018. – 250 б.

⁸ Абжалова М.А. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар таҳрири дастури учун): Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Фарғона, 2019. – 164 б.; Абжалова М. Матнларга автолингвистик ишлов бериш тизимлари // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

⁹ Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.;

¹⁰ Атабоев Н.Б. Инглиз тили корпусининг функционал хусусиятлари (СОСА мисолида): Филол. фан. докт. ... дисс. – Тошкент, 2020. – 189 б.

Тадқиқотнинг усуллари. Тадқиқот мавзусини ёритишда таснифлаш, тавсифлаш, қиёслаш, компонент, дифференциал-семантик, дистрибутив таҳлил усулларидан фойдаланилган.

Тадқиқотнинг илмий янгилиги қуйидагилардан иборат:

параллел корпуслар функционалининг объективлиги ва адекват моҳиятини таъминлашда аслиятдан таржима тилига бевосита ўгирилган матнларни жанрларга мувофиқ танлаш стратегияси ва матн элементларини параллел теглашда таржима трансформациялари чегараларини белгилаш устуворлик қилиши исботланган;

параллел корпуслар ва бошқа турдаги корпуслар матн танлаш, уни электрон шаклга киритиш, матнга метатавсиф бериш, қидирув тизими модулини яратиш, матнни корпусга юклаш, синтактик теглаш каби чизиқли жараёнда *умумийлик*; автоматик ишлов беришда корпус бирлиги сифатида аслият ва таржимадаги матннинг гап, сўз бирикмаси ва сўз даражасида разметкалаш жараёнида *тафовут* касб этиши далилланган;

параллел корпусларни тизимлаштиришда тил бирликлари шакл ва мазмун планлариаро таржима трансформациялари лисоний асимметрияни воқелантириши ва таржима жараёнида шакл трансформацияси ҳисобига мазмунни сақлаб қолиш тенденциясининг етакчилиги асосланган;

ўзбек-инглиз параллел корпусини яратишда экстралингвистик сатҳ сифатида намоён бўладиган маданий кодланган тил бирликлари – реалияларнинг таржимасини таъминлашда транслитерация усулининг самарали эканлиги аниқланган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

параллел корпус учун ўзбек тилидаги матнларни танлаш тамойиллари аниқланган, параллел корпус лингвистик таъминоти сифатида тарбиявий аҳамиятга эга бўлган матнлар тўпланган;

параллел корпуснинг лингвистик таъминоти сифатида инглиз тилидаги маънавий-ахлоқий аҳамиятга эга матнлар электрон кутубхонаси шакллантирилган;

ўзбек ва инглиз тили параллел корпуси лингвистик базаси шакллантирилган, дастурлаш учун лингвистик таъминот тайёрланган.

Тадқиқот натижаларининг ишончлилиги республика ва халқаро миқёсдаги илмий-услубий, илмий-амалий конференциялар материаллари, ОАК рўйхатидаги махсус журналлар ҳамда хорижий илмий журналларда чоп этилган мақолалар, назарий фикр ва хулосаларнинг амалиётга жорий этилганлиги, олинган натижаларнинг ваколатли ташкилотлар томонидан тасдиқланганлиги билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқотнинг илмий аҳамияти ўзбек тили параллел корпусларини яратишнинг назарий асосларини ишлаб чиқиш, компьютер лингвистикаси йўналишида тадқиқотлар яратишда илмий-назарий манба сифатида, шунингдек, ёзма таржима тамойилларини оптималлаштириш муаммосига янги ёндашувларни шакллантириш учун хизмат қилиши мумкинлиги билан белгиланади.

Тадқиқот натижаларининг амалий аҳамияти ишдаги назарий умумлашма ҳамда хулосалардан олий ва ўрта махсус таълим тизимида, умумтаълим

мактабларида ёшларга назарий маълумотлар беришда, ундан корпус лингвистикаси, компьютер лексикографиясининг фан сифатида ўқитилиш жараёнида дастур, режалар тузиш ҳамда мавзуларни баён этишда, профессионал таржимонларни тайёрлашда, хорижий тил таълимида услубий восита сифатида фойдаланиш мумкинлиги билан белгиланади.

Тадқиқот натижаларнинг жорий қилиниши. Ўзбек-инглиз параллел корпусини яратишнинг назарий асосларини ёритиш, унинг лингвистик базасини шакллантириш тамойилларини аниқлаш асосида:

Ўзбек-инглиз параллел корпуси учун ўзбек тилидаги матнларни танлаш принциплари, инглиз тилидаги матнларни танлаш афзалликлари, ўзбек-инглиз параллел корпуси матнларини лингвистик ва экстралингвистик теглаш принциплари билан боғлиқ натижалардан I-ОТ-2019-42 «Ўзбек ва инглиз тилларининг электрон (инсон қиёфаси, феъл-атвори, табиат ва миллий тимсоллар тасвири) поэтик луғатини яратиш» мавзусидаги инновацион лойиҳада фойдаланилган (Олий ва ўрта махсус таълим вазирлигининг 2020 йил 20 ноябрдаги 89-03-4852-сон маълумотномаси). Натижада ўзбек ва инглиз тилларида инсон қиёфаси, феъл-атвори, табиат ва миллий тимсоллар тасвирини билдирувчи сўзларнинг контекстдаги маъносини автоматик аниқлашга эришилган;

Ўзбек-инглиз параллел корпуси лингвистик базасини шакллантириш ҳамда дастурлаш учун лингвистик таъминот тайёрлаш билан боғлиқ хулосалардан Истанбул университети таълим жараёнида “Тилшунослик”, “Компьютер тилшунослиги”, “Корпус тилшунослиги” ҳамда “Лингвистик типология” фанларини дастурлаш учун лингвистик таъминот тайёрлашда фойдаланилган (Истамбул университетининг 2020 йил 20 майдаги маълумотномаси). Натижада ўзбек ва инглиз параллел корпус дастурлаш программаларини яратишдаги хусусиятларнинг умумий таҳлили, жаҳонга машҳур мутахассисларнинг бу борадаги мулоҳаза ва хулосалари изоҳи талабаларнинг ўзбек ва инглиз параллел корпус ҳақидаги назарий тушунчаларини бойитишга хизмат қилган;

Ўзбек-инглиз параллел корпусининг лингвистик ва таржима соҳасидаги афзалликлари, ўзбек-инглиз параллел корпусига материал танлаш тамойилларига доир хулосалардан Лал Баҳадур Шастри номидаги ҳинд маданият марказида «Ўзбек ва инглиз тили параллел корпуси лингвистик базасини шакллантириш ҳамда дастурлаш учун лингвистик таъминот муаммолари» мавзусида қилинган маърузада фойдаланилган (Лал Баҳадур Шастри номидаги ҳинд маданият марказининг 2020 йил 18 сентябрдаги маълумотномаси). Натижада марказ мутахассисларининг ўзбек ва инглиз тили параллел корпуси лингвистик базасини шакллантириш ҳамда дастурлаш учун лингвистик таъминот ҳақидаги билимлар базасини янада кенгайтиришга асос бўлиб хизмат қилган;

Ўзбек-инглиз параллел корпусининг умумий хусусиятлари, бошқа корпуслардан фарқли томонлари хорижий тилда берилишидаги факторларга доир хулосалардан Бухоро вилоят телерадиокомпаниясининг “Ассалом, Бухоро”, “Everyday English”, “Илм дунёси” туркум кўрсатувлари сценарийсини тайёрлашда фойдаланилган (Бухоро вилоят телерадиокомпаниясининг 2020 йил 18 августдаги 1/198-сон

маълумотномаси). Натижада ўзбек-инглиз параллел корпусининг ўзбек тилидаги матнларни тузишда кенг қўлланилиши, унинг интеграциялашув жараёни таҳлил қилинганлиги кўрсатув мухлисларининг миллий тилга бўлган ҳурмати ва чет тилини ўрганишга бўлган қизиқиши ошишига сабаб бўлган.

Тадқиқот натижаларининг апробацияси. Тадқиқот натижалари 7 та илмий-амалий анжуманда, шу жумладан, 2 та халқаро ҳамда 5 та республика анжуманларида апробациядан ўтказилган.

Тадқиқот натижаларининг эълон қилинганлиги. Диссертация мавзуси бўйича жами 18 та илмий иш нашр қилинган. Ўзбекистон Республикаси Олий аттестация комиссиясининг докторлик диссертациялари асосий илмий натижаларини чоп этиш тавсия этилган илмий нашрларда 8 та илмий мақола, шулардан, 4 таси республика ҳамда 4 таси хорижий журналларда эълон қилинган.

Диссертациянинг тузилиши ва ҳажми. Диссертация кириш, уч асосий боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловадан иборат. Умумий ҳажми 142 саҳифани ташкил этади.

ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ

Тадқиқотнинг **“Параллел корпус лингвистик ва лингводидактик база сифатида”** деб номланган биринчи бобида параллел корпуснинг умумий характеристикаси, тузилиши ва таркиби, параллел корпуснинг лингвистик, хорижий тил лингводидактикасидаги аҳамияти, таржимонлик соҳасидаги имкониятлари, мавжуд параллел корпуслар (рус-инглиз, рус-хитой, рус-француз поэтик параллел корпуси¹¹)нинг тавсифи, тузилиши ва аҳамияти ҳақида баҳс юритилади.

“Параллел корпуснинг умумий характеристикаси, тузилиши ва таркиби” деб номланган параграфда параллел корпуснинг умумий тавсифи келтирилади. Параллел корпус (Parallel Corpus) – параллел таржима матнларининг электрон аналоги; кўплаб «оригинал матн ва уларнинг бир ёки бир неча таржимаси» блокларидан иборат. Корпусдаги электрон матнлар оригинал матннинг ўзи ёки унинг бир қисми бўлиши мумкин¹². Шунингдек, параллел корпус оригинал матн ва унинг таржимаси йиғилган корпус саналади¹³. Параллел корпуснинг тузилиши, таркиби ва имкониятлари рус корпус лингвистикасида атрофлича ўрганилган¹⁴. Параллел корпус – асл ва таржимадаги электрон матнлар йиғиндисидан иборат корпус. Асл ва таржима

¹¹ <http://www.nevmenandr.net/fr/>

¹² Сосонина Е.П. Параллельные корпуса в обучении языку и переводу // http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_education_translation/

¹³ <https://translationrating.ru/glossary/parallelnyj-korpus/>

¹⁴ Добровольский Д. Какие задачи решают параллельные корпуса? // <https://postnauka.ru/video/>;
Добровольский Д.О. Корпус параллельных текстов в исследовании лексической семантики // <http://www.dialog-21.ru/media/2503/dobrovolskij.pdf>; Тао Ю., Захаров В.П. Разработка и использование параллельного корпуса русского и китайского языков // Автоматизация обработки текста. (Электрон ресурс): <https://www.researchgate.net/publication/306911853>; Кокорева А.А. Корпус параллельных текстов в обучении иностранному языку. – Вестник ТГУ (Тамбовский государственный университет), выпуск 2 (118), 2013. – С. 57-62.; Соснина Е.П. Параллельные корпуса в обучении языку и переводу (Электрон ресурс): http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_education_translation/

матнлар шунчаки ёнма-ён қўйилмайди, балки бу икки матндаги гап (синтактик бирлик) лар мазмунан мос келиш даражасида бўлакка бўлиниб, мана шу бирликлар ёнма-ён, бир-бирига ишора/ҳавола қилиш ҳолатида туради. Аслиятдаги матн фрагментига мос келувчи таржимадаги фрагмент белгилаб қўйилган бўлади. Айнан шу ҳолат бундай корпуслар ёрдамида турли лингвистик амаллар бажаришга имкон яратади¹⁵.

Замонавий корпус лингвистикасида параллел корпуснинг кўп тилли корпуслар (Multilingual Corpora) ҳамда таржима корпуслари (Translation Corpora) каби кўринишлари мавжуд. Бундай хусусиятга эга корпуснинг структур таркиби унинг мақсадидан келиб чиқиб: 1. таржимага ҳавола қилувчи одатий матн тарзида; 2. қиёслаш учун қулай бўлган «ойнадаги матнлар» шаклида; 3. маълумотлар базаси кўринишида бўлиши мумкин. Тузилиши ва матнларнинг жойлашуви, бирликларининг мослаштирилишига кўра параллел корпуснинг бир неча кўриниши мавжуд¹⁶: бир йўналишли, икки йўналишли, кўп йўналишли (мас., инглиз тилидан рус, немис, француз тилларига таржима қилинган матн).

Параллел корпус таржима тилининг структур тузилиши, ибора ва сўзларнинг маълум контекстдаги имкониятларини кузатишга хизмат қилади. Бундай корпуслар қиёсий тилшуносликда (икки тил структурасини таҳлил қилиш мақсадида); таржима соҳасида (оригинал матн эквивалентини бошқа тиллардан қидириш мақсадида); автоматик таржимада; лингводидактикада; лексикографияда кенг қўлланади.

Корпуслар таркибидаги материал тилига кўра, дастлаб, бир тилли ва кўп тилли корпус турларига бўлинади. Икки ёки кўп тилли корпуслар: 1) параллел ёки таржима корпуслари; 2) қиёсий корпус каби кўринишга эга.

Параллел корпуслар қандай мақсадни кўзлашидан қатъи назар таркибидаги бирликлар мослаштирилган бўлиши лозим. Мослаштириш битекст дастури асосида амалга оширилади. Параллел матнларни тўғрилаш – аслият ва таржима матнидаги бир-бирига тўғри келувчи гап, абзацларни идентификациялаш жараёни ҳисобланади. Матн қисмларини бундай мослаштириш турлича бўлиши мумкин. Қисмлар муаллиф мақсадига қараб ҳар хил бўлади¹⁷: боб, бўлим, абзац, гап ва ҳ. Фойдаланувчи учун энг қулай мослаштирилган шакл бу аслият ва таржима матнларининг гаплар бўйича мувофиқлаштирилиши ҳисобланади. Энг мукамал ишланган автоматик таржима дастури ҳам гапларнинг чегарасини аниқлашда аниқ ечим беролмайди. Асл матндаги фикр бир, икки ёки уч гап билан ифодаланиши мумкин ёки аксинча. Шунинг учун битта буйруқ билан аслият ва таржима матн қисмларини тенглаштиришнинг иложи йўқ. Параллел корпус таркибидаги матннинг гаплар бўйича мослигига эришиш турли лингвистик амалларни бажаришни тақозо этади. Таржима жараёнида аслиятдаги матннинг жумласи таржима матнда икки жумлага бўлиниши, аксинча,

¹⁵ Добровольский Д. Какие задачи решают параллельные корпусы // <https://postnauka.ru/video/54851>

¹⁶ <https://translationrating.ru/glossary/parallelnyj-korpus/>

¹⁷ Добровольский Д. Какие задачи решают параллельные корпусы. (Электрон ресурс): <https://postnauka.ru/video/54851>

оригинал матндаги бир неча гап таржимада битта гап ҳолида берилиши, гапларнинг ўрни алмаштирилиши мумкин.

Параллел корпусни қиёсий корпусдан фарқлаш лозим. Қиёсий корпусда турли тилда бир мавзуда яратилган матнларни қиёслаш мақсад қилинади. Бундай матнлар ҳам муайян тадқиқий масалаларнинг ечими учун хизмат қилади. Турли тилларда маълум бир дискурс вазияти, у ёки бу мавзудаги воқеа қандай ифодаланишини кузатиш лозим бўлганда, қиёсий корпусга мурожаат қилинади.

В.И.Мельник параллел ва қиёсий корпус фарқини шундай изоҳлайди¹⁸:

Параллел корпусга хос хусусият	Қиёсий корпусга хос хусусият
ПК қиёсий корпусдан фарқли равишда икки тил материални ўзида сақлайди.	ҚКдаги матнлар бир тилга мансуб бўлади.
ПК аниқ, мукамал разметкаланган бўлади.	ҚКда разметка назарда тутилмайди.
ПКда таржима изчиллигига эришилади.	ҚК таржима жараёни билан боғлиқ эмас.

“Параллел корпуснинг лингвистик, хорижий тил лингводидактикасидаги аҳамияти ва таржимонлик соҳасидаги имкониятлари” номли иккинчи бўлимда ушбу корпусларнинг таълимий аҳамияти ва таржимадаги имкониятлари тадқиқ этилади. Параллел корпуслар таълим жараёни учун қулай таълимий ресурс вазифасини ўтай олади. ПКнинг дидактик хусусияти сифатида (а) кўп тиллилик, (б) қидирув натижасининг контекст шаклида эканлиги, (в) тилнинг қўлланиш доирасини чегаралаб кўрсатиш имкониятларини кўрсатиш мумкин. У талабага махсус мақсадда қилинган таржимани ўргатиш, сўз маъносини яхшироқ аниқлаш, сўзнинг аниқ нутқий шароитда қўлланишини кузатиш асосида лексик кўникмани шакллантириш имкониятини беради. Параллел корпус таржима матнида синонимик қатордаги сўзлардан қайси бирини қўллаш ўринли эканлигини аниқ кўрсата олади¹⁹. Параллел корпусларнинг фақат таржима назарияси, икки тилли лексикография, тилларни қиёслаш билан боғлиқ тилшуносликнинг бошқа соҳалари учун фойдали эканлиги ҳақидаги хулоса билан чегараланиш ноўрин²⁰. Қиёсланаётган тиллар лексикасидаги сўзларнинг семантик таркибини янада чуқурроқ тавсифлашда ҳам муҳим аҳамият касб этади.

