

**АЛИШЕР НАВОИЙ НОМИДАГИ ТОШКЕНТ ДАВЛАТ ЎЗБЕК ТИЛИ
ВА АДАБИЁТИ УНИВЕРСИТЕТИ ҲУЗУРИДАГИ ИЛМИЙ
ДАРАЖАЛАР БЕРУВЧИ DSc.27.06.2017. Fil.19.01
РАҚАМЛИ ИЛМИЙ КЕНГАШ АСОСИДАГИ БИР МАРТАЛИК
ИЛМИЙ КЕНГАШ**

ТОШКЕНТ ДАВЛАТ ЎЗБЕК ТИЛИ ВА АДАБИЁТИ УНИВЕРСИТЕТИ

АБДУРАХМОНОВА НИЛУФАР ЗАЙНОБИДДИН ҚИЗИ

**ИНГЛИЗЧА МАТНЛАРНИ ЎЗБЕК ТИЛИГА ТАРЖИМА
ҚИЛИШ ДАСТУРИНИНГ ЛИНГВИСТИК ТАЪМИНОТИ
(СОДДА ГАПЛАР МИСОЛИДА)**

10.00.11 – Тил назарияси. Амалий ва компьютер лингвистикаси

**ФИЛОЛОГИЯ ФАНЛАРИ БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Тошкент – 2018 йил

Фалсафа доктори (PhD) диссертацияси автореферати мундарижаси

Оглавление автореферата диссертации доктора философии (PhD)

Contents of Dissertation Abstract of the Doctor of Philosophy (PhD)

Абдурахмонова Нилуфар Зайнобиддин қизи

Инглизча матнларни ўзбек тилига таржима қилиш дастурининг
лингвистик таъминоти (сода гаплар мисолида) 3

Абдурахмонова Нилуфар

Лингвистическое обеспечение программы перевода английских
текстов на узбекский язык (на примере простых предложений)..... 21

Abdurakhmonova Nilufar

Linguistic database of translating program from English texts into Uzbek
(as example of simple sentences) 41

Эълон қилинган ишлар рўйхати

Список опубликованных работ
List of published works 44

**АЛИШЕР НАВОИЙ НОМИДАГИ ТОШКЕНТ ДАВЛАТ ЎЗБЕК ТИЛИ
ВА АДАБИЁТИ УНИВЕРСИТЕТИ ҲУЗУРИДАГИ ИЛМИЙ
ДАРАЖАЛАР БЕРУВЧИ DSc.27.06.2017. Fil.19.01
РАҚАМЛИ ИЛМИЙ КЕНГАШ АСОСИДАГИ БИР МАРТАЛИК
ИЛМИЙ КЕНГАШ**

ТОШКЕНТ ДАВЛАТ ЎЗБЕК ТИЛИ ВА АДАБИЁТИ УНИВЕРСИТЕТИ

АБДУРАХМОНОВА НИЛУФАР ЗАЙНОБИДДИН ҚИЗИ

**ИНГЛИЗЧА МАТНЛАРНИ ЎЗБЕК ТИЛИГА ТАРЖИМА
ҚИЛИШ ДАСТУРИНИНГ ЛИНГВИСТИК ТАЪМИНОТИ
(СОДДА ГАПЛАР МИСОЛИДА)**

10.00.11 – Тил назарияси. Амалий ва компьютер лингвистикаси

**ФИЛОЛОГИЯ ФАНЛАРИ БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Тошкент – 2018 йил

Фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида № В2017.2.PhD/Fil220 рақам билан рўйхатга олинган.

Диссертация Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетида бажарилган.

Диссертация автореферати уч тилда (ўзбек, рус, инглиз (резюме)) Илмий кенгаш веб-саҳифасида (www.navoiy-uni.uz) ва «ZiyoNet» ахборот-таълим порталида (www.ziyounet.uz) жойлаштирилган.

Илмий раҳбар:	Дадабоев Ҳамидулла Арипович филология фанлари доктори, профессор
Расмий оппонентлар:	Каримов Суён Амирович филология фанлари доктори, профессор Маматов Абдумурод Эшонқулович филология фанлари доктори, профессор
Етакчи ташкилот:	Андижон давлат университети

Диссертация ҳимояси Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университети ҳузуридаги илмий даражалар берувчи DSc.27.06.2017.Fil.19.01 рақамли илмий кенгаш асосидаги бир марталик илмий кенгашнинг 2018 йил «__» _____ соат ____ даги мажлисида бўлиб ўтади (Манзил: 100174, Тошкент шаҳри, Юсуф Хос Ҳожиб кўчаси, 103-уй. Тел.: (99871) 246-63-58; fax: (99871) 2465321; e-mail: tdutau_nazorat@edu.uz.)

Диссертация билан Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетининг Ахборот-ресурс марказида танишиш мумкин (_____ рақами билан рўйхатга олинган). Манзил: 100174, Тошкент шаҳри, Юсуф Хос Ҳожиб кўчаси, 103-уй. Тел.: (99871) 246-63-58.

Диссертация автореферати 2018 йил «__» _____ куни тарқатилди.
(2018 йил «__» _____ даги _____ рақамли реестр баённомаси).

Ш.С.Сирожиддинов

Илмий даражалар берувчи илмий кенгаш асосидаги бир марталик илмий кенгаш раиси, филол.ф.д., профессор

Ё.Р.Одилов

Илмий даражалар берувчи илмий кенгаш асосидаги бир марталик илмий кенгаш илмий котиби, филол.ф.д.

Н. М.Маҳмудов

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

КИРИШ (фалсафа доктори (PhD) диссертацияси аннотацияси)

Тадқиқот мавзусининг долзарблиги ва зарурати. Бугунги кунда дунё компьютер лингвистикаси табиий тил имкониятларини компьютер тилига ўтказиш ҳамда у билан боғлиқ масалаларни компьютер технологияси ёрдамида ҳал қилиш каби муаммолар тадқиқи билан шуғулланмоқда. Машина таржимасининг иқтисодий тежамкор ва тезкор алоқа воситаси эканлиги бу йўналишда компьютер лингвистикаси фанининг тараққиёт истиқболлини белгилаш билан бирга, янги илмий муаммоларнинг туғилишига ҳам сабаб бўлмоқда.

Жаҳон компьютер лингвистикасида матнларни таҳрирловчи дастур, нутқий синтезатор, матн корпуси, матнларни лингвостатистик анализ қилувчи дастур, матнларни рефератлаш каби тадқиқотларнинг бажарилиши катта ҳажмли лингвистик базаларнинг яратилишига хизмат қилди. Шу маънода Интернет тизимидаги ижтимоий тармоқлар, илмий-техникавий матнлар ва турли конференцияларда ёзма ва оғзаки таржимон дастурларининг ўрни муҳим.

Мамлакатимизда жамиятнинг барча соҳаларида изчил ислохотлар амалга оширилиб, илм-фанни жадаллик билан ривожлантиришга алоҳида эътибор қаратилмоқда. Бошқа соҳаларда бўлгани сингари ўзбек тилшунослигида ҳам бошқа фанлар кесимида янги йўналишлар вужудга келди. Кейинги йилларда компьютер лингвистикасининг алоҳида фан сифатида шаклланиши ҳам компьютер ва тилшуносликнинг кесишувидир. Аммо ҳозиргача ўзбек компьютер лингвистикасида машина таржимасига оид тадқиқотларнинг етарли эмаслиги сабабли ўзбек тилидан хорижий тилга ва аксинча таржима дастури учун ўзбек тилининг лингвистик маълумотлар базасини яратиш зарур. Ўзбекистон Республикасини ривожлантиришнинг бешта устувор йўналиши бўйича 2017-2021 йилларга мўлжалланган Ҳаракатлар стратегиясида «...таълим тизимини такомиллаштириш, сифатли таълим хизматлари имкониятларини ошириш»¹га алоҳида эътибор қаратилган. Бу борада машина таржимасини тадқиқ қилиш, ўзбек ва инглиз тилларининг формал грамматикаси асосида таржимон дастур учун лингвистик таъминот, таржима алгоритми, гап структураларнинг математик моделларини яратиш муҳим аҳамият касб этади.

Ўзбекистон Республикаси Президентининг 2016 йил 13 майдаги ПФ-4797-сон «Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетини ташкил этиш тўғрисида»ги Фармони, 2012 йил 10 декабрдаги ПҚ-1875-сон «Чет тилларни ўрганиш тизимини янада такомиллаштириш чора-тадбирлари тўғрисида», 2013 йил 27 июндаги ПҚ-1989-сон «Ўзбекистон Республикаси Миллий ахборот-коммуникация тизимини янада ривожлантириш чора-тадбирлари тўғрисида», 2017 йил 17 февралдаги ПҚ-2789-сон «Фанлар академияси фаолияти, илмий тадқиқот ишларини ташкил этиш, бошқариш ва молиялаштиришни янада такомиллаш-

¹ Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги Фармони

тириш чора-тадбирлари тўғрисида»ги Қарорлари ҳамда мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу диссертация тадқиқоти 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 ривожлантириш, инновацион иқтисодиётни ривожлантириш» устувор йўналиши доирасида бажарилган.

Муаммонинг ўрганилганлик даражаси. Жаҳонда машина таржимаси бўйича кўпгина олимлар томонидан илмий изланишлар қилинган. Чунончи, Н.Д. Андреев, И.А. Мелчук, И.И. Ревзин, В.Ю. Розенсвейг, Ю.Н. Марчук, Р.Г. Пиотровский, Ю.А. Моторин, К.Б. Бектаев, А.Н. Беляев, И.К. Белский, А.В. Зубов, Г.Е. Мирам, Л.Л. Нелюбин, В.И. Перебийнос, В.А. Чижаковский, Е.А.Шингарев, Г.Г.Белоногов, Р.Г.Котов, Бабушкина Н.В, З.Шалаяпина, О.Ю. Мансурова, А.С. Панина, А.А. Хорошилов (Россия); А.Бут, Р. Ричанс, Ж. Хатчинс, Ж.Аллен, П. Браун (АҚШ); М. Нагао (Япония); А.Вахер (Эстония); Францияда Ж.Астроуни; Р.Синха, А.Жаин (Хиндистон); Б.Блейзер, У.Швол, А.Сторрер (Германия) сингари тадқиқотчиларнинг ишлари шулар жумласидандир. Машина таржимаси назариясига оид қоидага асосланган алгоритм, фразага асосланган таржима, статистикага асосланувчи дастурий таъминот яратиш бўйича етарли манбалар мавжуд². Хусусан, ушбу соҳада бир қатор номзодлик ва докторлик диссертациялари ёқланганини алоҳида таъкидлаш лозим. Бу тадқиқотлар рус тилининг назарий ва амалий жиҳатларини тадқиқ қилишга сезиларли улуш қўшган. Айниқса, ўтган асрнинг 60-80 йилларида машина таржимаси тизимларини ишлаб чиқишга йўналтирилган кўплаб илмий-амалий тадқиқотлар юзага келди³.

Республикада чет тилидаги матнни ўзбек тилига ва ўзбек тилидаги матнни чет тилига таржима қилувчи дастурнинг лингвистик таъминотини яратишга қаратилган жиддий тадқиқотлар амалга оширилган эмас. Бироқ дастурчи ва математиклар томонидан муайян доирада илмий изланишлар

² А.С. Панина проблемы лексические неоднозначности при переводе с японского языка на русский (В связи с задачами японско-русского автоматического перевода): Дисс. ... канд. филол. наук. – М., 2002; Хорошилов А.А. Теоретические основы и методы построения систем фразеологического машинного перевода: Дисс. ... д-ра тех. наук. – М., 2003; Douglas A., Lorna B., Siety M., R. Lee Humphreys, Louisa S. Machine translation: An Introductory Guide. – London, 1994; Марчук Ю. Модели переводы. Учебное пособие. – М.: Академия, 2010. – 174 с.; Christiane Fellbaum Wordnet: An Electronic Lexical Database. – London, 1998. – 447 p.; Мансурова О.Ю. Представление и передача пространственных отношений в многоязычной системе машинного перевода (на материале английского, русского и турецкого языков): Дисс. ... канд. филол. наук. – М., 2006; Кондратов А.М. Машинный перевод. – М., 1963; Ревзии И.И., Розенсвейг В.Ю. Основы общего и машинного перевода. – М.: Высшая школа, 1964; Ингве В. Язык для программирования задач машинного перевода. – М., 1967; Василевский А.Л., Повалочкая С.К. и др. Экспериментальный алгоритм автоматического перевода патентной документации с английского языка на русский // Информация, процессы и системы. – М., 1971. – №1. – С. 29-34.

³ Чижаковский В.А. Фразеология и машинный перевод (опыт составления работы немецко-русского автоматического словаря для публицистических и научных текстов). – Кишинов, 1971; Борисевич А.Д. Англо-русский автоматический словарь оборотов (к проблеме идиоматичности при обращении текста в системе «человек – машина – человек»). – Минск, 1972.

бажарилган⁴. Ўзбекистонда компьютер лингвистикаси шаклланишида ўзбек тили материаллари бўйича лингвостатистик тадқиқотлар олиб борган Н.Ёкутова, М.Айимбетов, С.Ризаев ва С.Мухамедовларнинг илмий изланишлари диққатга сазовор⁵. С.Мухамедов Р.Р.Пиотровский билан ҳаммуаллифликда ёзган «Инженерная лингвистика и опыт системно – статистического исследования узбекских текстов» номли китобида лингвистик моделлар, моделлаштириш ва унинг умумий тамойиллари ҳақида фикр юритилиб, ўзбекча матнларнинг квантатив моделлари келтирилган.⁶ 2001 йили ЎзМУда проф. А.Пулатов бошчилигида Компьютер лингвистикасига асос солинган. Бу давр мобайнида ўзбек тилшунослигидаги масалаларни компьютерда ҳал этишга қаратилган ўнлаб битирув малакавий ишлари ва магистрлик диссертациялари ҳимоя қилинди⁷. Хусусан, кўп тилли машина таржимаси тизимини яратишда Х.Аҳмедованинг⁸ магистрлик иши, компьютер лингвистикасининг назарий методологиясига доир бир неча адабиётлар⁹, ўзбек тилини моделлаштиришга оид тадқиқотлар¹⁰, С.Мухамедова¹¹, А.Раҳимовлар¹²нинг лингвистик жиҳатдан назарий қарашлари акс этган ўқув қўлланмалари эътиборга молик. Диссертацияни тайёрлаш жараёнида номлари саналган жаҳон ва ўзбек тадқиқотчиларининг изланишлари таҳлил этилди.

Тадқиқотнинг диссертация бажарилган олий таълим муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация тадқиқоти Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетининг «Тилнинг ижтимоий, тарихий ва замонавий тараққиётини ўрганиш» режаси доирасида бажарилган.

⁴ Хақимов М.Х. Расширяемый входной язык математического моделирования естественного языка для многоязычной ситуации машинного перевода // ЎзМУ хабарлари, 2009. – № 1. – С.75-80; Хақимов М.Х. Формальные системы машинного перевода в многоязычной ситуации / Материалы республиканской научно-технической конференции «Современные проблемы математики, механики и информационных технологий», НУУЗ, ИМ и ИТ АН РУз, 2008. – С. 297-300.

⁵ Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Филол. фан. д-ри. ... дисс. автореф. –Тошкент, 2008. – 50 б.; Мухамедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс. ...канд. филол. наук. –Ташкент, 1980. – 25 с.; Бектаев К.Б, Пиотровский Р.Г. Математическая лингвистика. – М.: Высшая школа, 1997. – 420 с.; Айимбетов М.К. Проблемы и методы квантитативно-типологического измерения близости тюркских языков (на материалах каракалпакского, казахского и узбекского языков): Автореф.дисс. ... д-ра филол.наук. –Ташкент, 1997. – 47 с.

⁶ Мухамедов С.А., Пиотровский Г.Г. Инженерная лингвистика и опыт системно – статистического исследования узбекских текстов. –Т.: Фан, 1986; Махмудов М.А., Пиотровская А.А., Садыков Т. Система машинного анализа и синтеза тюркской словоформы // Переработка текста методами инженерной лингвистики. –Минск, 1982.

⁷ Abduraxmonova N. Matnlarni kompyuterda ishlash muaommolari: Bitiruv malakaviy ishi. – Toshkent: O‘zMU 2009; Abdurahmanova N. Matlarni ovozlashtirish dasturining lingvistik ta’minoti: Magistrlik dissertatsiyasi. – Toshkent: O‘zMU, 2012. Abduraxmonova N. Inglizcha sodda gaplarni o‘zbek tiliga tarjima qilish dasturining lingvistik ta’minoti: – Toshkent: O‘zMU, 2011.

⁸ Ahmedova X. MVC platformasida ko‘p tillik tarjimon muhitini yaratish: Magistrlik dissertatsiyasi. – Toshkent: O‘zMU, 2015.

⁹ Pulatov A. Компьютер лингвистикаси. – Т.: Akademnashr, 2011. – 500 б.

¹⁰ Пулатов А.К., Джураева Н. Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал. – Т., 2002. – №1. – С.47-54; Пулатов А.К., Алиходжаев Б., Джураева Н. Разработка программы компьютерного анализа и синтеза глаголов узбекского языка // ЎзМУ хабарлари. –Ташкент, 2002. – № 2. – С.17-19.

¹¹ Мухамедова С. Компьютер лингвистикаси (методик қўлланма). – Т., 2007; Пулатов А, Мухамедова С. Компютер лингвистикаси (ўқув қўлланма) – Т., 2014; Мухамедова С. Ўзбек тилидаги ҳаракат феъллари асосида компьютер дастурлари учун лингвистик тахмин яратиш. – Тошкент, 2006. – 143 б.

