

**TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI  
HUZURIDAGI ILMIY DARAJALAR BERUVCHI  
DSc.13/30.12.2019.T.07.01 RAQAMLI ILMIY KENGASH**

---

**MUHAMMAD AL-XORAZMIY NOMIDAGI  
TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI**

**UZAKOVA MA'MURA ABDURAYIMOVNA**

**TURKIY TILLAR OILASIGA MANSUB UCH TILLI ELEKTRON  
TARJIMA UCHUN LINGVO-MATEMATIK MODELLAR VA  
HISOBLASH ALGORITMLARI**

05.01.07 – Matematik modellashtirish. Sonli usullar va  
dasturlar majmui

**TEXNIKA FANLARI BO'YICHA FALSAFA DOKTORI (PHD)  
DISSERTATSIYASI AVTOREFERATI**

**Toshkent – 2025**

**Texnika fanlari bo'yicha falsafa doktori (PhD) dissertatsiyasi  
avtoreferati mundarijasi**

**Оглавление автореферата диссертации доктора философии (PhD)  
по техническим наукам**

**Contents of dissertation abstract of the doctor of philosophy (PhD)  
on technical sciences**

**Uzakova Ma'mura Abdurayimovna**

Turkiy tillar oilasiga mansub uch tilli elektron tarjima uchun lingvo-  
matematik modellar va hisoblash algoritmlari..... 3

**Узакова Мамура Абдурайимовна**

Лингво-математические модели и вычислительные алгоритмы для  
трёхязычного электронного перевода относящихся к тюркским  
семействам..... 25

**Uzakova Mamura Abdurayimovna**

Linguistic-mathematical models and computational algorithms for electronic  
translation related to the Turkic languages family..... 50

**E'lon qilingan ishlar ro'yxati**

**Список опубликованных работ**

List of published works..... 52

**TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI  
HUZURIDAGI ILMIY DARAJALAR BERUVCHI  
DSc.13/30.12.2019.T.07.01 RAQAMLI ILMIY KENGASH**

---

**MUHAMMAD AL-XORAZMIY NOMIDAGI  
TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI**

**UZAKOVA MA'MURA ABDURAYIMOVNA**

**TURKIY TILLAR OILASIGA MANSUB UCH TILLI ELEKTRON TARJIMA  
UCHUN LINGVO-MATEMATIK MODELLAR VA HISOBLASH  
ALGORITMLARI**

05.01.07 – Matematik modellashtirish. Sonli usullar va  
dasturlar majmui

**Texnika fanlari bo'yicha falsafa doktori (PHD) dissertatsiyasi  
AVTOREFERATI**

**Toshkent – 2025**

Texnika fanlari bo'yicha falsafa doktori (PhD) dissertatsiyasi mavzusi O'zbekiston Respublikasi Oliy ta'lim, fan va innovatsiyalar vazirligi huzuridagi Oliy attestatsiya komissiyasida № B2025.1.PhD/T5300 raqam bilan ro'yxatga olingan.

Dissertatsiya Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universitetida bajarilgan.

Dissertatsiya avtoreferati uch tilda (o'zbek, rus, ingliz (rezyume)) Ilmiy kengashning veb sahifasida (www.tuit.uz) va "Ziyonet" Axborot ta'lim portalida (www.ziyonet.uz) joylashtirilgan.

**Ilmiy rahbar:** Nazirova Elmira Shodmonovna  
texnika fanlari doktori, professor

**Rasmiy opponentlar:** Muhammadiyeva Dildora Kabilovna  
texnika fanlari doktori, dotsent

Bakayev Ilhom Izatovich  
texnika fanlari falsafa doktori, dotsent

**Yetakchi tashkilot:** Qarshi davlat universiteti

Dissertatsiya himoyasi Toshkent axborot texnologiyalari universiteti huzuridagi DSc.13/30.12.2019.T.07.01 raqamli Ilmiy kengashning 2025-yil "30" 10 soat 16<sup>00</sup> dagi majlisida bo'lib o'tadi. (Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 238-64-43, e-mail: iktuit@tuit.uz).

Dissertatsiya bilan Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti Axborot-resurs markazida tanishish mumkin (№ ~~377~~ raqam bilan ro'yxatga olingan).  
Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 238-64-70.

Dissertatsiya avtoreferati 2025-yil "18" 10 kuni tarqatildi.  
(2025-yil "18" 10 dagi 31 raqamli ryestr bayonnomasi.)



*M. Musayev*

**M.M. Musayev**  
Ilmiy darajalar beruvchi  
Ilmiy kengash raisi,  
texnika fanlari doktori, professor

*J.X. Djumanov*

**J.X. Djumanov**  
Ilmiy darajalar beruvchi  
Ilmiy kengash ilmiy kotibi,  
texnika fanlari doktori, professor

*F.M. Nuraliyev*

**F.M. Nuraliyev**  
Ilmiy darajalar beruvchi  
Ilmiy kengash qoshidagi ilmiy seminar raisi,  
texnika fanlari doktori, professor

## **KIRISH (falsafa doktori (PhD) dissertatsiyasi annotatsiyasi)**

**Dissertatsiya mavzusining dolzarbligi va zarurati.** Zamonaviy dunyoda ko‘p tilli elektron tarjimonlariga ehtiyoj doimiy ravishda o‘shib bormoqda. Bunda ehtiyojlarning asosiy sabablari va jihatlaridan biri kompaniyalar xalqaro bozorlarda faoliyat yuritadi va turli mamlakatlardagi mijozlar, hamkorlar va yetkazib beruvchilar bilan tez va samarali muloqot qilishlari kerak. Ko‘p tilli veb-saytlar, marketing materiallari, ko‘rsatmalar, qonuniy hujjatlar to‘g‘ri va tezkor tarjimoni talab qiladi, bu esa shundan dalolat beradiki ushbu tadqiq qilingan ish hozirgi kunda dolzarb ahamiyat kasb etadi. Ko‘p tilli mashina tarjimonlariga bo‘lgan ehtiyoj nafaqat texnik talab, balki global axborot almashinuvi sharoitida strategik zaruratdir. Bunday tizimlar axborotga kirishni kengaytirish, xarajatlarni kamaytirish, aloqa samaradorligini oshirish va til to‘siqlarini bartaraf etishga yordam beradi. Elektron tarjimonlar yaratish, tabiiy tillarni matematik modellarini takomillashtirish asosida kompyuterlashtirish, omma orasida mukammal bo‘lgan axborot tizimlarini taqdim qilish murakkab masalalardan hisoblanadi. Jahonda yetakchi bo‘lgan AQSh, Buyuk Britaniya, Yaponiya, Xitoy, Rossiya, Kanada davlatlarda kompyuter lingvistikasi ustida bir qancha ishlar olib borilmoqda. Ushbu ishlarni esa ilmiy va amaliy jihatdan tadqiq qilish muhim ahamiyat kasb etadi. Bu mazkur tadqiqot ishning yanada dolzarbligini ko‘rsatadi.

Jahonda klassik nazariyalar bazasi, hamda matematik va funksional tahlilning zarur bo‘limlarini rasmiylashtirish, hisoblash matematikasiga asoslangan zamonaviy dasturlash texnologiyalari asosida teoremlarni isbotlash uchun algoritmik usullarni ishlab chiqish algoritmlashtirishning dolzarb muammolaridan biri hisoblanadi. Savol-javob tizimlarining algoritmik asoslari tabiiy tilda ma‘lumotlarni manipulyatsiya qilish bilan qurilgan. Bu shuningdek, bir tildan ikkinchisiga o‘tish bilan bog‘liq algoritmik dasturlash muammosi va aqlli tizimlarni yaratish uchun algoritmik asoslarni qurish va boshqalarni o‘z ichiga oladi. Algoritmik banklarda axborot qismlarini to‘ldirish va ularning operatsion qismlari uchun dasturiy ta‘minotni ishlab chiqish muhim ahamiyat kasb etmoqda. Turkiy tillar oilasiga mansub o‘zbek tilining kompyuterli lingvistikasini yaratish va elektron tarjimonlar faoliyatini optimallashtirish dolzarb vazifadir. Bu bizga o‘zbek tilini kompyuterga yo‘naltirilgan modellarini yaratishga imkon beradi, ya‘ni o‘zbek tilini o‘rganish, bilimlarni baholash, o‘zbek tilidagi matnlarni sinab ko‘rish, kompyuterga matnlarni ovozli kiritish, matnlarni tahrirlash masalasiga alohida e‘tibor qaratilmoqda.

O‘zbekiston Respublikasi Prezidentining 2019 yil 21 oktyabrdagi PF-5850-son “O‘zbek tilining davlat tili sifatidagi nufuzi va mavqei tubdan oshirish chora-tadbirlari to‘g‘risida”gi Farmonida “...davlat tilining axborot va kommunikatsiya texnologiyalari, xususan, Internet jahon axborot tarmog‘ida munosib o‘rin egallashini ta‘minlash, o‘zbek tilining kompyuter dasturlarini yaratish”<sup>1</sup> vazifalari belgilangan hamda mazkur faoliyatlarga tegishli me‘yoriy-huquqiy hujjatlarda belgilangan vazifalarni amalga oshirishda ushbu dissertatsiya tadqiqoti muayyan darajada xizmat qiladi. 2020 - 2030-yillarda o‘zbek tilini rivojlantirish va til siyosatini

---

<sup>1</sup> O‘zbekiston Respublikasi Prezidentining Farmoni, Mamlakatimizda o‘zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish chora-tadbirlari. <https://lex.uz/docs/-5058351>

takomillashtirish konsepsiyasi ishlab chiqildi. Bu vazifalarni amalga oshirishda o‘zbek lingvistikasining kompyuterga yo‘naltirilgan modellarini yaratish, hozirgacha o‘zbek kompyuter lingvistikasida mashina tarjimasiga oid tadqiqotlarning yetarli emasligi sababli o‘zbek tilidan xorijiy tilga va aksincha tarjima dasturi uchun o‘zbek tilining lingvistik ma’lumotlar bazasini yaratish zarur va bulardan amalda foydalanish uchun algoritmlar va dasturiy vositalarni yaratish muhim masalalardan biri hisoblanadi.

**Tadqiqotning respublika fan va texnologiyalar rivojlanishining ustuvor yo‘nalishlariga mosligi.** Mazkur tadqiqot Respublika fan va texnologiyalar rivojlanishining IV. “Axborotlashtirish va axborot-kommunikatsiya texnologiyalarini rivojlantirish” ustuvor yo‘nalishi doirasida bajarilgan.

**Muammoning o‘rganilganlik darajasi.** Jahonda elektron tarjimonlarni model va algoritmlarini ishlab chiqish bo‘yicha juda ko‘p tadqiqotlar olib borilgan bo‘lib, jumladan, elektron tarjimon asoschisi U.Uiver, R.Richens, A.A.Lyapunov, G.G.Belonogonov, V.G.Vasilev, L.A.Grashenko, Yu.N.Marchukva boshqalarning bu boradagi ishlari qiyosiy o‘rganib chiqildi.

Chet el olimlaridan Christopher Manning ishlari asosan sintaktik tahlil, ma’lumotni kiritish va ko‘p tilli tarjima tizimlarini takomillashtirishga qaratilgan bo‘lib shu jumladan, n-gram modellar, Hidden Markov Models (HMMs), part-of-speech tagging va syntactic parsing kabi masalalar yoritilgan. Bulardan tashqari Danqi Chen, Regina Barzilay, David Sontag, Djoshua Bendjio, Yan Lekun, Martin Djoens kabi mashhur olimlar semantik tushunish, ko‘p tilli tarjima ustida bir qancha ishlar olib bormoqdalar.

Respublikamizda o‘zbek elektron tarjimonlarni yaratish va ular ustida ish olib borish bo‘yicha A.Q.Po‘latov, M.X.Xakimov Sh.A. Nazirov, M. Aripov ko‘p tilli mashina tarjimasini muhitidagi eng dolzarb muammolar, elektron tarjimada rasmiy grammatikaning o‘rni, rasmiy grammatikadan algoritmlarga o‘tishda zarur bo‘lgan mexanizm sifatida tabiiy tillarni matematik modellashtirish masalasi va bu masalani hal qilish uchun yaratilgan so‘z turkumini hosil qilishning matematik modellari ustida ish borganlar. Yosh olimlarimizdan N.Abduraxmanova, S.Shukurov, X.Tursunov, B.Karshiyev, M.Komilovlar o‘zbek tili korpusi, parallel korpuslarning yaratish va boshqa kompyuter lingvistikasining rivojlanishi uchun bir qancha ishlar olib bormoqdalar.