Корпус хорижий тил таълимида мавжуд материални таҳлил қилиш ва камчиликларни аниқлашда фойдали. Бу умумметодологик характерга эга бўлиб, у ёки бу структура/лексика мавжуд типик дарслик, луғат ҳамда корпус асосида таҳлил қилинади; кейин иккала манбадаги натижа таққосланади. Демак, параллел корпус лингвистик тадқиқотлар учун материаллар массивини тақдим этувчи замонавий ўқув-услугий восита саналади.

Big, large ва *great* сўзларини корпусдан кузатамиз. Тезаурусларда бу сўзлар катталикни билдирувчи маънодош сўз сифатида қаралади. Рус-инглиз параллел корпусидан олинган мисолларда эса улар қуйидагича ишлатилган:

¹⁸ Мельник В.И. Переводческие корпусы – что они могут сказать нам о переводе и его преподавании? (Электрон ресурс): <https://cyberleninka.ru/article/n/perevodcheskie-korpusy-chto-oni-mogut-skazat-nam-o-perevode-i-ego-prepodavanii>

¹⁹ Добровольский Д. Какие задачи решают параллельные корпусы? (Электрон ресурс): <https://postnauka.ru/video/54851>

²⁰ Добровольский Д.О. Корпус параллельных текстов в исследовании лексической семантики. <http://www.dialog-21.ru/media/2503/dobrovolskij.pdf>

Следует сказать, что Дидро не верил вначале в возможность создания в Петербурге большой картинной галереи.	<i>It should be noted that Diderot first did not believe in the possibility of collecting a big picture gallery in Petersburg. [28]</i>
Но все больше фактов свидетельствуют о том, что самые большие опасности заключает в себе экипаж.	<i>But there is increasing evidence to show that one of the greatest hazards lies in the crew itself.[30]</i>
За годы своего существования он превратился из дворцовой галереи в музей, выделяющийся в ряду крупнейших музеев мира разносторонностью своего собрания.	<i>From a palace's picture gallery it went down to our time as one of the world's largest museums distinguished for versatility of its collection.[28]</i>

Келтирилган мисоллар тахлили шуни кўрсатадики, *large* сўзи кўпроқ *катталиқ* маъносини, аниқ предметни номловчи сўзнинг сифатини билдиради, *big* ва *great* сўзларида бу вазифа чегараланмаган. *Large* сифатига нисбатан *big* ва *great* сўзлари турғун бирикмалар таркибида кўпроқ учрайди. *Great* кўпроқ ҳис-туйғуни ифодалаш учун ишлатилади: *great weariness, great feeling*. Рус-инглиз параллел корпуси синоним сўзларнинг қўлланиш ҳолатини аниқ кўрсатиб беради. Луғатларда эса бундай маълумотлар мавжуд эмас. Улар орасидаги яна бир фарқ бу лексемаларнинг сўз бирикмаси таркибида учрашида. *Big noise, big head* каби бирикмаларда сифатни унинг синонимлари билан алмаштириш ўринсиз. Ҳамма вақт ҳам мазмунни сақлаган ҳолда сўзни синоними билан алмаштиришнинг иложи йўқ.

Хулоса ўрнида шуни таъкидлаш жоизки, замонавий таржима таълимига корпус методи асосидаги ёндашув янги-янги имкониятларни олиб киради. Миллий корпусдан фарқли равишда бундай корпус катта массивли материал асосида аниқ бир соҳадаги таржима хусусиятларини ўрганишга замин ҳозирлайди. Илмий-техник, ижтимоий-сиёсий, иқтисодий ва бошқа соҳаларга оид параллел корпуслар яратиш айнан шу соҳада тилнинг услубий имкониятларини махсус тадқиқ қилишга имкон беради.

Иккинчи боб “**Ўзбек-инглиз параллел корпусини тузишнинг назарий асослари**” деб номланган бўлиб, унда ўзбек-инглиз параллел корпуси учун материал танлаш тамойиллари, ўзбек-инглиз ПКни лингвистик ва экстралингвистик теглаш масалалари хусусида фикр юритилган.

“Ўзбек-инглиз параллел корпусига материал танлаш тамойиллари” деб номланган биринчи параграфда ПКга матн танлаш тамойиллари тадқиқ этилган. Рус-инглиз параллел корпусини яратишда Н.В.Гоголнинг “Петербург янгиликлари”²¹, Ф.М.Достоевскийнинг “Жиноят ва жазо”²², А.П.Чехов ҳикоялари²³нинг Констанс Гарнетт томонидан қилинган нисбатан эски таржималарининг электрон шаклидан фойдаланилган. Демак, корпус яратувчиси матн танлашда электрон форматга эга, тайёр таржима асарларига

²¹ The Collected Tales of Nikolai Gogol / translator Pevear R., Volokhonsky L. New York: Pantheon Books, 1998. – 435 p.

²² Crime and Punishment by Fyodor Dostoyevsky [Электронный ресурс] // Project Gutenberg: [сайт]. URL: <http://www.gutenberg.org/ebooks/2554>

²³ The Lady with the Dog and Other Stories by Anton Pavlovich Chekhov [Электронный ресурс] // Project Gutenberg: [сайт]. URL: <http://www.gutenberg.org/ebooks/13415>

таянади. Ю.М.Мухин, И.Янларнинг фикрига таянган ҳолда манба танлашда куйидаги омилларни ҳисобга олиш лозим деган хулосага келдик:

1. Таржима сифатини яхшилаш учун таржима матни машхур, малакали мутахассис ёки расмий ташкилот томонидан таржима қилинган бўлиши лозим. Ушбу корпусга Хитой Халқ Республикаси ҳукуматининг расмий сайти²⁴ материал манбаси сифатида хизмат қилган. Ушбу сайт материаллари таржимаси Central Compilation & Translation Bureau компанияси томонидан амалга оширилади; бу эса таржима сифатининг кафолатланганлигидан далолат беради.

2. Танланаётган материал структур-семантик жиҳатдан барқарор бўлиши лозим. Масалан, Бош вазирнинг ҳар йилги “Маъруза”ларида такрор учрайдиган сўз/жумла/бириқма, ҳатто матн структураси бўлади. Бу табиий ҳол. Бундай такрор бирликлар аслият ва таржима матни таржимаси ҳамда кейинги босқичда автоматик разметка учун муҳим.

3. Бадиий асар таржимасидан фарқли равишда расмий-идоравий, дипломатик, бошқарув ҳужжатлари таржимасида аслиятдаги матн мазмуни ҳамда структураси сақланиши лозим. Шу сабабли бундай оригинал ва таржима матни гап/жумлалари кетма-кетлиги бир хил бўлади; структур жаҳатдан ўзаро бир хил тузилади, ўхшаш бўлади. Шундай усул билан икки тилдаги матн бирликларини тенглаштириш вазифаси енгиллашади.

Бизнингча, ўзбек-инглиз параллел корпуси учун матнларни танлашда ҳам эл-юртни кўз қорачиғидай сақловчи ажойиб қаҳрамонлар мадҳ қилинган (1); аёлларнинг ҳақ-ҳуқуқлари ҳимоя қилинадиган (2); кишилар характеридаги ярамас одат, номаъқул иллатлар танқид остига олинадиган (3); мардлик, эпчиллик, довюраклик, меҳнатсеварлик, ҳалоллик, вафодорлик, саҳийлик ғоялари улуғланадиган (4) мазмундаги матнларга устуворлик бериш ўзбек-инглиз параллел корпусида тарбиявий, ахлоқий-маданий омилларнинг эътиборга олинишини таъминлайди; ПК материаллари тарбия/маданиятимизга зид бўлган мазмундаги контентдан холи бўлади.

“Ўзбек-инглиз параллел корпуси матнларини лингвистик теглаш муаммолари” деб номланган иккинчи параграфда лингвистик теглар муаммолари хусусида баҳс юритилган. Параллел корпусга автоматик ишлов бериш бошқа корпуслардан фарқ қилади: автоматик ишлов беришнинг асосий техник муаммоси корпус бирлиги сифатида аслият ва таржимадаги матнни гап, сўз бириқмаси ҳамда сўз даражасида разметкалаш ҳисобланади. НПдаги автоматик қидирув модели асосида буни амалга ошириш қийин²⁵. Европа тиллари, хитой ва япон каби катта лингвистик базага эга бўлган тилларнинг махсус параллел корпуси учун аслият ва таржима матни бирликларини тенглаштиришнинг қулай усуллари ишлаб чиқилган²⁶.

Корпус тузишдаги иккинчи муҳим босқич *матнни гап доирасида тенглаштириш*дир, яъни параллел корпусда икки тилдаги гап мазмунан ўз эквиваленти билан мослаштирилган бўлади. Матн бирликлари параллел

²⁴ <http://cn.theorychina.org/>

²⁵ Dzemyda G., Sakalauskas L. Optimization and knowledge-based technologies // Informatica. 2009. Vol. 20(2). – P. 165-172.

²⁶ Tiedemann J. News from OPUS – A Collection of Multilingual Parallel Corpora with Tools and Interfaces // Recent Advances in Natural Language Processing. 2009. Vol V. – P. 237-248.

корпусда сўз, сўз бирикмаси ва гап бўлиши мумкин. Таълимий мақсадда тузилаётган ПК учун гаплар чегарасида тенглаштиришнинг ўзи етарли бўлади. Чунки кейинги босқичда тенглаштириш ушбу дастлабки босқичсиз амалга ошмайди. Корпуснинг тенглаштириш бирлиги гап бўлганда, соф структур (гапнинг узунлиги, сўз миқдори) ҳамда статистик (ифодалаётган сўзи асосида) метод қўлланади. Бундай усул мукамал луғат базасини талаб этмайди ва катта луғат ресурсига эга бўлмаган тиллар учун қўлланиши мумкин. Параллел гапларни тенглаштиришда гапдан кичик бирликни машина таржимасининг статистик моделлари асосида амалга ошириш мумкин²⁷: аслиятдаги исталган гап элементи – сўзни иккинчи тилга таржима қилиш имкони мавжуд. Параллел корпусда оригиналдаги сўзнинг иккинчи тилга таржимаси унинг частотасига боғлиқ: параллел корпусда иккаласининг ҳам частотаси бир хил бўлиши зарур.

Аслият матнидаги битта гапнинг таржимада икки гап сифатида ифодаланишига типик мисол келтирамиз: “Нима гап, касалмисиз?” \diamond “What’s this?” “Are you ill?”²⁸. Бундай хатоликлар лексикани таҳлил қилиш жараёнида осонгина тузатилади. Аммо нисбатан мураккаб жараёни бор: бир неча гап 2-3 гап билан ифодаланади, аммо уларнинг чегараси мос келмайди: “Эҳ, у ухляпти-ку! – хитоб қилди у, – у барибир ухляпти!” \diamond “My goodness; how he sleeps! – she cried indignantly: And he is always asleep”. Бундай ҳолда гапларнинг чегарасини белгиламай тенглаштиришга тўғри келади. Қуйида Н.В.Гоголнинг “Петербурҗ ҳикоялари”²⁹ таржимасида сўзларни тенглаштириш босқичида учраган шундай “янги жуфтлик”ларни келтирамиз: jinoyat \diamond evil deed; ko'rsatma bermoq \diamond enjoin; muzlatilgan \diamond frozen; олиб бормоқ \diamond raise; қоровул \diamond watchmen; қабул қилмоқ \diamond conceive; ўз \diamond your; қолдирмоқ \diamond neglect; талаб қилмоқ \diamond compel; ҳаракат \diamond impulse; ўйланмоқ \diamond begin to think; таништирмоқ \diamond recur; рангпар \diamond poor; эшитди \diamond hearkened; вақтичоғлик \diamond divert; жойлашиши \diamond state; ёқимли \diamond delightful; маълумки \diamond as every one knows³⁰.

Синтактик таҳлил босқичи. Сўз даражасида тенглаштиришнинг кейинги босқичи аслият ва таржима гапда синтактик таҳлилни амалга оширишни талаб қилади. Сўнги пайтларда разметкаланган корпусларда синтактик таҳлилни йўлга қўйиш учун кўп ҳаракат қилинмоқда. Бундай таҳлил грамматик назарияни текширишга ёрдам беради, шунингдек, янги синтактик қонуниятларни очади. Бундай жараён формал грамматика “машқи” деб аталади: бунда таҳлил натижаси сифатли бўлади. Грамматика машқида «treebank» номи билан кенг қўлланивчи синтактик разметкаланган

²⁷ Brown P.F, Della Pietra V.J., Della Pietra S.A., Mercer R.L. The mathematics of statistical machine translation: parameter estimation // Computational Linguistics. 1993. Vol. 19(2).

²⁸ Қўштирноқ белгиси гапларнинг ажралган ҳолатини кўрсатади.

²⁹ The Collected Tales of Nikolai Gogol / translator Pevear R., Volokhonsky L. – New York: Pantheon Books, 1998. – P. 435.

³⁰ Кедрова Г.Е., Потемкин С.Б. Использование корпуса параллельных текстов для пополнения специализированного двуязычного словаря // Труды и материалы III Международного Конгресса исследователей русского языка «Русский язык: исторические судьбы и современность» (Москва, 20-23 марта 2007). – М.: МАКС Пресс, 2007. – С. 27-628.

корпус тузилади³¹. Инглиз, умуман, роман-герман тилларида модел ярата олувчи айрим мукамал парсерлар мавжуд. Афсуски, ўзбек тили параллел корпусини яратишда бу парсердан фойдаланиб бўлмайди, чунки бундай турдаги парсер ўзбек тили синтактик моделларни шакллантира олмайди.

Бобнинг “*Ўзбек-инглиз параллел корпуси матнларини экстралингвистик теглаш хусусида*” деб номланган учинчи бўлимида экстралингвистик теглаш масалалари тавсифланган. Тадқиқотларда экстралингвистик разметканинг қуйидаги турлари фарқланади³²:

1. Матн форматининг ўзига хослигини акс эттирувчи (боб, хатбоши, қисм ва ҳ.) тег.

2. Матн, унинг муаллифига тегишли маълумотни ифодаловчи тег. Чунки муаллиф ҳақидаги маълумот нафақат ном, балки ёш, жинс, у яшаган йил каби маълумотларни ҳам билдириши мумкин. Матн ҳақидаги маълумот эса матн (асар) номидан ташқари тили, ёзилган ҳамда нашр этилган йилини қамраб олади. Бундай маълумотларнинг мавжудлиги базада деталлаштирилган қидирувни амалга ошириш имконини беради. Экстралингвистик тег ёки метаахборот (русча: метаданные) – ташқи-интеллектуал маълумотларни қамраб олувчи, библиографик, типологик, тематик, социологик тавсифни; формал-структур тег (матн, бўлим, боб, қисм, абзац, гап)ни, техник-технологик тег (кодировка, манбанинг электрон версиясини қайта ишлаш санаси)ни бирлаштирувчи тег тури ҳисобланади.

Матнларга экстралингвистик изоҳ ёзишда ўзбек тилидаги аслият ва инглиз тилидаги таржима матнларига қўйидаги изоҳларни ёзиш мақсадга мувофиқ бўлади.

3. Ўзбек тилидаги матн форматининг боб, хатбоши, қисм ва ҳк. қисмлари ажратилади, белгиланади ҳамда инглиз тилидаги таржима матннинг айнан ўша боб, хатбоши, қисмлари белгиланиб, иккала тил учун матн қисмлари тенглаштирилади: аслиятда нечта қисм бўлса, таржима матнда қисмлар сони шунча бўлиши керак. Таржимада бирор-бир абзац тушиб қолган тақдирда, бундай қисмлар кўрсатиб берилади. Биз ўзбек-инглиз параллел корпуси учун матн танлаш жараёнида асар қисмлари (абзац)нинг тенглилигига алоҳида эътибор қаратамиз. Чунки ПК фойдаланувчиси икки тилдаги матнда барча қисмларни параллел кузатиб, ўрганиб боргани маъқул.

Матннинг экстралингвистик разметкасига тегишли кейинги изоҳ матн, унинг муаллифига тегишли маълумотни ифодаловчи тег бўлиши лозим. Ўзбек-инглиз параллел корпусининг ўзбек тили қисмида асар муаллифи ҳақидаги маълумот унинг номи, ёши, жинси, яшаган йили кабилардан иборат бўлса, инглиз тилидаги қисмида шу маълумотлар инглиз тилида берилади. Шу билан бирга, асар таржимони ҳақида ҳам худди шу параметрлар асосида маълумотлар бириктирилади.