¹² Раҳимов А. Компьютер лингвистикаси асослари. – Т.: Akademnashr, 2011.

Тадқиқотнинг мақсади инглизча содда гапларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминотини яратишга доир назарияларни таҳлил қилиш ва лингвистик асосларини яратишдан иборат.

Тадқиқот вазифалари:

лингвистик таъминот учун икки тилли луғатлар базасини шакллантириш;

инглизча-ўзбекча феълли фраземалар ҳамда икки тилнинг морфологик базасини яратиш;

машина таржимаси тизими методологиясини ўрганишда илғор методлардан фойдаланиш, таржиманинг лингвистик алгоритминини тузиш;

содда гаплар мисолида икки тилнинг лингвистик моделини шакллантиришда умумий ва фарқли жиҳатлари ҳамда таржимада лингвистик таҳлил босқичларни ўрганиш.

Тадқиқотнинг объекти сифатида инглиз ва ўзбек тилларидаги содда гап, сўз бирикмаси ҳамда унинг грамматик тавсифи танланган.

Тадқиқотнинг предметини инглизча-ўзбекча таржимон дастур учун лингвистик таъминот асосларини ишлаб чиқишда фойдаланиладиган тил ва нутқ бирликлари ташкил этади.

Тадқиқотнинг усуллари. Тадқиқот мавзусини ёритишда тавсифлаш, чоғиштириш, компонент таҳлил, статистик, моделлаштириш, трансформация, дарахтсимон синтактик таҳлил методларидан фойдаланилди.

Тадқиқотнинг илмий янгилиги қуйидагилардан иборат:

машина таржимасининг лингвистик таъминотини яратиш бўйича таржима аниқлигини таъминловчи хулосалар чиқарилган;

икки тилнинг морфологик ва синтактик моделлари, инглиз ва ўзбек тилларига оид аффикслар, морфологик луғат, феълли фраземаларнинг лингвистик таъминоти яратилган;

матнларни автоматик таржима қилишда инглизча содда гап моделларининг ўзаро мувофиқлиги аниқланган;

машина таржимаси учун матнларни морфологик, синтактик-семантик таҳлил қилиш тамойиллари асосланган;

электрон луғат ҳамда дастурий таъминот учун лингвистик таъминот яратиш тамойиллари бўйича парадигматик муносабатларни мувофиқлаштирувчи тавсиялар ишлаб чиқилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

машина таржимаси лингвистик таъминотини яратишга оид ўқув қўлланма¹³ ва таржима луғат¹⁴ чоп этилган;

инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминотига оид инглизча феълли фраземаларнинг ўзбекча таржима базаси ва ўзбекча сўзлар имлосини текширувчи мобил иловаси яратилган;

инглиз ва ўзбек тилларининг дериватив-морфотактик усуллари ҳамда статистикаси аниқланган;

¹³ Абдурахмонова Н. Машина таржимасининг лингвистик асослари. – Тошкент: Akademnashr, 2012. – 160 б.

¹⁴ Abduraxmonova N. English-Uzbek dictionary of phrasal verbs. – Toshkent.: Turonzamin, 2017.– 420 b.

сўз бирикмалари ва гап моделларининг инглиз ҳамда ўзбек тиллари бўйича мувофиқлиги таҳлил қилинган;

диссертацияда тадқиқ қилинган материаллар компьютер лингвистикаси ва машина таржимасида олиб бориладиган изланишларни бойитишда, «Компьютер лингвистикаси», «Машина таржимаси» фанларининг фан дастурларини тузишда манба сифатида фойдаланилган.

Тадқиқот натижаларининг ишончлилиги тилшунослик ва компьютер лингвистикасида назарий асосланган илмий хулосалар, инглиз тилидаги соҳавий нашрлар, қолаверса, диссертацияда қўйилган вазифаларни ҳал қилишнинг апробацияси, тавсияларнинг тажрибада синалганлиги, олинган натижаларнинг ваколатли тузилмалар томонидан тасдиқлангани билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқот натижаларининг илмий аҳамияти машина таржимасида лингвистик масалаларнинг тадқиқ этилиши, компьютер лингвистикасининг муайян йўналишларига доир методологик назарияларни таҳлил қилишга хизмат қилади, шунингдек, таржима дастури учун лингвистик таъминот тузиш, инглиз тилидан ўзбек тилига таржима қилишнинг меъёрий талаблари ҳамда методларини ўзбек тили табиатига мослаштирган ҳолда тадқиқ этилгани билан белгиланади.

Тадқиқот натижаларининг амалий аҳамияти таржима амалиётида муҳим манба бўлиши, тадқиқот материалларидан олий ўқув юртлари учун «Компьютер лингвистикаси», «Таржима назарияси», «Машина таржимаси» фанлари бўйича маърузалар ўқиш, дарслик ва қўлланмалар яратиш, луғатлар тузиш жараёнида фойдаланиш мумкинлиги билан изоҳланади.

Тадқиқот натижаларининг жорий қилиниши. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти бўйича олинган илмий ва амалий натижалар асосида:

тадқиқот натижалари асосида яратилган «Инглизча-ўзбекча феълли фраземалар луғати»да инглиз тилидаги барча услубларда кенг қўлланувчи феълли фраземаларнинг ўнг икки мингдан зиёди ҳамда уларнинг асосий феъллари жамланиб, ўзбекча муқобиллари аниқланган (Тошкент: Турон замин зиё, 2017. – 404 б.). Натижада уларнинг маълумотлар баъзаси яратилиб, машина таржимаси учун лисоний моделлари аниқланган, ўзбекча таржимадаги синоним ва омонимларнинг маънолари лингвистик таъминотда ўз ифодасини топган;

инглиз тилидан ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти учун яратилган инглизча-ўзбекча феълли фраземаларнинг лингвистик таъминоти ва инглизча феълли фраземалардаги илмий терминлардан Ф–4–02 рақамли «Дифференциал операторларнинг спектрал назарияси асосида математик физика ва оптимал бошқарув масалаларини ечишнинг янги усулларини ишлаб чиқиш» фундаментал лойиҳасида фойдаланилган (Фан ва технологияларни ривожлантиришни мувофиқлаштириш қўмитасининг 2017 йил 7 декабрдаги 02-11/1279-сон маълумотномаси). Натижада лойиҳага доир хорижий тадқиқотлар ўрганилган, назарий

адабиётлар билан бойитилган;

диссертацияда эришилган натижалар, жумладан, инглизча-ўзбекча таржима луғатининг ҳамда феълли фраземаларнинг лингвистик таъминотидан Ф-4-64 рақамли «Берилганларни интеллектуал таҳлилида объекларини умумлашган баҳоларини ва индивидуал метрикаларини ҳисоблаш усулларини яратиш ва асослаш» фундаментал лойиҳасида фойдаланилган (Фан ва технологияларни ривожлантиришни мувофиқлаштириш қўмитасининг 2017 йил 7 декабрдаги 02-11/1279-сон маълумотномаси). Илмий хулосаларнинг татбиқи орқали лойиҳа илмий натижавийлиги ошган.

Тадқиқот натижаларининг апробацияси. Мазкур тадқиқот натижалари 16 та халқаро, 25 та республика илмий-амалий анжуманида муҳокамадан ўтказилган.

Тадқиқот натижаларнинг эълон қилинганлиги. Диссертация мавзуси бўйича 62 та илмий иш, жумладан, 1 та таржима луғат, 1 та патент, 1 та илмий гувоҳнома, 1 та ўқув қўлланма, 1 та ўзбек тилига оид қўлланма (ҳаммуаллифликда), Ўзбекистон Республикаси Олий аттестация комиссиясининг докторлик диссертациялари асосий илмий натижаларини чоп этиш тавсия этилган илмий нашрларда 12 та мақола, жумладан, 11 таси республика, 1 таси хорижий журналда нашр қилинган.

Диссертациянинг тузилиши ва ҳажми. Диссертация кириш, тўрт асосий боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловалардан иборат бўлиб, ҳажми 160 саҳифани ташкил қилади.

ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ

Кириш қисмида диссертация мавзусининг долзарблиги ва зарурияти асосланган, тадқиқотнинг мақсад ва вазифалари, объекти ва предмети тавсифланган, республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги кўрсатилган, илмий янглиги ва амалий натижалари баён қилинган, олинган натижаларнинг илмий ва амалий аҳамияти очиб берилган, натижаларни амалиётга жорий қилиш, нашр этилган ишлар ва диссертация тузилиши бўйича маълумотлар келтирилган.

Диссертациянинг «**Машина таржимасининг назарий-методик асослари**» деб номланган биринчи бобида бу йўналишда олиб борилган тадқиқотлар таҳлил қилинган. Унда машина таржимасининг юзага келишига компьютер технологияларининг такомиллашуви омил сифатида кўрсатилган. Таржимани моделлаштириш тамойиллари ҳамда илмий-техникавий маълумотларга доир долзарб муаммоларни ҳал этишда тиллараро таржима мослашувидаги икки компонентли таржима модели, тилнинг статик тавсифи ҳамда динамик таржима жараёни акс этиши ва таҳрирлаш сифатини оширишда инсон ресурсидан фойдаланиш заруриятининг борлиги таъкидланган. Бу борада, рус олими Ю.Н. Марчук машина таржимасини тавсифланган таржима моделларини тушунтирувчи куч, шунингдек, тил моҳиятини айтиш пайтда текшириш имкониятини кўрсатувчи, таклиф этилган ижодий мулоҳазаларнинг самарадорлигини оширувчи восита сифатида

баҳолаган. Машина таржимаси учун лингвистик моделлаштириш назарияси ҳамда табиий тилнинг маълумотлар базасини яратиш масаласини тадқиқ этишда Россияда Г.Г.Белоногов, Р.Г.Котов, Н.В. Бабушкина, З.Шалыпина, О.Ю.Мансурова, А.С.Панина, А.А.Хорошилов, В.А.Чижаковский кабиларнинг машина таржимаси бўйича бажарилган докторлик ва номзодлик ишларида, қолаверса, АҚШда А.Бут, Р.Ричанс, Ж. Хатчинс, Ж.Аллен, П. Браун; Японияда М.Нагао; Эстонияда А.Вахер; Францияда Ж.Астроуни; Хиндистонда Р.Синха, А.Жаин; Германияда Б.Блейзер, У.Швол, А.Сторрер кабиларнинг назарий қарашлари бугунги кунда машина таржимасига оид турли тизимларнинг юзага келишида муҳим омил бўлди. Бундан ташқари, туркология мисолида ҳам машина таржимасига оид тадқиқотлар қилинган¹⁵.

Шунингдек, машина таржимаси назарий жиҳатдан ўзбек тилшунослигида илк бор тадқиқ этилётгани боис ушбу терминга берилган изоҳлар қайд этилган: 1) машина ёрдамидаги инсон таржимаси (machine-assisted human translation – МАНТ) – инсон томонидан матнни таржима қилишда компьютер электрон луғатлар ва унга киритилган кўрсатмаларидан фойдаланиши; 2) компьютер ёрдамидаги таржима (computer assisted-translation – САТ) – матнни компьютер воситасида таржима қилиш ва унда англашилган асосий тушунчаларни ажратишга хизмат қилиши; 3) инсон ёрдамидаги машина таржимаси (human-assisted machine translation – НАМТ) – компьютер дастурий таъминоти орқали амалга ошириладиган таржимага инсон ресурси томонидан таҳрир қилиш вазифасининг юкланиши; 4) тўлиқ автоматик машина таржимаси (fully automatic machine translation – ФАМТ) – таржимон дастури орқали матнни ҳам таржима, ҳам таҳрир қилиш вазифаларини амалга оширилиши назарда тутилади.

Машина таржимасини лингвистик аспектда ўрганиш Жон Хатчинс номи билан боғлиқ. Олим машина таржимаси тизимларида лингвистик таҳлил босқичларини қуйидагича таснифлайди¹⁶: 1. Қоидага асосланган машина таржимаси (MT): трансфер тизим (*transfer-based machine translation*); интерлингвал тизим (*interlingual machine translation*); луғат асосидаги тизим (*dictionary-based machine translation*). 2. Корпус асосидаги машина таржимаси- (*Corpus-based MT*): статистикага асосланган MT (*statistics-based*); намуна асосидаги MT (*example-based*). 3. Гибрид машина таржимаси (*Hybrid machine translation (HMT)*).

¹⁵ Рахимова Д. Исследование моделей и методов семантики машинного перевода с русского языка на казахский язык: Дисс. ...канд. техн. наук, – Казахстан, 2014; Махмудов М. Разработка системы формального морфологического анализа тюркской словоформы (на материале азербайджанского языка): Дисс... канд. филол. наук. – Баку, 1982; Kemal Altıntaş Turkish to Crimean Tatar machine translation system (a thesis submitted to the department of computer engineering and the institute of engineering and science of Bilkent university in partial fulfillment of the requirements for the degree of master of science) 2001. – P. 30; Sundetova A., Forcada M.L., Shormakova A., Aitkulova A. Structural transfer rules for English-to-Kazakh machine translation in the free/open-source platform apertium / Компьютерная обработка тюркских языков. Латинизация письменности (международная конференция.) – Астана, 2013. – P. 320-331; Сулейманов Д.Ш., Гатиатуллин А.Р., Гильмуллин Р.А., Аюпов М.М. К разработке татарско-тюркского машинного переводчика / Компьютерная обработка тюркских языков. Латинизация письменности (международная конференция.) –Астана, 2013. – P. 271-277; Мейрамбек кызы Ж., Хорошилов А. А.Синтактико-семантические представления для развития казахско-русского фразеологического машинного перевода / Компьютерная обработка тюркских языков. – Казань, 2017. – P. 181-195.

¹⁶ <http://www.hutchinsweb.me.uk/Valladolid-2003-1-ppt.pdf>

Диссертациянинг «**Таржима дастуридаги лингвистик таъминотнинг семантик ва мантиқий асослари**» деб номланган иккинчи бобида машина таржимасида тил ва нутқ бирликларининг ўрни, терминографиянинг автоматик таржимадаги аҳамияти, машина таржимаси тизимлари (LOGOS, VISTA, Sistran)да терминларни ўрганиш ва уларнинг маълумотлар базасини яратиш борасидаги таҳлиллар, терминологик луғатини лингвистик таъминотга киритиш, лексемаларнинг умумий ва хусусий маъноларини кодлаш орқали инглиз тилидаги матнларни ўзбек тилига таржима қилиш имкониятини кенгайтириш, терминларни ўзлаштириш ва уларнинг моделларини бериш масалалари таҳлилга тортилган.

Машина таржимасида тил ва нутқ бирликларининг бирдай аҳамият-лилиги таъкидланган. Машинада интуитив хусусиятнинг мавжуд эмаслиги таржимада иштирок этаётган гапдаги барча компонентлар қуршовининг ўрганилиши зарурияти кўрсатилган. Бундан ташқари, сўз шакллари таҳлил қилишда, уларнинг морфем асосларини аниқлашда дериватив имкониятларнинг инобатга олинishi, сўзларнинг структурал модели тавсифланиб, унинг базасини шакллантиришга доир фикрлар берилган. Унга кўра лингвистик таъминотда инглиз ва ўзбек тилларидаги лексемаларнинг ясалиш хусусиятлари ўрганилиб, аффикслар аранжировкаси бўйича моделлар тузилган: *(олд қўшимча+)* асос+сўз ясовчи қўшимча+ *луғавий шакл ясовчи қўшимча+синтактик шакл ясовчи қўшимча* ҳамда инглиз тилида *(Prefix+ root+lexical suffix+grammatical suffix)*.

Композиция ва аффиксация усулида сўз ясалишининг моделлари аниқланганда куйидагилар маълум бўлди: инглиз тилида композиция усулида от сўз туркумининг 13 тури, 60 турдаги ясовчи олд қўшимча, 122 турдаги ясовчи суффикс, 21 хил шакл ясовчи қўшимча аниқланди. Инглиз тилида композиция усулида сифат сўз туркумининг 20 тури, 28 та олд қўшимча, 53 та суффикс, 3 та шакл ясовчи суффикс сифат даражаларини шакллантиришга хизмат қилишини кўрсатди. Семантик таҳлил натижасига кўра, феъл ясовчи 22 та олд қўшимча, 10 та феъл ясовчи суффикс, шакл ҳосил қилувчи 5 та суффикс ҳамда композиция усулида феълнинг 9 турдаги модели аниқланди. Семантик таҳлил равишнинг композиция усулида 2 турдаги модел ҳамда 4 та равиш ясовчи олд қўшимча, 11 та суффикс, шакл ҳосил қилувчи 2 та суффикс – *er* ва *- est* аффикслари равишларга даражаланиш маъноси юклашини кўрсатди.

Грамматик категорияларни семантик таҳлил қилиш натижасида 52 турдаги предлог, 42 турдаги предлог-фраза, 18 турдаги равишдошли предлог, боғловчининг 70 тури, 91 турдаги модал сўзлар, 34 турдаги ундов сўзларнинг мавжудлиги аниқланди. Ўзбек тилининг композиция усулида 7 турдаги, асоси от туркумидан иборат сўзлардан ясовчи аффиксларнинг 38 тури, феъл асосли сўзлардан 52 та аффикс, асоси сифатдан иборат бўлган 9 кўринишдаги от ясовчи аффикс, асоси тақлид сўздан иборат 5 кўринишдаги от ясовчи аффикс турлари аниқланди. Сифат ясовчиларнинг жами 115, олд қўшимчаларнинг 8, озайтирма даража қўшимчаларининг 8, орттирма даража ясовчиларнинг 11, озайтирма даража ҳосил қилувчи қўшимчаларнинг

4 кўриниши ва композиция усулида 5 тури борлиги аниқланди.