Bu sohadagi tadqiqotlar tahlili shuni ko‘rsatadiki, jahon miqyosida elektron tarjimonlarda aynan turkiy tillar oilasiga mansub tillarga oid resurslar juda ham kam, ko‘p tilli tarjimonlarning ish faoliyatini tahlil qilinganda matnlar tarjima qilinganda aniq va sifatli chiqishi past ko‘rsatgichlarni bermoqda. Shundan kelib chiqadiki elektron tarjimonlarni yaratish masalasi dolzarb bo‘lib kelmoqda.

**Dissertatsiya tadqiqotining dissertatsiya bajarilgan oliy ta’lim yoki ilmiy-tadqiqot muassasasining ilmiy-tadqiqot ishlari rejalari bilan bog‘liqligi.** Dissertatsiya ishi Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti ilmiy-tadqiqot ishlari rejasining № IL-77-8023102181 – sonli “Hududlar va iqtisodiyot tarmoqlari bo‘yicha yangi ilmiy ishlanmalar va texnologiyalarni qamrab olgan “Innovatsion g‘oyalar banki” ni yaratish (2023-2025) mavzularidagi loyihasi doirasida bajarilgan.

**Tadqiqotning maqsadi** uch tilli o‘zbek, qirg‘iz va qoraqalpoq tillari uchun elektron tarjimani amalga oshiruvchi lingvo-matematik model va algoritmlarni hamda dasturiy majmuani ishlab chiqish.

**Tadqiqotning vazifalari:**

uch tilli elektron tarjimon dasturlarni tahlil qilish va mantiqiy-lingvistik modellar asosida turkiy tillar (o‘zbek, qirg‘iz va qoraqalpoq tillari) elektron lug‘at ma’lumotlar bazasini yaratish;

o‘zbek, qoraqalpoq va qirg‘iz tillarini leksik, morfologik, sintaktik tahlil qilish, murakkab so‘z birikmalarini hamda gap tuzilmasini hosil qilishning lingvo-matematik modeli va algoritmlarini ishlab chiqish;

sodda va murakkab so‘zlar hamda gap tuzilmasining lingvo-matematik modellari asosida o‘zbek, qoraqalpoq va qirg‘iz tillari uchun fe‘l zamonlarini inobatga oluvchi sodda va murakkab so‘zlar hamda gap tuzilmasini hosil qiluvchi jarayonlarning lingvo-matematik modellari va algoritmlari ishlab chiqish;

o‘zbek tilidagi so‘zlarni hosil qilishda fonetik tahlil olib borish hamda ushbu jarayonlarning lingvo-matematik modeli hamda algoritmini ishlab chiqish;

elektron tarjimani baholovchi kombinatsiyalanagan metrikani va uning hisoblash algoritmi ishlab chiqish;

o‘zbek, qoraqalpoq va qirg‘iz tillari elektron tarjimoni dasturiy majmuasining funksional modul tuzilmasini ishlab chiqish;

uch tilli o‘zbek, qoraqalpoq va qirg‘iz matnlarini tarjimasini amalga oshiradigan dasturiy majmua va foydalanuvchi interfeysini yaratishdan iborat.

**Tadqiqotning obyekti** o‘zbek, qoraqalpoq va qirg‘iz tillaridagi lingvistik korpuslari.

**Tadqiqotning predmeti** turkiy tillar uchun uch tilli elektron tarjimani amalga oshiruvchichi lingvo-matematik modellar, algoritmlar va dasturiy majmuani yaratish hamda elektron tarjimani baholovchi kombinatsiyalangan metrika va uning hisoblash algoritmini yaratishdan iborat.

**Tadqiqotning usullari.** Tadqiqot mavzusini yoritishda lingvistik va lingvo matematik modellashtirish, hisoblash algoritmlarini modulli tahlil qilish usullaridan hamda elektron tarjimani baholashda gibrid usullardan, zamonaviy dasturlash texnologiyalaridan foydalanildi.

**Tadqiqotning ilmiy yangiligi** quyidagilardan iborat:

o‘zbek, qoraqalpoq va qirg‘iz tillarining leksik, morfologik bazasi asosida o‘zbek, qoraqalpoq va qirg‘iz tillari uchun murakkab so‘z birikmalarini hamda gap tuzilmasini hosil qilishning lingvo-matematik modellari va algoritmlari ishlab chiqilgan;

sodda va murakkab so‘zlar hamda gap tuzilmasini hosil qilishning lingvo-matematik modellari asosida o‘zbek, qoraqalpoq va qirg‘iz tillarining fonetik tuzilmasi uchun sodda so‘z yasalishi jarayonlarining lingvo-matematik modellari hamda algoritmlari ishlab chiqilgan;

sodda va murakkab so‘zlar hamda gap tuzilmasining lingvo-matematik modellari asosida o‘zbek, qoraqalpoq va qirg‘iz tillari uchun fe‘l zamonlarini inobatga oluvchi sodda va murakkab so‘zlar hamda gap tuzilmasini hosil qiluvchi jarayonlarning lingvo-matematik modellari va algoritmlari ishlab chiqilgan;

o‘zbek, qoraqalpoq va qirg‘iz tillari uchun uch tilli elektron lug‘at hamda elektron tarjimon uchun ma‘lumotlar bazasi asosida ilmiy-texnik sohadagi so‘zlarni qidirish va matnlarning o‘zaro tarjimasini aniqlash, o‘zbek, qoraqalpoq va qirg‘iz tillarida elektron tarjimoni baholovchi kombinatsiyalangan metrika va uning hisoblash algoritmi ishlab chiqilgan.

**Tadqiqotning amaliy natijalari** quyidagilardan iborat:

uch tilli elektron lug‘at hamda elektron tarjimon ya‘ni lingvo.uz nomli ma‘lumotlar bazasi ishlab chiqilgan;

uch tilli elektron tarjimonni amalga oshirish uchun o‘zbek, qoraqalpoq va qirg‘iz tillarida murakkab so‘zlar va sodda gaplarni yasashning lingvo-matematik modeli ishlab chiqilgan;

uch tilli elektron tarjimonni amalga oshirish uchun o‘zbek, qoraqalpoq va qirg‘iz tillarida fe‘l zamonlari ishtirokida so‘z yasashining lingvo-matematik modeli ishlab chiqilgan;

uch tilli elektron tarjimonni amalga oshirish uchun o‘zbek tilida fonetik tahlil natijalarining lingvo-matematik modeli va algoritmi ishlab chiqilgan;

uch tilli elektron tarjimoni baholashda kombinatsiyalangan metrika uchun hisoblash algoritmi ishlab chiqilgan;

dasturiy majmuaning ishlash tamoyili va zaruriy talablar ishlab chiqilgan.

**Tadqiqot natijalarining ishonchliligi** tadqiqot natijalarining ishonchliligi uchta turkiy tillar orasida sodda gaplarni tarjima qilish maxsus dasturiy to‘plamdan foydalangan holda o‘tkazilgan tajribalar natijalarining ishonchliligini isbotlashga asoslangan yakuniy xulosalar bilan izohlanadi.

**Tadqiqot natijalarining ilmiy va amaliy ahamiyati.** Tadqiqot natijalarining ilmiy ahamiyati uch tilli turkiy tillar o‘zbek, qoraqalpoq va qirg‘iz tillarida tarjima jarayonini amalga oshirishda sodda va murakkab so‘zlarni hamda gaplarni tuzishning lingvo-matematik model va algoritmlarini, elektron tarjimoni baholashning kombinatsiyalangan metrika va uning hisoblash algoritmi yaratilishi bilan izohlanadi. Tadqiqot natijalarining amaliy ahamiyati shundan iboratki, yaratilgan lingvo-matematik model va algoritmlar sodda va murakkab so‘zlarni hamda gaplarni tuzishda ishlatiladigan tegishli dasturiy majmuani ishlab chiqish bilan izohlanadi.

**Tadqiqot natijalarining joriy qilinishi.** Dissertatsiya doirasida olib borilgan tadqiqotlar natijasida ishlab chiqilgan lingvo-matematik modellar, algoritmlar va yaratilgan dasturiy majmua asosida:

o‘zbek tilidagi so‘zlarni morfologik tahlil qiluvchi lingvo-matematik modeli va algoritmlari asosida yaratilgan dasturiy majmua “The computer workshop” MCHJ tahririyatidagi nashr materiallari va turli hujjatlar matnida morfem so‘zlarni avtomatik qidirish bo‘yicha kompyuter axborot texnologiyalaridan samarali foydalanish uchun joriy qilindi (O‘zbekiston Respublikasi Raqamli texnologiyalar vazirligi 07.07.2025 yil 33-8/4721-sonli ma‘lumotnomasi). Natijada nashriyot materiallari matnidagi omonim, sinonim, paronim so‘zlarni tahrir qilish, qidirish jarayonlari umumiy ish faoliyatining vaqtga nisbatan hajmini 1-2 barobarga qisqartirishga imkon berdi;

o‘zbek, qoraqalpoq va qirg‘iz tillarida murakkab so‘zlarni yasash algoritmi va lingvo-matematik modeli asosida ishlab chiqilgan dasturiy majmua “Big Makro

World” MChJ tahririyatidagi nashr materiallari va turli hujjatlarni, matnlarni tarjima qilishda yuzaga keladigan kamchiliklarni bartraf etishda joriy qilingan (O‘zbekiston Respublikasi Raqamli texnologiyalar vazirligi 07.07.2025 yil 33-8/4721-sonli ma‘lumotnomasi). Natijada taklif qilingan algoritmlar asosida ishlab chiqilgan dasturiy majmua “Big Makro World” tarjimalar byurosida elektron tarjima vaqtini 1-2 barobarga oshishiga imkon berdi;

o‘zbek, qoraqalpoq va qirg‘iz tillarida sodda va murakkab so‘zlar va gaplar hosil qilish, fonetik tahlil, gapda fe‘l zamonlari tahlillarining lingvo-matematik modellari va hisoblash algoritmi asosida yaratilgan dasturiy majmua “SHARQ” nashriyotida ilmiy texnik matnlarni yuqori sifat va kontekstual aniqlik bo‘yicha elektron tarjimini amalga oshirishga joriy qilingan (O‘zbekiston Respublikasi Raqamli texnologiyalar vazirligi 07.07.2025 yil 33-8/4721-sonli ma‘lumotnomasi). Natijada ilmiy-texnik matnlardagi sodda va murakkab so‘zlar va gaplarni hosil qilish, fonetik tahlil, gapda fe‘l zamonlarining tahlillari, so‘zlarni qidirish va matnlardagi kontekstual aniqlik bo‘yicha 78,2 % ga erishildi.

**Tadqiqot natijalarining aprobatsiyasi.** Mazkur tadqiqot natijalari 2 ta xalqaro va 9 ta respublika ilmiy-amaliy anjumanlarida muhokamadan o‘tkazilgan.

**Tadqiqot natijalarining e‘lon qilinganligi.** Tadqiqot mavzusi bo‘yicha jami 20 ta ilmiy ish chop etilgan, shulardan, O‘zbekiston Respublikasi Oliy Attestatsiya Komissiyasining doktorlik dissertatsiyalari asosiy ilmiy natijalarini chop etishda tavsiya etilgan ilmiy nashrlarda ilmiy maqola 4 ta, shulardan 1 ta xorijiy va 3 ta respublika jurnallarida nashr qilingan, 1 ta monografiya, scopus bazasida indekslangan konferensiyada 2 ta tezis, respublika ilmiy-amaliy anjumanlarda 8 ta ilmiy ishlar, shuningdek EHM uchun dasturiy mahsulot qayd etilganligi haqida 3 ta guvohnoma olingan.

**Dissertatsiyaning tuzilishi va hajmi.** Dissertatsiya ishi kirish, to‘rtta bob, xulosa, foydalanilgan adabiyotlar ro‘yxati va ilovalardan iborat bo‘lib, dissertatsiya hajmi 110 sahifadan tashkil etadi.