³¹ Large-Scale Induction and Evaluation of Lexical Resources from the Penn-II Treebank / R.O'Donovan, M.Burke, A.Cahill, J. van Genabith, A.Way // Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics. Barcelona, 2004. – С. 368-375.

³² Хамроева Ш. Ўзбек тили муаллифлик корпусининг тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа докт. (PhD) диссертацияси. – Бухоро, 2018. – Б. 72-73.

Ўзбек ва инглиз тили – иккала қисмда матн ҳақидаги маълумот матн (асар) номи, тили (ўзбек/инглиз), ёзилган ҳамда нашр этилган йили каби изоҳдан иборат бўлади. Бундай маълумотларнинг мавжудлиги корпуснинг иккала қисмида деталлаштирилган қидирувни бирдек амалга ошириш имконини беради. Агар шундай изоҳ фақат аслият матнида бўлса-ю, таржима матнга экстралингвистик изоҳ берилмаган бўлса, қидирув тенг натижа бермайди: фойдаланувчи фақат битта тилдаги матннинг маълумотларини ўқийди. Параллел корпуснинг энг асосий хусусияти иккала тилдаги матннинг барча хусусиятларини бирдек кузатиш имконини бериш экан, экстралингвистик маълумотлар ҳам шундай бириктирилиши, матнлар шу принцип асосида разметкаланиши лозим.

Демак, корпус аннотацияси (бирликлар разметкаси) ҳажми, даражаси тадқиқотчи мақсади билан белгиланади. Мавзу, хронология, жанр бўйича мутаносиб равишда танланган катта ҳажмли параллел корпус тилнинг умумий хусусиятларини ўрганишга хизмат қилади. Параллел корпусга автоматик ишлов бериш бошқа корпуслардан фарқ қилади: автоматик ишлов беришнинг асосий, техник муаммоси корпус бирлиги сифатида аслият ва таржимадаги матннинг гап, сўз бирикмаси ва сўз даражасида разметкалаш ҳисобланади. Бунинг учун турли хил статистик усуллар, частотани таққослаш зарур бўлади. Ўзбек-инглиз параллел корпуси экстралингвистик разметка мажмуи аслият матни учун матн муаллифи исми, фамилияси, тахаллуси, жинси, туғилган вақти, матн номи, ёзилиш вақти ва жойи, матн ҳажми: матннинг неча сўздан иборатлиги ҳақидаги маълумотдан иборат бўлади. Таржима матни учун ушбу маълумотларга матн таржимони, таржима вақти, жойи, матннинг таржимадаги ҳажми ҳақидаги маълумот қўшилади. Параллел корпуснинг иккала қисмида ҳам матн жанри, матн типи, матн хронотопининг кўрсатилиши экстралингвистик разметкани мукаммалаштиради.

Тадқиқотнинг **“Ўзбек-инглиз тили параллел корпусининг лингвистик базаси ва дастурий асослари”** деб номланган учинчи бобида параллел корпус учун лингвистик ва дастурий таъминот ишлаб чиқиш, таржимада реализация ва эквивалент бирликларнинг берилиши масаласи ҳақида фикр юритилади.

“Параллел корпус учун матн танлаш масаласи” номли биринчи бўлимда ПКнинг лингвистик таъминоти хусусида баҳс боради. Корпуснинг репрезентативлигини таъминлаш мақсадида корпус бирлиги сифатида турли жанрдаги матнлар танланади. ПКга қайси мавзу ва жанрдаги матнни танлаш тузувчининг мақсадига боғлиқ. ПКлар бир томонга йўналган (ўзбек тилидан инглиз тилига) ёки икки томонлама (ўзбек тилидан инглиз тилига, инглиз тилидан ўзбек тилига) таржимани акс эттириши мумкин. Ўзбек-инглиз тили параллел корпуси учун матн танлашда ўзбек тилидан инглиз тилига таржима қилинган ҳамда инглиз тилидан ўзбек тилига бевосита таржималарни йиғиш мақсадга мувофиқ.

Ўзбек-инглиз параллел корпуси икки блокдан иборат бўлади: ўзбек-инглиз тили блоги (1), инглиз-ўзбек тили блоги (2). Қидирув тизимида асарларнинг жанри, яратилган даври, муаллифи ва таржимони қидируви

ишга туширилади, шу сабабли асарлар ушбу параметрлар асосида саралашни тақозо этади. Таржима назарияси мутахассиси, моҳир таржимон М.Холбеков ХХ асрда хорижий тиллардан амалга оширилган таржималар тўғрисида қимматли маълумот беради³³. Агар асар таржимаси мавжуд бўлса, тайёр маҳсулотдан қайта параллел корпус тузиш қанчалик аҳамиятли? Корпус таржима қилишга эмас, тайёр таржималарни йиғишга мўлжалланган экан, у илм-фан ва таржима назарияси учун нима беради? Юқоридаги ўринли саволларга қуйидагича жавоб берамиз:

1) мавжуд жаҳон параллел корпуслари тайёр таржималарга асосланади. Биз ҳам ўзбек-инглиз параллел корпусига тайёр таржима матнларини олишимиз лозим;

2) таржимани параллел корпусга бирлаштириш бутун бошли асарни аслият ва таржимада ёнма-ён ўқиш имконини беради;

3) энг асосийси, параллел корпус – у ёки бу бирлик (содда гап, қўшма гап, сўз бирикмаси, фразеологик бирлик, сленг, метафора) аслият ва таржимада қандай ифодаланишини кузатишнинг энг қулай усули саналади.

Шу сабабли ҳар қандай параллел корпус учун материал танлашда таржиманинг аслиятдан бевосита амалга оширилган вариантини олиш принципи устунлик қилади. Чунки бирликлар билвосита таржимада ўз қимматини йўқотади, бундай ҳолатда параллел корпус ўз вазифасини тўла бажармайди. Бир неча марта қайта таржима қилинган асарларнинг фақат бевосита таржима қилинган нусхасини олиш ёки қиёслаш учун билвосита таржималарни ҳам киритиш мумкин. Бу тузувчининг тадқиқий ёки иллюстратив корпус тузиш мақсадидан келиб чиқади.

М.Холбеков таржималарни қиёслар экан, аслиятда “To be, or not to be: that is the question” сатрлари билан бошланувчи ҳаяжонли монологнинг русча таржималарида қарийб бир хил қолипда ўгирилганлигини асослайди³⁴:

1	М.Вронченко, (1829)	“Быть или не быть – таков вопрос”
2	А.Кронеберг, (1844)	“Быть или не быть? Вот в чем вопрос!”
3	М.Лозинский, (1933)	“Быть или не быть – таков вопрос”
4	Б. Пастернак, (1940)	“Быть или не быть: вот в чем вопрос”

Қиёслар ўзбек таржимонлари худди шу мисрани турлича кўринишда таржима қилганликларини кўрсатади.

1	Чўлпон таржимасида	Ё ҳаёт, ё ўлим, масала шунда.
2	М.Шайхзода таржимасида	Тирик қолмоқ ё ўлмоқ? Шудир масала!
3	Жамол Камол таржимаси	Ё ҳаёт, ё мамот: масала шундоқ!

М.Холбековнинг таъкидлашича³⁵, таржималар аслият билан солиштирилганида, Чўлпон таржимаси русчадаги М.Лозинский нусхасига, М.Шайхзода таржимаси А.Кронеберг матнига яқинлиги ойдинлашади. Жамол Камол таржимаси эса асл нусхага монанд ўгирилгани кўринади.

“Ўзбек-инглиз параллел корпусда бирликларнинг таржимада берилиши” номли иккинчи бўлимда таржимада трансформацияга учраган

³³ Холбеков М. Ўзбекистонда таржима ва таржимашунослик (1991-2016 йиллар) // Хорижий филология. – 2016 йил. – №3. – Б. 16-42.

³⁴ Кўрсатилган асар.

³⁵ Кўрсатилган асар.

морфемаларнинг параллел корпусда берилиши масаласи, реалияларнинг таржимада берилиши ҳамда унинг параллел корпусда ифодаланиши ҳақида баҳс юритилади. Амалий жиҳатдан таржимада морфемани морфема билан ифодалаш ҳолати кам учрайди. Назарий жиҳатдан инглиз тилидаги *-er, -or, -su, -man, -ship, -ment, -hood* каби морфеманинг ўзбек тилида турли луғавий эквиваленти мавжуд бўлса-да, таржимада доим ҳам эквивалентликка эришиш имкони бўлмайди. Чунки изоҳли ва икки тилли луғатларда гапда иштирок этмаган сўзлар, яъни системадаги сўзларнинг эквиваленти доим мавжуд бўлади. Нутқда қўлланиши билан уларнинг шакли, услуби, матн турлари ҳамда гапда қўлланган структуралари ўзгаради³⁶. Таржимада морфемик трансформация усулларининг қайси бири қай даражада қўлланиши матнга, ундаги тил бирликларига, таржима муносабати билан контактга кирган тиллар хусусиятларига бевосита боғлиқ бўлади.

Параллел корпус таржима матнларни параллел кузатиш имконини беради, аммо юқорида кўрсатилган трансформациялар каби ҳолатни ажратиб кўрсатиш даражаси чегараланган. Чунки мавжуд параллел корпусларнинг кўпчилиги икки (аслият ва таржима) тилдаги мантни фақат гап кўринишида параллел кўрсатади. Аммо назарий жиҳатдан ПКларда бирикма ва сўзларни параллел кўрсатиш имкони ҳам мавжуд. Агар ПК тузувчиси фақат гапни эмас, балки сўз бирикма ҳамда сўзни параллел ҳолатда теглашни мақсад қилса, унда юқорида кўрсатилган трансформацияларга алоҳида эътибор қаратиш зарур бўлади.

Таржимада трансформацияга учраган эркин бирикма ёки ясама сўзни аслият ва таржимада параллел қилиш мумкин. Бу вазифа қуйидаги кетма-кетликда амалга оширилади:

- 1) таржима ва аслият матнлари қийсланиб, шундай birlik /ҳодисадалар рўйхати тузилади;
- 2) трансформацияга учраган birlikларнинг ўзбекча таржимаси келтирилади;
- 3) birlikлар ўзаро параллеллаштирилади, тегланади;
- 4) birlikларнинг маълумотлар базаси – луғати шакллантирилади.

Миллий тилнинг ўзига хос томонларини ифодалашда реалиянинг ўрни беқиёс, чунки реалия тил бирлиги сифатида муқобилсиз лексика сирасига киради. У лингвистик билимнинг ажралмас қисми бўлиб, тил ва маданиятнинг ўзаро алоқасини тадқиқ этишда муҳим қизиқиш уйғотади³⁷. Инглиз тили реалиялари ҳақида Л.С.Бархударов, Г.Д.Томахин, А.Д.Щвейцер, А.О.Ивановларнинг ишларида ўқиймиз³⁸. М.Жуманиёзова, А.Атажановалар Г.Д.Томахин³⁹нинг реалия – “маълум бир халқ ва миллатга тегишли бўлган моддий маданият нарсалари, тарих фактлари, давлат ташкилоти, миллий ва фольклор қахрамонлари номи, мифологик мавжудотлар ва ҳоказоларга оид

³⁶ Ўша манба.

³⁷ Жуманиёзова М., Атажанова А. Реалия ҳақидаги қарашлар талқини // Таржима, ахборот, мулоқот – сиёсий ва ижтимоий кўприк. Халқаро илмий-амалий анжуман материаллари. – Самарқанд: СамДЧТИ, 2019. – Б.158-159.

³⁸ Бархударов Л.С. Язык и перевод. – М: Международное отношения, 1975.; Томахин Г.Д. Реалии-американизмы. – М.: Высшая школа, 1988.

³⁹ Томахин Г.Д. Реалии в языке и культура // ИЯШ. 1997. – №3. – С.13.

номлар” деган фикрни маъқуллайди. Демак, турмуш тушунчаларини акс эттирадиган реалияларнинг таржима тилида учрамаслиги транслитерация усулидан фойдаланишни тақозо этади; таржима орқали реалияларнинг этномаданиятдаги ўзига хослигини сақлаб қолиш учун замин яратади. Параллел корпуслар икки халқ тилини қиёслаш ва ўрганиш асносида уларнинг маданиятини ҳам намоён этади, демак, ПКларда реалияларни параллел ҳолатда ажратиб кўрсатиш ниҳоятда муҳим.

Бобнинг “*Ўзбек-инглиз параллел корпуси лингвистик базасини шакллантириш асослари*” деб номланувчи учинчи бўлимида ўзбек-инглиз параллел корпуси лингвистик базасини шакллантириш масалалари таҳлил қилинади. Материаллар таҳлили учун Ў.Ҳошимовнинг “Дунёнинг ишлари” қиссаси⁴⁰, бу асарнинг О.Мўминов томонидан инглиз тилига амалга оширилган “Affairs of life. Stories” таржима нашри⁴¹, жумладан, “Қарз” (“A debt”), “Гилам пайпоқ” (“A sock”), “Бола йиғиси” (“Baby’s cry”) ҳикоялари матн танлаб олинди.

ПК тузишда ўзбек тили матнларини таҳлил қила олувчи дастур тузиш мақсадга мувофиқ. Бу дастур <нуктадан нуктагача>, <нуктадан сўроқгача>, <нуктадан ундовгача> ва ҳ. принциплар асосида содда, қўшма гапларни ажратиб чиқишга мўлжалланган. Аммо биз матнларни қўлда параллеллаб чиқдик, чунки ишнинг мақсади фақат матнни параллеллаш эмас, балки матнларни параллел теглашда ундаги бирликларнинг берилишини аниқлашдан иборат.

“Қарз” ҳикояси 30 та гап (содда ва қўшма)га ажратилиб, ўзбек ва инглиз тилига қилинган таржимаси параллел ҳолда келтирилди. Жадвалда ўзбек ва инглиз тилида таржима жараёнида махсус қоида ва истиснолар ажратиб кўрсатилган, бир-бирига сўзма-сўз мос келмаган бирликлар белгиланган. Таржима жараёнида содда гап қўшма гап билан (ёки аксинча) ифодаланиши, маълум бир абзацлар тушиб қолиши ёки қўшилиши, таржимада сленг ишлатилиши оқибатида гап чегараларининг тенг келмай қолиши ҳолатлари учраб туради. Параллел корпус тузувчиси матнни автоматик параллел ҳолатда теглагандан сўнг шундай вазиятларни қўлда тўғрилаб чиқишига тўғри келади. Биз ҳам шундай ҳолатларни кузатдик ва уларни алоҳида белгиладик. Зеро, бу амалиётлар кейинчалик ПК учун лингвистик базани шакллантириш воситаси бўладиган луғатлар яратишга асос бўлади.

АТПЖнинг 19-сатридаги *Ия, юбиляр экансан-да, шошмай тур-чи, ҳозир...* бирлиги *Oh, congratulations! Wait, just a minute...* шаклида таржима қилинган бўлиб, битта қўшма гап икки содда гап билан ифодаланган. Шунингдек, *Ия, юбиляр экансан-да* гапи *Oh, congratulations* тарзида берилиб, гапларнинг ифода мақсадига кўра тури ўзгарган. Таржимада *юбиляр экансан* ифодаси *табриклайман* деган мазмунни ифодаловчи сўз билан берилган. Агар бу каби бирликлар тегланмаса, корпус сўз қидируви натижаси пасаяди.

⁴⁰ Hoshimov O`. Dunyoning ishlari: qissa. – Toshkent: “Sharq” nashriyot-matbaa aksiyadorlik kompaniyasi bosh tahririyati, 2005. – B. 10-11.

⁴¹ Utkir Hoshimov. Affairs of life. Stories (Translator: O.M.Muminov). – Tashkent, 2013. – P.9-10.

“Гилам пайпоқ” ҳикояси АТПЖнинг 7-сатрида келган *As soon as the weather got cold she used to have a pain in her legs* гапи *Салқин тушиши билан шишиб кетади, оғрийди* қўшма гапи воситасида ифодаланган.

АТПЖнинг 20-21-сатрида келган *Терлаб туриб муз едим. Бир кун акаларим билан яхмалак ўйнаб терлаб кетдим* содда гаплари *I played on an ice-rink with my brothers and sweated up and I ate an ice then* тарзида қўшма гап воситасида ифодаланган.

АТПЖнинг 45-46-сатрида келган *У томоғимга нимадир қилди. Дод солиб қўлини тишлаб олдим* иккита содда гап *She did something with my throat and I bit her finger* тарзида битта қўшма гап кўринишида таржима қилинган.

Бундан ташқари яна бир типик ҳолат учраб туради. Ўзбек тилидаги матнда мавжуд бўлмаган гап таржима матнда пайдо бўлган. Улар матн мазмунини аниқлаштириш мақсадида келтирилади.