Инглиз ва ўзбек тилларидаги дериватив имкониятнинг ўрганилиши морфотактик ҳолатни аниқлашда, яъни ҳар бир морфеманинг семантик ва мантикий чегарасини топишга ёрдам беради. Шунингдек, кўшма феълнинг 5, ясовчи аффиксларнинг 60, модал маъно ифодаловчи 6 тури, феъл нисбат кўшимчаларининг 26, инкор шаклини ифодаловчи 4, майл кўшимчаларининг 30, замон кўшимчаларининг 31, шахс-сон кўшимчаларининг 23 кўриниши мавжудлиги маълум бўлди. Равиш ясовчи кўшимчаларнинг умумий 50, олд кўшимчаларнинг 2, икки асосли равишларнинг 4 тури аниқланди.

Лингвистик таъминотда турғун бирикмаларнинг семантик ва мантикий имкониятларини баҳолашда намуна сифатида инглиз тилида мавжуд феълли фраземалар таҳлилга тортилди. Тадқиқотда феълли фраземаларнинг Access маълумотлар базасида тузилган 12 мингдан зиёди жамланиб, 3 мингдан ортиқ асосий феълларнинг маънолари билан ўзбек тилидаги муқобиллари берилди. Фраземалар таркибий жиҳатдан турли хил бўлгани боис уларнинг семантик майдони алоҳида эътиборга олинган. Ушбу тузилган маълумотлар базаси инглизча-ўзбекча-инглизча йўналишга мўлжалланган. Феълли фраземаларнинг алоҳида маъноси Access майдонларида IDси билан киритилган. Фраземанинг ҳар бир маъноси матн услуги ва контекст ҳолатига биноан уларнинг синонимлари билан алоҳида тарзда махсус белгиларда берилди. Бу ҳолат дастурий таъминот тузиш ва маълумот қидирувида қулай имконият яратади, шу сабабли феълли фраземаларнинг барча маънолари ҳисобга олиниб, 80 мингга яқин семалар базаси жамланди.

Ушбу лингвистик таъминотнинг аҳамиятли жиҳати шундаки, ундан ҳам база сифатида матнни таржима қилувчи дастур учун, ҳам асосий феълдан кейинги майдонда бирликларнинг транскрипцияси киритилганлиги туфайли электрон луғат сифатида фойдаланиш мумкин. Матнларни таржима қилишда фраземаларнинг моделлари алоҳида моделлар базасига киритилди. Зеро, таркиб жиҳатдан ўзаро ажраладиган ва ажралмайдиган феъллар алоҳида тарзда берилди. Ушбу феълли фраземаларнинг базаси қуйидаги моделларга асосланган: (*V* -*феъл*; *P*-*particles* (*предлог ёки равиш*); ↓-*бўлиши мумкин ёки мумкин эмас*; *smth.*-*something*, *smb.*-*somebody*): 1) *V* + *oneself* + *P*+*smth.* => *align oneself with smb. or smth.*; 2) *V* + *oneself* + *P*+*smb.* => *attach oneself to smb*; 3) *V* + *P*+ *smb. or smth.*+*P smth.* => *arrange with smb. about smth.*; 4) *V*+ *smb. or smth.* + *P*+ *smb. or smth.* => *associate smb. or smth. with smb. or smth.*; 5) *V*+ *oneself* +*P* => *arch (oneself) over* каби қирққа яқин моделлар аниқланди. Ушбу моделлардан келиб чиқиб, қуйидаги белгилар базага киритилди: 1) *P*₁, *P*₂, ... , *P*_{*N*} – *N* асосий феъл маънолари; 2) *N*₁, *N*₂, ... , *N*_{*k*} – *k* феълли фразема шакллари; 3) *V*₁, *V*₂, ... , *V*₁ –*l* феълли фраземалар маънолари; 4) *P*₁_, *P*₂_, ... , *P*_{*N*} – асосий феъл синонимлари; 5) *V*₁_, *V*₂_, ... , *V*₁_ – феълли фразема маъноларининг синонимлари; 6) *P* – ҳеч қайси феълга тегишли эмас. Инглиз тилидаги феълли фраземалар ўзбек тилида ибора, кўшма феъл, кўмакчи феълли сўз кўшилмаси, содда феъл каби категорияларга мос келади. Машина таржимаси учун феълли фраземаларнинг таҳлил қилиш алгоритми ишлаб чиқилди:

Шунингдек, ўзбек тили сўз бирикмалари модели орқали инглиз тилидаги моделларининг мослиги тадқиқ этилди. Натижада нафақат инглиз тили, балки ўзбек тилидан инглиз тилига ўгириш имкониятини ҳам яратилади. Тобе боғланишли бирикмада ҳоким сўзнинг қайси сўз туркуми билан ифодаланишига кўра: 1) отли бирикма; 2) феълли бирикма тарзида бўлиниши синтактик муносабат шакллари аниқлашда енгиллик яратади. Битишув муносабатли отли бирикмаларнинг *Noun+Noun=> темир ускуна <=> Noun+Noun=> iron equipment; Adj.+Noun=> қулай имконият <=> Adj.+Noun=> suitable opportunity* каби моделлари тузилиб, бу турга мансуб 20; битишув муносабатли феълли бирикмаларнинг: *Adj.+Verb => яхши ўқимоқ- reading well* каби 3; бошқарув муносабатли отли бирикманинг *Noun + дан+ Noun => Андижондан хат <=> Noun + Prep.+ Noun => the letter from Andijan* каби 20; бошқарув муносабатли феълли бирикмалар: *Noun +za+ Verb => мактабга бормоқ<=> Infinitive+Prep.+Noun=>to go to school* кабиларнинг 10; мослашув муносабатли отли бирикмалар (туркий изофа) *Noun/Pronoun+ning+Noun=> китобнинг муқоваси <=> Noun+of+Noun=> the cover of the book* каби 3 тури аниқланди.

Терминларга хос таржимадаги мураккабликлар: 1) у ёки бу терминнинг иккинчи бир тилда айнан мавжуд бўлмаслиги; 2) терминларнинг сўз бирикмалари тарзида ифодаланиши; 3) матннинг услубини аниқлаш ва илмий терминларнинг маъноларини услубга кўра матн мазмунига боғлаш; 4) муайян соҳавий лексема турли маънода қўлланишининг лингвистик таъминот яратишдаги аҳамияти мисоллар орқали очиб берилган.

Диссертациянинг «**Машина таржимасида матн таҳлили**» номли учинчи бобида таржиманинг уч асосий компоненти мавжудлиги ва булар асос матн, унга қўйилган масалаларнинг нечоғлик тўғрилиги ҳамда терминология масалалари муҳим жиҳатлардан эканлиги таъкидланган. Матнни таржима қилишда қуйидаги масалалар ўрганилди: таҳлил жараёни, сўзларнинг матндаги маъноларини тушуниш ҳамда матнни тиклаш жараёни. Машина таржимасининг таҳлил босқичларига токенизация (сўз шакллари аниқлаш), лемматизация (сўзни аниқлаш) ва стеммизация (етакчи ва кўмакчи морфемаларни аниқлаш)лар киради.

Лексик сатҳнинг ўзига хослиги автоматик таржимада бир қатор мураккабликларни юзага келтиради. Жумладан, инглиз тилидаги аффиксал ва лексик омонимия, тил бирликларининг лингвокултурологик хусусияти, от ва сифат сўз туркумига кирувчи лексемаларда синонимик градуонимиянинг мавжудлиги ва услубий томони, терминларнинг турли соҳада бир неча маънода қўлланиши, уларни таҳлил қилишда юқорида таъкидланган бирликларнинг алоҳида базасини тузиш муҳим деган хулосага келинди.

Инглиз ва ўзбек тилларидаги грамматик категориялар, уларнинг умумий ва фарқли жиҳатлари чоғиштирма аспектда таҳлилга тортилди. Машина таржимасида морфологик ва синтактик таҳлилнинг аҳамияти, морфологик анализда котегориялар таснифлашида нимани эмас, балки қандай таснифлаш турли жиҳатларга кўра амалга оширилиши қайд этилган. Морфологик тавсифнинг компонентлари сифатида луғатнинг махсус

объектли модели, луғат, морфологик қоидалар ҳамда фонологик ва морфофонологик қоидалар тушунилади. Базадаги маълумотларни дастурий таъминотда тўғри топиш учун қуйидаги белгилар киритилди: N_W – муайян сўз фақат бир сўз туркуми доирасида омонимлик ҳосил қилса; N_L – муайян сўз иккидан ортиқ сўз туркуми доирасида омонимликни воқеалантиса; N_Adj . { N_Adv .} – сўз икки сўз туркумига оид бўлганда, иккинчи сўз туркуми аниқ кўрсатилган.

Морфологик анализ фақат у ёки бу категорияларни таҳлил этибгина қолмай, Родолфе Делмонтенинг фикрича, лексикон учун қуйидаги лингвистик категориялар зарур¹⁷: 1) грамматик категориялар; 2) семантик категориялар; 3) дискурс даражасидаги категориялар; 4) синтактик категориялар; 5) тобели категориялар; 6) семантик концептуал категориялар; 7) чегараланган лексик бирликлар; 8) грамматик чекловлар. Исталган тизим (ахборот қидирув, машина таржимаси, автоматик таҳрир)да сўзларнинг грамматик (морфологик ва синтактик) таҳлиliga мурожаат этилади. Шу боис кўп тадқиқотларда морфологик таҳлилни оптималлаштиришга оид қатор фикрлар билдирилган¹⁸. Лингвистик таъминотда инглиз ва ўзбек тилининг морфологик луғати тузилди. Морфологик луғатда мустақил сўз туркумларига қўшилиши мумкин бўлган аффиксларнинг базаси Accessда яратилиб, қўшимча олиши мумкин бўлганларга 1 рақами киритилди.

Ўзбек тилида феълнинг ўзак шакли нутқда ифода нуктаи назаридан омонимлик ҳосил қилиши мумкин: $n_1 \rightarrow$ тур (тип) \leftrightarrow $n_2 \rightarrow$ тур (ўрнidan турмоқ) $n_3 \rightarrow$ тур (уйғонмоқ) \rightarrow $n_4 \rightarrow$ тур (нархда турмоқ). Бу ҳолатда унинг омонимлик ҳосил қилувчи варианты бир сўз туркумига тегишли бўлса, N_W (homonym \rightarrow words), агар икки сўз туркуми доирасида омонимлик ҳосил қилса, иккинчи омонимик туркум кўрсатилади: N_Adj (homonym \rightarrow adjective), агар бир неча сўз туркуми доирасида кузатилса, N_L (homonym \rightarrow lists) каби белгилар қўйилди. Бундай белгилар морфоанализни янада осонлаштиришга хизмат қилади. Матннинг синтактик таҳлилида Н.Чомскийнинг шажарасимон таҳлил методига таянилади. Ҳозир статистик ва нейро машина таржимаси синтаксисга асосланган тизим билан ривожланмоқда. Матндаги семантик алоқани ўрганиш учун йўналтирилган лингвистик таҳлилда грамматик категорияларни таҳлил этиш жараёни давом этмоқда, зеро,

¹⁷ Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, Inc. – New York, 2008. – P. 4-5.

¹⁸ Belonogov G. G. On the use of a similarity method for automatic processing of textual information [Об использовании метода аналогии при автоматической обработке текстовой информации] Problems of cybernetics [Problemy kibernetiki]. Issue 28. – Moscow: Nauka Publ., 1974; Bolshakov I. A. CrossLexica: A large electronic dictionary of collocations and semantic links between Russian words [КроссЛексика – большой электронный словарь сочетаний и смысловых связей русских слов]. Компьютерная лингвистика и Интеллектуальные Технологии: Труды Международной Конференции «Диалог 2009» [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference «Dialogue 2009»]. – Moscow, 2009. – P. 45–50; Gelbukh A. F. Effectively realizable morphologic model of inflective language [Эффективно реализуемая модель морфологии и эктивного языка]. Научно-техническая Информация [Scientific and Technical Information], series 2, #1, 1992. – P. 24–31; Sokirko A. V. Bystroslavar': morphological prediction of new Russian words using very large corpora [Быстрословарь: предсказание морфологии русских слов с использованием больших лингвистических ресурсов]. Компьютерная лингвистика и Интеллектуальные Технологии: Труды Международной Конференции «Dialog 2010» [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference «Dialogue 2010»]. – Moscow, 2010. – P. 450.

морфологик структура барча лингвистик имкониятларни қамраб ололмайди, аммо объектни ўрганишда етарли тасаввур беради: *read somebody like a book* – бирор кишини жуда яхши тушунмоқ; *book club* – китобхонлар тўғараги; *speak by the book* – аниқ маълумотга таяниб гапирмоқ.

Диссертациянинг «**Машина таржимаси тизимида матнларни моделлаштириш**» номли тўртинчи бобида содда гапларнинг ифода мақсадига кўра турлари моделлаштирилган ва қуйидаги турларга кўра синтактик моделлари ўрганилган: инглиз тилидаги от-кесимли ва феъл-кесимли содда гапларнинг ифода мақсадига кўра дарак гапнинг 12 типи: 1. Noun ⊕ Verb ⊕ ↓ Adverb; 2. Article ⊕ ↓ Adjective ⊕ Noun ⊕ ↓ Verb ⊕ ↓ Adjective каби; сўроқ гапнинг 10 тури: 1. Helping verb ⊕ Noun ⊕ ↓ Verb ⊕ ↓ Noun; 2. Helping verb ⊕ Noun ⊕ ↓ Pronoun ⊕ ↓ Verb ⊕ ↓ Noun; ундов гапнинг 12 тури: 1. Noun ⊕ Helping verb ⊕ Verb; 2. Noun ⊕ ↓ Noun ⊕ Helping Verb ⊕ ↓ Verb аниқланди.

Матндаги ҳар бир фрагмент (гап бўлаги сифатида сўз ёки сўзлар йиғиндиси)ни ажратиш учун автоматик буйруқ берилганда ҳамда компьютер хотирасидан ўрин олган лингвистик таъминотга бевосита мурожаат этилганда, трансформациянинг аҳамиятлилиги таъкидланади. Шу жиҳатдан гап моделининг нечоғлик аниқлиги ва киритилган лингвистик базанинг мукамаллиги ишнинг самарадорлигини оширишга хизмат қилади. Бу ўринда инглиз ва ўзбек тиллари мисолида матнни автоматик таржима қилиш учун тилларнинг ўзаро лингвистик моделлардаги вазиятини аниқлаш зарурияти мавжуд. Бунда инглиз ва ўзбек тилларидаги мувофиқлашув муҳим. Жумладан, [*Neither of {of Prn.} Noun*]+*PS+Obj.*] => [{*Ol. | Ot*} +*келишик*{-дан,-нинг}] *Ҳеч қайси* {*бир*{+и |-миз |-имиз|-нгиз |-ингиз|-лари} } + {*Ol. | Ot*} +*Ob.*+{*эмас*} *KB|K*] (*Neither of the students has passed exam.* => *Талабаларнинг ҳеч қайси бири имтиҳон топширмади*) каби гапларнинг хослик нуқтаи назардан алоҳида таҳлилга тортилиши таржиманда зарурийдир.

Диссертацияда Java дастурлаш тилида таржиманинг алгоритмик жиҳатдан қай тарзда амалга оширилиши кўрсатилган. Унга кўра икки тилли таржимон муҳитида алгоритмлар ва дастурий таъминотни яратиш учун маълумотлар базасига қуйидаги белгиларни киритиш ўринли.

Базаларнинг номи	Функцияси
R_i	Предмет соҳалари бўйича терминлар ва иборалар базаси
Q1	Тилдаги барча ўзак сўзлар базаси
K1	Тилдаги барча ясама сўзлар базаси
V2	Гап бўлаклари базаси
V3	Сўз туркумлари базаси

Ҳар икки тил учун юқоридаги жадваллар яратилди. Матнни ташкил қилувчи ҳар бир гапнинг тузилишини, муайян сўз туркуми қандай гап бўлаги вазифасида келишини аниқлашда, албатта, ҳар бир тилнинг грамматикасига мурожаат қилинади. Таржимон муҳитининг ушбу шакли ўзбек-инглиз, инглиз-ўзбек йўналишларидаги таржимани амалга оширади: киритилган матнни (*Z*) гап бўлақларига ажратиб, ҳар бир сўзнинг бошқа тилдаги таржимасини терминлар базасидан олиб, шу тилнинг грамматикаси асосида

гап бўлаклари жойлаштирилади.

Табиий тилларни математик моделлаштиришдаги кенгаювчи кириш тилига¹⁹ қуйидаги белгилашлар берилади:

- T3_{i1}-сўзнинг бошқа тилдаги таржимаси ва унинг гапдаги вазифасидан иборат бўлган тўплам, $1 \leq i1 \leq m$;

- T4_{j1}- сўзнинг бошқа тилдаги таржимаси, $1 \leq j1 \leq m1$;

- T2-таржима қилинган матн;

- Э4-эга; Э2 –кесим; Э5-аниқловчи; Э6-тўлдирувчи; Э7- ҳол.

- Муайян тилдаги гап тузилишининг бошқа тилдаги мос гап тузилишлари моделини яратамиз. Қуйида ўзбек тилидаги гап тузилишларига мос инглиз тилидаги гап моделларини кўриб чиқамиз.