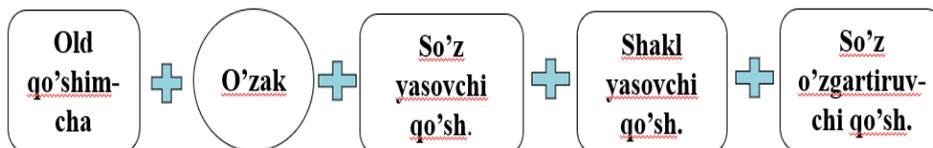
## DISSERTATSIYANING ASOSIY MAZMUNI

**Kirish** qismida dissertatsiya mavzusining dolzarbligi va zarurati asoslangan, tadqiqotning maqsadi va vazifalari, obyekt va predmeti tavsiflangan, respublika fan va texnologiyalari rivojlanishining ustuvor yo‘nalishlariga mosligi ko‘rsatilgan, tadqiqotning ilmiy yangiligi va amaliy natijalari bayon qilingan, olingan natijalarning ishonchligi, ilmiy va amaliy ahamiyati ochib berilgan, tadqiqot natijalarini amaliyotga joriy qilish, ishning aprobatsiyasi, nashr etilgan ishlar va dissertatsiya tuzilishi bo‘yicha ma‘lumotlar keltirilgan.

Dissertatsiya ishining “**Elektron tarjimonlarning rivojlanishi va ularning hozirgi holati**” deb nomlangan birinchi bobida tadqiqot usullari tahlili va dissertatsiya mavzularga oid mavjud ishlar tahlili, hamda hozirda mavjud bo‘lgan xalqaro va respublikamizdagi elektron tarjimonlarning qiyosiy tahlili, ahamiyati, roli va muammolari haqida fikrlar bayon etilgan. Shu bilan birga elektron tarjimonlarning rivojlanish bosqichlari, muamoni yechishning asosiy yondashuvlari keltirilgan. Turkiy tillar oilasiga mansub o‘zbek, qoraqalpoq va qirg‘iz tillarining kelib chiqish tavsifi, sxemasi ko‘rib chiqilgan.

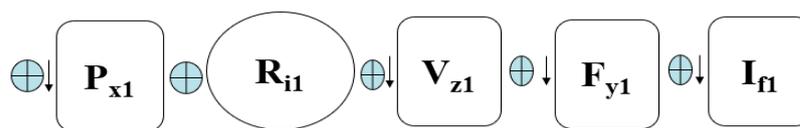
Dissertatsiyaning “Turkiy tillar oilasiga mansub uch tilli elektron tarjima uchun sodda va murakkab soʻzlarni hosil qilishning lingvo-matematik modellari va algoritmlari” deb nomlangan ikkinchi bobida oʻzbek, qoraqalpoq va qirgʻiz tillarida sodda va murakkab soʻzlarni hosil qilishning lingvo-matematik modellari va algoritmlari keltirilgan.

Turkiy tillar oilasi agglyutinativ tillar guruhiga kirganligi sababli soʻz yasaliishi jarayoni ketma-ket va morfemalarning chegarasi aniq koʻrinib turishi bilan boshqa til guruhlaridan ajralib turadi. Buni quyidagicha ifodalaymiz:



1-rasm. Oʻzbek tilida soʻz yasaliishining mantiqiy-lingvistik modeli

Ushbu rasmda oʻzbek tilida soʻz yasaliishining ketma-ketligi ifodalangan boʻlib, oʻzakka old va ort qoʻshimchalar qoʻshilishi asosida soʻz hosil boʻlishining lingvistik koʻrinishi ifodalangan. 1-rasmda tabiiy tilda soʻz yasaliishini formallashtirib elektron tarjimaning birinchi qadamini ifodalaymiz:



2-rasm. Oʻzbek tilida soʻz yasaliish sxemasining formal modeli

Bu yerda 2-rasmga aniqliklar kiritamiz:

$P_x$  – old qoʻshimcha,  $R_i$  - soʻz asosi,  $V_z$  – soʻz yasovchi qoʻshimcha,

$F_y$  – shakl yasovchi qoʻshimcha,  $I_f$  – soʻz oʻzgartiruvchi qoʻshimcha.

$\oplus$  - birikish belgisi,  $\downarrow$  - kelishi yoki kelmasligi mumkin boʻlgan belgi.

Formal modelni koʻrinishda ifodalaydigan boʻlsak, quyidagini olamiz:

$$W_{uzb} = \oplus \downarrow P_x \oplus R_i \oplus \downarrow V_z \oplus \downarrow F_y \oplus \downarrow I_f \quad (1)$$

Bu yerda,

$W_{uzb}$ - soʻz,  $P_{x1}$ - old qoʻshimcha,

$V_{z1}$ - soʻz yasovchi qoʻshimcha,

$F_{y1}$  - shakl yasovchi qoʻshimcha,

$I_{f1}$ - soʻz oʻzgartiruvchi qoʻshimcha

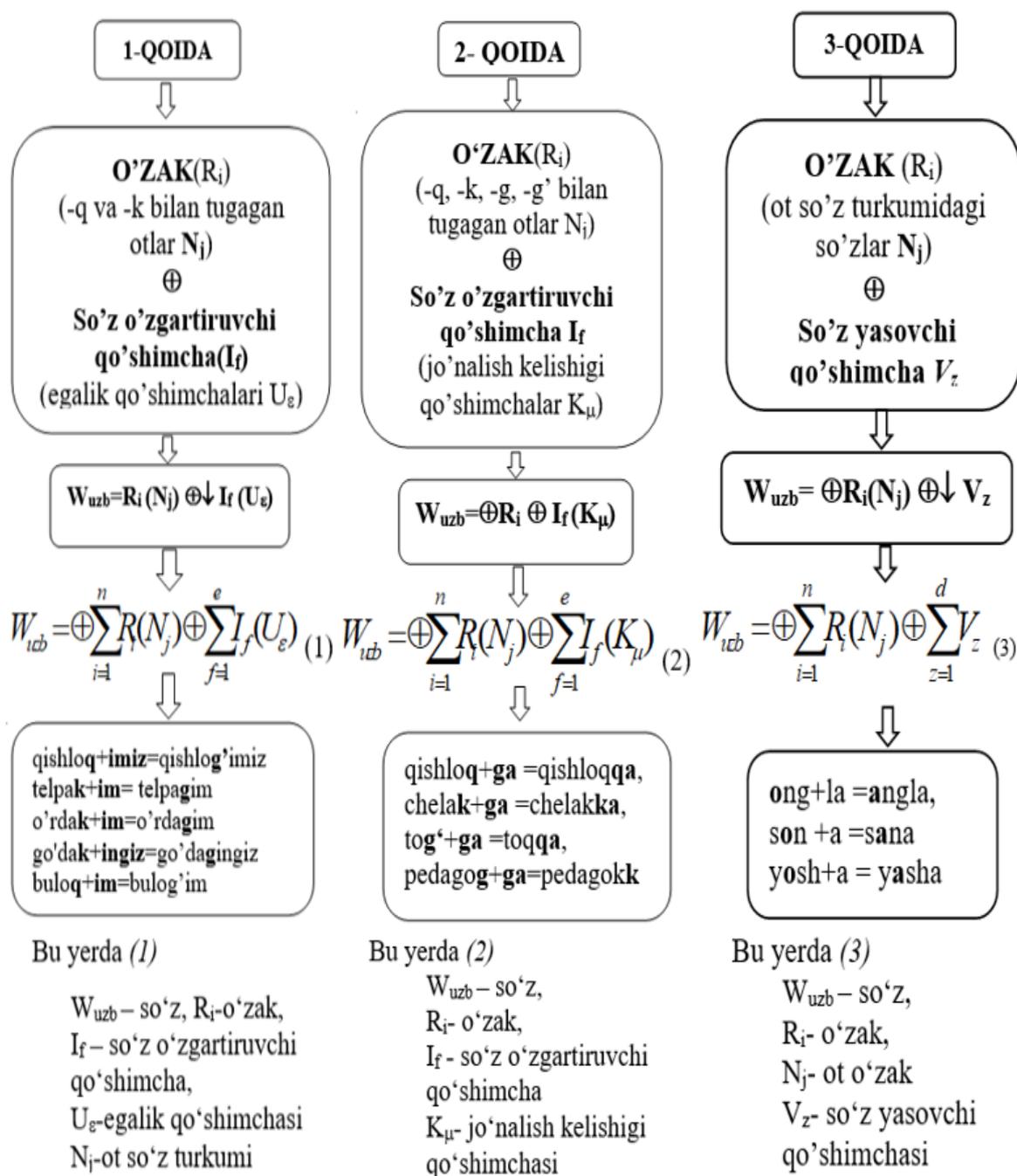
(1) Ifodani yoyib chiqadigan boʻlsak, quyidagini olamiz:

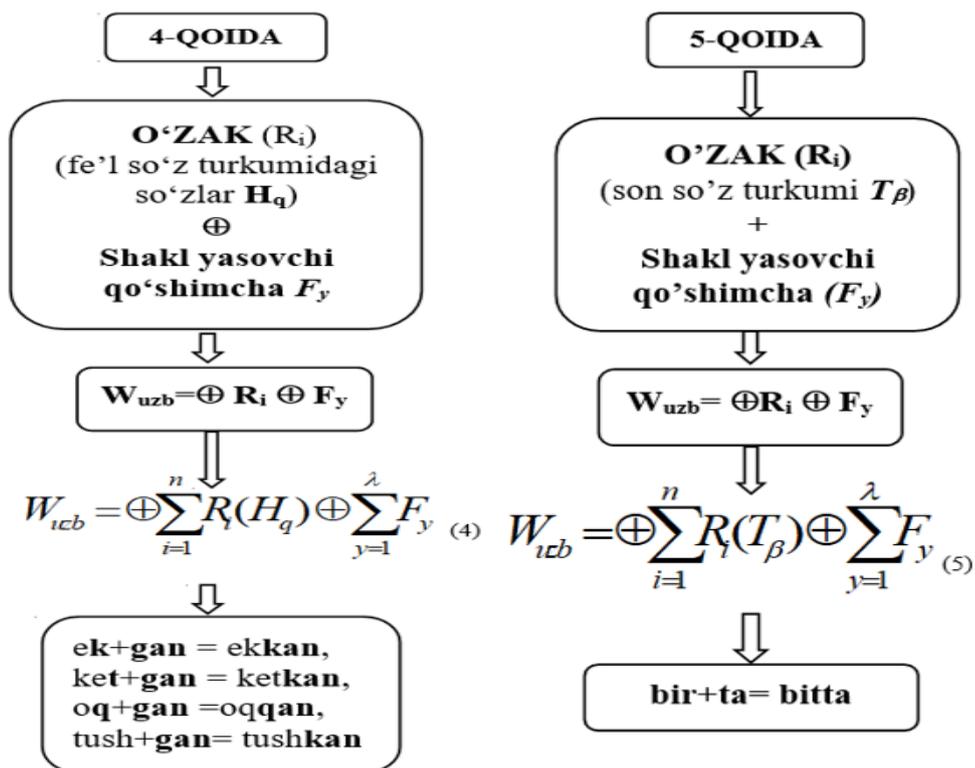
$$W_{uzb} = \oplus \downarrow \sum_{x_1=1}^m P_{x_1} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{z_1=1}^d V_{z_1} \oplus \downarrow \sum_{y_1=1}^v F_{y_1} \oplus \downarrow \sum_{f_1=1}^e I_{f_1} \quad (2)$$

Bu yerda,

$W_{uzb}$  –murakkab soʻz,  $P_{x1}$ -old qoʻshimcha,  $R_{i1}$ -soʻzning oʻzagi,  $V_{z1}$ -soʻz yasovchi qoʻshimcha,  $F_{1y}$ -shakl yasovchi qoʻshimcha,  $I_{f1}$ -soʻz oʻzgartiruvchi qoʻshimcha, i, m, z, d, y, v, f, e - oʻzak va qoʻshimchalar uzunligi,  $\oplus$  birikish belgisi,  $\downarrow$ -kelishi yoki kelmasligi mumkin boʻlgan belgi.