АТПЖнинг 62-сатридаги *Сарпойчан келавердингизми?! – деди Ҳожи буви ҳамон ўша ваҳимали оҳангда* контексти *Did you come like this? – said Hoji grandmother* тарзида ўгирилганини кўриш мумкин. Кўриняптики, *сарпойчан* сўзига мос эквивалент топилмаган ҳамда *шундай келавердингизми? (Did you come like this?)* кабилида умумий қилиб таржима қилинган, аммо китобхон маънони контекстдан англаши мумкин.

АТПЖнинг 63-сатридан шуни кўриш мумкинки, ўзбек тилидаги матнда мавжуд бўлмаган гап таржима матнда пайдо бўлган: *Still in the same tone.*

“Қарз” ҳикояси АТПЖнинг 16-сатрида келтирилган *Онноқ бантик тақиб, атлас кўйлак кийиб олибди* гапи *She wore bright white band and atlas dress* шаклида таржима қилинган. Эътибор берадиган бўлсак, *атлас* сўзи реалия ҳисобланиб, у инглиз тилига *atlas* шаклида берилган; ипак кўйлак (*silk dress*) кўринишида эмас.

“Гилам пайпоқ” ҳикояси АТПЖнинг 24-сатрида келган *Ойим томоғимни аччиқтош билан чайиб кўрди, бўлмади, туршак қайнатиб сувини ичирди, бўлмади...* қўшма гапидаги ажратиб кўрсатилган сўз *Mother washed my throat with alum, but it was useless. She gave me water of boiled dry apricot fruit but there was no result* гапида бирикма сифатида берилган. Чунки инглиз тилида *туршак* сўзининг эквиваленти мавжуд эмас, демак, у реалия.

Ўзбек-инглиз параллел корпусида асл ҳолида бериш имкони бўлмайдиган бирликлар ҳамда уларни қидирув учун мослаш муаммолари хусусида алоҳида тўхталиш лозим.

“Қарз” ҳикояси АТПЖнинг 12-сатрида келтирилган *Менинг ишимга аралашиб нима қиласан?* гапи сўзма-сўз таржима қилинмаган; инглиз тилидаги турғун *It is not your business* ибораси билан берилган. Демак, бундай таржималарда сўзнинг эквивалентини қидириш, кузатиш имкони бўлмайди.

“Гилам пайпоқ” ҳикоясининг сарлавҳаси “A sock” деб таржима қилинган. Бунинг сабаби шундаки, инглиз тилида шу бирликни ифодаловчи эквивалент сўз топилмаган. Бизнингча, бирикма тарзида таржима қилиш мумкин, аммо таржимон шуни маъқул топган. Демак, бу бирликлар ўз эквиваленти билан таржима қилинмаган; бу бирлик ҳам параллел корпус луғатида таржимада эквиваленти мавжуд бўлмаган бирликлар қаторидан ўрин олишга эҳтиёж сезади.

Хулоса ўрнида шуни айтиш мумкинки, таржима жараёнида содда гап қўшма гап билан (ёки аксинча) ифодаланиши, маълум бир абзацлар тушиб қолиши ёки қўшилиши, таржимада сленг ишлатилиши оқибатида гап чегараларининг тенг келмай қолиш ҳолатлари учраб туради. Параллел корпус тузувчиси матнни автоматик параллел ҳолатда теглагандан сўнг шундай ҳолатларни қўлда тўғрилаб чиқишига тўғри келади. Бу амалиётлар кейинчалик ПК учун лингвистик базани шакллантириш воситаси бўладиган луғатлар яратишга асос бўлади.

ХУЛОСА

1. Параллел корпус – оригинал матн ва унинг бир/бир неча таржимасидан иборат корпус. Корпусда оригинал матн ёки унинг бир қисми бўлиши мумкин. Асл ва таржима матни шунчаки ёнма-ён қўйилмайди, балки икки матндаги синтактик бирликлар мазмунан мос келиш даражасида бўлакка бўлиниб, бир-бирига ишора/ҳавола қилиш ҳолатида туради. Аслиятдаги матн фрагментига мос келувчи таржимадаги фрагмент белгилаб қўйилади.

2. Параллел корпуснинг кўп тилли ҳамда таржима корпус каби кўринишлари мавжуд. Тузилиши, матннинг жойлашуви, бирликларининг мослаштирилишига кўра параллел корпуснинг бир йўналишли, икки йўналишли, кўп йўналишли каби кўринишлари фарқланади. ПК таркиби, мақсадига кўра таржимага ҳавола қилувчи одатий матн; қиёслаш учун қулай бўлган «ойнадаги матн»; маълумотлар базаси кўринишида бўлиши мумкин.

3. Параллел корпус таржима тилининг структур тузилиши, ибора ва сўзларнинг маълум контекстдаги имкониятини кузатишга хизмат қилади. ПКдан қиёсий тилшунослик, таржима, машина таржимаси, лингводидактика, лексикография соҳаларида унумли фойдаланиш мумкин.

4. ПК материали таҳлили асосида таржима сифати, даражасини кузатиш мумкин: матн эркин бадий таржима ёки айнан таржима қилингани параллел корпусда яққол кўриниб туради. Миллий корпусдан фарқли равишда бундай корпус катта массивли материал асосида аниқ бир соҳадаги таржима хусусиятини ўрганишга замин ҳозирлайди. Илмий-техникавий, ижтимоий-сиёсий, иқтисод ва бошқа соҳага оид ПК яратиш айнан шу соҳада тилнинг услубий имкониятларини махсус тадқиқ қилишга имкон беради.

5. Дунё компьютер лингвистикасида қатор конкорданс дастури, корпус менежери, улар асосида махсус мақсадларни кўзловчи кичик (параллел ва қиёсий) корпус яратиш имкони мавжуд. Ўзбек тили учун ҳам шундай конкорданс ҳамда корпус менежерлари ишлаб чиқиш талаб этилади. Мавжуд параллел корпуслар тили, ҳажми, дастурий таъминоти билан бир қанча фарқларга эга бўлса-да, ПКларнинг умумий жиҳати разметкалаш корпус тилининг морфологик анализатори ёрдамида бажарилиши, барча ПКларда морфосинтактик теглаш ҳам амалга оширилиши билан изоҳланади.

6. Корпуслар орасидаги дастлабки фарқ унинг интерфейси ёки корпус менежериде сезилади. Айрим ПКда интерфейс қидирув натижасини наشر йили, муаллиф, жанр каби параметрлар бўйича чегаралайди, саралайди. Бошқа ПК эса корпус менежериде биринчи тил учун статистик ишлов ойнаси

самарали натижа берса, иккинчи – таржима тили учун шу имкониятга эга эмаслиги билан фарқ қилади. Қидирув тизими сўз, сўз бирикмаси, дизъюнкция, конъюнкция, инкор қидирувини амалга оширади; таржима матнидан шу шаклларнинг эквивалентини тақдим қилади; қидирувнинг турли морфологик шакли ҳам мавжуд.

7. ПК тузишнинг технологик жараёни матн танлаш ва уни дастлабки электрон шаклга киритиш; матнга метатавсиф бериш; қидирув тизими модулини яратиш; матнни корпусга юклаш; корпус бирликларини синтактик таҳлил қилиш; экспериментал эксплуатация ўтказиш; натижани таҳлил қилиш каби умумий тамойилларга амал қилади. Параллел корпусга автоматик ишлов бериш бошқа корпуслардан фарқ қилади: автоматик ишлов беришнинг асосий, техник муаммоси корпус бирлиги сифатида аслият ва таржимадаги матннинг гап, сўз бирикмаси ва сўз даражасида разметкаланиш ҳисобланади. Бунинг учун турли хил статистик усуллар, частотани таққослаш зарур бўлади.

8. Корпус тузишдаги кейинги муҳим босқич матнни гап доирасида тенглаштириш, яъни параллел корпусда икки тилдаги гап мазмунан ўз эквиваленти билан мослаштирилган бўлади. ПКнинг икки тилдаги матнларини тенглаштириш матндаги гаплар чегарасини аниқлаш; аслият ва таржима матндаги бир-бирини ифодалаётган гаплар чегарасини кўрсатиш билан белгиланади.

9. Гапнинг узунлигига қараб тенглаштириш интервал ёки тиниш белгига нисбатан жуда сезувчан бўлади: битта ўринсиз пробел/тиниш белгининг қўлланиши тенглаштиришда хатоликка олиб келади. Фавқулодда ҳодисалар статистик таҳлил учун етарли маълумот бермайди. Сўз даражасида тенглаштиришнинг яна бир муҳим босқичи аслият ва тажима гапда синтактик таҳлилни амалга оширишни талаб қилади. Разметкаланган корпусда синтактик таҳлилни йўлга қўйиш грамматик назарияни текширишга ёрдам беради, янги синтактик қонуниятларни очади: ПКдаги синтактик таҳлил шуниси билан аҳамиятли.

10. Корпусда таржима сифати жуда муҳим: таржима матни машҳур, малакали мутахассис ёки расмий ташкилот томонидан таржима қилинган бўлиши лозим. Бадиий асар таржимасидан фарқли равишда расмий-идоравий, дипломатик, бошқарув ҳужжатлари таржимасида аслиятдаги матн мазмун ҳамда структураси сақланиши лозим. Шу сабабли бундай оригинал ва таржима матн гап/жумлалари кетма-кетлиги бир хил бўлади; структур жаҳатдан ўзаро бир хил тузилади; ўхшаш бўлади. Бу усул билан икки тилдаги матн бирликларини тенглаштириш вазифаси енгиллашади.

11. ПКни тузишда матнни автоматик таҳлил қилувчи дастур зарурий восита саналади. Тиниш белги, график безак элементи асосида гапларнинг чегарасини аниқловчи дастур – ПК тузиш учун бирламчи иш қуроли. Ўзбек-инглиз параллел корпуси экстралингвистик разметка мажмуи аслият матни учун матн муаллифи исми, фамилияси, тахаллуси, жинси, туғилган вақти, матн номи, ёзилиш вақти ва жойи, матн ҳажми: матннинг неча сўздан иборатлиги ҳақидаги маълумотдан иборат бўлади. Таржима матни учун ушбу

маълумотларга матн таржимони, таржима вақти, жойи, матннинг таржимадаги ҳажми ҳақидаги маълумот қўшилади.

12. Ўзбек-инглиз параллел корпусида аслият ва таржимадаги мақолларни теглаш, фойдаланувчига қидирув натижасини параллел ҳолатда кўрсатиш корпуснинг қимматини, аҳамиятлилик даражасини яна ҳам оширади. Ўзбек-инглиз ПК матнларини танлашда қаҳрамонлар мадҳ қилинадиган; аёлларнинг ҳақ-ҳуқуқлари ҳимоя қилинадиган; кишилар характеридаги ярамас одат, номаъқул иллатлар танқид остига олинадиган; мардлик, эпчиллик, довюраклик, меҳнатсеварлик, ҳалоллик, вафодорлик, саҳийлик ғоялари улуғланадиган мазмундаги матнларнинг танлаб олиниши ўзбек-инглиз параллел корпусида тарбиявий, ахлоқий-маданий омилларнинг эътиборга олинишини таъминлаган бўлади; ПК материаллари тарбия/маданиятимизга зид бўлган мазмундаги материаллардан холи бўлади.

13. Реалияларда тил ва маданият ўртасидаги боғлиқлик янада яққол кўзга ташланади. Маиший тушунчани ифодаладиган реалиянинг таржимада учрамаслиги транслитерация усулидан фойдаланишни тақозо этади; таржима орқали реалиянинг этномаданиятдаги ўзига хосликни сақлаб қолиши учун замин яратади. ПК икки халқ тилини қиёслаш ва ўрганиш асносида уларнинг маданиятини ҳам намоён этар экан, демак, ПКда реалияларни параллел ҳолатда ажратиб кўрсатиш ниҳоятда муҳим.

14. Ўзбек-инглиз параллел корпуси лингвистик базасини шакллантириш учун танлаб олинган ва таҳлил қилинган матнларнинг таржимаси таҳлилида содда гап қўшма гап билан (ёки аксинча) ифодаланиши, маълум бир абзацлар тушиб қолиши ёки қўшилиши, таржимада сленг ишлатилиши оқибатида гап чегараларининг тенг келмай қолиш ҳолатлари учради. Гапларнинг ифода мақсадига кўра тури ўзгариши, қўшма гапдаги биричи содда гап таржима қилиниб, иккинчиси тушириб қолдирилиши, ўзбек тилидаги матнда мавжуд бўлмаган гапнинг таржима матнда пайдо бўлиши кузатилади. Ўзбек тилидаги матнда бўлмаган гапнинг таржима матнда пайдо бўлишини матн мазмунини аниқлаштириш мақсадида келтирилиши билан изоҳлаш мумкин. Параллел корпус тузувчиси матнни автоматик параллел ҳолатда теглагандан сўнг шундай ҳолатларни қўлда тўғрилаб чиқишига тўғри келади. Бу амалиётлар кейинчалик ПК учун лингвистик базани шакллантириш воситаси бўладиган луғатлар яратишга асос бўлади.

**НАУЧНЫЙ СОВЕТ DSc.03/04.06.2021.Fil.72.03
ПО ПРИСУЖДЕНИЮ НАУЧНЫХ СТЕПЕНЕЙ ПРИ
БУХАРСКОМ ГОСУДАРСТВЕННОМ УНИВЕРСИТЕТЕ**

БУХАРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

КАРИМОВ РУСТАМ АБДУРАСУЛОВИЧ

**ЛИНГВИСТИЧЕСКИЕ И ПРОГРАММНЫЕ ВОПРОСЫ
СОСТАВЛЕНИЯ УЗБЕКСКО-АНГЛИЙСКОГО ПАРАЛЛЕЛЬНОГО
КОРПУСА**

**10.00.06 – Сравнительное литературоведение, сопоставительное
языкознание и переводоведение**

**АВТОРЕФЕРАТ
диссертации доктора философии (PhD) по ФИЛОЛОГИЧЕСКИМ НАУКАМ**

Бухара – 2022

Тема диссертации доктора философии (PhD) зарегистрирована под номером B2020.2.PhD/Fil587 в Высшей Аттестационной комиссии при Кабинете Министров Республики Узбекистан.

Диссертация выполнена в Бухарском государственном университете.

Автореферат диссертации на трёх языках (узбекском, русском, английском (резюме)) размещён на веб-странице Бухарского государственного университета (www.buxdu.uz) и информационно-образовательном портале «ZiyoNet» (www.ziyo.net.uz).

Научный руководитель: Менглиев Бахтиёр Ражабович
доктор филологических наук, профессор

Официальные оппоненты: Ахмедов Ойбек Сапорбаевич
доктор филологических наук, профессор

Атабоев Нозимжон Бобожон ўгли
доктор философии по филологическим наукам (PhD)

Ведущая организация: Термезский государственный университет

Защита диссертации состоится «22» апреля 2022 года в 10⁰⁰ часов на заседании Научного совета DSc.03/04.06.2021.Fil.72.03 по присуждению учёных степеней при Бухарском государственном университете. (Адрес: 200118, г. Бухара, улица М.Икбол, 11. Тел.: +99865 221-29-14; факс: +99865 221-27-57, e-mail: buxdu_rektor@buxdu.uz.)

С диссертацией можно ознакомиться в Информационно-ресурсном центре Бухарского государственного университета (зарегистрировано за № 125). (Адрес: 200118, г. Бухара, улица М.Икбол, 11. Тел.: +99865 221-25-87.)

Автореферат диссертации разослан «06» апреля 2022 года.
(Протокол-реестр рассылки № 22-01 от «06» апреля 2022 года.)



Д.С. Ураева
Председатель Научного совета по присуждению ученых степеней, доктор филол. наук, профессор

З.И. Расулов
Ученый секретарь Научного совета по присуждению ученых степеней, кандидат филол. наук, доцент

М.М. Жураева
Председатель Научного семинара при Научном совете по присуждению ученых степеней, доктор филол. наук, доцент

ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))

Актуальность и востребованность темы диссертации. Изучение проблем компьютерной и корпусной лингвистики в мировой лингвистике интенсифицировалось в 60-е годы XX века и в начале XXI века появились сотни языков корпусов. Расширился доступ к автоматическому переводу, тезаурусу, электронным словарям, созданы их научно-теоретические основы, стали применяться на практике, появилось много видов языковых корпусов, которые различались по назначению. Эти обновления послужили толчком к появлению новых научных направлений, связанных с внедрением информационных технологий в языкознание.

В мировой лингвистике движение к научно-теоретическому изучению корпусной лингвистики достигло своего апогея в XXI веке. В стремительно развивающемся направлении компьютерной лингвистики повышение качества автоматического перевода, использование параллельных корпусов в переводческих программах стало актуальной проблемой переводоведения. На сегодняшний день в области компьютерной лингвистики одними из актуальных вопросов являются выбор единиц параллельного корпуса, лингвистическое и экстралингвистическое тегирование, разработка алгоритма параллельного корпуса, создание лингвистической базы для параллельного корпуса.