а) ўзбек тилидаги дарак гап тузилишининг математик моделлари қуйидагича бўлсин:

- I. 1. $\langle E4 \rangle \downarrow \oplus \langle E5 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma2 \rangle$.
2. $\downarrow \langle E5 \rangle \oplus \langle E4 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma2 \rangle$.
3. $\downarrow \langle E5 \rangle \downarrow \oplus \langle E5 \rangle \oplus \langle E4 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma2 \rangle$.
4. $\langle E4 \rangle \downarrow \oplus \langle E5 \rangle \downarrow \oplus \langle E6 \rangle \oplus \langle \Gamma2 \rangle$.
5. $\langle E4 \rangle \oplus \langle \Gamma2 \rangle$.
6. $\langle E4 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma2 \rangle$.
7. $\langle E4 \rangle \downarrow \oplus \langle E6 \rangle \oplus \langle \Gamma2 \rangle$.

[1,3,4] дан келтирилган математик моделларда гап таркиби сўз туркумларидан ташкил топган, шу математик моделларга бироз қўшимчалар киритамиз. Яъни моделлардаги сўз туркумларини қайси гап бўлагига оидлиги аниқланиб, кейин унинг бошқа тилдаги таржимаси олинади ва гапнинг тузилиш тартибига кўра жойлаштирилади. Моделлардан энг оптимал ва маъноли таржимани ҳосил қилиш алгоритмининг вазифасига юклатилади. Юқорида келтирилган ўзбек тилидаги дарак гап тузилишларига мос инглиз тилидаги дарак гапларнинг математик моделлари қуйидагича шакллантирилади:

- II. 1. $\langle E4 \rangle \downarrow \oplus \langle E5 \rangle \oplus \langle \Gamma2 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle$.
2. $\downarrow \langle E5 \rangle \oplus \langle E4 \rangle \oplus \langle \Gamma2 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle$.
3. $\downarrow \langle E5 \rangle \downarrow \oplus \langle E5 \rangle \oplus \langle E4 \rangle \oplus \langle \Gamma2 \rangle \downarrow \oplus \langle E6 \rangle$.
4. $\downarrow \langle E5 \rangle \oplus \langle E4 \rangle \oplus \langle \Gamma2 \rangle \downarrow \oplus \langle E6 \rangle$.
5. $\langle E4 \rangle \oplus \langle \Gamma2 \rangle$.
6. $\langle E4 \rangle \oplus \langle \Gamma2 \rangle \downarrow \oplus \langle E7 \rangle$.
7. $\langle E4 \rangle \oplus \langle \Gamma2 \rangle \downarrow \oplus \langle E6 \rangle$.

Юқорида келтирилган гап тузилишлари ва терминлар баъзасидан фойдаланган ҳолда таржима алгоритмини келтирамиз.

Қуйидаги белгилашлардан фойдаланамиз:

Q1_uz=« SELECT * FROM `Q1_uz` »-ўзбек тиладиги барча ўзак сўзлар;

K1_uz=« SELECT * FROM `K1_uz` »-ўзбек тилидаги барча ясама сўзлар;

¹⁹ Хақимов М.Х. Расширяемый входной язык математического моделирования естественного языка для многоязычной ситуации машинного перевода // ЎЗМУ хабарлари. – Тошкент, 2009. – № 1. – С.75-80.

$Q1_eng = \langle \text{SELECT } * \text{ FROM } `Q1_eng` \rangle$ -инглиз тилидаги барча ўзак сўзлар;
 $K1_eng = \langle \text{SELECT } * \text{ FROM } `K1_eng` \rangle$ -инглиз тилидаги барча ясама сўзлар;

E_i - Z матндан ажратиб олинган гаплардан, $1 \leq i \leq n$;

$L1_j$ - E_i -дан ажратиб олинган сўзлар, $1 \leq j \leq n1$;

[2] да келтирилган алгоритм бажарилгандан сўнг қуйидаги «сўз излаш» алгоритми Z гапларга, гаплар эса сўзларга ажратиблиб олингач, ҳар бир сўз дастлаб ўзак сўзлар баъзасидан изланади, агар у базада йўқ бўлса, ясама сўзлар базасидан қидирилади. Сўз топилгач, унинг бошқа тилдаги таржимаси олинади. Масалан, ўзбек-инглиз йўналиши учун 1-таржима алгоритми қуйидагича бўлади.

1. $L1_j$ даги ҳар бир сўз $Q1_uz$ дан излансин. Топилса, 2- қадамга акс ҳолда 4-қадамга ўтилсин;

2. $Q1_uz$ дан шу ўзак сўзнинг инглиз тилидаги тартиби (ID) олинсин;

3. $Q1_eng$ дан ўзак сўзнинг таржимаси олинсин ва 7- қадамга ўтилсин;

4. $L1_j$ даги ҳар бир сўзни $K1_uz$ дан излансин;

5. $K1_uz$ дан шу ясама сўзнинг $K1_eng$ даги тартиби (ID) олинсин;

6. $K1_eng$ дан ясама сўзнинг таржимаси олинсин;

7. Сўзнинг гапдаги вазифаси ҳам аниқлансин ва $T3_{i1}$ тўпламга жойлаштирилсин;

8. $T3_{i1}$ тўлдирилган тўпламни $UzbekIngliz (T3_{i1})$ функциясига узатилсин;

$UzbekIngliz (T3_{i1})$ функциясининг натижалари $T2$ га жойлаштирилсин;

$UzbekIngliz (T3_{i1})$ [2] да келтирилган ўзбек-инглиз йўналишининг таржима алгоритми асосида ёзилган функция қуйидаги кўринишда ёзилади. Функцияни ёзишда қуйидаги белгилашлардан фойдаланамиз:

1. $E_{T3_{k1}}$ - ўзбек тилидаги гап тузилишларига мос бўлган инглиз тилидаги гап тузилишлари $1 \leq k1 \leq m2$;

2. $T3_{i1}$ тўпламга юклатилган сўзларнинг гапдаги вазифалари E_{8_k} га юклансин;

3. E_{8_k} га мос гап тузилиши $E_{T3_{k1}}$ дан излаб топилсин;

4. $E_{T3_{k1}}$ дан излаб топилган гап бўлаклари вазифасида келган сўзлар олинсин ва $T2$ га юклатилсин;

Инглиз-ўзбек йўналиши учун таржима алгоритми қуйидагича бўлади:

1. $L1_j$ даги ҳар бир сўз $Q1_eng$ дан излансин. Топилса, 2- қадамга акс ҳолда 4-қадамга ўтилсин;

2. $Q1_eng$ дан шу ўзак сўзнинг инглиз тилидаги тартиби (ID) олинсин;

3. $Q1_uz$ дан ўзак сўзнинг таржимаси олинсин ва 7-қадамга ўтилсин;

4. $L1_j$ даги ҳар бир сўзни $K1_eng$ дан излансин;

5. $K1_eng$ дан шу ясама сўзнинг $K1_uz$ даги тартиби (ID) олинсин;

6. $K1_uz$ дан ясама сўзнинг таржимаси олинсин;

7. Сўзнинг гапдаги вазифаси аниқлансин ва $T3_{i1}$ тўпламга жойлаштирилсин;

8. $T3_{i1}$ тўлдирилган тўпламни $InglizUzbek (T3_{i1})$ функциясига узатилсин;

$UzbekIngliz (T3_{i1})$ функциясининг натижаларини $T2$ га жойлаштирилсин. Бу ерда $InglizUzbek (T3_{i1})$ [2] да келтирилган инглиз-ўзбек йўнали-

шининг таржима алгоритми асосида ёзилган функция. InglizUzbek (T_{3i1}) функцияси куйидаги кўринишда ёзилади. Функцияни ёзишда куйидаги белгилашлардан фойдаланамиз:

$E_{T_{4k1}}$ -инглиз тилидаги гап тузилишларига мос бўлган ўзбек тилидаги гап тузилишлари $1 \leq k_1 \leq m_2$;

1. T_{3i1} тўпламга юклатилган сўзларнинг гапдаги вазифалари E_{8k} га юклансин;

2. E_{8k} га мос гап тузилишини $E_{T_{4k1}}$ дан излаб топилсин;

3. $E_{T_{4k1}}$ дан излаб топилган гап бўлаклари вазифасида келган сўзлар олинсин ва T_2 га юклатилсин;

Ушбу функциянинг Java дастурлаш тилидаги тегини куйидагича келтирамиз:

```
private String EngUzb(String suz) throws ObjectNotFoundException {
    int uzakId=0;
    String uzbSuz =«»;
    int gapBulagiId=0;
    UzakEnglish ue=uzakSuzEnglishDao.getUzakEnglishByword(suz);
    if(ue.getUzakEnglish().equals(suz)){
        uzakId=ue.getUzakSuzlarId();
        List<UzakSuzlar> usList=uzakSuzUzbekDao.getuzakSuzlarListByRId(uzakId)
        for (UzakSuzlar us : usList) {
            uzbSuz=us.getUzakSuzlar();
        }
    }else{
        YasamaEnglish e=yasamaSuzEnglishDao.getYasamaEnglishByWord(suz)
        if(suz.equals(ye.getYasamaEnglish())){
            uzakId=ye.getYasamaSuzlarId();
            YasamaSuzlar yu=(YasamaSuzlar) yasamaSuzUzbekDao.getYasamaSuzlarListByRId(uzakId);
            uzbSuz=yu.getYasamaSuzlar();
        }else {
            uzbSuz=suz;
        }
    }
    return uzbSuz;
}
```

Хуллас, ўзбек тили билан боғлиқ таржима дастурларининг яратилиши муҳим вазифалардан биридир. Бизнинг тадқиқотимиз, аввало, грамматикага асосланувчи (Rule based translation) икки тилли таржима муҳитида яратилади. Ўрганилган ва эришилган ижобий натижалардан фойдаланиб, кўп тилли таржимон муҳитини яратиш истиқболдаги режалардан саналади.

ХУЛОСА

Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминотини яратиш борасида олиб борилган тадқиқот натижасида куйидаги хулосаларга келинди:

1. Инглиз ва ўзбек тилларидаги ясовчи кўшимчаларни статистик жиҳатдан ҳисоблаш ва уларнинг семантик базасини яратиш дастурий таъминот тузишда ёрдам беради. Компьютер учун кўшимчаларни алоҳида базалар тарзида ажратиш, моделларини аниқлаш таҳлил жараёнида муҳим.

2. Моделлаштириш машина таржимасида энг самарали метод саналади. Бунда лисоний ҳодисаларни муайян даражада таҳлил қилиш, текшириш ва солиштириш имконияти мавжуд. Нутқий ҳодисаларнинг барчасини ўзида тўлиқ акс эттирмаса-да, улар ҳақида муайян хулоса беришда аҳамиятли. Тилда универсал қолипни яратиш мумкин эмас, чунки тил нутқий ҳодиса сифатида ҳар бир индивидда ўз аксини топади. Масаланинг оптимал ечимини топишда ҳар бир тилнинг таҳлили учун ўнлаб алгоритмлар яратилганини ҳисобга олиб, таъкидлаш лозимки, тилни формал даражада моделлаштириш самарали усулдир.

3. Компьютер лингвистикасида моделлаштириш содда гаплар мисолида амалга оширилади. Қўшма ва мураккаб гаплар синтактик тузилма сифатида фраза ҳисобланади. Содда гапларни формаллаштириш мураккаб гапларнинг таркибий қисмларини таҳлил қилишда муайян даражада самаралидир.

4. Агглютинатив ва флектив тиллар ҳар жиҳатдан фарқлангани сабабли уларнинг ягона таҳлил қилиш методини яратиш ўта муҳим. Шу боис тилларнинг ўзига хос жиҳатларига таянган ҳолда гап қурилмаларини моделлаштиришда нутқий ҳодисаларнинг эҳтимолга тўғри келувчи моделлари ҳам инобатга олиниши зарур.

5. Машина таржимасида синтаксисга асосланган таҳлил қилишнинг самарали усул эканлиги тасдиқланди. Чунки синтаксис орқали гапда англашилаётган мазмунни ҳар жиҳатдан текшириш имкони туғилади. Қолаверса, от ва феъл-кесим муносабатида гапларни ажратиб олиш грамматик категорияларнинг иштирокини яхшироқ кузатиш ва таҳлил қилишда қулай усул ҳисобланади.

6. Турғун бирикмалар ва сўз бирикмаларини бир-биридан ажратиш, уларнинг моделларини аниқ кўрсатиш ҳамда семантик хусусиятларини белгилаш компьютер лингвистикасининг муҳим вазифаларидан биридир. Шу боис таржима дастурининг лингвистик базасида феълли фраземаларнинг электрон луғати тузилди.

7. Жами 12 мингдан зиёд инглиз тилидаги феълли фраземалар ва уларнинг 80 мингга яқин маънолари базага киритилди. Улардаги кўп маънолик ҳамда ифодалайдиган маъноларининг синонимлари учун махсус белгилар қўйилди. Бу кейинги тадқиқотлар учун лингвистик таъминот вазифасини ўтайди.

8. Машина таржимасида лингвистик таъминот мукамал даражада бўлсагина унинг дастурий таъминоти яратилади. Кейинги тадқиқотларда лингвистик таъминотнинг семантик қатлами, тил ва нутқ бирликларининг алоҳида яратилган базалари машина таржимаси олдида турган омонимия ва синонимия масаласини ҳал қилишда ёрдам беради.

9. Икки тилли машина таржимаси фақат инглизча матнларни ўзбек тилига таржима қилмай, ундан аксинча ҳолатдаги таржимада ҳам фойдаланиш мумкин. Шу ўринда ўзбек тили формал грамматикасининг қатъий мезонларини белгилаш келажакда кўп тилли таржима технологиясини ишлаб чиқишга хизмат қилади.

**РАЗОВЫЙ НАУЧНЫЙ СОВЕТ НА ОСНОВЕ НАУЧНОГО СОВЕТА
DSc.27.06.2017. Fil.19.01 ПО ПРИСУЖДЕНИЮ УЧЕНОЙ СТЕПЕНИ
ПРИ ТАШКЕНТСКОМ ГОСУДАРСТВЕННОМ УНИВЕРСИТЕТЕ
УЗБЕКСКОГО ЯЗЫКА И ЛИТЕРАТУРЫ ИМЕНИ АЛИШЕРА НАВОИ**

**ТАШКЕНТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
УЗБЕКСКОГО ЯЗЫКА И ЛИТЕРАТУРЫ**

АБДУРАХМОНОВА НИЛУФАР

**ЛИНГВИСТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММЫ ПЕРЕВОДА
АНГЛИЙСКИХ ТЕКСТОВ НА УЗБЕКСКИЙ ЯЗЫК
(НА ПРИМЕРЕ ПРОСТЫХ ПРЕДЛОЖЕНИЙ)**

**10.00.11 – Теория языка.
Прикладная и компьютерная лингвистика**

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ ДОКТОРА ФИЛОСОФИИ (PhD)
ПО ФИЛОЛОГИЧЕСКИМ НАУКАМ**

Тема диссертации доктора философии (PhD) по филологическим наукам зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан под номером № B2017.2.PhD/Fil220.

Диссертация выполнена в Ташкентском государственном университете узбекского языка и литературы имени Алишера Навои.

Автореферат диссертации на трех языках (узбекский, русский, английский (резюме)) размещен на веб-странице Научного совета по адресу: www.navoiy-uni.uz и Информационно-образовательном портале «ZiyoNet» (www.ziyo.net).

Научный консультант: **Дадабоев Хамидулла Арипович**
доктор филологических наук, профессор

Официальные оппоненты: **Суюн Каримов Амирович**
доктор филологических наук, профессор

Абдумурод Маматов Эшанкулович
доктор филологических наук, профессор

Ведущая организация: **Андижанский государственный университет**

Защита диссертации состоится «_____» _____ 2018 года в _____ часов на заседании разовой научный совет на основе научного совета DSc.27.06.2017. Fil.19.01 по присуждению ученой степени при Ташкентском государственном университете узбекского языка и литературы имени Алишера Навои. Адрес: г.Ташкент, Яккасарайский район, улица Юсуф Хос Хожиб, д. 103. Тел.: (+99871) 246-63-58; факс: (+99871) 246-53-21; e-mail: tdutau_nazorat@edu.uz.

С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои (зарегистрирован за №_____). Адрес: 100174, г.Ташкент, Яккасарайский район, улица Юсуф Хос Хожиб, д. 103. Тел: (+99871) 246-63-58.

Автореферат диссертации разослан «_____» _____ 2018 г.
(реестр протокола рассылки № _____ «_____» _____ 2018 г.).

Ш.С. Сиражиддинов
Председатель разового научного совета на основе научного совета по присуждению ученых степеней, доктор филол. наук, профессор

Ё.Р.Одилов
Учёный секретарь разового научного совета на основе научного совета по присуждению ученых степеней, доктор филол. наук

Н.М.Махмудов
Председатель научного семинара при научном совете по присуждению ученых степеней, доктор филол. наук, профессор

ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))

Актуальность и востребованность темы диссертации. На сегодняшний день компьютерная лингвистика занимается исследованием возможностей перевода естественного языка на компьютерный язык с помощью компьютерных технологий, а также решение вопросов, связанных с этим. Машинный перевод является экономичным и быстрым средством связи и это не только определяет путь развития науки компьютерная лингвистика, но и способствует порождению новых научных задач.

Исследования в мировой компьютерной лингвистики – программа рецензирования текстов, речевой синтезатор, корпус текста, программа лингвостатистического анализа текстов, реферат текстов – послужили фактором создания большого объёма лингвистических баз в языкознании. В связи с этим важна роль письменных и устных переводческих программ в научно-технических текстах и разных конференциях социальной сети Интернет.