Oʻzbek tilida soʻz yasash jarayonini oʻrganib chiqib, soʻz yasash jarayoni murakkab jaryon boʻlganligini va elektron tarjimaning sifatli chiqishi uchun bu jarayonni chuqur oʻrganib fonetik tahlil bosqichini alohida formallashtirdik. Oʻzbek tilida soʻz yasashida fonetik tahlil asosan 5 ta qoidaga binoan yasashini koʻrib chiqamiz (misollar bilan):

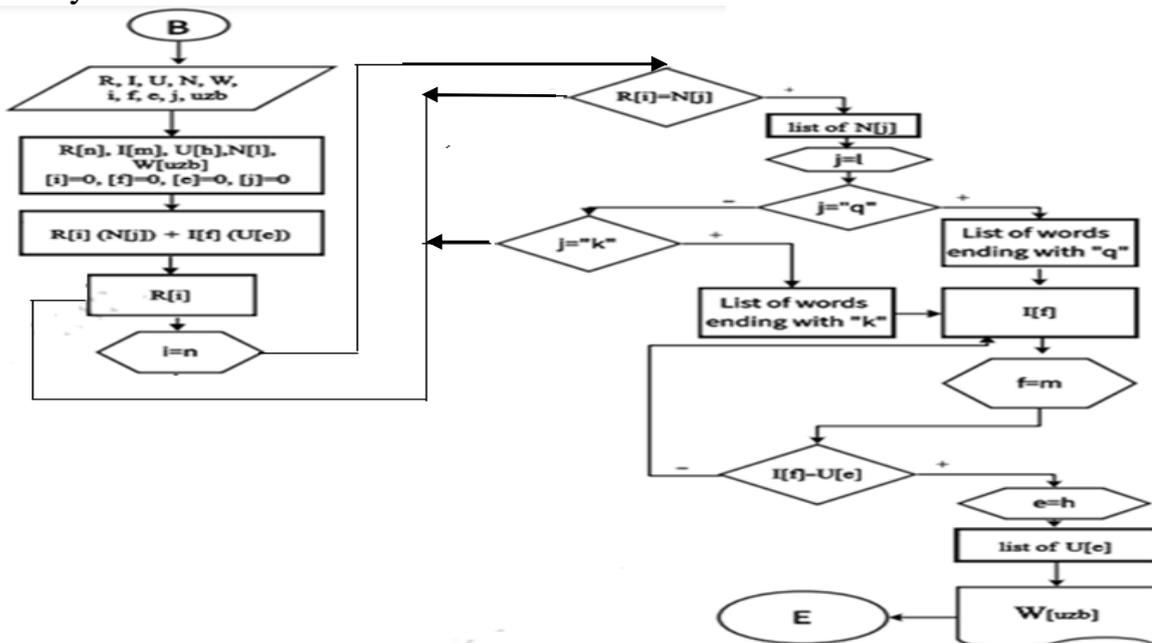




Bu yerda (4)  
W<sub>uzb</sub> – soʻz,  
R<sub>i</sub> – oʻzak,  
H<sub>q</sub> – feʼl oʻzak  
F<sub>y</sub> – shakl yasovchi qoʻshimchasi

Bu yerda (5)  
W<sub>uzb</sub> – soʻz,  
R<sub>i</sub> – oʻzak,  
T<sub>β</sub> – son oʻzak  
F<sub>y</sub> – shakl yasovchi qoʻshimchasi

Oʻzbek tilida koʻp boʻgʻinli soʻzlarda fonetik tahlilning lingvo-matematik modellari (1), (2), (3), (4) va (5) ifodalarda aks ettirilgan. Ishlab chiqilgan lingvo-matematik modellarni umumlashtirib quyidagi blok-sxemasini ifodalaymiz:



3- rasm. Oʻzbek tilida soʻz yasalishida fonetik tahlili blok-sxemasi

Yuqoridagilarni inobatga olgan holda elektron tarjimini amalga oshirishda aniq va sifatli natijalarga erishish o‘z natijalarini ko‘rsatadi.

Tadqiqot ishimizda qaralayotgan turkiy tillar oilasiga mansub qoraqalpoq va qirg‘iz tillarida so‘z yasashining lingvo-matematik modellarini ham ishlab chiqamiz. Bunda (1) ifodaga mos ravishda (8) va (9) larni qoraqalpoq tili uchun ishlab chiqamiz.

$$W_{kk} = \oplus \downarrow P_{x_2} \oplus \downarrow R_{i_2} \oplus \downarrow V_{z_2} \oplus \downarrow F_{y_2} \oplus \downarrow I_{f_2} \quad (8)$$

$$W_{kk} = \oplus \downarrow \sum_{x_2=1}^j P_{x_2} \oplus \downarrow \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{z_2=1}^d V_{z_2} \oplus \downarrow \sum_{y_2=1}^v F_{y_2} \oplus \downarrow \sum_{f_2=1}^e I_{f_2} \quad (9)$$

Bu yerda yuqoridagilar uchun aniqliklar kiritamiz:

$W_{kk}$ -qoraqalpoqcha so‘z,  $R_{i_2}$ – so‘z o‘zagi,  $V_{z_2}$ – so‘z yasovchi qo‘shimcha,  $F_{y_2}$ – shakl yasovchi qo‘shimcha,  $I_{f_2}$ – so‘z o‘zgartiruvchi qo‘shimcha,  $\oplus$  - bog‘lash belgisi,  $\downarrow$  - kelishi mumkin yoki kelmasligi mumkin,  $x, i, z, y, f$  – morfemalar uzunligi. Bu yerda indeksdagi 2 ikkinchi til deb belgilandi, ya’ni 2 -til qoraqalpoq tili.

Qirg‘iz tili uchun ham (1) ifodani mos ravishda quyidagicha ishlab chiqamiz:

$$W_{kg} = \oplus \downarrow P_{x_3} \oplus \downarrow R_{i_3} \oplus \downarrow V_{z_3} \oplus \downarrow F_{y_3} \oplus \downarrow I_{f_3} \quad (10)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=1}^m P_{x_3} \oplus \downarrow \sum_{i_3=1}^n R_{i_3} \oplus \downarrow \sum_{z_3=1}^d V_{z_3} \oplus \downarrow \sum_{y_3=1}^v F_{y_3} \oplus \downarrow \sum_{f_3=1}^e I_{f_3} \quad (11)$$

Bu yerda yuqoridagilar uchun aniqliklar kiritamiz:

$W_{kg}$ -qirg‘izcha so‘z,  $R_{i_3}$  – so‘z o‘zagi,  $V_{z_3}$ – so‘z yasovchi qo‘shimcha,  $F_{y_3}$  – shakl yasovchi qo‘shimcha,  $I_{f_3}$  – so‘z o‘zgartiruvchi qo‘shimcha.  $\oplus$  - bog‘lash belgisi,  $\downarrow$  - kelishi mumkin yoki kelmasligi mumkin,  $x, i, z, y, f$  – morfemalar uzunligi

Qirg‘iz tilida murakkab so‘z yasash strukturasi o‘ziga xos tomonlaridan biri bu tarjima vaqtida 1 va 2 shaxs egalik qo‘shimchalari so‘z oldiga chiqadi. Shu yerda aniqlik kiritish uchun ta’rifga yuzlanamiz:

**Ta’rif:** Agar qirg‘iz tilida murakkab so‘z tarjima qilinayotganda (masalan o‘zbekchadan qirg‘izchaga ( $W_{uzb} \rightarrow W_{kg}$ )) so‘z o‘zgartiruvchi qo‘shimcha ( $I_f$ ) egalik qo‘shimcha sifatida uchrasa ( $\tilde{I}_{f_2}$ ), ya’ni egalik qo‘shimchalarning aynan I va II shaxslaridagi qo‘shimchalari (*-im, -ing, -imiz, -ingiz*) kelsa, u holda qirg‘iz tili tarjimasida egalik qo‘shimchalari *menin, senin, anin, bizdin, siznin, alanin* egalik olmoshi ( $G_t$ )ga aylanib, alohida so‘z birikmasi sifatida so‘zning boshida keladi.

$$\tilde{W}_{kg} = \oplus G_{t_3} \oplus R_{i_3} \oplus \downarrow V_{z_3} \oplus \downarrow F_{y_3} \oplus \downarrow \tilde{I}_f \quad (12)$$

Bu yerda (12)

$\tilde{W}_{kg}$  - tarjimadan keyin qirg‘iz tilidagi murakkab so‘z,  $G_t$  – egalik olmoshi.

Masalan, O‘zbekchada: bog‘+bon+lar+imiz+ning

Qirg'izchada : ***bizdin*** bag+ban+dar+din

Qirg'iz tilida so'z yasaliş ketma-ketliklarni **sonli usulda (sistema ko'rinishida)** keltrib o'tilgan. Murakkab so'zlarni qismlarga bo'lib o'rganilganda qirg'iz tili uchun quyidagi holatlarga duch kelindi:

$$\begin{aligned}
 W_{kg1} &= \sum_{i_3=1}^{\xi} G_{i_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y_3=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f_3=1}^e I_{f_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg2} &= \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y_3=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg3} &= \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f_3=1}^e I_{f_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg4} &= \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y_3=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f_3=1}^e I_{f_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg5} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg6} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.
 \end{aligned}$$

$$W_{kg7} = \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_{f_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.$$

$$W_{kg8} = \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.$$

$$W_{kg9} = \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.$$

$$W_{kg10} = \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.$$

$$W_{kg11} = \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.$$

$$W_{kg12} = \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.$$

$$W_{kg13} = \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ bu yerda } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.$$

Bu yerda  $W_{kg1}, W_{kg2}, W_{kg3}, W_{kg4}, W_{kg5}, W_{kg6}, W_{kg7}, W_{kg8}, W_{kg9}, W_{kg10}, W_{kg11}, W_{kg12}, W_{kg13}$  – murakkab so‘z,  $N(\omega)$  – ixtiyoriy bitta predikat, har bir  $\omega$  so‘zi ikki yoki undan ortiq so‘z birikmasi sifatida ifodalanadi,  $l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1}$ , bu yerda  $k - 1$  affiks uzunligi.

Turkiy tillar oilasiga mansub o‘zbek, qoraqalpoq va qirg‘iz tillarida gap yasalişining lingvo-matematik modellari xususiylashtirib ishlab chiqildi va ushbu holatlarda fe‘l zamonlari kesimida ham lingvo-matematik modellar ishlab chiqildi:

O‘zbek, qoraqalpoq va qirg‘iz tillarida fe‘l zamonlari kesimida so‘z yasalişining lingvo-matematik modellari (o‘tgan zamon uchun)

$$W_{uzb} = \oplus \downarrow \sum_{x_1=1}^m P_{x_1} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{w_1=1}^{\tau} B_{w_1} \oplus \downarrow \sum_{c_1=1}^l L_{c_1} \oplus \downarrow \sum_{e_1=1}^o K_{e_1} \oplus \downarrow \sum_{d_1=1}^p O_{d_1} \quad (13)$$

$$W_{qq} = \oplus \downarrow \sum_{x_2=1}^m P_{x_2} \oplus \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{w_2=1}^{\tau} B_{w_2} \oplus \downarrow \sum_{c_2=1}^l L_{c_2} \oplus \downarrow \sum_{e_2=1}^o K_{e_2} \oplus \downarrow \sum_{d_2=1}^p O_{d_2} \quad (14)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=1}^m P_{x_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \downarrow \sum_{w_3=1}^{\tau} B_{w_3} \oplus \downarrow \sum_{c_3=1}^l L_{c_3} \oplus \downarrow \sum_{e_3=1}^o K_{e_3} \oplus \downarrow \sum_{d_3=1}^p O_{d_3} \quad (15)$$

Bu yerda  $P_x$ -old qo‘shimcha,  $R_i$ -o‘zak,  $B_w$ -o‘tgan zamondagi qo‘shimcha,  $L_c$ -1-shaxs-son qo‘shimchasi,  $K_e$ -2-shaxs-son qo‘shimchasi,  $O_d$ -3-shaxs-son qo‘shimchasi

O‘zbek, qoraqalpoq va qirg‘iz tillarida fe‘l zamonlari kesimida hozirgi zamon fe‘l zamonlari qo‘shimchalari asosida so‘z yasalişining lingvo-matematik modellari

$$W_{uzb} = \oplus \downarrow \sum_{x_1=1}^m P_{x_1} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{g_1=1}^b Q_{g_1} \oplus \downarrow \sum_{c_1=1}^l L_{c_1} \oplus \downarrow \sum_{e_1=1}^o K_{e_1} \oplus \downarrow \sum_{d_1=1}^p O_{d_1} \quad (16)$$

$$W_{qq} = \oplus \downarrow \sum_{x_2=1}^m P_{x_2} \oplus \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{g_2=1}^b Q_{g_2} \oplus \downarrow \sum_{c_2=1}^l L_{c_2} \oplus \downarrow \sum_{e_2=1}^o K_{e_2} \oplus \downarrow \sum_{d_2=1}^p O_{d_2} \quad (17)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=1}^m P_{x_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \downarrow \sum_{g_3=1}^b Q_{g_3} \oplus \downarrow \sum_{c_3=1}^l L_{c_3} \oplus \downarrow \sum_{e_3=1}^o K_{e_3} \oplus \downarrow \sum_{d_3=1}^p O_{d_3} \quad (18)$$