В годы независимости в компьютерной лингвистике была проделана большая работа по автоматическому переводу, достижению понимания и обработки узбекского языка, искусственным интеллектом. Потому как "...сохранение чистоты государственного языка, его обогащение и повышение культуры речи населения; обеспечение активной интеграции государственного языка в современные информационные технологии и коммуникации"¹ является актуальной задачей, стоящей сегодня перед узбекской компьютерной лингвистикой.

Данное диссертационное исследование в определенной степени окажет практическое содействие выполнению задач, отраженных в Указах Президента Республики Узбекистан №УП-4797 «О создании Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои» от 13 мая 2016 года, № УП-4947 «О Стратегии дальнейшего развития Республики Узбекистан» от 7 февраля 2017 года, №УП-2789 «О мерах по дальнейшему совершенствованию деятельности Академии наук, организации, управления и финансирования научно-исследовательской деятельности» от 17 февраля 2017 года, Постановлении Президента Узбекистан №ПП-4479 «О широком праздновании тридцатилетия принятия закона Республики Узбекистан «О государственном языке» от 4 октября 2019 года, Указе №УП-5850 «О мерах по кардинальному повышению роли и авторитета узбекского языка в качестве государственного

¹Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 20 октябрдаги «Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сўёсатини такомиллаштириш чора-тадбирлари тўғрисида»ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>

языка» от 21 октября 2019 года, Послании Президента Республики Узбекистан к Олий Мажлису от 24 января 2020 года, Указе Президента Республики Узбекистан №УП-6084 «О мерах по дальнейшему развитию узбекского языка и совершенствованию языковой политики в стране» от 20 октября и в других нормативно-правовых документах имеющих отношение к данной сфере деятельности

Соответствие исследования приоритетным направлениям развития науки и технологий республики. Диссертационное исследование выполнено в соответствии с приоритетным направлением развития науки и технологий республики I. «Формирование системы инновационных идей и пути их реализации в социальном, правовом, экономическом, культурном, духовно-просветительском развитии информатизированного общества и демократического государства»

Степень изученности проблемы. Несмотря на то, что корпусная лингвистика в мировой лингвистике является новым направлением, целенаправленных исследований в корпусной области достаточно много. Размышления составителей корпуса Брауна Н.Френсиса, Г.Кучеры о принципах корпусообразования, теоретические и практические предложения автора проекта банка английского языка Дж. Синклера можно рассматривать как предварительные исследования² корпуса. В русской лингвистике В.П.Захаров, А.Б.Кутузов, Е.В.Недошивина, В.В.Рыков и В.Плунгян проводили исследования относительно корпуса, его типов, специфики, социального значения корпуса, принципов корпусообразования³. Параллельные корпуса были исследованы В.В.Мамонтовой, Д.В.Сычиновой⁴. В узбекском языкознании проведено некоторое количество исследований в области компьютерной и корпусной лингвистики. Хотя создание узбекско-английского параллельного корпуса не являлось объектом специального исследования, в монографическом плане были изучены некоторые проблемы компьютерной лингвистики, в частности, создание лингвистического обеспечения для компьютерных программ на основе глаголов действия⁵, лингвистическое обеспечение машинного перевода

² Френсис Н., Кучера Г. Вычислительный анализ современного американского варианта английского языка. – М., 1967.; Синклер Д. Предисловие к книге «Как использовать корпуса в преподавании иностранного языка»/ Д. Синклер [Электронный ресурс]. – Режим доступа: <http://www.ruscorpora.ru/corpora-info.html>, свободный.

³ Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс)- //lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf; Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. – Санкт-Петербург. –2006. 26 с.; Рыков В.В. Курс лекций по корпусной лингвистике. URL: <http://rykov-cl.narod.ru/c.html>; Плунгян В. Зачем мы делаем Национальный корпус русского языка? «Отечественные записки» 2005, №2. http://magazines.russ.ru/oz/2005/2/2005_2_20-pr.html

⁴ Мамонтова В.В. корпус параллельных текстов и база данных для исследования переводческих соответствий: проблемы и процедуры формирования. – Альманах современной науки и образования Тамбов: Грамота, 2008. № 8 (15): в 2-хч. Ч.II. С.117-118. – Адрес статьи: www.gramota.net/materials/1/2008/8-2/50.html; Сычина Д.В. Европейский перфект сквозь призму параллельного корпуса. – *Acta Linguistica Petropolitana. Труды института лингвистических исследований*. – 2016, № 2. – Т.12. – С. 85-114.

⁵ Мухаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. – Тошкент, 2006.;

узбекско-английского языка⁶, принципы построения авторского корпуса узбекского языка⁷, проблемы графематического анализа единиц узбекского языка⁸, принципы построения лингвистической базы языкового корпуса⁹ и функциональные характеристики английского корпуса¹⁰. Но в вышеупомянутых исследованиях вопрос о создании параллельных корпусов не стоит на повестке дня. В процессе подготовки диссертации были всесторонне изучены вышеперечисленные научные исследования, местами выражено отношение к ним и использованы в ходе исследования.

Связь диссертационного исследования с планами научно-исследовательских работ высшего образовательного учреждения, где выполнена диссертация. Исследование выполнено в соответствии с концепцией “Проблема оптимизации обучения иностранным языкам”, изучаемой на межфакультетской кафедре иностранных языков Бухарского государственного университета.

Целью исследования является изучение теоретических основ построения узбекско-английского параллельного корпуса, разработка принципов построения лингвистической базы узбекско-английского параллельного корпуса.

Задачи исследования:

описание общих свойств параллельного корпуса, раскрытие аспектов, отличающих его от других типов корпуса;

демонстрация преимуществ параллельного корпуса в области лингводидактики иностранного языка, лингвистики и перевода;

выявление принципов отбора материала для узбекско-английского параллельного корпуса и сбор материала для фрагмента узбекско-английского параллельного корпуса;

определение путей лингвистического и экстра-лингвистического тегирования текстов узбекско-английского параллельного корпуса и разработка предложений для программ параллельного тегирования текста;

В качестве **объекта исследования** были выбраны узбекско-английские параллельные корпуса.

Предмет исследования составляет лингвистическая база узбекско-английского параллельного корпуса и вопросы программирования.

⁶ 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. – 52 б.

⁷ Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2018. – 250 б.

⁸ Абжалова М.А. Ўзбек тилидаги матнларни тахрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар тахрири дастури учун): Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Фарғона, 2019. – 164 б.; Абжалова М. Матнларга автолингвистик ишлов бериш тизимлари // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

⁹ Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.;

¹⁰ Атабоев Н.Б. Инглиз тили корпусининг функционал хусусиятлари (СОСА мисолида): Филол. фан. докт. ... дисс. – Тошкент, 2020. – 189 б.

Методы исследования. При освещении темы исследования использовались методы классификационного, описательного, сравнительного, компонентного, дифференциально-семантического, дистрибутивного анализа.

Научная новизна исследования состоит из следующих:

описаны общие особенности параллельного корпуса, раскрыты его отличия от других корпусов; доказаны преимущества параллельного корпуса в области лингводидактики иностранного языка, лингвистики и перевода;

разработаны принципы отбора текстов на узбекском языке; обоснованы преимущества выбора текстов на английском языке; определены принципы отбора материала для узбекско-английского параллельного корпуса;

определены проблемы лингвистического тегирования текстов узбекско-английского параллельного корпуса и предложены пути тегирования;

определены принципы экстралингвистического тегирования текстов узбекско-английского параллельного корпуса.

Практические результаты исследования:

определены принципы отбора текстов на узбекском языке для параллельного корпуса, собраны тексты, имеющие воспитательное значение в качестве языкового обеспечения параллельного корпуса;

в качестве лингвистического обеспечения параллельного корпуса сформирована библиотека духовно-нравственных текстов на английском языке;

сформирована лингвистическая база параллельного корпуса узбекского и английского языков, подготовлено лингвистическое обеспечение для программирования.

Достоверность результатов исследования подтверждается материалами научно-методических и научно-практических конференций республиканского и международного уровня, публикациями в специализированных журналах и зарубежных научных журналах из списка ВАК, внедрением в практику теоретических идей и выводов, подтверждением полученных результатов уполномоченными организациями.

Научная и практическая значимость результатов исследования. Исследование служит научно-теоретическим источником для разработки теоретических основ создания параллельных корпусов узбекского языка, осуществления исследований в направлении компьютерной лингвистики.

Практическая значимость результатов исследования определяется тем, что теоретические обобщения и выводы работы могут быть использованы при предоставлении теоретической информации молодежи в системе высшего и среднего специального образования, в общеобразовательных школах, как источник при составлении программ, планов и изложении тем в процессе обучения корпусной лингвистике и компьютерной лексикографии как предмету, в качестве методического средства при обучении иностранному языку.

Внедрение результатов исследования. Из выводов об общих чертах параллельного корпуса, его отличиях от других корпусов, преимуществах

параллельного корпуса в области лингводидактики иностранного языка, лингвистики и перевода; На основе принципов отбора материала для узбекско-английского параллельного корпуса научные результаты, полученные в процессе определения принципов лингвистической и экстралингвистической маркировки текстов узбекско-английского параллельного корпуса:

на основе результатов анализа принципов подбора текстов на узбекском языке для узбекско-английского параллельного корпуса, преимуществ подбора текстов на английском языке, принципов лингвистического и экстралингвистического тегирования текстов узбекско-английского параллельного корпуса в Ташкентском государственном университете узбекского языка и литературы имени Алишера Навои была подготовлена электронная (человеческий облик, темперамент) книга узбекского и английского языков «I-OT-2019-42"., образ природы и национальные символы))» (справочник Министерства высшего и среднего специального образования от 20 ноября 2020 г. № 89-03-4852). В результате на узбекском и английском языках было достигнуто автоматическое определение контекстуального значения слов, обозначающих образ человека, характер, природу и национальные символы;

выводы, связанные с формированием лингвистической базы параллельного корпуса узбекского и английского языков и подготовкой лингвистического программного обеспечения для программирования, были использованы в образовательном процессе Стамбульского университета, то есть по дисциплинам “Лингвистика”, “компьютерное языкознание”, “корпусное языкознание” и “лингвистическая типология”, а также результаты научно-теоретических заключений по подготовке лингвистического программного обеспечения для программирования (ректор Стамбульского университета проф. справочник доктора Махмуда Ак от 20 мая 2020 года). В результате общий анализ особенностей создания узбекских и английских программ параллельного корпусного программирования, комментарии и выводы всемирно известных специалистов по этому вопросу послужили обогащению теоретических представлений студентов об узбекском и английском параллельном корпусе;

выводы о преимуществах узбекского и английского параллельных корпусов в области лингвистики и перевода, а также принципы отбора материала для узбекско-английского параллельного корпуса были использованы в докладе, сделанном в индийском культурном центре имени Лала Бахадура Шастри на тему «Проблемы лингвистического обеспечения для формирования лингвистической базы и программирования параллельного корпуса узбекского и английского языка» (справочник индийского культурного центра имени Лала Бахадура Шастри от 18 сентября 2020 года). Это послужило основой для формирования лингвистической базы параллельного корпуса узбекского и английского языков специалистов центра и дальнейшего расширения базы знаний о лингвистическом обеспечении для программирования;

Научные взгляды на общие черты узбекско-английского параллельного корпуса, факторы, отличающие его от других корпусов на иностранном языке, были использованы при подготовке сценария Бухарской областной телерадиокомпании “Ассалом, Бухара”, “Everyday English”, “Илм дунеси” (справочник Бухарской областной телерадиокомпании от 18 августа 2020 года № 1/198). В результате широкое использование узбекско-английского параллельного корпуса для составления текстов на узбекском языке, анализ процесса интеграции узбекско-английского параллельного корпуса привели к повышению уважения к национальному языку и повышению мотивации к изучению иностранного языка.

Апробация результатов исследования. Результаты данного исследования были обсуждены на 2 международных и 5 республиканских конференциях.

Опубликованность результатов исследования. Всего по теме диссертации опубликовано 18 научных работ. В частности, в научных изданиях Высшей аттестационной комиссии Республики Узбекистан рекомендовано к публикации основные научные результаты докторских диссертаций 8 статей, из них 4 в республиканских и 4 в зарубежных журналах.

Структура и объем диссертации. Диссертация состоит из введения, трех основных глав, заключения, списка использованной литературы и приложения. Общий объем 142 страницы.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснована актуальность и необходимость темы, показана зависимость исследования от приоритетных направлений развития науки и техники республики, даны цели и задачи, описаны объект и предмет исследования, изложена научная новизна и практические результаты, раскрыта научная и практическая значимость результатов, проанализированы, апробированы, представлены сведения о структуре опубликованных работ и диссертации.

В первой главе исследования под названием “**Параллельный корпус как лингвистическая и лингводидактическая база**” рассмотрены общая характеристика, структура и состав параллельного корпуса, значение параллельного корпуса в лингвистике, лингводидактике иностранного языка и возможности в области переводоведения, характеристика существующих параллельных корпусов (русско-английский, русско-китайский, русско-французский поэтический параллельный корпус¹¹): исследованы структура и возможности параллельного корпуса.

В первом разделе первой главы, озаглавленной как “Общая характеристика, структура и состав параллельного корпуса” приводится общая характеристика параллельного корпуса. Параллельный корпус (Parallel Corp) – электронный аналог текстов параллельного перевода; состоит из

¹¹ <http://www.nevmenandr.net/fr/>

множества блоков «оригинальный текст и один/несколько их переводов». Электронные тексты в корпусе могут быть как самим оригинальным текстом, так и его часть¹². Также параллельным корпусом считается корпус, в котором собран исходный текст и его перевод¹³. Структура, состав и возможности параллельного корпуса всесторонне изучены в отечественной корпусной лингвистике¹⁴. Параллельный корпус-корпус, состоящий из совокупности электронных текстов в оригинале и переводе. Оригинальные и переводные тексты не просто ставятся рядом, а разбиваются на части на уровне соответствия по содержанию высказываний (систактических единиц) в этих двух текстах, причем эти единицы стоят рядом, в положении ссылки/ссылки друг на друга. Фрагмент в переводе, соответствующий фрагменту текста в оригинале, будет помечен. Именно это обстоятельство позволяет с помощью таких корпусов выполнять различные языковые операции¹⁵.

В современной корпусной лингвистике параллельный корпус представлен как многоязычный корпус (Multilingual Corp.) и корпус перевода (Translation Corp.). Структурное содержание корпуса с таким свойством, исходя из его назначения, может быть¹⁶ (1) в виде обычного текста, ссылающегося на перевод; (2) в виде «текстов в окне», удобных для сравнения; (3) в виде базы данных. По структуре и расположению текстов, адаптации их единиц различают несколько видов параллельного корпуса: однонаправленный, двунаправленный, многонаправленный (напр., перевод текста с английского на русский, немецкий, французский языки).

Параллельный корпус служит для наблюдения за структурным содержанием языка перевода, возможностями словосочетаний и слов в определенном контексте. Такие корпуса широко используются в сравнительной лингвистике (с целью анализа структуры двух языков); в области перевода (с целью поиска эквивалента исходного текста на других языках); в автоматическом переводе; в лингводидактике; лексикографии.

По материальному языку, из которого состоят корпуса, они изначально подразделяются на одноязычные и многоязычные типы корпусов. Двух-или многоязычные корпуса выглядят как 1) параллельные или переводческие корпуса; 2) сравнительные корпуса.

Независимо от того, какую цель преследуют параллельные корпуса, составные части должны быть адаптированы. Адаптация осуществляется на основе программы bitext. Параллельная коррекция текста-процесс

¹² Соснина Е.П. Параллельные корпуса в обучении языку и переводу // http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_education_translation/

¹³ <https://translationrating.ru/glossary/parallelnyj-korpus/>

¹⁴ Добровольский Д. Какие задачи решают параллельные корпуса? // <https://postnauka.ru/video/>; Добровольский Д.О. Корпус параллельных текстов в исследовании лексической семантики // <http://www.dialog-21.ru/media/2503/dobrovolskij.pdf>; Тао Ю., Захаров В.П. Разработка и использование параллельного корпуса русского и китайского языков // Автоматизация обработки текста. (Электрон ресурс): <https://www.researchgate.net/publication/306911853>; Кокорева А.А. Корпус параллельных текстов в обучении иностранному языку. – Вестник ТГУ (Тамбовский государственный университет), выпуск 2 (118), 2013. – С. 57-62.; Соснина Е.П. Параллельные корпуса в обучении языку и переводу (Электрон ресурс): http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_education_translation/

¹⁵ Добровольский Д. Какие задачи решают параллельные корпуса // <https://postnauka.ru/video/54851>

¹⁶ <https://translationrating.ru/glossary/parallelnyj-korpus/>

идентификации совпадающих предложений, абзацев в тексте оригинала и перевода. Такое выравнивание частей текста может быть разным. Части различаются в зависимости от цели автора¹⁷: глава, раздел, абзац, предложение и т. д. Наиболее удобной для пользователя адаптированной формой является согласование оригинальных и переводных текстов по пунктам. Даже самая совершенная программа автоматического перевода не может дать точного решения для определения границ предложений. Мысль в оригинальном тексте может быть выражена одним предложением, двумя-тремя предложениями или наоборот. Поэтому одной командой невозможно выровнять части текста оригинала и перевода. Достижение соответствия текста в параллельном корпусном содержании предложениям открывает путь к выполнению различных языковых операций. В процессе перевода предложение текста в оригинале может быть разделено в тексте перевода на два предложения, и наоборот, несколько предложений в оригинальном тексте могут быть даны в качестве одного предложения в переводе, роль предложений может быть заменена.