В нашей стране во всех сферах общества проводятся последовательные реформы, и большое внимание уделяется развитию наук. Как и в других сферах в узбекском языкознании на пересечении с другими предметами и науками возникли новые направления. В связи с этим нужно подчеркнуть, что в последние годы формирование компьютерной лингвистики – пересечение компьютера и языкознания, преподаётся в высших учебных заведениях как отдельный предмет. Однако исследований относительно машинного перевода в узбекской компьютерной лингвистике недостаточно, необходимо создать базу лингвистических сведений узбекского языка для программы перевода с узбекского языка на иностранный язык и наоборот.

В Стратегии действий в 2017–2021 годы по пяти приоритетным направлениям развития Республики Узбекистан уделено особенное внимание «...совершенствованию обучения, повышению возможностей качественной службы обучения»¹. В связи с этим, является важным исследование машинного перевода, создание алгоритма перевода, математических моделей структур предложения, лингвистического снабжения, базы для переводческой программы на основе формальной грамматики узбекского и английского языков.

Данное диссертационное исследование в определённой мере служит выполнению задач, предусмотренных в указах Президента Республики Узбекистан № УП-4797 «Об организации Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои» от 13 мая 2016 года, № УП-4947 «О Стратегии действий по дальнейшему развитию Республики Узбекистан» от 7 февраля 2017 года, в Постановлениях Президента Республики Узбекистан № ПП-1875 «О мерах дальнейшего совершенствования изучения иностранных языков» от 10 декабря 2012 года,

¹ Указ Президента Республики Узбекистан № УП-4947 «О Стратегии действий по дальнейшему развитию Республики Узбекистан на 2017-2021 гг.» от 7 февраля 2017 года

№ ПП-1989 «О мерах дальнейшего развития Национальной информационно-коммуникационной сферы Республики Узбекистан» от 27 июня 2013 года, № ПП-2789 «О мерах дальнейшего совершенствования деятельности Академии наук, организации, руководства и финансирования научных исследовательских работ» от 17 февраля 2017 года, а также в других нормативно-правовых актах, имеющих отношение к данной области деятельности.

Соответствие исследования приоритетным направлениям развития науки и технологий республики. Диссертационная работа выполнена согласно приоритетным направлениям развития науки и технологий Республики Узбекистан: I. «Социальное, правовое, экономическое, культурное и морально-духовное развитие демократического государства и информационного общества, развитие инновационной экономики».

Степень изученности проблемы. В мире проведены исследования по машинному переводу многими учёными, такими как: Н.Д.Андреев, И.А.Мелчук, И.И.Ревзин, В.Ю.Розенсвейг, Ю.Н.Марчук, Р.Г.Пиотровский, Ю.А.Моторин, К.Б.Бектаев, А.Н.Беляев, И.К.Белский, А.В.Зубов, Г.Е.Мирам, Л.Л.Нелюбин, В.И.Перебийнос, В.А.Чижаковский, Е.А.Шингарев, Г.Г.Белоногов, Р.Г.Котов, Н.В.Бабушкина, З.Шалыпина, О.Ю.Мансурова, А.С.Панина, А.А.Хорошилов (Россия); А.Бут, Р.Ричанс, Ж.Хатчинс, Ж.Аллен, П.Браун (США); М.Нагао (Япония); А.Вахер (Эстония); Ж.Астроуни (Франция); Р.Синха, А.Жаин (Индия); Б.Блейзер, У.Швол, А.Сторрер (Германия). Существуют достаточно источников по созданию программного обеспечения основанных на статистике; перевод, основанный на фразе; алгоритм, основанный на правилах теории машинного перевода². В частности, нужно подчеркнуть, что в этой сфере защищены ряд кандидатских и докторских диссертаций. Данные исследования значительно послужили изучению теоретических и прикладных свойств русского языка. А именно, в 60-80 годы прошлого века возникли множество научно-прикладных исследований, направленные на произведение систем машинного перевода³.

В нашей республике, к сожалению, не проведены серьёзные исследования, направленные на создание лингвистического обеспечения

² Панина А.С. Проблемы лексические неоднозначности при переводе с японского языка на русский (В связи с задачами японско-русского автоматического перевода): Дисс. ... канд. филол. наук. – М., 2002; Хорошилов А.А. Теоретические основы и методы построения систем фразеологического машинного перевода: Дисс. ... д-ра тех. наук. – М., 2003; Douglas A., Lorna B., Siety M., Lee R. Humphreys, Louisa S. Machine translation: An Introductory Guide. – London, 1994; Марчук Ю. Модели переводы. Учебное пособие. – М.: Академия, 2010. – 174 с.; Christiane Fellbaum Wordnet: An Electronic Lexical Database. – London, 1998. – 447 p.; Мансурова О.Ю. Представление и передача пространственных отношений в многоязычной системе машинного перевода (на материале английского, русского и турецкого языков): Дисс. ... канд. филол. наук. – М., 2006; Кондратов А.М. Машинный перевод. – М., 1963; Ревзии И.И., Розенсвейг В.Ю. Основы общего и машинного перевода. – М.: Вышс. шк, 1964; Ингве В. Язык для программирования задач машинного перевода. – М., 1967; Василевский А.Л., Повалоцкая С.К. и др. Экспериментальный алгоритм автоматического перевода патентной документации с английского языка на русский // Информация, процессы и системы. – М., 1971. – №1. – С. 29-34.

³ Чижаковский В.А. Фразеология и машинный перевод (опыт составления работы немецко-русского автоматического словаря для публицистических и научных текстов). – Кишинёв, 1971; Борисевич А.Д. Англо-русский автоматический словарь оборотов (к проблеме идиоматичности при обращении текста в системе «человек – машина – человек»). – Минск, 1972.

переводческой программы текста с иностранного языка на узбекский язык и, наоборот: с узбекского языка на иностранный язык. Однако со стороны программистов и математиков в некоторой степени проведены научные исследования в этом направлении⁴. В Узбекистане в формировании компьютерной лингвистики достойны вниманию научные работы Н.Ёкутовой, М.Айимбетова, С.Ризаева и С.Мухамедова, которые провели лингвостатистические исследования на материалах узбекского языка⁵. В книге С.Мухамедова и Р.Р.Пиотровского «Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов» речь идёт о лингвистических моделях, о моделировании и его общих тенденциях, даются количественные модели узбекских текстов⁶. В 2001 году под руководством профессора А.Пулатова в Национальном университете Узбекистана создана Компьютерная лингвистика. За этот период защищены десятки дипломных работ и магистерских диссертаций, направленные на решение вопросов узбекского языкознания с помощью компьютера⁷. В частности, имеют важное значение в создании многоязычного машинного перевода магистерская диссертация Х.Ахмедов⁸, несколько литератур по теоретической методологии компьютерной лингвистики⁹, исследования по моделированию узбекского языка¹⁰, учебные пособия, в которых приводятся теоретические взгляды С.Мухамедовой¹¹, А.Рахимова¹² по данной проблеме. В течение написания диссертации проанализированы перечисленные выше исследования учёных мира и Узбекистана.

⁴ Хакимов М.Х. Расширяемый входной язык математического моделирования естественного языка для многоязычной ситуации машинного перевода // ЎзМУ хабарлари, 2009. – № 1. – С.75-80; Хакимов М.Х. Формальные системы машинного перевода в многоязычной ситуации / Материалы республиканской научно-технической конференции «Современные проблемы математики, механики и информационных технологий», НУУз, ИМ и ИТ АН РУз. 2008. – С. 297-300.

⁵ Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: Филол. фан. д-ри. ... дисс. автореф. –Тошкент, 2008. – 50 б.; Мухамедов С.А. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс. ... канд. филол. наук. – Ташкент, 1980. – 25 с.; Бектаев К.Б., Пиотровский Р.Г. Математическая лингвистика. – М.: Высшая школа, 1997. – 420 с.; Айимбетов М.К. Проблемы и методы количественно-типологического измерения близости тюркских языков (на материалах каракалпакского, казахского и узбекского языков): Автореф.дисс. ... д-ра филол. наук. – Ташкент, 1997. – 47 с.

⁶ Мухамедов С.А., Пиотровский Г.Г. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. – Т.: Фан, 1986; Махмудов М.А., Пиотровская А.А., Садыков Т. Система машинного анализа и синтеза тюркской словоформы // Переработка текста методами инженерной лингвистики. – Минск, 1982.

⁷ Abduraxmonova N. Matnlarni kompyuterda ishlash muaommolari: Bitiruv malakaviy ishi. – Toshkent: O‘zMU 2009; Abdurahmanova N. Matlarni ovozlashtirish dasturining lingvistik ta’minoti: Magistrlik dissertatsiyasi. – Toshkent: O‘zMU, 2012. Abduraxmonova N. Inglizcha sodda gaplarni o‘zbek tiliga tarjima qilish dasturining lingvistik ta’minoti: – Toshkent: O‘zMU, 2011.

⁸ Axmedova X. MVC platformasida ko‘p tillik tarjimon muhitini yaratish: Magistrlik dissertatsiyasi. – Toshkent: O‘zMU, 2015.

⁹ Pulatov A. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011. – 500 б.

¹⁰ Пулатов А.К., Джураева Н. Разработка формальной модели грамматики узбекского языка // Узбекский математический журнал. – Ташкент, 2002. – № 1. – С.47-54; Пулатов А.К., Алиходжаев Б., Джураева Н. Разработка программы компьютерного анализа и синтеза глаголов узбекского языка // ЎзМУ хабарлари. – Ташкент, 2002. – № 2. – В.17-19.

¹¹ Мухамедова С. Компьютер лингвистикаси (методик қўлланма). – Тошкент, 2007; Пулатов А., Мухамедова С. Компьютер лингвистикаси (ўқув қўлланма) – Тошкент, 2014; Мухамедова С. Ўзбек тилидаги ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таҳмин яратиш. – Тошкент, 2006. – 143 Б.

¹² Раҳимов А. Компьютер лингвистикаси асослари. – Тошкент: Akademnashr, 2011.

Связь диссертационного исследования с планами научно-исследовательских работ высшего образовательного или научно-исследовательского учреждения, где выполнено диссертационное исследование. Диссертационная работа выполнена в соответствии с планами научно-исследовательских работ Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои по теме «Изучение социальное, историческое и современное развитие языка».

Целью исследования является анализ теорий по созданию лингвистического обеспечения программы перевода простых предложений с английского языка на узбекский язык и создание лингвистической основы данной программы.

Задачи исследования:

формирование базы двух язычных словарей для лингвистического обеспечения;

формирование базы англо-узбекских глагольных фразем и морфологической базы двух языков;

использовать ведущие методы при изучении методологии системы машинного перевода и составить лингвистический алгоритм перевода;

на примере простых предложений выявить общие и различные свойства формирования лингвистических моделей двух языков, а также изучение этапов лингвистических анализов в переводе.

Объект исследования простые предложения, а также словосочетания и их грамматические характеристики в английском и узбекском языках.

Предмет исследования составляют языковые и речевые единицы при создании основ лингвистического обеспечения для программы англо-узбекского перевода.

Методы исследования. При освещении темы диссертации использованы методы классификации, сопоставление, компонентный анализ, статистический метод, а также методы моделирования, трансформации, деривационный синтаксический анализ.

Научная новизна исследования заключается в следующем:

сделаны выводы, обеспечивающие точность перевода, по созданию лингвистического обеспечения машинного перевода;

созданы морфологические и синтаксические модели двух языков, также база глагольных фразем, аффиксов, морфологический словарь английского и узбекского языков;

определено соответствие моделей простых предложений английского языка в процессе автоматического перевода текстов;

обоснованы принципы морфологического, синтаксического и семантического анализа текстов для машинного перевода;

разработаны рекомендации, приспособляющие парадигматические соотношения, по принципам создания лингвистического обеспечения для программного обеспечения и электронного словаря.

Практический результат исследования заключается в следующем:

напечатаны учебные пособия¹³ по созданию лингвистического обеспечения для машинного перевода и словарь-переводчик¹⁴;

созданы узбекская база перевода английских глагольных фразем относящихся лингвистическому обеспечению программы перевода английских текстов на узбекский язык, а также мобильное приложение проверки правописания узбекских слов;

выявлены синтаксис и деривационные морфотаксические способы английского и узбекского языков;

проанализировано соответствие моделей словосочетаний и предложений английского и узбекского языков;

материалы, исследованные в диссертации, использованы как источник обогащения исследований по машинному переводу и компьютерной лингвистике, в частности при создании предметных программ «Компьютерная лингвистика», «Машинный перевод».

Достоверность результатов исследования объясняется теоретически обоснованностью научных выводов языкознания и компьютерной лингвистики, изданиями на английском языке относительно данной области, апробацией решения задач поставленных в диссертации, проверка рекомендаций на практике, утверждением полученных результатов уполномоченными структурами.

Научная и практическая значимость результатов исследования.

Научная значимость результатов исследования заключается в том, что исследование лингвистических проблем машинного перевода служат анализу методологических теорий относительно определённых направлений компьютерной лингвистики. Научная ценность работы определяется созданием лингвистического обеспечения для программы перевода, исследованием нормативных востребований и методов перевода с английского языка на узбекский язык и их приспособлением к природе узбекского языка.

Практическая значимость результатов исследования объясняется возможностью использования материалов исследования в высших учебных заведениях в процессе чтения лекций по предметам «Компьютерная лингвистика», «Теория перевода», «Машинный перевод», а также создания учебников, пособий и словарей.

Внедрение результатов исследования. На основе научных и практических результатов, полученных по лингвистическому обеспечению программы перевода с английского языка на узбекский язык:

в «Словаре англо-узбекских глагольных фразем», созданный на основе результатов исследования, собрано более 12 тысяч глагольных фразем, а также основных глаголов широко использующихся во всех стилях

¹³ Абдурахмонова Н. Машина таржимасининг лингвистик асослари. – Тошкент: Akademnashr, 2012. – 160 б.

¹⁴ Abduraxmonova N. English-Uzbek dictionary of phrasal verbs. – Toshkent.: Turonzamin, 2017. – 420 b.

английского языка и определены узбекские эквиваленты. В результате чего создана их база сведений, определены их лингвистические модели для машинного перевода, отражено в лингвистическом обеспечении значение синонимов и омонимов в узбекском переводе;

в фундаментальном проекте по теме «Разработка новых методов решения задач математической физики и оптимального руководства на основе спектральной теории дифференциальных операторов» (справка № 02-11/1279 Комитета по координации развития науки и технологий от 7 декабря 2017 года) использовано лингвистическое обеспечение глагольных фразем, созданная для лингвистического обеспечения программы перевода с английского языка на узбекский язык, научные термины английских глагольных фразем. В результате изучены зарубежные исследования относительно проекта, пополнено теоретическими литературами, исследования по результатам проекта изданы как научные статьи;

в фундаментальном проекте под номером Ф-4-64 по теме «Разработка и обоснование методов вычисления обобщённых оценок и индивидуальных метрик объектов и интеллектуальном анализе данных» (справка № 02-11/1279 Комитета по координации развития науки и технологий от 7 декабря 2017 года) использованы результаты исследования, лингвистическое обеспечение англо-узбекского переводческого словаря и глагольных фразем. В результате научных исследований в зарубежных странах изданы статьи на английском языке, изучены иностранные литературы по проекту, в результате использования научных терминов повышена научная результативность проекта.

Апробация результатов исследования. Результаты данного исследования объявлены на 16 международных, 25 республиканских научно-практических конференциях.

Публикации результатов исследования. По теме диссертации всего опубликовано 62 работы, из них 1 переводческий словарь, 1 патент, 1 научное свидетельство, 1 учебное пособие, 1 пособие по узбекскому языку (в соавторстве), 12 статей в научных журналах, рекомендованных Высшей аттестационной комиссии Республики Узбекистан, для публикации основных результатов докторских диссертаций (одна статья в зарубежном журнале).

Структура и объем диссертации. Диссертация состоит из введения, четырёх основных глав, заключений, списка использованной литературы и приложений. Общий объем работы составляет 160 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обосновывается актуальность и востребованность темы диссертации, излагаются цель и задачи, объект и предмет исследования, показано соответствие приоритетным направлениям развития науки и технологий, интерпретируется научная новизна и практическое значение исследования, раскрыто значение научных и практических результатов, предоставляется информация о внедрении результатов на практике, опубликованных работ и структуре диссертации.

В первой главе **«Теоретические-методические основы машинного перевода»** анализируются исследования проведенные в этом направлении, усовершенствование компьютерных технологий указывается как фактор осуществления машинного перевода. Утверждается необходимость использования двухкомпонентной модели перевода в согласовании межъязыкового перевода, статической характеристики языка в процессе решения актуальных проблем по научно-техническим сведениям и принципам моделирования перевода, а также человеческих ресурсов в отображении динамического перевода и в повышении рецензирования. По этому поводу русский ученый Ю.Н.Марчук машинный перевод оценивает, как силу объясняющую классифицированные модели перевода, а также как средство показа возможности в данное время проверки смысла языка, повышения результативности предложенных творческих размышлений. При исследовании создания базы сведений естественного языка и теории лингвистического моделирования для машинного перевода послужили важным фактором докторские и кандидатские работы, выполненные в России следующими учеными Г.Г.Белоногов, Р.Г.Котов, Н.В.Бабушкина, З.Шалыпина, О.Ю.Мансурова, А.С.Панина, А.А.Хорошилов, В.А.Чижаковский; а также научные взгляды ученых, как А.Бут, Р.Ричанс, Ж.Хатчинс, Ж.Аллен, П.Браун (США); М.Нагао (Япония); А.Вахер (Эстония); Ж.Астроуни (Франция); Р.Синха, А.Жаин (Индия); Б.Блейзер, У.Швол, А.Сторрер (Германия) на сегодняшний день послужили появлению разных систем по машинному переводу. Кроме этого проведены исследования по машинному переводу и в тюркологии¹⁵.