Bu yerda  $Q_g$ - hozirgi zamondagi fe‘l qo‘shimchasi

O'zbek, qoraqalpoq va qirg'iz tillarida fe'l zamonlari kesimida kelasi zamon qo'shimchalari asosida so'z yasalishining lingvo-matematik modellari:

$$W_{uzb} = \oplus \downarrow \sum_{x_2=1}^m P_{x_1} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{u_1=1}^s T_{u_1} \oplus \downarrow \sum_{c_1=1}^l L_{c_1} \oplus \downarrow \sum_{e_1=1}^o K_{e_1} \oplus \downarrow \sum_{d_1=1}^p O_{d_1} \quad (19)$$

$$W_{qq} = \oplus \downarrow \sum_{x_2=1}^m P_{x_2} \oplus \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{u_2=1}^s T_{u_2} \oplus \downarrow \sum_{c_2=1}^l L_{c_2} \oplus \downarrow \sum_{e_2=1}^o K_{e_2} \oplus \downarrow \sum_{d_2=1}^p O_{d_2} \quad (20)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=1}^m P_{x_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \downarrow \sum_{u_3=1}^s T_{u_3} \oplus \downarrow \sum_{c_3=1}^l L_{c_3} \oplus \downarrow \sum_{e_3=1}^o K_{e_3} \oplus \downarrow \sum_{d_3=1}^p O_{d_3} \quad (21)$$

*Bu yerda  $T_u$ - hozirgi zamondagi fe'l qo'shimchasi*

Yuqorida turkiy tillar oilasiga mansub o'zbek, qoraqalpoq va qirg'iz tillaridagi fe'l zamonlaridagi so'z yasalishining tahlili va bu tahlillar asosida lingvo-matematik modellar ishlab chiqildi.

Dissertatsiya ishining **“Turkiy tillar oilasiga mansub uch tilli elektron tarjima uchun gaplarni hosil qilishning lingvo-matematik modellari va algoritmlari”** deb nomlangan uchinchi bobida o'zbek, qoraqalpoq va qirg'iz tillarida gap yasalishining lingvistik tasnifi, formal modellariva lingvo-matematik modellari keltirilgan. O'zbek tilida so'z yasalish bosqichidan keyin gap yasalish bosqichiga o'tiladi. Gap yasalish bosqichida birinchi navbatda gap yasalishi uchun kerak bo'ladigan komponentalarni o'rganib chiqishimiz kerak. Gap yasalishining asosini ega va kesim tashkil qiladi. Bunda ega gapning egasi va kesim esa gapda ish harakatni bildiruvchi komponentalar hisoblanadi. Lingvistik nuqtai nazarda ega va kesimni gapning bosh bo'laklari deb ataladi. Bundan tashqari gapning ikkinchi darajali bo'laklari ham mavjud. Quyidagi 4-rasmda ushbu ma'lumotlarning sxema ko'rinishini ko'rishimiz mumkin:



4-rasm. O'zbek tilida gap asosini tashkil etuvchi bolaklar tasnifi lingvistik tomondan

O'zbek tilida so'z yasalishi bosqichidan keyin gap ( $S_{uzbsimple}$ ) yasalishiga o'tadigan bo'lsak, o'zbek tili grammatikasi juda boy ekanligiga amin bo'lamiz.

O'zbek tilida gaplar tuzilishiga ko'ra sodda gap va qo'shma gapga ajratiladi.

Sodda gap tuzilishiga ko'ra:

1. S o d d a y i g' i q g a p,

2. Sodda yoyiq gaplarga ajratiladi (4-rasm). Bunday tasniflashda ikkinchi darajali bo'laklarning ishtirok etish-etmasligi e'tiborga olinadi.



5-rasm. O‘zbek tilida sodda gaplarning lingvistik tomondan turlari

Faqat bosh bo‘laklar (ega va kesim)dan tuzilgan gap sodda yig‘iq gap deyiladi. Sodda yig‘iq gaplarga misollarni quyidagicha ko‘rishimiz mumkin:

$\underbrace{\text{Axborot}}_{\text{ega}} \underbrace{\text{keltirildi}}_{\text{kesim}}, \underbrace{\text{Kompyuter}}_{\text{ega}} \underbrace{\text{ishlayapti}}_{\text{kesim}}. \dots$

Ushbu misolga izoh kiritadigan bo‘lsak, o‘zbek tilida gap yasaliş strukturasi har doim ega birinchi navbatda, kesim esa egadan keyin gapning oxirida keladi. Bu tartib qat‘iy hisoblanadi.

Bosh va ikkinchi darajali bo‘laklarning ishtirokida tuzilgan gap sodda yoyiq gap deyiladi. Masalan:

$\underbrace{\text{Bu yilgi}}_{\text{aniqlovchi}} \underbrace{\text{seminar}}_{\text{ega}} \underbrace{\text{yanvar oyining boshida}}_{\text{hol}} \underbrace{\text{o‘tdi}}_{\text{kesim}}$

$\underbrace{\text{Jismoniy mashqlar}}_{\text{aniqlovchi}} \underbrace{\text{organizmni}}_{\text{ega}} \underbrace{\text{to‘ldiruvchi}}_{\text{aniqlovchi}} \underbrace{\text{har xil kasalliklardan}}_{\text{to‘ldiruvchi}} \underbrace{\text{saqlaydi}}_{\text{kesim}}$

O‘zbek tilida bosh va ikkinchi darajali bo‘laklar ishtirokida gapning tuzilishi sxemasi.

Yuqorida berilgan sxemalar asosida o‘zbek tilida sodda gaplarni hosil qilishning lingvo-matematik modellari ishlab chiqildi.

$$S_{uzbsimple1} = E_1(x_i, \dots, x_n) \cup K_1(x_i, \dots, x_n) \quad (22)$$

Bu yerda,  $S_{uzbsimple1}$ -sodda yig‘iq gap,  $E_1$ -ega,  $K_1$ -kesim,  $i$  va  $n$  – har bir predikatning  $x$  to‘plam ichida o‘zgarish qadamlari,  $\cup$  - birlashtirish belgisi.

Izoh,  $E_1 \in x, x = \{i \text{ dan } n \text{ gacha o‘zgaradi}\}$

$K_1 \in x, x = \{i \text{ dan } n \text{ gacha o‘zgaradi}\}$

Masalan:

*Axborot keltirildi (kel+tir+il+di).*

Ushbu gapni tahlil qilish davomida, birinchi bo‘lib gap bo‘laklar(tokenlar)ga ajratiladi va turi aniqlanadi, har bir tokenlar morfologik tahlil asosida morfemalarga ya‘ni o‘zak( $R_{i1}$ ) va suffixlarga ajratiladi. Bu yerda ega tub so‘z va morfemalarga ajralmaydi, kesim esa morfemalarga ajratiladi. Quyidagi lingvo-matematik model asosida bitta gap yasalişining misolini sonli usulda (sistema ko‘rinishida) keltirganmiz:

$$S_{uzbsimple1} = \begin{cases} E1(x_i, \dots, x_n) \cup \\ K1(x_i, \dots, x_n) \rightarrow R_{i1} \oplus F_{y1} \oplus F_{y1} \oplus F_{y1} \end{cases}$$

Bu yerda,

$S_{uzbsimple1}$ -sodda yig'iq gap,  $E1$ -ega,  $K1$ -kesim,  $i$  va  $n$  – har bir predikatning  $x$  to'plam ichida o'zgarish qadamlari,  $R_{i1}$ -o'zak,  $F_{y1}$ -shakl yasovchi qo'shimcha,  $\cup$  - birlashtirish belgisi.

Izoh,  $E1 \in x, x = \{i \text{ dan } n \text{ gacha o'zgaradi}\} i \geq 1, n \geq 1$

$K1 \in x, x = \{i \text{ dan } n \text{ gacha o'zgaradi}\} i \geq 1, n \geq 1$

(22) ifodadan ko'rinib turibdiki, sodda yig'iq gaplarda faqatgina ikkita bo'lak, ya'ni ega va kesim ishtirik etyapti.

Keyingi bosqichda sodda yoyiq gaplar uchun lingvo matematik model qurib chiqildi va quyidagicha ifodalandi:

$$S_{uzbsimple2} = E1(x_i, \dots, x_n) \cup \downarrow O1(x_i, \dots, x_n) \cup \downarrow H1(x_i, \dots, x_n) \cup \\ \cup \downarrow D1(x_i, \dots, x_n) \cup K1(x_i, \dots, x_n) \quad (23)$$

Bu yerda,  $S_{uzbsimple2}$ -sodda yoyiq gap,  $E1$ -ega,  $K1$ -kesim,  $O1$ -to'ldiruvchi,  $D1$ -aniqlovchi,  $H1$ -hol,  $i$  va  $n$  – har bir predikatning  $x$  to'plam ichida o'zgarish qadamlari,  $\cup$  - birlashtirish belgisi,  $\downarrow$  - kelishi yoki kelmasligi mumkin bo'lgan belgi.

Izoh,  $E1 \in x, x = \{i \text{ dan } n \text{ gacha o'zgaradi}\} i \geq 1, n \geq 1$

$K1 \in x, x = \{i \text{ dan } n \text{ gacha o'zgaradi}\} i \geq 1, n \geq 1$

Masalan:

Bu yilgi seminar yanvar oyining boshida o'tdi,  
aniqlovchi ega hol kesim

Bu yil+gi seminar yanvar oy+i+ning bosh+i+da o't+di.

Bu gapda ega, kesim, aniqlovchi hamda hol ishtirok etyapti.

Quyida gap yasalishining sonli usulda (sistema ko'rinishida) ifodalanilgan:

$$S_{uzbsimple2} = \begin{cases} D1(x_i, \dots, x_n) \rightarrow R_{i1} \_R_{i1} \oplus F_{y1} \cup \\ \cup E1(x_i, \dots, x_n) \cup \\ H1(x_i, \dots, x_n) \rightarrow R_{i1} \_R_{i1} \oplus F_{y1} \oplus I_{f1} \_R_{i1} \oplus F_{y1} \oplus I_{f1} \cup \\ K1(x_i, \dots, x_n) \rightarrow R_{i1} \oplus F_{y1} \end{cases}$$

Bu yerda,

$S_{uzbsimple2}$ -sodda yoyiq gap,  $E1$ -ega,  $K1$ -kesim,  $R_{i1}$ -o'zak,  $F_{y1}$ -shakl yasovchi qo'shimcha,  $I_{f1}$  – so'z o'zgartiruvchi qo'shimcha,  $i$  va  $n$  – har bir predikatning  $x$  to'plam ichida o'zgarish qadamlari,  $\cup$  - birlashtirish belgisi.

Turkiy tillar oilasiga kiruvchi tillar agglyutinatив tillar sinfiga kirganligi uchun yuqoridagi lingvo matematik modellar mos ravishda qoraqalpoq va qirg'iz tillari uchun ham qo'llaniladi.

O'zbek tilining matematik modellari M. Xakimov ishlarida ham keltirilgan bo'lib, unda o'zbek tilidagi gaplarning maqsadga ko'ra turlari(darak, so'roq, undov gaplar) bo'yicha matematik modellar ishlab chiqilgan.