Следует отличать параллельный корпус от сравнительного. В сравнительном корпусе тексты, созданные по одной теме на разных языках, собираются в один корпус для целей компаративистики. Такие тексты также служат для решения определенных организационных вопросов. К сравнительному корпусу прибегают в разных языках, когда необходимо проследить, как выражается та или иная ситуация дискурса, событие в том или ином предмете.

В.И.Мельник объясняет разницу между параллельным и сравнительным корпусом следующим образом¹⁸ :

Характеристика параллельного корпуса	Характеристика сравнительного корпуса
ПК хранит в себе два языковых материала, в отличие от сравнительного корпуса.	Тексты в СК будут принадлежать одному языку.
ПК будет четким, идеально отфильтрованным.	В СК разметка не предусмотрена.
В ПК достигается согласованность перевода.	СК не связан с процессом перевода.

Во втором разделе этой главы, озаглавленном как “Значение параллельного корпуса в лингвистике, лингводидактике иностранного языка и возможности в области переводоведения”, исследуется образовательное значение этих корпусов и возможности в переводе. Параллельные корпуса могут служить удобным образовательным ресурсом для образовательного процесса. В качестве дидактической характеристики языка можно указать (а) многоязычность, (б) то, что результат поиска имеет контекстную форму, (в) возможности языка, ограничивающие его область применения. Она дает возможность обучить школьника переводу, выполненному с особой целью,

¹⁷ Добровольский Д. Какие задачи решают параллельные корпуса. (Электрон ресурс): <https://postnauka.ru/video/54851>

¹⁸ Мельник В.И. Переводческие корпуса – что они могут сказать нам о переводе и его преподавании? (Электрон ресурс): <https://cyberleninka.ru/article/n/perevodcheskie-korpusy-chto-oni-mogut-skazat-nam-o-perevode-i-ego-prepodavanii>

лучше определить значение слова, сформировать лексические навыки на основе наблюдения за употреблением слова в конкретных речевых условиях. Параллельный корпус наглядно показывает, какое из слов синонимического ряда уместно использовать в тексте перевода¹⁹. Нецелесообразно ограничиваться выводом о том, что параллельные корпуса полезны только для теории перевода, двуязычной лексикографии, других направлений лингвистики, связанных с сравнением языков²⁰. Немаловажное значение имеет и более глубокое описание семантического состава слов в лексике сравниваемых языков.

Корпус полезен при анализе имеющегося материала и выявлении недостатков в обучении иностранному языку. Он носит общетеоретический характер и та или иная структура/лексика анализируется на основе имеющегося типового учебника, словаря и корпуса; затем сравнивается результат из обоих источников. Таким образом, параллельный корпус представляет собой современный учебно-методический инструмент, предоставляющий объем материалов для лингвистических исследований. Слова *big*, *large* и *great* наблюдаем из корпуса. В тезаурусе эти слова рассматриваются как синонимичные слова, обозначающие величину. А в примерах из русско-английского параллельного корпуса они используются следующим образом:

<i>Следует сказать, что Дидро не верил вначале в возможность создания в Петербурге большой картинной галереи.</i>	<i>It should be noted that Diderot first did not believe in the possibility of collecting a big picture gallery in Petersburg. [28]</i>
<i>Но все больше фактов свидетельствуют о том, что самые большие опасности заключает в себе экипаж.</i>	<i>But there is increasing evidence to show that one of the greatest hazards lies in the crew itself.[30]</i>
<i>За годы своего существования он превратился из дворцовой галереи в музей, выделяющийся в ряду крупнейших музеев мира разносторонностью своего собрания.</i>	<i>From a palace's picture gallery it went down to our time as one of the world's largest museums distinguished for versatility of its collection.[28]</i>

Анализ приведенных примеров показывает, что слово *large* в большей степени выражает значение величины, качества слова, называющего конкретный предмет, в то время как в словах *big* и *great* эта функция не ограничена. Слова *большой* и *большой* чаще встречаются в устойчивых сочетаниях по сравнению с прилагательным *большой*. *Great* чаще используется для выражения эмоций: *большая* изношенность, *большое* чувство. Русско-английский параллельный корпус наглядно показывает ситуацию употребления слов-синонимов. А в словарях такой информации нет. Другое различие между ними заключается в том, что эти лексемы

¹⁹ Добровольский Д. Какие задачи решают параллельные корпуса? (Электрон ресурс): <https://postnauka.ru/video/54851>

²⁰ Добровольский Д.О. Корпус параллельных текстов в исследовании лексической семантики. <http://www.dialog-21.ru/media/2503/dobrovolskij.pdf>

встречаются в структуре словарного запаса. В таких сочетаниях, как *big noise*, *big head*, неуместно заменять прилагательное его синонимами. Не всегда удается заменить слово синонимом, сохранив при этом смысл.

В заключение следует отметить, что корпусный подход к современному переводческому образованию приносит все новые и новые возможности. В отличие от национального корпуса, такой корпус обеспечивает основу для изучения особенностей перевода в конкретной области на основе большого массивного материала. Создание параллельных корпусов, касающихся научно-технической, социально-политической, экономической и других сфер, позволяет проводить специальные исследования методологических возможностей языка именно в этой области.

Вторая глава называется **“Теоретические основы построения параллельного корпуса узбекско-английского языка”**. В ней рассмотрены принципы отбора материала для узбекско-английского параллельного корпуса, вопросы лингвистической и экстралингвистической маркировки узбекско-английского ПК.

В первом разделе этой главы, озаглавленном как “Принципы подбора материала для узбекско-английского параллельного корпуса”, рассмотрены принципы подбора текста для ПК. Материалом для составления русско-английского параллельного корпуса послужили электронные формы относительно старого перевода Констансии Гарнетт произведений русских классиков Н.В.Гоголя “Петербургские вести”²¹, Ф.М.Достоевского “Преступление и наказание”²², рассказы А.П.Чехова²³.

Это означает, что при выборе текста у автора есть электронный формат и она опирается на готовый перевод на произведени. Опираясь на мнение специалистов Ю.М.Мухина и И.Яна мы пришли к выводу, что при выборе источника необходимо учитывать следующие факторы:

1. Для повышения качества перевода текст перевода должен быть переведен известным, квалифицированным специалистом или официальной организацией. Источником материала для этого корпуса послужил официальный сайт правительства Китайской Народной Республики²⁴. Перевод материалов данного сайта осуществляется компанией Central Compilation & Translation Bureau; что свидетельствует о гарантированном качестве перевода.

2. Выбранный материал должен быть структурно-семантически устойчивым. Например, ежегодное “Выступление” премьер-министра будет иметь повторяющееся слово/предложение/соединение, даже текстовую структуру. Это естественно. Такие повторяющиеся единицы важны для

²¹ The Collected Tales of Nikolai Gogol / translator Pevear R., Volokhonsky L. New York: Pantheon Books, 1998. – 435 p.

²² Crime and Punishment by Fyodor Dostoyevsky [Электронный ресурс] // Project Gutenberg: [сайт]. URL: <http://www.gutenberg.org/ebooks/2554>

²³ The Lady with the Dog and Other Stories by Anton Pavlovich Chekhov [Электронный ресурс] // Project Gutenberg: [сайт]. URL: <http://www.gutenberg.org/ebooks/13415>

²⁴ <http://cn.theorychina.org/>

перевода текста оригинала и перевода, а также для автоматической разметки на следующем этапе.

3. В отличие от перевода художественного произведения, при переводе официально-ведомственных, дипломатических, управленческих документов содержание и структура оригинального текста должны быть сохранены. По этой причине последовательность предложений/предложений в таком оригинальном и переведенном тексте будет одинаковой; структурно единой; похожим. Таким способом облегчается задача выравнивания единиц текста на двух языках.

На наш взгляд, при отборе текстов в узбекско-английском параллельном корпусе будут приняты во внимание воспитательные, нравственно-культурные факторы, если осуществляется отбор текстов следующего содержания: превозношение выдающихся героев, защищавших как зеницу ока свой народ и землю (1); защита прав женщин (2); критика пороков в характере людей (3); превозношение идей мужества, смелости, трудолюбия, честности, верности, щедрости (4).; Материалы для ПК будут свободны от контента, который противоречит нашему воспитанию/культуре.

Во втором разделе, озаглавленном как "Проблемы лингвистической маркировки текстов узбекско-английского параллельного корпуса", рассмотрены проблемы лингвистических тегов. Параллельная корпусная автоматическая обработка отличается от других корпусов следующим: основной, технической проблемой автоматической обработки является разметка на уровне предложения, словосочетания и словосочетания текста в оригинале и переводе как корпусных единиц. Это сделать сложно на основе модели автоматического поиска в НП²⁵. Для специального параллельного корпуса языков с большой лингвистической базой, таких как европейские языки, китайский и японский, разработаны удобные способы выравнивания единиц текста оригинала и перевода²⁶.

Вторым важным этапом корпусообразования является выравнивание текста в рамках предложения, т. е. в параллельном корпусе предложение на двух языках выравнивается по содержанию со своим эквивалентом. Единицами текста могут быть предложение, словосочетание и слово в параллельном корпусе. Для ПК, составляемого в учебных целях, достаточно будет выравнивания в пределах предложения. Ведь выравнивание на следующем этапе не обходится без этого начального этапа. Когда говорят о единице выравнивания корпуса, применяют чисто структурный (длина предложения, количество слов) и статистический (на основе выражаемого слова) методы²⁷. Такой способ не требует отличной словарной базы и может быть применен для языков, не обладающих большим словарным ресурсом. При выравнивании параллельных предложений меньшая единица

²⁵ Dzemyda G., Sakalauskas L. Optimization and knowledge-based technologies // Informatica. 2009. Vol. 20(2) – P. 165-172.

²⁶ Tiedemann J. News from OPUS – A Collection of Multilingual Parallel Corpora with Tools and Interfaces // Recent Advances in Natural Language Processing. 2009. Vol V. – P. 237-248.

²⁷ Brown P.F, Della Pietra V.J., Della Pietra S.A., Mercer R.L. The mathematics of statistical machine translation: parameter estimation // Computational Linguistics. 1993. Vol. 19(2).

предложения может быть реализована на основе статистических моделей машинного перевода: есть возможность перевести на второй язык любой элемент предложения – слово в оригинале. В параллельном корпусе перевод слова в оригинале на второй язык зависит от его частоты: в параллельном корпусе необходимо, чтобы частота обоих была одинаковой.

Приведем типичный пример того, как одно предложение в тексте оригинала выражается в переводе как два предложения: “Что случилось, вы больны?” <> “Что это?” “What’s this? ” “Are you ill?”²⁸. Такие ошибки легко исправляются в процессе лексического разбора. Но есть и относительно сложный процесс: несколько предложений представлены 2-3 предложениями, но граница между ними не совпадает: Он воскликнул: «Он еще спит!» "My goodness; how he sleeps! – she cried indignantly: And he is always this". <> «Боже мой; как он спит! - возмущенно воскликнула она: А он всегда спит». В этом случае вы должны уравнивать, не устанавливая границ предложений. Вот некоторые из «новых пар», встречающихся в переводе «Петербургских рассказов» Н. В. Гоголя²⁹ на стадии уравнивания слов: преступление <> злое мертвое; поручить <> enjoin; застывший <> frozen; взять <> председатель; сторож <> watchmen; принять <> concive; свой <> свой; оставить <> neglect; требовать <> compel; действовать <> импульсивно; жениться <> begin to think; ввести <> recur; бледный <> бедный; слышал <> hearkened; времяпрепровождение <> diverty; расположение <> State; приятно <> delightful; известно <> as every one knows³⁰.

Этап синтаксического анализа. Следующий этап выравнивания на уровне слов оригинальность и перевод требует проведения синтаксического разбора в предложении. В последнее время предпринимаются большие усилия по созданию синтаксического разбора в размеченных корпусах. Такой анализ помогает проверить грамматическую теорию, а также открывает новые синтаксические закономерности. Такой процесс называется «упражнением» формальной грамматики: при нем результат анализа становится качественным. В грамматическом упражнении создается корпус с синтаксической разметкой, который широко используется под названием «treebank»³¹. Есть несколько отличных парсеров, которые могут моделировать английский, романский и германский языки в целом. К сожалению, этот парсер не может быть использован для создания параллельного корпуса узбекского языка, поскольку этот тип парсера не может формировать синтаксические модели узбекского языка.

²⁸ қўштирноқ белгиси гапларнинг ажралган ҳолатини кўрсатади.

²⁹ The Collected Tales of Nikolai Gogol / translator Pevear R., Volokhonsky L. New York: Pantheon Books, 1998. 435 p.

³⁰ Кедрова Г.Е., Потемкин С.Б. Использование корпуса параллельных текстов для пополнения специализированного двуязычного словаря // Труды и материалы III Международного Конгресса исследователей русского языка «Русский язык: исторические судьбы и современность» (Москва, 20-23 марта 2007): М.: МАКС Пресс, 2007. С. 27-628.

³¹ Large-Scale Induction and Evaluation of Lexical Resources from the Penn-II Treebank / R.O'Donovan, M.Burke, A.Cahill, J. van Genabith, A.Way // Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics. Barcelona, 2004. – С. 368-375.

В третьем разделе главы под названием “Об экстралингвистическом тегировании текстов узбекско-английского параллельного корпуса” описаны вопросы экстралингвистического тегирования. В исследованиях выделяют следующие виды экстралингвистической разметки³²:

1. Отражающие специфику текстового формата (глава, абзац, раздел и т.д.))-.

2. Текст, тег, представляющий информацию, относящуюся к его автору. Потому что информация об авторе может указывать не только имя, но и такие данные, как возраст, пол, год в котором он жил. Текстовая информация, кроме названия текста (произведения), охватывает язык, год написания и издания. Наличие такой информации позволяет осуществлять детальный поиск по базе. Экстралингвистический тег или метаданные (метаданные) – вид тега, включающий внешнюю-интеллектуальную информацию, библиографическое, типологическое, тематическое, социологическое описание; формальный-структурный тег (текст, раздел, глава, часть, абзац, предложение), технико-технологический тег (кодировка, дата обработки электронной версии источника).

При написании экстралингвистических комментариев к текстам рекомендуется писать следующие комментарии к исходным текстам на узбекском языке и к переведенным текстам на английском языке.

Глава, абзац, раздел и т. д. текстового формата на узбекском языке. части выделяются, обозначаются, а также обозначаются те же главы, абзацы, части текста перевода на английский язык и выравниваются части текста для обоих языков: чем больше частей в оригинале, тем больше частей в тексте перевода. В случае выпадения абзаца в переводе указываются такие части. В процессе подбора текста для узбекско-английского параллельного корпуса мы обращаем особое внимание на равенство частей произведения (абзацев). Потому что пользователю ПК лучше параллельно наблюдать и изучать все части в двуязычном тексте.

Следующим комментарием к экстралингвистической отметке текста должен быть текст, тег, представляющий информацию, принадлежащую его автору. В узбекской части узбекско-английского параллельного корпуса информация об авторе произведения состоит из его имени, возраста, пола, года проживания и т. д., В то время как в английской части эта информация дается на английском языке. При этом информация о переводчике работы прилагается исходя из тех же параметров.

Узбекский и английский языки – в обеих частях информация о тексте состоит из комментария, такого как название текста (произведения), язык (узбекский/английский), год написания и издания. Наличие такой информации позволяет осуществлять детальный поиск в обеих частях корпуса одинаково. Если такое замечание содержится только в тексте оригинала, но перевод не содержит экстралингвистического комментария к

³² Хамроева Ш. Ўзбек тили муаллифлик корпусининг тузишининг лингвистик асослари: Филол. фан. бўйича фалсафа докт. (PhD) диссертацияси. – Бухоро, 2018. – Б. 72-73.

тексту, поиск не даст равного результата: пользователь прочитает только информацию текста на одном языке. Поскольку главной особенностью параллельного корпуса является возможность одинаково наблюдать все особенности текста на обоих языках, то и экстралингвистическая информация должна быть закреплена таким образом, что тексты должны быть маркированы по этому принципу.