¹⁵ Рахимова Д. Исследование моделей и методов семантики машинного перевода с русского языка на казахский язык: Дисс. ...канд. техн. наук, – Казахстан, 2014; Махмудов М. Разработка системы формального морфологического анализа тюркской словоформы (на материале азербайджанского языка): Дисс. ... канд. филол. наук. – Баку, 1982; Kemal Altıntaş Turkish to Crimean Tatar machine translation system (a thesis submitted to the department of computer engineering and the institute of engineering and science of Bilkent university in partial fulfillment of the requirements for the degree of master of science) 2001. – P. 30; Sundetova A., Forcada M.L., Shormakova A., Aitkulova A. Structural transfer rules for English-to-Kazakh machine translation in the free/open-source platform apertium / Компьютерная обработка тюркских языков. Латинизация письменности (международная конференция) – Астана, 2013. – P. 320-331; Сулейманов Д.Ш., Гатиатуллин А.Р., Гильмуллин Р.А., Аюпов М.М. К разработке татарско-турецкого машинного переводчика / Компьютерная обработка тюркских языков. Латинизация письменности (международная конференция) – Астана, 2013. – P. 271-277; Мейрамбек кызы Ж., Хорошилов А.А. Синтактико-семантические представления для развития казахско-русского фразеологического машинного перевода. / Компьютерная обработка тюркских языков. – Казань, 2017. – P. 181-195.

А также, предоставляются определения данного термина, так как машинный перевод в узбекском языкознании с теоретической точки зрения исследуется впервые: 1) перевод человека с помощью компьютера (*machine-assisted human translation – МАНТ*) – в процессе перевода текстов использование человеком компьютерных электронных словарей и показаний приведенных в нем; 2) перевод с помощью компьютера (*computer assisted-translation – САТ*) – перевод текста посредством компьютера и его служба при выборе основных понятий осознанных в нем; 3) машинный перевод с помощью человека (*human-assisted machine translation – НАМТ*) – человек участвует в качестве редактора перевода, осуществленного средством компьютерного программного обеспечения; 4) полный автоматический машинный перевод (*fully automatic machine translation – FAMT*) – имеется ввиду полный перевод и редактирование посредством программы перевода. Изучение машинного перевода в лингвистическом аспекте связано с именем Жона Хатчинса. Ученый классифицирует лингвистические этапы анализа систем машинного перевода¹⁶: 1. Машинный перевод основанный на правиле (МП): трансфертная система (*transfer-based machine translation*); интерлингвальная система (*interlingual machine translation*); система, основанная на словаре (*dictionary-based machine translation*); 2. Машинный перевод, основанный на корпусе (*Corpus-based MT*): машинный перевод, основанный на статистике (*statistics-based*); машинный перевод, основанный на примере, образце (*example-based*); 3. Гибридный машинный перевод (*Hybrid machine translation (HMT)*).

Во второй главе диссертации «**Семантические и логические основы лингвистического обеспечения программы перевода**» анализируются следующие вопросы: место языковых и речевых единиц для машинного перевода; важность терминографии в автоматическом переводе; анализ по изучению терминов в системах машинного перевода (*LOGOS, VISTA, Sistran*) и созданию баз сведений для них; введение в лингвистическое обеспечение терминологический словарь; расширение возможностей перевода текстов с английского языка на узбекский язык путем кодирования общих и частных значений лексем; освоение терминов и предоставление их моделей и т.д.

Утверждается роль не только языковых единиц, но и важность речевых единиц для машинного перевода. Указана необходимость изучения окружения всех компонентов предложения, участвующего в переводе, так как у машины нет интуитивных свойств. Кроме этого, рекомендуется принять во внимание деривационные возможности при выделении морфемных основ формы слов, а также идеи по созданию баз слов посредством характеристики их структуральных моделей. В связи с этим, изучены свойства словообразования английского и узбекского языков и составлены модели по аранжировки аффиксов для лингвистического

¹⁶ <http://www.hutchinsweb.me.uk/Valladolid-2003-1-ppt.pdf>

обеспечения: в узбекском языке (*префикс+*) *основа+словообразовательный аффикс + аффикс, образующий лексическую форму + аффикс, образующий синтаксическую форму* и в английском языке (*Prefix+*) *root +lexical suffix +grammatical suffix*).

В процессе определения словообразовательных моделей способами композиция и аффиксация, выявлены следующие положения: 13 видов существительного, образованного способом композиция, 60 видов словообразовательных префиксов, 122 вида словообразовательных суффиксов, 21 вид формаобразующих аффиксов в английском языке, 20 видов прилагательного, образованного способом композиция, 28 префиксов, 53 суффикса, 3 формаобразующих суффиксов служат формированию степеней прилагательного. В результате семантического анализа выявлено 22 префикса, 10 суффиксов образующие глагол, 5 формаобразующих суффиксов, а также 9 видов модели глагола способом композиция. Семантический анализ показал то, что в английском языке существуют 2 вида модели наречия способом композиция, 4 префикса образующие наречие, 11 суффиксов и 2 суффикса образующие форму *-er* и *-est* придают наречию значение степени.

В результате семантического анализа грамматических категорий выявлено существование 52 вида предлога, 42 вида предлог-фразы, 18 видов деепричастного предлога, 70 видов союза, 91 вид модальных слов, 34 вида междометий. Выявлено что в узбекском языке существуют 7 видов словообразования способом композиция, основа – существительное + 38 видов словообразовательного аффикса, основа – глагол + 52 словообразовательного аффикса, основа – прилагательное + 9 видов словообразовательных аффиксов образующие существительное, основа – слова подражания + 5 видов словообразовательных аффиксов образующие существительное. Определено общее число 115 видов аффиксов образующие прилагательные, 8 видов суффиксов, 8 видов аффиксов уменьшительной степени прилагательного, 11 видов превосходной степени прилагательных, 4 вида средств формирующие уменьшительную степень прилагательных, а также 5 видов способа композиции.

Изучение деривационных возможностей английского и узбекского языков помогает определить морфотактическое положение, то есть увидеть семантическую и логическую границу каждой морфемы. Кроме этого наблюдаются 5 видов сложного глагола, 60 видов словообразующих аффиксов, 6 видов выражения модальности, 26 видов аффиксов формирующие залого глагола, 4 вида выражения отрицания глагола, 30 аффиксов наклонения глагола, 31 аффиксов времени глагола, 23 вида аффиксов лиц. Определено общее 50 видов аффиксов образующие наречие, 2 вида суффиксов, 4 вида двух коренных наречий.

В процессе оценки семантических и логических возможностей устойчивых сочетаний в лингвистическом обеспечении в качестве образца привлечены к анализу глагольные фраземы существующие в английском

языке. В исследовании собрано более 12 тысяч глагольных фразем составленных на базе Access, а также предоставлены значения более 3 тысяч основных глаголов и даны эквиваленты на узбекском языке. Так как состав фразем отличается друг от друга, уделяется особое внимание на их семантическое поле.

База сведений, составленная нами, намечена по направлению английский – узбекский – английский. В полях Access отдельные значения глагольных фразем даются с ID. Взяв во внимание стиль текста и состояния контекста каждое значение фраземы даётся в специальных знаках, с отдельными дополнениями синонимами. Это создает удобную возможность при создании программного обеспечения и при поиске сведений. Поэтому приняв во внимание все значения глагольных фразем, собрана база около 80 тысяч сем.

Важное значение данного лингвистического обеспечения состоит в том, что его можно использовать и как базу для программы перевода текстов, и как электронный словарь, так как в поле глагола даются транскрипция единиц. При переводе модели фразем внесены в базу отдельных моделей. Ибо глаголы разделяемые и неразделяемые по структуре даны отдельно, база этих глагольных фразем основывается на следующих моделях: (*V* – глагол; *P*-particles (предлог или наречие); ↓ – можно разделить или нет; *smth.*-something, *smb.*-somebody): 1) *V* + *oneself* + *P*+*smth.* => *align oneself with smb. or smth.*; 2) *V* + *oneself* + *P*+*smb.* => *attach oneself to smb.*; 3) *V* + *P*+ *smb. or smth.*+*P smth.* => *arrange with smb. about smth.*; 4) *V*+ *smb. or smth.* + *P*+ *smb. or smth.* => *associate smb. or smth. with smb. or smth.*; 5) *V*+ *oneself* +*P* => *arch (oneself) over.* Определено общие число таких моделей около сорока. На основе данных моделей в базу добавлены следующие знаки: 1) *P1, P2, ... , PN* – *N* значения основного глагола; 2) *N1, N2, ... , Nk* – *k* формы глагольных фразем; 3) *V1, V2, ... , V1* – *l* значения глагольных фразем; 4) *P1_, P2_, ... , PN* – синонимы основного глагола; 5) *V1_, V2_, ... , V1_* – синонимы значений глагольных фразем; 6) *P* – не относится ни к какому глаголу. Глагольные фраземы английского языка соответствуют таким категориям узбекского языка как фразеологизмы, сложные глаголы, слово соединения с вспомогательным глаголом, простой глагол. Для машинного перевода разработан алгоритм анализа глагольных фразем:

А также, посредством модели словосочетаний узбекского языка исследовано соответствие моделей английского языка. В результате создана возможность перевода не только с английского на узбекский язык, но и наоборот. Разделение подчиненных сочетаний на 1) существительное сочетание (сочетаний имен); 2) глагольное сочетание, в зависимости от части речи передового слова, облегчило определение форм синтаксической связи, отношения. Созданы 20 видов модели сочетаний имен примыкании на подобие: *Noun+Noun* => *темир ускуна* <=> *Noun+Noun* => *iron equipment* (железное оборудование); *Adj.+Noun* => *қулай имконият* <=> *Adj.+Noun* => *suitable opportunity* (удобная возможность); 3 вида модели глагольных

сочетаний в примыкании: *Adj.+Verb => яхши ўқимоқ– reading well (хорошо читать)*. Созданы 20 видов модели сочетаний имен в отношении руководство: *Noun + дан + Noun => Андижондан хат <=> Noun + Prep.+ Noun => the letter from Andijan (письмо из Андижана)*; 10 видов модели глагольных сочетаний в отношении руководство: *Noun +га + Verb => мактабга бормоқ <=> Infinitive+Prep.+Noun=>to go to school (идти в школу)*; 3 вида модели определительных сочетаний имен (турецкий изофет) *Noun|Pronoun+ning+Noun=>китобнинг муқоваси <=> Noun+of+Noun=> the cover of the book (обложка книги)*.

Посредством некоторых примеров обоснованы значение сложностей перевода терминов: 1) не существование того или иного термина во втором языке; 2) выражение терминов в словосочетаниях; 3) определение стиля текста и связывание значение научного термина к смыслу текста; 4) использование одной лексемы в разных значениях в определенной сфере и т.д. в процессе создания лингвистического обеспечения.

В третьей главе работы под названием «Анализ текста в машинном переводе» утверждается существование трех важных компонентов перевода: текст, степень правильности поставленных задач, вопросы терминологии и их роль. В процессе перевода текста изучены следующие вопросы: процесс анализа, процесс понимания смыслов слов в тексте и процесс восстановления текста. К этапам анализа машинного перевода относятся токенизация (определение формы слов), лемматизация (определение слова), стеммизация (определение ведущих и вспомогательных морфем).

Своеобразность лексического пласта создало ряд сложностей в автоматическом переводе. В том числе, лингвокультурологические свойства языковых единиц, существование синонимической градуонимии лексем в кругу существительных и прилагательных, аффиксальная и лексическая омонимия в английском языке. А также со стилистической точки зрения, употребление терминов в разных сферах в нескольких значениях. В процессе анализов пришли к выводу, что необходимо создать отдельные базы указанных выше единиц.

Грамматические категории английского и узбекского языков, их общие и различные черты привлечены к анализу в сопоставительном аспекте. Отмечены роль морфологического и синтаксического анализа в машинном переводе. Кроме этого подчеркнуто то, что в морфологическом анализе осуществляется классификация категорий по принципам ни *что* классифицируется, а *как* классифицируется. Как компоненты морфологического анализа понимаются: специально объектный модель словаря, словарь, морфологические, фонологические и морфофонологические правила. Что бы правильно найти сведения данные в программном обеспечении, введены следующие знаки: *H_W* – определённое слово создает омонимию только в кругу одной части речи; *H_L* – определенное слово создает омонимию в кругу более двух частей речи; *H_Adj.* {*H_Adv.*} –

определенное слово создает омонимию только в кругу двух частей части речи, вторая часть речи указывается точно.

Морфологический анализ анализирует не только те или иные категории, но по мнению Родолфе Делмонте для лексикона важны и следующие лингвистические категории¹⁷: 1) грамматические категории; 2) семантические категории; 3) категории степени дискурса; 4) синтаксические категории; 5) подчинительные категории; 6) семантические концептуальные категории; 7) разграниченные лексические единицы; 8) грамматические ограничения. В любой системе (поиск информации, машинный перевод, автоматическое рецензирование) обращается к грамматическому (морфологическому и синтаксическому) анализу слов. Поэтому во многих исследованиях даются мнения об оптимизации морфологического анализа¹⁸. В лингвистическом обеспечении создан морфологический словарь английского и узбекского языка. В морфологическом словаре база аффиксов, которые могут быть включены к самостоятельным частям речи, создана в Accessce, и к которым могут присоединяться окончания (аффиксы) внесена цифра 1.

В узбекском языке на базе данных если даются основа глагола, могут появиться омонимия: $n_1 \rightarrow$ тур (*встань*) (тип) \leftrightarrow $n_2 \rightarrow$ тур (*ўрнидан турмоқ*) (*вставить с места*) $n_3 \rightarrow$ тур (*уйғонмоқ*) (*проснись*) \rightarrow $n_4 \rightarrow$ тур (*нархда турмоқ*) (стоит (цена)). В данном случае если омонимия возникает в кругу одной части речи: H_W (homonym- \rightarrow words), если омонимия возникает в кругу двух частей речи, указывается вторая часть речи омонимической пары: H_Adj (homonym- \rightarrow adjective), если омонимия наблюдается в кругу нескольких частей речи, ставится знак: H_L (homonym- \rightarrow lists). Такие знаки служат облегчению морфологического анализа. При синтаксическом анализе текста воспользовались генеалогическим (родословным) методом анализа Н.Чомского. В настоящее время статистический и нейромашинный перевод развиваются по системе основанной на синтаксисе. В лингвистическом анализе, направленном на изучение семантической связи в тексте, продолжается процесс анализа грамматических категорий, ибо морфологическая структура не может охватить всю лингвистическую

¹⁷ Rodolfe D. Computational Linguistic Text Processing: Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, Inc. – New York, 2008. – P. 4-5.

¹⁸ Belonogov G.G. On the use of a similarity method for automatic processing of textual information [Об использовании метода аналогии при автоматической обработке текстовой информации] Problems of cybernetics [Problemy kibernetiki]. Issue 28. – Moscow: Nauka Publ., 1974; Bolshakov I.A. CrossLexica: A large electronic dictionary of collocations and semantic links between Russian words [КроссЛексика – большой электронный словарь сочетаний и смысловых связей русских слов]. Компьютерная лингвистика и Интеллектуальный Технологии: Труды Международной Конференции «Диалог 2009» [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference «Dialogue 2009»]. – Moscow, 2009. – P. 45–50; Gelbukh A. F. Effectively realizable morphologic model of inflective language [Эффективно реализуемая модель морфологии и активного языка]. Научно-техническая Информация [Scientific and Technical Information], series 2, #1, 1992. – P. 24–31; Sokirko A. V. Bystroslovar': morphological prediction of new Russian words using very large corpora [Быстростловарь: предсказание морфологии русских слов с использованием больших лингвистических ресурсов]. Компьютерная лингвистика и Интеллектуальный Технологии: Труды Международной Конференции «Dialog 2010» [Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference «Dialogue 2010»]. – Moscow, 2010. – P. 450.

возможность, но может дать достаточное представление для изучения объекта: *read somebody like a book* – бирор кишини жуда яхши тушунмоқ (очень хорошо понять какую-нибудь книгу); *book club* – китобхонлар тўгараги (кружок читателей); *speak by the book* – аниқ маълумотга таяниб гапирмоқ (говорит опираясь на точный факт).

В четвертой главе диссертации «**Моделирование текстов в системе машинного перевода**» моделированы виды по цели выражения простых предложений и изучены синтаксические модели по следующим видам: 12 типов повествовательного предложения простых предложений с существительным – сказуемом и глагольным – сказуемом в английском языке: Noun ⊕ Verb ⊕↓ Adverb; 2. Article ⊕↓ Adjective ⊕ Noun ⊕↓ Verb ⊕↓ Adjective; 10 видов вопросительного предложения: 1. Helping verb ⊕ Noun ⊕↓ Verb ⊕↓ Noun; 2. Helping verb ⊕ Noun ⊕↓ Pronoun ⊕↓ Verb ⊕↓ Noun; 12 видов восклицательного предложения: 1. Noun ⊕ Helping verb ⊕ Verb; 2. Noun ⊕↓ Noun ⊕ Helping Verb ⊕↓ Verb.