O‘zbek tilida gaplarning yasalishi (22) va (23) larga mos ravishda qoraqalpoq va qirg‘iz tillari uchun ham quyidagilarni (24,25) keltiramiz:

$$S_{kksimple1} = E_2(x_i, \dots, x_n) \cup K_2(x_i, \dots, x_n) \quad (24)$$

$$S_{kksimple2} = \downarrow E_2(x_i, \dots, x_n) \cup O_2(x_i, \dots, x_n) \cup \downarrow H_2(x_i, \dots, x_n) \cup \downarrow D_2(x_i, \dots, x_n) \cup \downarrow K_2(x_i, \dots, x_n) \quad (25)$$

*Bu yerda*

$S_{kksimple1}$  – qoraqalpoq tilida sodda yig‘iq gap tuzilishi (24)

$S_{kksimple2}$  – qoraqalpoq tilida sodda yoyiq gap tuzilishi (25)

$E_2$ - ega,  $K_2$ - kesim,

$O_2$ -to‘ldiruvchi,

$D_2$ -aniqllovchi,

$H_2$ -hol,  $i$  va  $n$  – har bir predikatning  $x$  to‘plam ichida o‘zgarish qadamlari,

$\cup$  - birlashtirish belgisi,  $\downarrow$  - kelishi yoki kelmasligi mumkin bo‘lgan belgi.

*Izoh,  $E1 \in x, x = \{i \text{ dan } n \text{ gacha o‘zgaradi}\} i \geq 1, n \geq 1$*

*$K1 \in x, x = \{i \text{ dan } n \text{ gacha o‘zgaradi}\} i \geq 1, n \geq 1$*

$$S_{kgsimple1} = E_3(x_i, \dots, x_n) \cup K_3(x_i, \dots, x_n) \quad (26)$$

$$S_{kgsimple2} = \downarrow E_3(x_i, \dots, x_n) \cup \downarrow O_3(x_i, \dots, x_n) \cup \downarrow H_3(x_i, \dots, x_n) \cup \downarrow D_3(x_i, \dots, x_n) \cup \downarrow K_3(x_i, \dots, x_n) \quad (27)$$

*Bu yerda*

$S_{kgsimple1}$  – qoraqalpoq tilida sodda yig‘iq gap tuzilishi (26)

$S_{kgsimple2}$  – qoraqalpoq tilida sodda yoyiq gap tuzilishi (27)

$E_3$ - ega,  $K_3$ - kesim,

$O_3$ -to‘ldiruvchi,

$D_3$ -aniqllovchi,

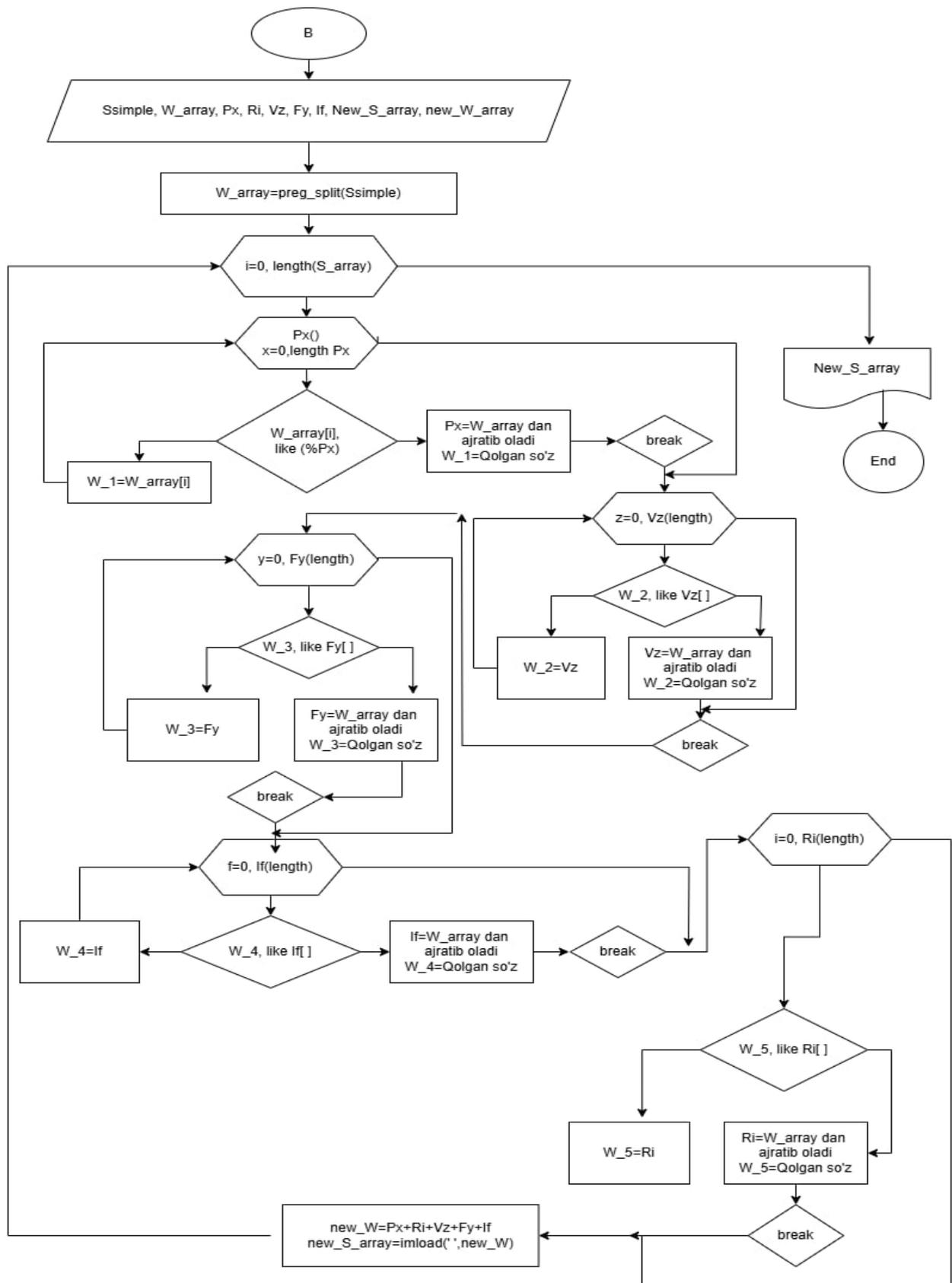
$H_3$ -hol,  $i$  va  $n$  – har bir predikatning  $x$  to‘plam ichida o‘zgarish qadamlari,

$\cup$  - birlashtirish belgisi,  $\downarrow$  - kelishi yoki kelmasligi mumkin bo‘lgan belgi.

*Izoh,  $E1 \in x, x = \{i \text{ dan } n \text{ gacha o‘zgaradi}\} i \geq 1, n \geq 1$*

*$K1 \in x, x = \{i \text{ dan } n \text{ gacha o‘zgaradi}\} i \geq 1, n \geq 1$*

Yuqorida ishlab chiqilgan lingvo-matematik modellar asosida uch tilli elektron tarjimini amalga oshiruvchi algoritm ham ishlab chiqildi (6-rasm):



6-rasm. Uch tilli elektron tarjimoni analga oshirishni tafsiflovchi blok-sxemasi (A algoritmi)

Uch tilli elektron tarjima sifatini baholovchi kombinatsiyalangan metrikaning matematik formulasini keltiradigan bo‘lsak, METEORE metrikasini hisoblash formulasi quyidagicha :

$$\text{METEOR} = \text{GM} \times (1 - \alpha) \quad (28)$$

$$\text{GM} = \frac{P \cdot R}{\alpha \cdot P + (1 - \alpha) \cdot R}$$

Bu yerda, ga teng va quyidagicha tavsiflanadi:

$\text{GM}$  (Harmonic Mean) – aniqlik ( $P$ ) va to‘liqlik ( $R$ ) o‘rtasida o‘rtacha garmonik qiymat;

$\alpha$ -tarjimadan keyin olingan matnning tartibiga beriladigan qiymat (tartib buzilgan yoki buzilmaganligi bo‘yicha)

ROUGE-1 ni hisoblash formulasi quyidagicha :

$$\text{ROUGE} - 1 = 2 \cdot \frac{\text{ROUGE} - 1 \cdot \text{Precision} \cdot \text{ROUGE} - 1 \cdot \text{Recall}}{\text{ROUGE} - 1 \cdot \text{Precision} + \text{ROUGE} - 1 \cdot \text{Recall}} \quad (29)$$

Ushbu usullarni birlashtirib quyidagi kombinatsiyalangan metrikani olamiz:

$$\text{MRS}_{\text{hb}} = \text{METEORE} + \text{ROUGE-1F1score}$$

$$\text{MRS}_{\text{hb}} = \frac{P \cdot R}{\alpha \cdot P + (1 - \alpha) \cdot R} + 2 \frac{P \cdot R}{P + R} + \text{Suffix}_{\text{PR}} \quad (30)$$

(30) ni izohlaydigan bo‘lsak aniqlik va to‘liqlik parametrlari borligini ko‘rishimiz mumkin, tadqiqot ishimizdan kelib chiqib Suffix parametrini kiritamiz.

Bu yerda  $M$ -METEOR,  $R$ -ROUGE-1,  $S$ -Suffix,  $hb$ -hybrid,  $P$ -aniqlik,  $R$ -to‘liqlik ya‘ni qo‘shimchalarning mos tushishlari hamda sonini hisobga olib tarjima sifatini baholaydi.

Elektron tarjimani baholovchi hisoblash algoritmi qadamma qadam berildi:

1-qadam. Dastur boshi;

2-qadam. Matn ( $T \downarrow$ ) kiritiladi;

3-qadam. A algoritm amalga oshiriladi;

4-qadam. Olingan natija METEOR asosida tekshiriladi, ya‘ni:

$\text{METEOR} = \text{GM} \times (1 - \alpha)$  orqali etalon matn bilan solishtiriladi

( $T \downarrow = T \uparrow$ ), natijalar hisoblanadi va keyingi qadamga o‘tiladi;

5-qadam.  $T \downarrow$  ROUGE-1 asosida hisoblanadi, ya‘ni 1-grammlar hisoblanadi va keyingi qadamga o‘tiladi;

6-qadam.  $T \downarrow = T \uparrow$  da suffixlar taqqoslanadi va o‘rta arifmetigi hisoblanadi;

7-qadam.  $\text{METEOR} + \text{ROUGE} + \text{SUFFIX}_{\text{PR}}$  qiymatlarining o‘rta arifmetigi hisoblanib  $\text{MRS}_{\text{hb}}$  ga o‘zlashtiriladi va keyingi qadamga o‘tiladi;

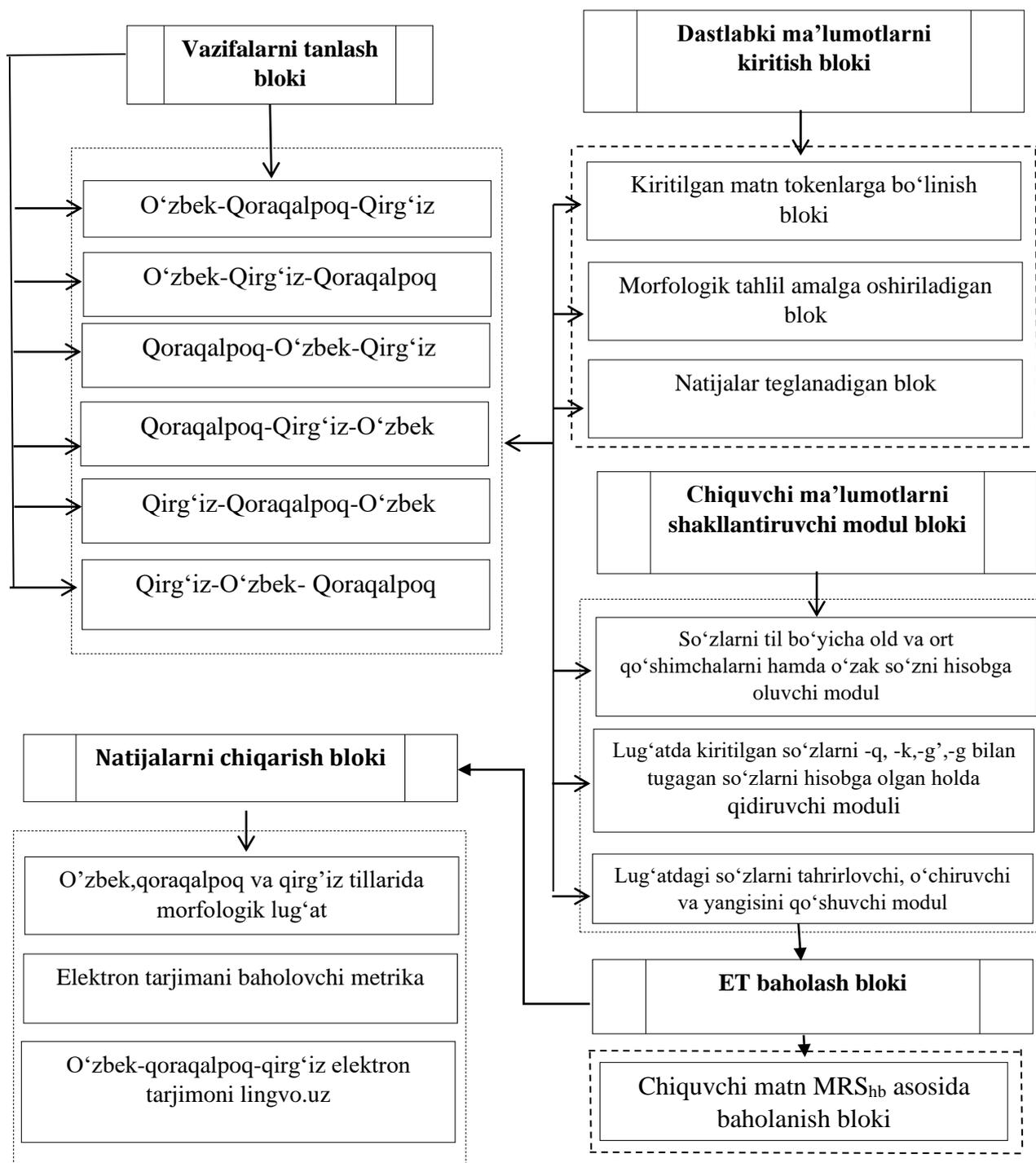
8-qadam.  $\text{MRS}_{\text{PR}}$  chop qilinadi;

9-qadam. Dastur tugashi.