Следовательно, объем, степень аннотации корпуса (разметки агрегатов) определяется целью исследователя. Большой объемный параллельный корпус, пропорционально подобранный по тематике, хронологии, жанру, служит для изучения общих черт языка. Параллельная корпусная автоматическая обработка отличается от других корпусных: основной, технической проблемой автоматической обработки является разметка на уровне предложения, словосочетания и словосочетания текста в оригинале и переводе как корпусных единиц. Для этого понадобятся различные статистические методы, сравнение частоты. Узбекско-английский параллельный корпус набор экстралингвистической разметки для оригинального текста будет содержать информацию о том, сколько слов содержит текст: имя, фамилия, псевдоним, Пол, дата рождения, название текста, время и место написания, объем текста. Для текста перевода к этим данным добавляется информация о переводчике текста, времени, месте перевода, объеме текста в переводе. Отображение жанра текста, типа текста, хронотопа текста в обеих частях параллельного корпуса совершенствует экстралингвистическую разметку.

В третьей главе исследования, озаглавленной как **«Лингвистическая база и основы программирования узбекско-английского параллельного корпуса»**, рассмотрена разработка лингвистического и программного обеспечения для параллельного корпуса, передача реальности и эквивалентных единиц при переводе.

В первом разделе этой главы, озаглавленной как "Вопросы подбора текста для параллельного корпуса", рассмотрено лингвистическое обеспечение ПК. В целях обеспечения репрезентативности корпуса в качестве корпусной единицы отбираются тексты различных жанров. Какую тематику и жанр текста выбрать для ПК, зависит от целей составителя. ПК могут отображать односторонний (с узбекского на английский) или двусторонний (с узбекского на английский, с английского на узбекский) перевод. При подборе текста для параллельного корпуса узбекско-английского языка рекомендуется собирать переводы с узбекского на английский язык и прямые переводы с английского на узбекский язык.

Узбекско-английский параллельный корпус будет состоять из двух блоков: узбекско-английского языкового блока (1) и англо-узбекского языкового блока (2). В поисковой системе запускается поиск по жанру, периоду создания, автору и переводчику произведений, в связи с чем требуется сортировка произведений по этим параметрам. Специалист по теории перевода, опытный переводчик М.Холбеков дает ценную информацию о переводах, осуществленных в XX веке с иностранных

языков³³.

Насколько актуально воссоздание корпуса из готового продукта, если есть перевод произведения? Поскольку корпус предназначен не для перевода, а для сбора готовых переводов, что он дает науке и теории перевода? Мы ответим на приведенные выше уместные вопросы следующим образом:

1) Существующие параллельные корпуса мира основаны на готовых переводах. Мы также должны получить готовые тексты переводов на узбекско-английский параллельный корпус;

2) Объединение перевода в параллельный текст позволяет читать все произведение рядом в оригинале и в переводе;

3) Самое главное, что параллельный корпус - это наиболее удобный способ наблюдать, как та или иная единица (простое предложение, составное предложение, фраза, фразеологизм, сленг, метафора) выражается в оригинале и в переводе.

Поэтому при выборе материала для любого параллельного корпуса преобладает принцип получения варианта перевода, выполненного непосредственно из оригинала. Потому что единицы теряют свою ценность в косвенном переводе, параллельный корпус в этом случае не выполняет свою функцию полностью. Можно также включить косвенные переводы, чтобы получить только непосредственно переведенную копию произведений, которые были переведены несколько раз, или для сравнения. Это происходит из намерения составителя создать исследовательский или иллюстративный корпус.

Сравнивая переводы, М. Холбеков обосновывает тот факт, что в оригинале захватывающий монолог, начинающийся строками “To be, or not to be: that is the question”, обращен примерно в той же форме, что и русские переводы³⁴:

1	М.Вронченко, (1829)	<i>“Быть или не быть – таков вопрос”</i>
2	А.Кронеберг, (1844)	<i>“Быть или не быть? Вот в чем вопрос!”</i>
3	М.Лозинский, (1933)	<i>“Быть или не быть – таков вопрос”</i>
4	Б. Пастернак, (1940)	<i>“Быть или не быть: вот в чем вопрос”</i>

Сравнение показывает, что узбекские переводчики переводили один и тот же стих в разных формах.

1	В переводе Чулпана	<i>Ё ҳаёт, ё ўлим, масала шунда.</i>
2	Перевод М. Шайхзода.	<i>Тирик қолмоқ ё ўлмоқ? Шудир масала!</i>
3	Перевод Джамалы Кемаля	<i>Ё ҳаёт, ё мамот: масала шундоқ!</i>

По словам М. Холбекова³⁵, при сравнении переводов с оригиналом становится ясно, что перевод Чолпона близок к русской копии М. Лозинского, перевод М. Шайхзода к тексту А. Кронеберга. Перевод Джамалы Кемаля, похоже, был переведен так же, как и оригинал.

³³ Холбеков М. Ўзбекистонда таржима ва таржимашунослик (1991-2016 йиллар) // Хорижий филология. – 2016 йил. – №3. – Б. 16-42.

³⁴ Кўрсатилган асар.

³⁵ Кўрсатилган асар.

Во втором разделе под названием “Перевод единиц в узбекско-английском параллельном корпусе” рассмотрен вопрос о передаче морфем, подвергшихся трансформации в переводе, в параллельном корпусе, о передаче реалий в переводе и их представлении в параллельном корпусе. В практическом смысле явление представления морфемы морфемой в переводе не редкость. Хотя теоретически существуют различные языковые эквиваленты морфем в английском языке, такие как *-er, -or, -ki, -man, -ship, -ment, -hood*, в узбекском языке не всегда удается достичь эквивалентности в переводе. Потому что в толковых и двуязычных словарях всегда присутствуют слова, не участвующие в предложении, то есть эквиваленты слов в системе. По мере использования в речи изменяются их форма, стиль, типы текста и структуры, используемые в предложении³⁶. В какой степени какой-либо из способов морфемного преобразования будет применен в переводе, будет напрямую зависеть от текста, языковых единиц в нем, особенностей языков, вступивших в контакт в связи с переводом.

Параллельный корпусный перевод позволяет параллельное наблюдение за текстами, но имеет ограниченную степень выделения состояния, как и вышеприведенные преобразования. Потому что большинство существующих параллельных корпусов отображают логику двух (оригинала и перевода) языков параллельно только в форме предложения. Но теоретически в ПК также есть возможность параллельного отображения сочетаний и слов. Если компилятор ПК нацелен на то, чтобы пометить не только предложение, но и словосочетание, а также слово в параллельном состоянии, то необходимо обратить особое внимание на вышеупомянутые преобразования.

Произвольное составное или искусственное слово, подвергшееся трансформации в переводе, может быть параллельным как в оригинале, так и в переводе. Эта задача выполняется в следующей последовательности:

- 1) сравниваются тексты перевода и оригинала и составляется список таких единиц /явлений;
- 2) приводится узбекский перевод единиц, подвергшихся трансформации;
- 3) единицы распараллелены между собой, помечены;
- 4) формируется база данных – словарь единиц измерения.

Роль реальности в выражении специфических аспектов национального языка уникальна, поскольку реальность как единица языка является одним из безальтернативных лексиконов. Это неотъемлемая часть лингвистических знаний, которая представляет большой интерес для изучения взаимосвязей языка и культуры³⁷. О реалиях английского языка мы можем узнать из произведений Л.С. Бархударова, Г.Д. Томахина, А.Д. Швейцера, А.О.

³⁶ Ўша манба.

³⁷ Жуманиёзова М., Атажанова А. Реалия ҳақидаги қарашлар талқини // Таржима, ахборот, мулоқот – сиёсий ва ижтимоий кўприк. Халқаро илмий-амалий анжуман материаллари. – Самарқанд: СамДЧТИ, 2019. – 360 б. – Б.158-159.

Иванова³⁸. М.Джуманиёзова, А.Атаджанова согласны с мнением Г.Д. Томахина³⁹ о том, что реальность - это «названия предметов материальной культуры, принадлежащие определенному народу и нации, исторические факты, государственное устройство, имена национальных и фольклорных героев, мифологических существ и т. Д.» Следовательно, отсутствие реалий, отражающих концепции жизни на языке перевода, требует использования метода транслитерации; через перевод закладывает основу для сохранения самобытности реалий этнокультуры. Параллельные корпуса также отражают свою культуру при сравнении и изучении двух местных языков, а это означает, что чрезвычайно важно параллельно различать реальности на ПК.

В третьем разделе главы, который называется “Основы формирования лингвистической базы узбекско-английского параллельного корпуса” анализируются вопросы формирования лингвистической базы узбекско-английского параллельного корпуса. Для анализа материалов отобраны повесть У. Хашимова "Мирские дела"⁴⁰, перевод данного произведения на английском языке под названием “Affairs of life. Stories” выполненный О.Муминовым⁴¹, в частности рассказы “Долг” (“A debt”), “Валенки” (“A sock”), “Детский плач” (“Baby’s cry”).

При создании ПК желательно создать программу, способную анализировать тексты на узбекском языке. Это программа <из точки в точку>, <из точки в вопрос>, <из точки в восклицательный знак> и т. д. он предназначен для разделения простых, составных предложений на основе принципов. Но мы распараллеливали тексты вручную, так как целью работы является не только распараллеливание текста, но и определение того, какие единицы в нем даны при параллельном тегировании текстов.

Рассказ "Долг" разбит на 30 предложений (простое и составное), и приведен параллельный перевод на узбекский и английский языки. В таблице выделены особые правила и исключения в процессе перевода на узбекский и английский языки, выделены единицы, которые не соответствуют друг другу дословно. В процессе перевода встречаются случаи, когда простое предложение выражается слитным предложением (или наоборот), выпадение или добавление определенных абзацев, несоответствие границ предложения в результате использования сленга в переводе. Параллельному корпусовщику придется вручную исправлять такие ситуации после автоматического параллельного тегирования текста. Мы тоже наблюдали подобные случаи и выделили их отдельно. В конце концов, эти практики лежат в основе создания словарей, которые впоследствии станут средством формирования языковой базы для ПК.

³⁸ Бархударов Л.С. Язык и перевод. – М: Международное отношения, 1975.; Томахин Г.Д. Реалии-америкнизмы. – М. Высшая школа, 1988.

³⁹ Томахин Г.Д. Реалии в языке и культура // ИЯШ. 1997. – №3. – С.13.

⁴⁰ Hoshimov O`. Dunyoning ishlari: qissa. – Toshkent: “Sharq” nashriyot-matbaa aksiyadorlik kompaniyasi bosh tahririyati, 2005. – 208 b. – B. 10-11.

⁴¹ Utkir Hoshimov. Affairs of life. Stories (Translator: O.M.Muminov). – Tashkent, 2013. – 164 p. – P.9-10.

Единица “*Ия, юбиляр экансан-да, шошмай тур-чи, ҳозир...*” в 19 строке АТПЖ переводится в форме *Oh, congratulations! Wait, just a minute...* и выражается односложным предложением или двумя простыми предложениями. Кроме того предложение *Ия, юбиляр экансан-да* передано в форме с *Oh, congratulations* и тип предложения изменился в соответствии с целью выражения. Выражение *юбиляр экансан* в переводе дается словом, выражающим смысл *поздравляю*. Если такие единицы не тегированы, то корпус поиска по ключевым словам будет уменьшен.

Предложение *As soon as the weather got cold she used to have a pain in her legs* из рассказ в «Валенки» приведенный 7 строке АТПЖ выражается четверз сложное предложение *Салқин тушиши билан шишиб кетади, оғрийди*

Простые предложения *Терлаб туриб муз едим. Бир куни акаларим билан яхмалак ўйнаб терлаб кетдим* приведенные в 20-21 строках АТПЖ выражены с помощью сложных предложений в форме *I played on an is-rink with my brothers and sweated up and I ate an is then.*

Два простых предложения *У томогимга нимадир қилди. Дод солиб қўлини тишлаб олдим* приведенные в 45-46-й строках АТПЖ переведены как одно сложное предложение в форме *She did something with my throat and I bit her finger*

Также встречается еще одна типичная ситуация. В тексте перевода появилось предложение, отсутствующее в тексте на узбекском языке, Они приводятся с целью уточнения содержания текста.

В 62 строке АТПЖ модно увидеть, что контекст *Сарпойчан келавердингизми?! – деди Ҳожжи буви ҳамон ўша ваҳимали оҳангда* переведен как *Did you come like this? – said Hoji grandmother.* Видно, что что не нашлось подходящего эквивалента слову *сарпойчан*, и сочетание *шундай келавердингизми? (Did you come like this?)* переводится как общий смысл, но читатель может понять значение из контекста.

Из строки 63 АТАЖ видно, что в тексте перевода появилось предложение, отсутствующее в тексте на узбекском языке: *Still in the same tone.*

Предложение *Оппоқ бантик тақиб, атлас кўйлак кийиб олибди* из рассказа "Долг" представленной в строке 16 АТПЖ переводится как *She wore bright white band and atlas dress.* Обратите внимание, что слово *атлас* считается реалием, и оно было переведено на английский в форме *atlas* а не в форме шелкового платья (*silk dress*).

Выделенное слово из сложного предложения *Ойим томогимни аччиқтош билан чайиб кўрди, бўлмади, туршак қайнатиб сувини ичирди, бўлмади...* в рассказе “Валенки”, приведенного в 24 строке АТПЖ переведен в качестве сочетания в предложении *Mother washed my throat with alum, but it was useless. She gave me water of boiled **dry apricot fruit** but there was no result.* Потому что в английском языке нет эквивалента слова *туршак*, соответственно, это реалия.

Особо следует остановиться на тех единицах, которые в оригинальном виде не могут быть переданы в узбекско-английском параллельном корпусе, а также на проблемах их приспособления к поискам.

Предложение *Менинг ишимга аралашиб нима қиласан* из рассказа “Долг” приведенное в 12 строке слово не переведено дословно а предоставлено устойчивой фразой *it is not your business* на английском языке Следовательно, нет возможности искать, отслеживать эквивалент перевода таких слов.

Заголовок рассказа «Валенок» переводится как “A sock” Это связано с тем, что в английском языке не найдено эквивалентного слова для обозначения этой единицы. На наш взгляд, можно переводить в форме сочетания, но переводчик предпочел это. Следовательно, эти единицы не переводятся их эквивалентами; эту единицу также необходимо включить в словарь параллельного корпуса среди единиц, не имеющих эквивалента в переводе.

В заключение можно сказать, что в процессе перевода простое предложение представляется сложным предложением (или наоборот), некоторые абзацы опускаются или добавляются, а использование сленга в переводе приводит к несостыковке границ предложения. Составитель параллельного корпуса должен вручную исправить такие случаи после автоматического параллельного тегирования текста. Эти методы затем лягут в основу создания словарей, которые будут средством формирования лингвистической базы для ПК.

ЗАКЛЮЧЕНИЕ

1. Параллельный корпус - это корпус, состоящий из оригинального текста и одного / нескольких его переводов. Корпус может содержать исходный текст или его часть. Оригинальный и переведенный текст не просто помещаются рядом, а синтаксические единицы в двух текстах разделены на части на уровне соответствия содержанию, в позиции указания / ссылки друг на друга. Отмечен фрагмент в переводе, соответствующий фрагменту исходного текста.

2. Параллельный корпус выглядит как многоязычный переведенный корпус. В зависимости от структуры, расположения текста, расположения блоков внешний вид параллельного корпуса различается как односторонний, двусторонний, многосторонний. Контент ПК - это обычный текст, который ссылается на перевод в соответствии с его назначением; «Текст в окне», который легко сравнивать; может быть в виде базы данных.

3. Параллельный корпус служит для наблюдения за структурным содержанием языка перевода, возможностью словосочетаний и фраз в определенном контексте. ПК может быть эффективно использован в области сравнительного языкознания, перевода, машинного перевода, лингводидактики, лексикографии.

4. На основе анализа материала ПК можно наблюдать качество, уровень перевода: текст является вольным художественным переводом или точный перевод хорошо виден в параллельном корпусе. В отличие от национального корпуса, такой корпус закладывает основу для изучения характера перевода в той или иной области на основе большого массива материала. Создание ПК в научной, технической, социально-политической, экономической и других областях позволяет специально изучить методологические возможности языка в этой области.

5. В мировой компьютерной лингвистике существует ряд согласованных программ corpus manager, на основе которых можно создать небольшой (параллельный и сравнительный) корпус специального назначения. Для узбекского языка необходимо разработать таких конкодарсных и корпусных менеджеров. Хотя существующий язык параллельного корпуса имеет много различий в языке, размере, программном обеспечении, общий аспект ПК объясняется тем, что маркировка выполняется с помощью морфологического анализатора языка корпуса, а морфосинтаксическая маркировка также выполняется на всех ПК.