Утверждается важность трансформации при обращении к лингвистическому обеспечению, размещенному в памяти компьютера и при автоматическом приказе для выделения каждого фрагмента (слово или словосочетания в качестве части предложения) текста. В связи с этим, точность модели предложения и совершенство внесенной лингвистической базы служит повышению эффективности работы. Поэтому, в этом обстоятельстве существует необходимость определения позиции между собой лингвистических моделей языков для автоматического перевода на примере английского и узбекского языков. Здесь соответствие, согласованность в английском и узбекском языках является одним из важных сторон. В том числе, является важным в переводе привлечение к отдельному анализу с точки зрения специфики предложений как [*Neither of {of Prn./ Noun}+PS+Obj.*] => [{*Ol. | Ot*} +*келишик {-дан,-нинг}*] *Ҳеч қайси {бир}{+u |-миз | -имиз|-нгиз |-ингиз|-лари}* } + {*Ol. | Ot*} +*Ob.*+{*эмас*} *KB/K*] (*Neither of the students has passed exam.* => *Талабаларнинг ҳеч қайси бири имтиҳон топширмади (Ни один из студентов не сдал экзамены).*

Показано как с алгоритмической стороны осуществляется результат исследования в программном языке Java. По нему в двух языковом переводческом пространстве для создания программного обеспечения целесообразно внесение следующих знаков в базу сведений.

Наименование баз	Функция
R_i	База терминов и фразеологизмов по предметным сферам
Q1	База всех коренных слов языка
K1	База всех производных слов языка
V2	База членов предложения
V3	База частей речи

Для обоих языков внесены указанные выше таблицы. Данное транспортирование переводческой среды служит переводу научных текстов. При определении структуры каждого предложения, составляющие текст,

каким членом предложения выступает определенная часть речи, конечно, необходимо обращаться к грамматике каждого языка. Данное транспортирование переводческой среды осуществляет перевод по направлению узбекский – английский, английский – узбекский: внесенный текст (Z) разделяется на члены предложения, перевод каждого слова берется из базы терминов и размещаются члены предложения по грамматике этого языка.

Расширяемому входному языку математического моделирования естественных языков¹⁹ даются следующие обозначения:

$T3_{i1}$ – сборник состоящий из перевода и функций в предложении слова в другом языке, $1 \leq i1 \leq m$;

$T4_{j1}$ – перевод слова в другом языке, $1 \leq j1 \leq m1$;

$T2$ – переведенный текст;

$\mathcal{E}4$ – подлежащее; $\mathcal{E}2$ – сказуемое; $\mathcal{E}5$ – определение; $\mathcal{E}6$ – дополнение; $\mathcal{E}7$ – обстоятельство.

Создаём модели подходящих структур предложения в другом языке, структуру предложения определенного языка. Ниже представлены модели предложения в английском языке подходящие структуре предложений узбекского языка.

а) математические модели предложений повествовательной структуры в узбекском языке обозначаем так:

- I.
 1. $\langle E4 \rangle \downarrow \oplus \langle E5 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma 2 \rangle$.
 2. $\downarrow \langle E5 \rangle \oplus \langle E4 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma 2 \rangle$.
 3. $\downarrow \langle E5 \rangle \downarrow \oplus \langle E5 \rangle \oplus \langle E4 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma 2 \rangle$.
 4. $\langle E4 \rangle \downarrow \oplus \langle E5 \rangle \downarrow \oplus \langle E6 \rangle \oplus \langle \Gamma 2 \rangle$.
 5. $\langle E4 \rangle \oplus \langle \Gamma 2 \rangle$.
 6. $\langle E4 \rangle \downarrow \oplus \langle E7 \rangle \oplus \langle \Gamma 2 \rangle$.
 7. $\langle E4 \rangle \downarrow \oplus \langle E6 \rangle \oplus \langle \Gamma 2 \rangle$.

В [1, 3, 4] математических моделях структура предложения состоит из частей речи, к данным математическим моделям вносим некоторые дополнения. То есть, определяем каким членом предложения участвует часть речи, берется их перевод в другом языке и размещается в структуре предложения. Создание самого оптимального и содержательного перевода из моделей возлагается обязанностям алгоритма. Математические модели предложений повествовательной структуры в английском языке подходящие к указанным выше математическим моделям предложений повествовательной структуры в узбекском языке формируются так:

- II.
 1. $\langle E4 \rangle \downarrow \oplus \langle E5 \rangle \oplus \langle \Gamma 2 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle$.
 2. $\downarrow \langle E5 \rangle \oplus \langle E4 \rangle \oplus \langle \Gamma 2 \rangle \downarrow \oplus \langle E6 \rangle \downarrow \oplus \langle E7 \rangle$.
 3. $\downarrow \langle E5 \rangle \downarrow \oplus \langle E5 \rangle \oplus \langle E4 \rangle \oplus \langle \Gamma 2 \rangle \downarrow \oplus \langle E6 \rangle$.
 4. $\downarrow \langle E5 \rangle \oplus \langle E4 \rangle \oplus \langle \Gamma 2 \rangle \downarrow \oplus \langle E6 \rangle$.

¹⁹ Хакимов М.Х. Расширяемый входной язык математического моделирования естественного языка для многоязычной ситуации машинного перевода // ЎзМУ хабарлари. – Ташкент, 2009. – № 1. – С.75-80.

5. $\langle E4 \rangle \oplus \langle \Gamma 2 \rangle$.
6. $\langle E4 \rangle \oplus \langle \Gamma 2 \rangle \downarrow \oplus \langle E7 \rangle$.
7. $\langle E4 \rangle \oplus \langle \Gamma 2 \rangle \downarrow \oplus \langle E6 \rangle$.

Воспользовавшись указанными структурами предложения и базой терминов приводим алгоритм перевода.

Используем следующие определения:

$Q1_uz = \langle \text{SELECT * FROM `Q1_uz`} \rangle$ – все коренные слова узбекского языка;

$K1_uz = \langle \text{SELECT * FROM `K1_uz`} \rangle$ – все производные слова узбекского языка;

$Q1_eng = \langle \text{SELECT * FROM `Q1_eng`} \rangle$ – все коренные слова английского языка;

$K1_eng = \langle \text{SELECT * FROM `K1_eng`} \rangle$ – все производные слова английского языка;

E_i – Z из предложений, выбранных из текста, $1 \leq i \leq n$;

$L1_j$ – E_i – слова выбранные из, $1 \leq j \leq n1$.

После выполнения алгоритма в [2], следующий алгоритм «поиск слова» разделяется на предложения Z , а предложения на слова. Сначала выполняется поиск каждого слова из базы коренных слов, если не находится – выполняется поиск из базы производных слов. Когда находится слово, получается его перевод в другом языке. Например, первый алгоритм перевода по направлению узбекский – английский выглядит так:

1. Каждое слово из $L1_j$ найти из $Q1_uz$. Если найдётся, переходим на шаг 2, если нет переходим на шаг 4;

2. (ID) в английском языке данное слово взять из $Q1_uz$;

3. Взять перевод коренного слова из $Q1_eng$ и перейти на шаг 7;

4. Каждое слово из $L1_j$ найти из $K1_uz$;

5. (ID) в $K1_eng$ данного производного слова взять из $K1_uz$;

6. Взять перевод производного слова из $K1_eng$;

7. Определить каким членом предложения является слово и разместить его в сборнике $T3_{i1}$;

8. Пополненный сборник $T3_{i1}$ отправить в функцию $UzbekIngliz(T3_{i1})$;

9. Результаты функции $UzbekIngliz(T3_{i1})$ разместить в $T2$.

Здесь $UzbekIngliz(T3_{i1})$ функция, написанная на основе алгоритма перевода по направлению узбекский–английский, приведенный в [2]. Функция $UzbekIngliz(T3_{i1})$ пишется в следующем порядке. Для написания функции используем следующие обозначения:

1. $ET3_{k1}$ – структура предложений английского языка подходящие к структурам предложения в узбекском языке $1 \leq k1 \leq m2$;

2. Обязанности в предложении слов, загруженных в сборник $T3_{i1}$ загрузить в $E8_k$;

3. Подходящую структуру предложения $E8_k$ найти из $ET3_{k1}$;

4. Подобрать слова, выполняющие функцию членов предложения, найденных из $ET3_{k1}$ загрузить в $T2$;

Алгоритм перевода по направлению английский–узбекский:

1. Каждое слово из $L1_j$ найти из $Q1_eng$. Если найдётся, переходим на шаг 2, если нет, переходим на шаг 4;
2. (ID) в английском языке данного слова взять из $Q1_eng$;
3. Взять перевод коренного слова из $Q1_uz$ и перейти на шаг 7;
4. Каждое слово из $L1_j$ найти из $K1_eng$;
5. (ID) в $K1_uz$ данного производного слова взять из $K1_eng$;
6. Взять перевод производного слова из $K1_uz$;
7. Определить каким членом предложения является слово и разместить его в сборнике $T3_{i1}$;
8. Пополненный сборник $T3_{i1}$ отправить в функцию `InglizUzbek` ($T3_{i1}$);
9. Результаты функции `ИнглизУзбек` ($T3_{i1}$) разместить в $T2$.

Здесь `ИнглизУзбек` ($T3_{i1}$) функция, написанная на основе алгоритма перевода по направлению английский–узбекский, приведённый в [2]. Функция `InglizUzbek` ($T3_{i1}$) пишется в следующем порядке. Для написания функции используем следующие обозначения:

$ET4_{k1}$ – структура предложений узбекского языка подходящие к структурам предложения в английском языке $1 \leq k1 \leq m2$;

1. Обязанности в предложении слов, загруженных в сборник $T3_{i1}$ загрузить в $E8_k$;
2. Подходящую структуру предложения $E8_k$ найти $ET4_{k1}$;
3. Подобрать слова, выполняющие функцию членов предложения, найденных из $ET4_{k1}$ и загрузить в $T2$;

Тег в программном языке Java следующей функции приведен ниже:

```
private String EngUzb(String suz) throws ObjectNotFoundException {
    int uzakId=0;
    String uzbSuz =««;
    int gapBulagiId=0;
    UzakEnglish ue=uzakSuzEnglishDao.getUzakEnglishByword(suz);
    if(ue.getUzakEnglish().equals(suz)){
        uzakId=ue.getUzakSuzlarId();
        List<UzakSuzlar> usList=uzakSuzUzbekDao.getuzakSuzlarListByRId(uzakId)
        for (UzakSuzlar us : usList) {
            uzbSuz=us.getUzakSuzlar();
        }
    }else{
        YasamaEnglish e=yasamaSuzEnglishDao.getYasamaEnglishByWord(suz)
        if(suz.equals(ye.getYasamaEnglish())){
            uzakId=ye.getYasamaSuzlarId();
            YasamaSuzlar yu=(YasamaSuzlar) yasamaSuzUzbekDao.getYasamaSuzlarListByRId(uzakId);
            uzbSuz=yu.getYasamaSuzlar();
        }else {
            uzbSuz=suz;
        }
    }
    return uzbSuz;
}
```

Таким образом, создание программных переводов, связанных с узбекским языком является одной из важных задач. Наше исследование в

первую очередь создается в двуязычном переводческом пространстве, основанном на грамматике (Rule based translation). Создание многоязычного переводческого пространства на базе с изученными и достигнутыми положительными результатами, является одним из будущих проектов, перспективных исследований.

ЗАКЛЮЧЕНИЕ

В результате исследования проведенного по созданию лингвистического обеспечения программы перевода английских текстов на узбекский язык пришли к следующим выводам:

Проведение статистического подсчета и создание семантической базы словообразующих аффиксов английского и узбекского языков помогает составлению программного обеспечения. В процессе анализа важным является распределение аффиксов в отдельные базы для компьютера и определение их моделей.

Моделирование является самым эффективным методом в машинном переводе. Здесь существует возможность в определенной степени анализировать, проверять и сопоставлять лингвистические явления. Хотя оно полностью не отражает в себе все речевые явления, но имеет значение при определенном выводе о них. Не возможно создать универсальный шаблон в языке, потому что язык как речевое явление отражается в каждом индивидууме. Для оптимального решения вопроса, учитывая, что для анализа каждого языка созданы десятки алгоритмов отметим, что моделирование языка в формальной степени является продуктивным, эффективным способом.

В компьютерной лингвистике моделирование осуществляется по простым предложениям. Сложные и составные предложения как синтаксические структуры являются фразами. Формулирование простых предложений в некоторой степени эффективно при анализе структурных частей составных предложений.

Так как агглютинативные и флективные языки по многим параметрам отличаются друг от друга, очень важно создание единственного общего метода их анализа. Поэтому при моделировании сооружений предложений опираясь на своеобразные свойства языков, необходимо учитывать модели речевых явлений, подходящих к вероятностям.

В машинном переводе анализ основанный на синтаксисе подтвержден как эффективный способ, потому что посредством синтаксиса возникает возможность проверки смысла предложения с разных ракурсов. Кроме того, разделение предложений в соотношении подлежащие – имя и подлежащее – глагол считается удобным способом прослеживания участия грамматических категорий и анализа.

Отделяется особой важностью различение друг от друга устойчивых сочетаний и словосочетаний, конкретный показ их моделей, а также

обозначение их семантических свойств. В лингвистической базе программы перевода создан электронный словарь глагольных фразем.

В диссертации в базу внесены более 12 тысяч глагольных фразем английского языка и их значения около 80 тысяч. Внесены специальные знаки для их многозначности и синонимии выраженных значений. Это служит лингвистическим обеспечением для последующих исследований.

Только тогда, когда лингвистическое обеспечение в машинном переводе является совершенным в полной мере, можно создать его программное обеспечение. В дальнейших исследованиях семантический слой лингвистического обеспечения, отдельные базы языковых и речевых единиц поможет решению вопросов омонимии и синонимии, стоящих перед машинным переводом.

Двухязычный машинный перевод можно использовать не только для перевода английских текстов на узбекский язык, но и наоборот. Надо заметить, что определение окончательных критерий формальной грамматики узбекского языка послужит в будущем разработке технологии многоязычного перевода.

**ONE-TIME SCIENTIFIC COUNCIL BASED ON SCIENTIFIC COUNCIL
WITH THE NUMBER OF DSc. 27.06.2017. Fil. 19.01 AWARDED
SCIENTIFIC DEGREES UNDER TASHKENT STATE UNIVERSITY OF
UZBEK LANGUAGE AND LITERATURE NAMED AFTER
ALISHER NAVAI**

**TASHKENT STATE UNIVERSITY OF UZBEK LANGUAGE AND
LITERATURE**

ABDURAKHMONOVA NILUFAR

**LINGUISTIC DATABASE OF TRANSLATING PROGRAM FROM
ENGLISH TEXTS INTO UZBEK (AS EXAMPLE OF SIMPLE
SENTENCES)**

10.00.11 – Theory of language. Applied and computational linguistics

**DISSERTATION ABSTRACT FOR THE DOCTOR OF PHILOSOPHY (PhD)
OF PHILOLOGICAL SCIENCES**

Tashkent – 2018

The theme of dissertation of doctor of philosophy (PhD) was registered with № B2017.2.PhD/Fil220 at the Supreme Attestation Commission under the Cabinet of Ministers of the Republic of Uzbekistan.

The dissertation has been done at the Tashkent State University of Uzbek Language and Literature named Alisher Navai.

The abstract of dissertation has been posted on website in the web-link <http://navoiy-uni.uz> and on the information-educational portal of «ZiyoNet» under the website address www.ziynet.uz in three languages (Uzbek, Russian and English).

Scientific consultant: **Khamidulla Dadaboyev Aripovich**
Doctor of Philology, professor

Official opponents: **Suyun Karimov Amirovich**
Doctor of Philology, professor

Abdumurod Mamatov Eshonkulovich
Doctor of Philology, professor

Leading organization: **Andijan State University**

The defense will take place on «_____» _____ 2018 _____ at _____ a meeting of a one-time Scientific Council based on the Scientific Council DSc. 27.06.2017. Fil. 19.01 under the Tashkent State university of Uzbek Language and Literature named after Alisher Navai. Address: 100100, Tashkent, Yakkasaroy district, st.Yusuf Xos Hojib, 103. Tel.: (99871) 233-36-50; Fax: (99871) 233-71-44; e-mail: (www.navoiy-uni.uz).

The dissertation can be reviewed at the information-resource centre of Tashkent State university of Uzbek Language and Literature named Alisher Navai. Address: 100100, Tashkent, Yakkasaroy district, st.Yusuf Xos Hojib, 103. Tel.: (99871) 233-36-50; Fax: (99871) 233-71-44; e-mail: (www.navoiy-uni.uz).

The abstract of dissertation is distributed on: «_____» _____ 2018.
(Protocol at the register № on _____ of «_____» _____ 2018).

Sh.S.Sirojiddinov

Chief of a one-time Scientific Council based on the Scientific Council awarding scientific degrees, Doctor of Philological sciences, Professor

Y.R.Odilov

Scientific secretary a one-time Scientific Council based on Scientific Council awarding scientific degrees, Doctor of Philological sciences

N.M.Makhmudov

Chairman of the Scientific Council dealing with Doctor of Philological sciences, Professor

INTRODUCTION (abstract of PhD thesis)

The aim of the research work is to analyze theories on building linguistic database of translating program from English simple texts into Uzbek and to create linguistic foundations.