Izoh: A algoritm - yuqorida berilgan uch tilli tarjimani amalga oshiruvchi algoritm(6-rasm);  $T \downarrow$  - kiruvchi matn;  $T \uparrow$  - chiquvchi matn;  $\text{MRS}_{\text{hb}}$  - METEOR, ROUGE,  $\text{SUFFIX}_{\text{Precision, Recall}}$ , hb-hybrid so‘zining qisqartmasi.

Ushbu keltirilgan hisoblash algoritmi olingan tarjimaning kontekstual mazmuni, leksik va grammatik jihatdan qay darajada to‘g‘riligini foizlarda chiqarib beradi.

Dissertatsiya ishining **“Turkiy tillar oilasiga mansub uch tilli elektron tarjimaning funksional moduli va dasturiy majmuasi”** deb nomlangan to‘rtinchi bobida elektron tarjimon ma’lumotlar bazasining infologik modeli, elektron lug‘at va tarjimonning funksional moduli hamda dasturiy ta’minotdan foydalanish yo‘riqnomasi ishlab chiqilgan.



7- rasm. Dasturiy ta’minot ishlashining funksional sxemasi

Bizning tadqiqot ishimizda (7-rasm) elektron tarjimon quyidagi modul bloklardan tashkil topgan:

1. Vazifalarni tanlash bloki;
2. Dastlabki ma'lumotlarni kiritish bloki;
3. Chiquvchi ma'lumotlarni shakllantiruvchi modul bloki
4. Elektron tarjimani baholash bloki;
5. Natijalarni chiqarish bloki.

So'zlarni til bo'yicha old va ort qo'shimchalarni hamda o'zak so'zni hisobga oluvchi modul. Turkiy tillar uchun tarjimani amalga oshirishda qo'shimchalar va o'zak so'zlarni e'tiborga olish juda ham zarur. Biz ham shuni hisobga olgan holda quyidagi jadvallarni shakllantirib oldik.

Agarda tarjima qilinishda foydalanilayotgan so'zlar ma'lumotlar bazasida uchramasa foydalanuvchiga ogohlantirish beradi yoki shu so'zni o'zini chiqaradi.

## XULOSA

“Turkiy tillar oilasiga mansub uch tilli elektron tarjima uchun lingvo-matematik modellar va hisoblash algoritmlari” mavzusidagi dissertatsiya ishi bo'yicha olib borilgan tadqiqotlar natijasida quyidagi xulosalar taqdim etildi:

O'zbek, qoraqalpoq va qirg'iz tillarining leksik, morfologik tahlil qilindi, sodda va murakkab so'z birikmalarini yasash lingvo-matematik modeli ishlab chiqildi. Natijada o'zbek, qoraqalpoq va qirg'iz tillarida sodda va murakkab so'zlarni kontekstual mos tuzishga asos bo'ladi;

O'zbek, qoraqalpoq va qirg'iz tillari uchun sodda va murakkab gap yasash jarayonining lingvo-matematik modellari ishlab chiqildi. Natijada ushbu lingvo-matematik modellar asosida o'zbek, qoraqalpoq va qirg'iz tillarida gaplarni yasash ketma-ketligini elektron tarjima sifatini baholashga xizmat qildi;

Sodda va murakkab so'zlar hamda gap tuzilmasining lingvo-matematik modellari asosida o'zbek, qoraqalpoq va qirg'iz tillari uchun fe'l zamonlarini inobatga oluvchi sodda va murakkab so'zlar hamda gap tuzilmasini hosil qiluvchi jarayonlarning lingvo-matematik modellari va algoritmlari ishlab chiqildi. Natijada elektron tarjimada matnning kontekstual to'g'ri tarjimasining 1-2 barobar oshirilishiga xizmat qildi;

O'zbek, qoraqalpoq va qirg'iz tillari uchun uch tilli elektron lug'at hamda elektron tarjimon uchun ma'lumotlar bazasi asosida ilmiy-texnik sohadagi so'zlarni qidirish va matnlarning o'zaro tarjimasini aniqlash, o'zbek, qoraqalpoq va qirg'iz tillarida elektron tarjimani baholovchi kombinatsiyalangan metrika va uning hisoblash algoritmi ishlab chiqildi. Natijada uch tilli elektron tarjimani baholash sifatini 78,2% ga oshirilishga xizmat qildi;

Yuqorida bajarilgan ilmiy tadqiqot ishi asosida “BMW” MChJ, “TCW” MChJ va “SHARQ” aksiyadorlik kompaniyasida dalolatnomalar tuzildi. Keltirilgan dalolatnomalar asosida tadqiqot ishi natijalari UzINFOCOM, UNICON hamda RTSIRITI soha korxonalarida amaliyotga joriy qilindi va natijada uch tilli elektron tarjimaning aniqligi 86,7% ga oshirilishga xizmat qildi.

**НАУЧНЫЙ СОВЕТ DSc.13/30.12.2019.Т.07.01  
ПО ПРИСУЖДЕНИЮ УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ  
УНИВЕРСИТЕТЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

---

**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ  
ТЕХНОЛОГИЙ ИМЕНИ МУХАММАДА АЛЬ-ХАРЕЗМИЙ**

**УЗАКОВА МАМУРА АБДУРАЙИМОВНА**

**ЛИНГВО-МАТЕМАТИЧЕСКИЕ МОДЕЛИ И ВЫЧИСЛИТЕЛЬНЫЕ  
АЛГОРИТМЫ ДЛЯ ТРЁХЪЯЗЫЧНОГО ЭЛЕКТРОННОГО ПЕРЕВОДА  
ОТНОСЯЩЕГО К ТЮРКСКОЙ ЯЗЫКОВОЙ СЕМЬЕ**

05.01.07 – Математическое моделирование. Численные методы и комплекс программ

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

**Ташкент – 2025**

Тема диссертации доктора философии (PhD) по техническим наукам зарегистрирована в Высшей аттестационной комиссии при Министерстве высшего образования, науки и инноваций Республики Узбекистан за номером № B2025.1.PhD/T5300

Диссертация выполнена в Ташкентском университете информационных технологий.

Автореферат диссертации на трех языках (узбекский, русский, английский (резюме)) размещен на веб-странице научного совета ([www.tuit.uz](http://www.tuit.uz)) и на Информационно-образовательном портале "ZiyoNet" ([www.ziynet.uz](http://www.ziynet.uz)).

Научный руководитель:	Назирова Элмира Шодмоновна доктор технических наук, профессор
Официальные опоненты:	Мухаммадиева Дилдора Кабиловна доктор технических наук, доцент Бакаев Илхом Изатович доктор философии технических наук, доцент
Ведущая организация:	Каршинский государственный университет

Защита диссертации состоится "30" 10 2025 г. в 16<sup>00</sup> часов на заседании Научного совета DSc.13/30.12.2019.T.07.01 при Ташкентском университете информационных технологий. (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-43; e-mail: [iktuit@tuit.uz](mailto:iktuit@tuit.uz)).

С диссертацией можно ознакомиться в Ташкентском университете информационных технологий (регистрационный номер № 377). (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-70).

Автореферат диссертации разослан "18" 10 2025 года.

(протокол рассылки № 31 от "15" 10 2025 г.).



*M. Musaev*

**М.М. Мусаев**  
Председатель Научного совета по  
присуждению учёных степеней, д.т.н., профессор

*J. X. Djumanov*

**Ж.Х. Джуманов**  
Ученый секретарь Научного совета по  
присуждению учёных степеней, д.т.н. профессор

*F. M. Nuralliev*

**Ф.М. Нуралиев**  
Председатель Научного семинара при  
Научном совете по присуждению учёных  
степеней, д.т.н. профессор

## Введение (аннотация к диссертации доктора философии (PhD))

**Актуальность и востребованность темы диссертации.** Развитие современного общества в мире, внедрение компьютерных технологий во все сферы жизни, широкое использование оцифрованной информации приобретают все большее значение. Поскольку компьютерные технологии глубоко проникли во все сферы деятельности, они также набирают силу в области лингвистики. Начиная с XX века мировая лингвистика уделяет все большее внимание созданию системы терминов, относящихся к области компьютерной лингвистики, и их научному изучению. В связи с колоссальными изменениями, произошедшими в социально-политической, экономической и духовной жизни общества, изучение терминов тюркской языковой семьи с точки зрения машинного обучения и компьютерной лингвистики, а также реализация процессов их компьютерного моделирования является одной из наиболее актуальных проблем современности. Действительно, создание компьютерно-ориентированных моделей естественного языка и решение на их основе ряда задач, автоматический анализ естественных языков в морфологическом, синтаксическом и лексическом плане, выяснение их специфических аспектов имеют большое значение в компьютерной лингвистике. Особое внимание уделяется разработке методов, моделей и программных технологий, используемых в процессе многоязыкового перевода в странах, достигших развитого уровня в области компьютерной лингвистики и электронного перевода, в том числе в США, Великобритания, Японии, Китай, России, Канада и других странах.

Глобальный многоязычный перевод упрощает общение. Он устраняет барьеры между языками, что способствует развитию торговли, образования, научных исследований и дипломатических отношений. Многоязычные системы полезны для изучения языка. Перевод информации с одного языка на другой позволяет пользователям получать доступ к контенту на разных языках. Используя Интернет и мобильные технологии, он позволяет обслуживать пользователей на разных языках, гарантируя работу услуг в глобальном масштабе. С помощью машинного перевода многоязычные переводческие услуги могут быть быстрыми и эффективными, позволяя переводить тексты на многие языки за короткое время.

В Указе Президента Республики Узбекистан от 21 октября 2019 года № ПФ-5850 «О мерах по коренному повышению престижа и статуса узбекского языка как государственного» обозначены задачи «...обеспечения государственного языка достойным местом в глобальной информационной сети с использованием информационно-коммуникационных технологий, в частности сети Интернет, и создания компьютерных программ для узбекского языка»<sup>2</sup>, и данное диссертационное исследование в определенной мере послужит реализации задач, обозначенных в соответствующих нормативно-правовых документах. Разработана Концепция развития узбекского языка и

---

<sup>1</sup> Указ Президента Республики Узбекистан «О мерах по дальнейшему развитию узбекского языка и совершенствованию языковой политики в нашей стране». <https://lex.uz/docs/-5058351>

совершенствования языковой политики в 2020 – 2030 годы. Для реализации этих задач необходимо создание компьютерно-ориентированных моделей узбекского языкознания, а поскольку в настоящее время в узбекской компьютерной лингвистике недостаточно исследований по машинному переводу, необходимо создание лингвистической базы данных узбекского языка для программ перевода с узбекского на иностранный язык и обратно, и одним из важных вопросов является создание алгоритмов и программных средств для их практического использования.