6. Первое различие между корпусами чувствуется в его интерфейсе или диспетчере корпусов. На некоторых ПК интерфейс ограничивает результаты поиска такими параметрами, как год публикации, автор, жанр. Другой ПК отличается тем, что в диспетчере корпуса окно статистической обработки для первого языка дает эффективный результат, а для второго - языка перевода у него нет этой опции. Поисковая машина выполняет поиск слова, фразы, дизъюнкции, союза, отрицания; представляет собой эквивалент этих форм из переведенного текста; Также существуют разные морфологические формы поиска.

7. Технологический процесс создания ПК - это выделение текста и ввод его в исходную электронную форму; дать мета-описание текста; создать модуль поисковой системы; загрузить текст в тело; синтаксический анализ корпусных единиц; ведение экспериментальной эксплуатации; следует общим принципам, таким как анализ результатов. Автоматическая обработка параллельного корпуса отличается от других корпусов: основная техническая проблема автоматической обработки - это разметка текста в оригинале и переводится как единица корпуса на уровне речи, фразы и слова. Для этого необходимо сравнить разные статистические методы, периодичность.

8. Следующим важным шагом в формировании корпуса является выравнивание текста внутри предложения, то есть в параллельном корпусе двуязычное предложение адаптируется к своему содержанию из своего собственного эквивалента. Приравнивание двуязычных текстов ПК для определения границ предложений в тексте; оригинал и перевод помечаются границами предложений, которые выражают друг друга в тексте.

9. В зависимости от длины предложения уравнение очень чувствительно к интервалу или пунктуации: использование одного неуместного пробела / пунктуации приводит к ошибке в уравнении. Чрезвычайные ситуации не предоставляют достаточной информации для

статистического анализа. Другой важный шаг в выравнивании на уровне слов требует реализации синтаксического анализа в исходных и переходных предложениях. Создание синтаксического анализа в помеченном корпусе помогает изучить грамматическую теорию, выявляя новые синтаксические законы: поэтому синтаксический анализ на ПК важен.

10. Очень важно качество перевода в корпусе: переведенный текст должен быть переведен известным квалифицированным специалистом или официальной организацией. В отличие от перевода произведения искусства, перевод официально-административных, дипломатических, административных документов должен сохранять содержание и структуру исходного текста. Поэтому последовательность предложений / предложений оригинального и переведенного текста будет одинаковой; структурно взаимно идентичны; будет похоже. Этот метод упрощает задачу выравнивания текстовых единиц на двух языках.

11. Необходимым инструментом является программа, которая автоматически анализирует текст при создании ПК. Пунктуация, программа, определяющая границы предложений на основе графического элемента декора - основного рабочего инструмента для создания ПК. Узбекско-английский параллельный корпус экстралингвистической разметки исходного текста состоит из информации об имени, фамилии, прозвище, поле, дате рождения, названии текста, времени и месте написания, размере текста: из какого количества слов состоит текст. . Для переведенного текста эта информация включает информацию о переводчике текста, времени, месте перевода, размере текста в переводе.

12. В узбекско-английском параллельном корпусе тегирование оригинальных и переведенных статей с параллельным отображением результатов поиска для пользователя еще больше увеличивает ценность и значимость корпуса. Герои хвалят подборку узбекско-английских текстов ПК; права женщин защищены; порочная привычка человеческого характера, в которой критикуются гнусные пороки; подборка текстов, воспевающих идеи мужества, ловкости, отваги, трудолюбия, честности, верности, щедрости, обеспечит учет образовательных, моральных и культурных факторов в узбекско-английском параллельном корпусе; Материалы для ПК не содержат контента, противоречащего нашему воспитанию / культуре.

13. В реальности связь между языком и культурой становится более явной. Отсутствие реальности в переводе, представляющем понятие жизни, требует использования метода транслитерации; через перевод закладывает основу для сохранения этнокультурной идентичности действительности. Поскольку ПК также отражает их культуру в процессе сравнения и изучения языков двух народов, поэтому чрезвычайно важно параллельно различать реалии ПК.

14. При анализе перевода текстов, отобранных и проанализированных для формирования лингвистической базы узбекско-английского параллельного корпуса, бывают случаи, когда в переводе используется сленг, простые предложения представлены составным предложением (или

наоборот), отдельные абзацы опускаются или добавляются. Замечено, что тип предложений меняется в зависимости от цели выражения, первое простое предложение в составном предложении переводится, а второе опускается, предложение, которого нет в узбекском тексте, появляется в переведенном тексте. Тот факт, что предложение, которого нет в узбекском тексте, появляется в переведенном тексте, можно объяснить тем, что оно цитируется для пояснения содержания текста. Компилятор параллельного корпуса должен вручную исправить такие случаи после автоматического параллельного нажатия текста. Эти методы затем лягут в основу создания словарей, которые будут средством формирования лингвистической базы для ПК.

**SCIENTIFIC COUNCIL AWARDING
SCIENTIFIC DEGREES DSc.03/04.06.2021.Fil.72.03
AT BUKHARA STATE UNIVERSITY**

BUKHARA STATE UNIVERSITY

KARIMOV RUSTAM ABDURASULOVICH

**LINGUISTIC AND SOFTWARE ISSUES OF COMPILATION OF UZBEK-
ENGLISH PARALLEL CORPUS**

10.00.06 - Comparative literature, comparative linguistics and translation studies

**DISSERTATION ABSTRACT
for a Doctor of Philosophy scientific degree (PhD) in Philology**

The theme of the dissertation of Doctor of Philosophy (PhD) was registered at the Supreme Attestation Commission at the Cabinet of Ministers of the Republic of Uzbekistan under number B2020.2.PhD/Fil587.

The doctoral dissertation has been carried out at the Bukhara State university.

The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (abstract)) on the Scientific Council website (www.buxdu.uz) and on the website "ZiyoNet" information and educational portal (www.ziynet.uz).

Scientific supervisor: **Mengliev Bakhtiyor Rajabovich**
Doctor of Philological Sciences, Professor

Official opponents: **Akhmedov Oybek Saparbayevich**
Doctor of Philological Sciences, Professor

Ataboev Nozimjon Bobojon ugli
Doctor of Philosophy in Philology

Leading organization: **Termez State university**

The dissertation defense will be held on "22" April 2022 at 10⁰⁰ at the meeting of the Scientific Council awarding scientific degrees DSc.03/04.06.2021.Fil.72.03 at Bukhara State University (Address: 200118, Bukhara city, M.Ikbol street, 11. Phone: (0 365) 221-29-14; Fax: (0 365) 221-27-57; e-mail: buxdu_rektor@buxdu.uz).

The dissertation is available at the Information-resource centre of Bukhara State University, (registration number № 125). (Address: 200118, Bukhara, M. Ikbol street, 11. Phone: (0 365221-25-87.)

Dissertation abstract is distributed on "06" April 2022.
(Mailing report № 22-01 on "06" April 2022).



D.S.Uraeva
Chairman of the Scientific Council awarding scientific degrees, Doctor of Science in Philology, Professor

Z.I.Rasulov
Scientific Secretary of the Scientific Council awarding scientific degrees, Doctor of Philosophy in Philology, Associate Professor

M.M.Juraeva
Chairman of the Scientific Seminar under Scientific Council awarding scientific degrees, Doctor of Science in Philology, Associate Professor

INTRODUCTION (abstract of the PhD dissertation)

The aim of the research is to study the theoretical characteristics of the compilation of the Uzbek-English parallel corpus, to develop the principles for the linguistic base of compiling the Uzbek-English parallel corpus.

The object of the research work is Uzbek-English parallel corpora.

Scientific novelty of the research work:

the general features of the parallel corpora have been described, their peculiarities comparing to the other types of corpora have been revealed; the advantages of the parallel corpus in the field of foreign language linguodidactics, linguistics and translation have been proved;

the principles of selection of texts in the Uzbek language have been worked out; the advantages of selecting English texts have been demonstrated; the principles of material selection for the Uzbek-English parallel body have been defined;

problems of linguistic tagging of Uzbek-English parallel corpus texts have been identified and ways of tagging have been proposed;

the principles of extralinguistic tagging of Uzbek-English parallel corpus texts have been distinguished.

Implementation of the research results. On the basis of research results and practical outcomes about the general features of the parallel corpus, its differences from other corpus, the advantages of the parallel corpus in the field of foreign language linguodidactics, linguistics and translation; principles of material selection for the Uzbek-English parallel corpus, based on the scientific results obtained in the process of determining the principles of linguistic and extralinguistic tagging of the Uzbek-English parallel corpus:

the results of the analysis of the principles of selecting texts in the Uzbek language for the Uzbek-English parallel corpus, the advantages of selecting texts in English, the principles of linguistic and extralinguistic tagging of texts in the Uzbek-English parallel corpus were implemented at the innovational grand project № OT-2019-42 concerning the creation of an electronic poetic dictionary the Uzbek and English languages (On human image, character, nature and national symbols) issued by Tashkent State University of the Uzbek Language and Literature named after Alisher Navoi (Reference of the Ministry of Higher and Secondary Specialized Education registered with № 89-03-4852 as of November 20, 2020). As a result, an automatic definition of the contextual meaning of words denoting a person's image, character, nature and national symbols was achieved in the Uzbek and English languages;

the conclusions on the preparation of linguistic software for the formation and programming of the parallel corpus of Uzbek and English were implemented at Istanbul University's educational process, i.e. at the classes on "Linguistics", "Computer Linguistics", "Corpus Linguistics" and "Linguistic Typology" (Reference of the Rector of Istanbul University, Prof. Dr. Mahmoud Ak, as of May 20, 2020). As a result, a general analysis of the features of the Uzbek and English parallel corpus programming programs, outcomes and conclusions of world-

renowned experts in this area served to enrich students' theoretical comprehension of the Uzbek and English parallel corpus.

conclusions about the advantages of the Uzbek and English parallel corpus in the field of linguistics and translation, as well as the principles of selection of material for the Uzbek-English parallel corpus were used in a report made at the Indian Cultural Center named after Lal Bahadur Shastri on the topic "Problems of linguistic support for the formation of a linguistic base and programming the parallel corpus of the Uzbek and English languages" (Reference of the Indian Cultural Center named after Lal Bahadur Shastri as of September 18, 2020). This served as a basis for further expansion of the knowledge base of linguistic software for the formation and programming of the linguistic base of the parallel corpus of Uzbek and English for the language specialists of the center;

scientific views on the general features of the Uzbek-English parallel corpus, the factors that distinguish it from other corpora in a foreign language, were used in the preparation of the script of the Bukhara Regional TV and Radio Company's programs as "Assalom, Bukhara", "Everyday English", "Ilm dunyosi (World of science)" (Reference of Bukhara regional television and radio company registered with № 1/1988 as of August 18, 2020). Consequently, the widespread use of the Uzbek-English parallel corpus in collecting texts in Uzbek, the analysis of the process of integration of the Uzbek-English parallel corpus has led to an increase in respect for the national language and motivation to learn a foreign language.

The outline of the thesis. The thesis consists of an introduction, three chapters, conclusions, a list of used literature and appendix. The total volume is 142 pages of the text.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (I часть; Part I)

1. Karimov R.A. Text Selection Issue for Parallel Corpus // The American Journal of Social Science and Education Innovations. –USA, 2020. – Vol. 2 (09), – P. 311-316. (SJIF: 7.13).

2. Karimov R.A. Problems of Linguistic Tagging of Uzbek-English Parallel Corpus Texts // Theoretical and Applied Science. –USA, 2020. – Vol. 86, Issue 06. – P. 137-140. (SJIF: 7.13).

3. Karimov R.A., Mengliyev B.R. The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of Interpretation // International Journal of Engineering and Advanced Technology (IJEAT). – July, 2020. – Vol 8, Issue S53. – P. 388-391. (Scopus)

4. Karimov R.A., Mengliyev B.R. Theoretical Fundamentals of Uzbek-English Parallel Corpus // Journal of Critical Reviews. – Malaysia, 2020. – Vol 7, Issue 17. – P. 73-76. (Scopus)

5. Karimov R.A. On Extralinguistic Tightening of Uzbek-English Texts // Education and Training. – Urganch, 2020. – № October. – P. 92-96. (10.00.00; №22).

6. Каримов Р.А. Параллел корпуснинг лингвистик, хорижий тил лингводидактикасидаги ахамияти ва таржимонлик соҳасидаги имкониятлари // Илм сарчашмалари. – Урганч, 2021. – № 1. – Б. 150-153. (10.00.00; №3)

7. Каримов Р.А. Ўзбек ва Инглиз параллел корпуси матнларини лингвистик теглаш муаммолари // «Сўз санъати» халқаро журнали. – Тошкент, 2021. – № 1, 4-жилд. – Б. 81-86. (10.00.00; № 31)

8. Karimov R.A. On Extralinguistic Tagging of Uzbek-English Parallel Corpus texts // Ўзбекистон Миллий университети хабарлари, Тошкент, 2020. – № 4. – Б. 150-153. (10.00.00; № 15)

II бўлим (II часть; part II)

9. Karimov R.A. Using Synonyms in Translated Texts // Proceedings of Online International Conference on Technological Developments in Systematic Research. – India. August 14th, 2020. – P. 76-77.

10. Karimov R.A. The Importance of Using a Parallel Corpus in Translation // Science, Research, Development. International Conference Proceedings. – France. September, 2020. – P. 103-105.

11. Каримов Р.А. Реалияларнинг таржимада берилиши ҳамда унинг параллел корпусда ифодаланиши // «Таълим тизимида чет тилларни ўрганишнинг замонавий муаммолари ва истиқболлари» мавзусидаги халқаро илмий-амалий анжуман материаллари тўплами. – Бухоро, 2020. – Б. 253-255.

12. Каримов Р.А. Трансформацияга учраган морфемаларнинг параллел корпусда берилиши // «Таълим тизимида чет тилларни ўрганишнинг

замонавий муаммолари ва истиқболлари» мавзусидаги халқаро илмий-амалий анжуман материаллари тўплами. – Бухоро, 2020. – Б. 251-253.

13. Каримов Р.А. Инглиз ва Ўзбек параллел корпусида мақолларнинг ифодаланиши // «Ўзбек миллий ва таълимий корпусларни яратишнинг назарий ҳамда амалий масалалари» мавзусидаги халқаро илмий-амалий анжуман материаллари тўплами. – Тошкент, 2021. – Б. 220-225.

14. Каримов Р.А. Параллел корпуснинг лингвистик, хорижий тил лингводидактикасидаги аҳамияти ва таржимонлик соҳасидаги имкониятлари // «Ўзбек миллий ва таълимий корпусларни яратишнинг назарий ҳамда амалий масалалари» мавзусидаги халқаро илмий-амалий анжуман материаллари тўплами. – Тошкент, 2021. – Б. 220-225.

15. Каримов Р.А. Дунё корпус тилшунослиги ривожидида параллел корпуслар тавсифи // «Тил, адабиёт, таржима, адабий танқидчилик халқаро илмий форуми: замонавий ёндашувлар ва истиқболлар» мавзусидаги халқаро миқёсидаги илмий-амалий анжуман материаллари. – Бухоро, 2021. – Б. 36-39.

16. Каримов Р.А. Параллел корпусларнинг тузилиши ва имконияти тадқиқи // «Тил, адабиёт, таржима, адабий танқидчилик халқаро илмий форуми: замонавий ёндашувлар ва истиқболлар» мавзусидаги халқаро миқёсидаги илмий-амалий анжуман материаллари. – Бухоро, 2021. – Б. 74-80.

17. Каримов Р.А. Ўзбек ва Инглиз параллел корпусида халқ мақолларининг ифодаланиши // «Тилнинг лексик-семантик тизими, қиёсий типологик изланишлар ва адабиётшунослик муаммолари» мавзусидаги халқаро миқёсидаги илмий-амалий анжуман материаллари. – Бухоро, 12-сон, 2021. – Б. 111-115.

18. Каримов Р.А. Таржимада Трансформацияга учраган морфемаларнинг параллел корпусда берилиши масаласи // «Тилнинг лексик-семантик тизими, қиёсий типологик изланишлар ва адабиётшунослик муаммолари» мавзусидаги халқаро миқёсидаги илмий-амалий анжуман материаллари. – Бухоро, 12-сон, 2021. – Б. 106-108.

Автореферат «Дурдона» нашриётида таҳрирдан ўтказилди ва ўзбек, рус ҳамда инглиз тилларидаги матнларнинг мослиги текширилди.

Босишга рухсат этилди: 18.03.2022. Бичими 60x84 1/16. Рақамли босма усулида босилди. Times New Roman гарнитураси. Шартли босма тобоғи: 3.2. Адади 100 нусха. Буюртма №127.

Гувоҳнома АИ № 178. 08.12.2010.
«Sadriddin Salim Buxoriy» МЧЖ босмахонасида чоп этилди.
Бухоро шаҳри, М.Иқбол кўчаси, 11-уй. Тел.: 0(365) 221-26-45.