The object of research work is word combinations and simple sentences of English and Uzbek languages, additionally grammatical descriptions

Scientific novelty of the research work is as follows:

Drawn conclusions which provide exactness of translation on creating linguistic database of machine translation.

created linguistic database of phrasal verbs, morphological lexicon, affixes of English and Uzbek languages and their morphological and syntactic models;

identified the coordination of models of simple sentences for automated translation;

substantiated principles that are for analyzing morphological, syntactic-semantic texts of machine translation;

worked out recommendations of coordinating paradigmatic attitudes on creating principles for electron dictionary and software for linguistic database

Implementation of the research results:

Results based on scientific and applied achievements on linguistic database of translating program of English texts into Uzbek are:

collected more than twelve thousand phrasal verbs and their notion verbs which are used in all styles of English in «English-Uzbek phrasal verbs dictionary» and identified their Uzbek equivalencies. As a result, linguistic database was created, lingual models for machine translation were determined, and in linguistic database synonyms and homonyms in Uzbek translations were represented;

linguistic database of English-Uzbek phrasal verbs for linguistic database of translating program from English into Uzbek, scientific terms in English phrasal verbs have been used in the Project of Fundemantal Researches F-4-02 «Devising new methods of solution of optimal directive tasks and mathematic physics of differential operators» (Reference of Commitee for the Coordination of the development of Science and Technology of December 7, 2017 No. FTA-02-11/1279). As a result, foreign investigations connected with the project were studied, enriched with theoretical literatures, scietific researches were published as articles;

linguistic database of English-Uzbek dictionary and phrasal verbs have been used in the Project of Fundemantal Researches F-4-02 «Foundation and creation of computable methods of individual metrics and generelized merits in the intelectual database analyzing» (Reference of Commitee for the Coordination of the development of Science and Technology of December 7, 2017 No. FTA-02-11/1279). As a result, articles have been published in the foreign countries, foreign literatures were adopted to the project, scientific effectiveness of the project was improved by usage of scientific terms.

The structure and outline of the research work. The thesis consists of an introduction, four chapters with outcomes, general conclusions, appendix and the list of the used literatures. The volume of the thesis is 160 pages.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (I часть; I part)

1. Абдурахмонова Н. Матнни автоматик таржима қилиш жараёни ва унга таъсир этувчи лингвистик омиллар // ЎзМУ хабарлари. – Тошкент, 2012. –№1. –Б. 76-79. (10.00.00 № 7)

2. Абдурахмонова Н. Инглизча-ўзбекча машина таржимаси тизимининг семантик-синтактик таҳлил қилиш муаммоси // Хорижий филология. – Самарқанд, 2012. –№3 (44). –Б. 8-12. (10.00.00 № 19)

3. Абдурахмонова Н. Машина таржимаси тизимида сўз бирикмаларининг семантик базаси // Таълим тизимида ижтимоий-гуманитар фанлар. –Тошкент: ТДТУ, 2012. – №3. –Б. 140-145. (10.00.00 №16)

4. Абдурахмонова Н. Инглизча матнлардаги эга ва кесим мослашувини ўзбек тилига моделлаштириш // ЎзМУ хабарлари. –Тошкент, 2013. – №4. –Б. 155-158. (10.00.00 №1)

5. Абдурахмонова Н. Матн машина таржимасининг ўрганиш объекти сифатида // ЎзМУ хабарлари.–Тошкент, 2014. –№1/2. –Б. 279-281. (10.00.00 № 1)

6. Абдурахмонова Н. Компьютерда илмий матнларни таржима қилишнинг асосий масалалари // Илм сарчашмалари. –Хоразм, 2013. – № 5. – Б. 40-45. (10.00.00 № 11)

7. Abdurakhmonova N. Modeling analytic forms of verb in Uzbek as stage of morphological analysis in machine translation // Iranian Journal of Social Sciences and Humanities Research. UCT. J. Soc. Human. Resear. (UJSSHR). –Takestan, 2017. Volume 5, Issue 4. – 20 p. (№ 5 Global Impact factor, Impact factor-1,87)

8. Абдурахмонова Н. Морфологик таҳлилнинг машина таржимасида берилиши (от сўз туркуми модели мисолида) // Ўзбекистонда хорижий тиллар (электрон илмий-методик журнал). –Тошкент: ЎзДЖТУ, 2017. –№ 2. (16) –Б. 40-51. (10.00.00 №17)

9. Абдурахманова Н., Хакимов М. Логико-лингвистические модели слов и предложений английского языка для многоязычных ситуаций компьютерного перевода. / Компьютерная обработка тюркских языков. Латинизация письменности. 1-я Международная конференция. –Астана, 2013.– С. 297-302.

10. Abdurakhmova N. Modeling word combinations in terms of parts of speech in the process English-Uzbek machine translation / TurkLang. 2017. Vth international conference on Computer processing of Turkic languages. – Kazan, 2017. – P. 153-162.

11. Абдурахмонова Н., Исматуллаева Н. Машина таржимаси учун илмий матнлардаги барқарор бирикмаларнинг алгоритмик тавсифи / Ўзбек тилшунослигининг долзарб масалалари (Республика илмий-назарий анжумани материаллари), 2016. – Б. 189-196.

12. Абдурахмонова Н. Ўзбекча матнларни морфологик таҳлил қилувчи автоматик анализаторни яратиш масалалари / Ёш тилшунос ва адабиёт-шуносларнинг анъанавий республика илмий-амалий анжумани. –Тошкент: ТошДЎТАУ, 2016. –Б. 34-37.

II бўлим (II часть; II part)

13. Абдурахмонова Н. Таржимон дастур учун инглизча-ўзбекча феъл ва феълли фраземалар базаси. № ВГУ 00341 рақамли илмий гувоҳнома, 2017.

14. Абдурахмонова Н., Каримбердиева Ш. Имлони текширувчи мобил илова. № DGU 04591 рақамли патент, 2017.

15. Абдурахмонова Н. Машина таржимасининг лингвистик асослари. Ўқув қўлланма. – Тошкент: Akademnashr, 2012. –160 б.

16. Абдурахмонова Н. Матн, тасвир, овоз кўринишидаги ахборотларни рақамлаштириш. Компьютер лингвистикаси дарслигининг 5-параграфи. – Тошкент: Академнашр, 2011. –Б. 27-35.

17. Abdurakhmonova N. English-Uzbek dictionary of phrasal verbs. – Tashkent: Turon zamin ziyo, 2017. – 420 b.

18. Менглиев Б., Ўролов Х., Абдурахмонова Н. Ўзбек тилидан универсал қўлланма. – Тошкент: Akademnashr, 2014. –250 б.

19. Abdurakhmonova N. Automatic morphological analyze for English-Uzbek system // Известия Кыргызский государственный технический университет им. И.Раззакова. Теоретический и прикладной научно-технический журнал. – Бишкек, 2016. – № 2 (38). –С. 12-18.

20. Абдурахмонова Н. Ўзакдош феълларнинг кетма-кет қўлланишига доир / Тилшуносликка илк қадам. Талабаларнинг илмий мақолалар тўплами. – Тошкент, 2007. –Б. 55-59.

21. Абдурахмонова Н.Тил–яшаш калити / Магистратура ва бакалавриат талабалари илмий мақолалари тўплами. –Тошкент, 2007. –Б. 2-4.

22. Абдурахмонова Н.Ўзбек ва инглиз тилларида шарт эргаш гаплар / Тилшуносликка илк қадам. Талабаларнинг илмий мақолалар тўплами. – Тошкент, 2008. –Б. 121-124.

23. Абдурахмонова Н. Ўзбек ва инглиз тилларидаги ёрдамчи феълларнинг қиёсий таҳлили / Ўзбек филологиясига оид тадқиқотлар. – Тошкент, 2009. –Б. 54-55.

24. Абдурахмонова Н. Инглиз тилидан ўзбек тилига таржима дастурининг лингвистик таъмини / Ўзбек филологиясининг долзарб масалалари. –Тошкент, 2010. –Б. 280-285.

25. Абдурахмонова Н.Таржима дастурлари учун лингвистик таъмин яратишда предлог масаласига доир баъзи мулоҳазалар / Лингвист I. Илмий мақолалар тўплами. –Тошкент, 2010. –Б. 95-97.

26. Абдурахмонова Н., Тошимов Р. Матнни автоматик таржима қилишда «сунъий интеллект» масаласи / Ўзбекистоннинг мустақил тараққиёт йўли ва ижтимоий-фалсафий тафаккур янгиланиши. II Ёш олимлар конференцияси материаллари. –Тошкент, 2011. –Б. 241-243.

27. Абдурахмонова Н. Ахборотлашган жамиятда автоматик таржиманинг ўрни / Лингвист. II. Илмий мақолалар тўплами. – Тошкент, 2011. –Б. 132-135.

28. Абдурахмонова Н., Хакимов М. Семантические базы английского языка для многоязычной ситуации компьютерного перевода / Проблемы современной математики. – Карши, 2011. –С. 311-313.

29. Абдурахмонова Н.Таржимон дастурининг таъминотини яратишда лингвистик моделларнинг ўрни (инглиз ва ўзбек тиллари мисолида). / Филологиянинг долзарб муаммолари. Илмий-амалий анжуман материаллари (мақолалар тўплами). – Фарғона, 2012. –Б. 199-202.

30. Абдурахмонова Н. Машина таржимаси тизимида терминология масаласи / Низомий номидаги ТДПУ ЮНИСЕФ доирасида ўтказилган Республика илмий-амалий конференция. – Тошкент, 2012. –Б. 69-71.

31. Абдурахмонова Н. Матнни автоматик таржима қилишда шакл ва мазмун муносабати / Замоनावий тилшунослик ва таржимашуносликнинг долзарб масалалари Республика илмий-назарий конференция. – Тошкент, 2012. –Б. 49-52.

32. Абдурахмонова Н. Машина таржимаси тизимида лингвистик таҳлил масаласи / Республика илмий-амалий конференция мақолалар тўплами. – Тошкент, 2012. –Б. 23-26.

33. Абдурахмонова Н. Матнни автоматик таржима қилишда электрон луғатларнинг ўрни / Лингвист III. Илмий мақолалар тўплами. – Тошкент, 2012. –Б.137-141.

34. Абдурахмонова Н. Матнни синтактик таҳлил қилиш хусусида / XXI аср - интеллектуал инновацион ғоялар асри. Республика илмий ва илмий-техник конференция. –Тошкент: ЎЗМУ, 2012. –Б 200-201.

35. Абдурахмонова Н. Машина таржимасининг назарий лингвистик илдизлари // Ёш олимлар хабарномаси. – Тошкент: ЎЗМ, 2012. – №1. – Б. 104-107.

36. Абдурахмонова Н. Инглизча-ўзбекча машина таржимаси модули / Ўзбек филологиясининг долзарб масалалари. –Тошкент, 2013. –Б. 212-216.

37. Abdurakhmonova N. Grammatical analyze in machine translation / Компьютерная обработка тюркских языков. Латинизация письменности. 1-я Международная конференция. –Астана, 2013. – С. 294-297.

38. Abdurakhmonova N. The problem of word combinations in English-Uzbek machine translation // XIV всероссийской научно-практической конференции молодых, аспирантов и студентов (с международным участием). Северо-Восточный федеральный университет имени М.К.Аммосова –Нерюнгри, 2013. –С. 329-332.

39. Абдурахмонова Н. Машина таржимаси тизимлари / Тилшуносликнинг долзарб масалалари. VI илмий мақолалар тўплами.–Тошкент: Akadernashr, 2013. –Б. 212-214.

40. Дадабоев Ҳ., Абдурахмонова Н. Инглизча матнларни ўзбек тилига таржима қилиш алгоритми / Ахборот коммуникация технологияларидан фойдаланиш–ўзбек тили таълим самарадорлигини ошириш омили. «Ўзбек

тили» доимий анжумани ўн иккинчи йиғини материаллари. – Гулистон: ГДУ, 2013. –Б. 209-210.

41. Абдурахмонова Н. Принципы машинного перевода с английского языка на узбекский / Современная наука: тенденции развития, проблемы и перспективы. Международный аспирантский форум. –Ереван, 2013. –С. 20-24.

42. Абдурахмонова Н. Машина таржимасида таржима бирликлари / XXI аср - интеллектуал инновацион ғоялар асри. Республика илмий-амалий семинар материаллари. –Тошкент, 2014. –Б. 267-269.

43. Абдурахмонова Н. Машина таржимасида лингвистик таҳлил масаласи / Лингвист IV. Илмий мақолалар тўплами. –Тошкент, 2014. –Б.148-152.

44. Абдурахмонова Н. «Машина таржимаси»нинг илмий амалий тадқиқи / Мустақиллик йилларида Ўзбекистон таълим тизими: ислохотлар, ютуқлар ва истиқболлар. Республика илмий-амалий конференция материаллари. –Тошкент, 2016. – Б. 434-437.

45. Абдурахмонова Н., Каримбердиева Ш. База данных англо-узбекского словаря фразеологических глаголов. Ўзбекистон: тил ва маданият. 3-серия. – Тошкент, 2017. – Б. 122-124.

46. Абдурахмонова Н., Холиқова З. Ўзбек тилини ўргатиш дастурини яратишга доир айрим масалалар. / Ёш тилшунос ва адабиётшуносларнинг анъанавий республика илмий-амалий анжумани. –Тошкент: ТошДЎТАУ, 2016. –Б. 31-34.

47. Abdurakhmova N. Semantic issues of machine translation // O‘zbek terminologiyasi: bugungi holati va istiqboli nomli ilmiy-nazariy anjuman materiallari. –Tashkent, 2017. –B. 177-179.

48. Абдурахмонова Н. Машина таржимасида моделлаштириш методи / Problems of theoretical assumptions of linguistics and methodology in the period of globalization. Халқаро масофавий конференция. – Тошкент: ТошДЎТАУ, 2017. – Б. 141-148.

49. Abdurakhmonova N. Database of English-Uzbek verbs and phrasal verbs for machine translation / English is key to Dialogue. International conference for ELT specialists in higher education. – Tashkent: Uzbekistan State World Languages University, 2017. – 4 p.

50. Абдурахмонова Н. Компьютер лингвистикасининг назарий асослари ва йўналишлари / Нутқ маданияти ва ўзбек тилшунослигининг долзарб муаммолари мавзусидаги Республика илмий-амалий конференция. – Андижон: АДУ, 2017. – Б. 10-13.

51. Абдурахмонова Н. Компьютер лексикографиясига оид айрим масалалар / Хорижий тилларни ўқитишда она тилининг аҳамияти: муаммолар, ечимлар ва истиқболли режалар. Масофавий Республика илмий-амалий конференция. – Тошкент: ТошДЎТАУ, 2017. (диск, сертификат)

52. Абдурахмонова Н., Каримбердиева Ш. Имлони текширувчи мобил илованинг семантик базаси / Ўзбек миллий лексикографиясининг долзарб масалалари мавзусидаги Вазирлик миқёсида ўтказилган илмий-назарий конференция. – Тошкент: ТошДЎТАУ, 2017. – Б. 233-236.

53. Абдурахмонова Н. Компьютер лексикографияси: WordNet / Лингвист.

Илмий мақолалар тўплами. – Тошкент: ТошДЎТАУ, 2017. –Б. 86-89.

54. Абдурахмонова Н. Компьютер лингвистикасининг назарий асослари ва йўналишлари / Лисон илмининг чин яловбардори (Профессор Ҳамид Фуломович Неъматовнинг 75 йиллигига бағишланган), – Тошкент: БухДУ, 2017. –Б. 51-53.

55. Abdurakhmova N. The translation model of the type of simple sentences in the present simple (in active voice) from English into Uzbek / Россия ва Ўзбекистон миллий касб-хунар тизимида ЭСВЕТ ни тажриба синовдан ўтказиш лойиҳаси аъзолари ташрифи доирасида Миллий малакавий рамкаларнинг касбий таълим ислохотлардаги ўрни. Халқаро конференция (Қўқон давлат педагогика институти). –Қўқон, 2017– Б. 93-101.

56. Исроилов Ж., Абдурахмонова Н. Матн муҳаррирларида ўзбек тилида ёзилган матнларнинг имло хатоларини ўзбек алифбосига асосланган клавиатура орқали бартараф этиш йўллари / Ўзбек тилшунослигининг долзарб масалари. Республика илмий анжуман материаллари. –Тошкент: ТошДЎТАУ, 2017. –Б. 57-60.

57. Abdurakhmova N., Axmedova X. Algorithm based on linguistic models in machine translation between English and Uzbek / TurkLang. 2017. Vth international conference on Computer processing of Turkic languages. – Kazan, 2017. – P. 162-170. (certificate)

58. Абдурахмонова Н., Абдуллаева О. Корпус лингвистикасига оид тадқиқотлар / Ўзбек тилшунослигининг долзарб масалари. Республика илмий анжуман материаллари. –Тошкент: ТошДЎТАУ, 2017. –Б. 175-180.

59. Мухамедова С., Абдурахмонова Н. Ўзбек компьютер лингвистикасининг бугуни ва истиқболи / Ўзбек тилшунослигининг долзарб масалари. Республика илмий анжуман материаллари. –Тошкент: ТошДЎТАУ, 2017. –Б. 38-41.

60. Абдурахмонова Н. Моделирование грамматических категорий и аналитических глаголов узбекского языка как этап морфологического анализа в машинном переводе / TurkLang. 2017. Vth international conference on Computer processing of Turkic languages. – Kazan, 2017. – P. 156-180.

61. Abdurakhmonova N. Homonymy in machine translation / Modern trends in teaching foreign languages: achievements, challenges, and solutions International conference, –Tashkent: USWLU, 2017. – P. 17.

62. Abdurakhmonova N. Problems and perspectives of computational linguistics in Uzbekistan / The great steppe. International Turkic academy. Papers II nd forum of social sciences. – Astana, 2017. – B. 406-409.

Автореферат «Тил ва адабиёт таълими» журнали таҳририятида таҳрирдан ўтказилиб, ўзбек, рус, инглиз (резюме) тилларидаги матнлар ўзаро мувофиқлаштирилди (02.03.2018 йил).

Босишга рухсат этилди: 03.03.2018 йил.
Бичими 60x44 ¹/₁₆, «Times New Roman»
гарнитурда рақамли босма усулида босилди.
Шартли босма табағи 3,1. Адади: 100. Буюртма: № 70.

Ўзбекистон Республикаси ИИВ Академияси,
100197, Тошкент, Интизор кўчаси, 68.

«АКАДЕМИЯ НОШИРЛИК МАРКАЗИ»
Давлат унитар корхонасида чоп этилди.