Решения и инициативы нашего Президента, направленные на развитие узбекского языка, направлены на широкое использование узбекского языка в обществе и укрепление его позиций как государственного языка. Эти решения создают важную основу для продвижения, изучения и эффективного использования языка

**Соответствие исследования приоритетным направлениям развития науки и технологий республики.** Данное исследование является частью IV Программы Республики Узбекистан по развитию науки и технологий. Реализуется в рамках приоритетного направления «Развитие информатизации и информационно-коммуникационных технологий».

**Степень изученности проблемы.** В мире проведено множество исследований по разработке моделей и алгоритмов электронных переводчиков, в том числе работы основоположников электронных переводчиков У. Уивера, Р. Риченса, А.А. Ляпунов, Г.Г. Белоногонов, В.Г. Васильев, Л.А. Гращенко, Ю.Н. Марчук и другие.

Работа зарубежного ученого Кристофера Мэннинга сосредоточена в основном на совершенствовании синтаксического анализа, ввода данных и многоязычных систем перевода, включая такие вопросы, как модели n-грамм, скрытые марковские модели Hidden Markov Model (НММ), разметка частей речи и синтаксический анализ. Кроме того, такие известные ученые, как Даньци Чен, Регина Барзилай, Дэвид Зонтаг, Джошуа Бенжио, Янн Лекун и Мартин Джоенс, проводят ряд работ по семантическому пониманию и многоязычному переводу.

А.К. Пулатов, М.Х. Хакимов, Ш.А.Назиров отвечают за создание и работу узбекских электронных переводчиков в нашей республике. Ш.А.Назиров, М.Арипов работали над наиболее актуальными проблемами в среде многоязычного машинного перевода, ролью формальной грамматики в электронном переводе, вопросом математического моделирования естественных языков как необходимого механизма перехода от формальной грамматики к алгоритмам, а также математическими моделями словообразования, созданными для решения этой проблемы. Молодые ученые Н. Абдурахманова, С. Шукуров, Х. Турсунов, Б. Каршиев, М. Комилов ведут ряд работ по созданию корпуса узбекского языка, параллельных корпусов и других направлений компьютерной лингвистики.

Анализ исследований в этой области показывает, что во всем мире электронные переводчики имеют крайне мало ресурсов для языков, относящихся к тюркской языковой семье, а при анализе работы многоязычных переводчиков они показывают низкую точность и высокое качество вывода при

переводе текстов. Из этого следует, что вопрос создания электронных переводчиков становится все более актуальным.

**Связь диссертационного исследования с планами научно исследовательских работ высшего образовательного учреждения, где выполнена диссертация.** Диссертационная работа выполнена в рамках плана НИР Ташкентского университета информационных технологий имени Мухаммада ал-Хоразмий по проекту № ПЛ-77-8023102181 по теме «Создание «Банка инновационных идей», охватывающего новые научные разработки и технологии в регионах и отраслях экономики (2023-2025 годы)».

**Цель исследования** построение лингво-математических моделей и алгоритмов перевода с узбекского языка на каракалпакский и киргизский, базы данных трёхязычного электронного словаря и разработка программного комплекса.

**Задачи исследования:**

анализ трёхязычных программ электронного переводчика и создание электронной словарной базы тюркских языков таких как узбекский, каракалпакский и киргизский языки;

лексико-морфологический анализ узбекского, каракалпакского и киргизского языков, формирование сложных фраз и простой структуры предложений, разработка лингво-математических моделей и вычислительных алгоритмов;

разработка лингво-математической модели и алгоритма для слов, образованных при помощи временных глаголов в узбекском, каракалпакском и кыргызском языках;

проведение фонетического анализа для словообразования слов в узбекском языке и разработка лингво-математической модели, алгоритма;

разработка математической модели и вычислительного алгоритма комбинированной метрики для оценки трёхязычного электронного перевода;

формализация процессов программирования на основе модульного анализа вычислительных алгоритмов электронного перевода тюркских языков;

создание программного комплекса и пользовательского интерфейса, основанного на предложенных моделях и алгоритмах, который переводит узбекско-каракалпакско-киргизский и каракалпакско – узбекско -киргизский тексты на три языка.

**Объектом исследования** является лингвистические корпуса узбекского, каракалпакского и киргизского языков.

**Предметом исследования** является разработка лингво-математических моделей, алгоритмов и программного комплекса для реализации трёхязычного электронного перевода для тюркских языков, а также разработка комбинированной метрики и вычислительного алгоритма для оценки трёхязычного электронного перевода.

**Методы исследования.** При освещении темы исследования использовались лингво-математическое моделирование, методы модульного анализа вычислительных алгоритмов, современные технологии программирования.

**Научная новизна исследования** заключается в следующем:

на основе лексико-морфологической базы узбекского, каракалпакского и кыргызского языков разработаны лингво-математические модели и алгоритмы образования сложных словосочетаний и структур предложений для узбекского, каракалпакского и кыргызского языков;

на основе лингво-математических моделей образования простых и сложных слов также предложений разработаны лингво-математические модели и алгоритмы процессов простого словообразования для фонетического строя узбекского, каракалпакского и кыргызского языков;

на основе лингво-математических моделей простых и сложных слов также предложений разработаны лингво-математические модели и алгоритмы процессов простого и сложного словообразования и структур предложений для узбекского, каракалпакского и кыргызского языков с учетом времен глаголов; разработаны трехязычный электронный словарь для узбекского, каракалпакского и кыргызского языков и база данных электронного переводчика для поиска слов научно-технического направления и определения взаимоперевода текстов, а также комбинированная метрика и вычислительный алгоритм для оценки электронного перевода на узбекский, каракалпакский и кыргызский языки.

**Практические результаты исследования** заключаются в следующем:

разработана база данных для трехязычного электронного словаря и переводчика, метод поиска слов и определения взаимного перевода текстов на основе семантического анализа;

для реализации трехязычного электронного переводчика разработана лингво-математическая модель формирования сложных слов на узбекском, каракалпакском и киргизском языках;

для реализации трехязычного электронного переводчика разработана лингво-математическая модель словообразования с участием глагольных времен на узбекском, каракалпакском и киргизском языках;

для реализации трехязычного электронного переводчика разработана лингво-математическая модель для построения простых предложений на узбекском, каракалпакском и киргизском языках;

для реализации трехязычного электронного переводчика разработана лингво-математическая модель и алгоритм фонетического анализа на узбекском языке;

для оценки трехязычного электронного перевода усовершенствована математическая формула для комбинированной метрики;

разработаны принципы работы программного комплекса и необходимые требования.

**Достоверность результатов исследования** основана на проведенных опытах по переводу простых предложений между тремя тюркскими языками с использованием специального программного комплекса, что подтверждается окончательными выводами о надежности полученных результатов.

**Научная и практическая значимость результатов исследования.**

Научная значимость результатов исследования заключается в создании

лингво-математических моделей и алгоритмов построения простых и сложных слов также предложений при переводе на трёхязычные тюркские языки таких как узбекский, каракалпакский и кыргызский, комбинированная метрика и вычислительный алгоритма для оценки электронного перевода. Практическая значимость результатов исследования заключается в разработке соответствующего программного комплекса, используемого для построения простых и сложных слов и предложений.

#### **Внедрение результатов исследования.**

На основе лингво-математических моделей, алгоритмов и программного обеспечения, разработанных в результате исследований, проведенных в рамках диссертации:

разработанное на основе алгоритма образования сложных слов узбекского, каракалпакского и кыргызского языков и лингвистико-математической модели программное обеспечение было внедрено в Акционерном обществе издательство «SHARQ» для устранения недостатков при переводе текстов (Справка Министерства цифровых технологий Республики Узбекистан от 7 июля 2025 года № 33-8/4723). В результате достигнуто повышение общей эффективности работы на 78,2%;

разработанный программный комплекс на основе алгоритма и лингво-математической модели построения простых предложений на узбекском, каракалпакском и кыргызском языках был внедрён в редакцию перевода научно-технических текстов ООО «Big Makro Word» (справка Министерства цифровых технологий Республики Узбекистан № 33-8/4723 от 7 июля 2025 года). В результате общий объём работ по времени сокращён в 1-2 раза;

разработанный программный комплекс на основе алгоритма и математической модели построения сложных слов на узбекском языке был внедрён в издательство ООО «The Computer Workshop» (справка Министерства цифровых технологий Республики Узбекистан № 33-8/4723 от 7 июля 2025 года). Полученные результаты были использованы для анализа узбекских слов в программе «Uzmorphoanalyzer – программа для морфологического анализа узбекских слов».

**Апробация результатов исследования.** Результаты данного исследования были обсуждены на 9 международных (2 из которых — Scopus) и 5 республиканских научно-практических конференциях.

**Опубликованность результатов исследования.** По теме исследования было опубликовано 20 научных работ: 4 статей опубликованы в научных изданиях, рекомендованы Высшей Аттестационной Комиссией Республики 1 опубликованы в зарубежных журналах, а 3 — в республиканских журналах. Также результаты исследования были представлены на 2 международных научно-практических конференциях, 8 республиканских научно-практических конференциях. Была издана 1 монография, а также зарегистрированы программные продукты для ЭВМ (о чем получены 2 свидетельства) и создана база данных (о чем получено 1 свидетельство).

**Структура и объем диссертации.** Диссертационная работа состоит из введения, четырех глав, заключения, списка использованных источников и приложений. Объем диссертации составляет 110 страниц.

## ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

**Во введении** обоснована актуальность и необходимость темы диссертации, описаны цель и задачи исследования, объект и предмет исследования, указано соответствие исследования приоритетным направлениям развития науки и технологий Республики, изложены научная новизна и практические результаты исследования, приведены данные о надежности полученных результатов, научной и практической значимости, внедрении результатов исследования в практику, апробации работы, опубликованных работах и структуре диссертации.

В первой главе диссертации: **«Развитие электронных переводчиков и их текущее состояние»**, проводится анализ методов исследования и анализ существующих работ по теме диссертации, а также сравнительный анализ международных и республиканских электронных переводчиков, их значения, роли и проблем. Кроме того, приведены этапы развития электронных переводчиков, основные подходы к решению проблемы. Рассматривается происхождение и схема тюркских языков, таких как узбекский, каракалпакский и киргизский.

Во второй главе диссертации: **«Лингво-математические модели и алгоритмы формирования простых и сложных слов для трехязычного электронного перевода, относящегося к тюркским языкам»**, приведены лингво-математические модели и алгоритмы формирования простых и сложных слов в узбекском, каракалпакском и киргизском языках. Поскольку тюркские языки принадлежат к агглютинативной группе языков, процесс словообразования в них отличается четким определением границ морфем и последовательным характером, что отличает их от других языковых групп. Это можно выразить следующим образом:

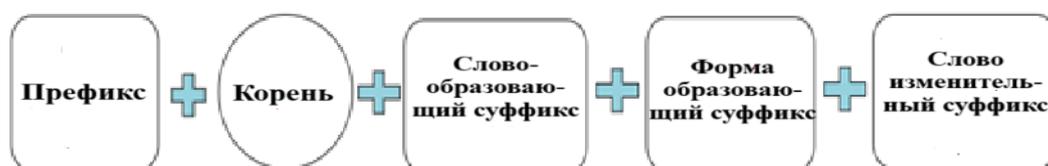


Рисунок 1. Логическо-лингвистическая модель словообразования в узбекском языке

На данной схеме показана последовательность словообразования в узбекском языке, где основа слова получает префиксы и суффиксы, что приводит к образованию нового слова. На 1- рисунке формализуется процесс словообразования в естественном языке, что отражает первый шаг электронного перевода.

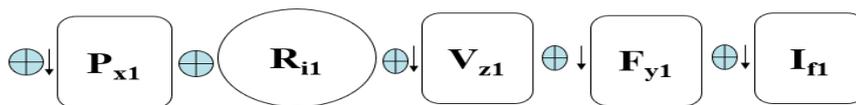


Рисунок 2. Формальная модель схемы словообразования в узбекском языке  
Здесь, чтобы внести ясность в рисунок 2, уточним следующее:

$P_x$  — приставка,  $R_i$  — основа слова,  $V_z$  — словообразовательный суффикс,  $F_y$  — формаобразовательный суффикс,  $I_f$  — словоизменятельный суффикс,  $\oplus$  — символ соединения,  $\downarrow$  — символ, который может быть или не быть. Если выразить формальную модель в виде формулы, то получаем следующее:

$$W_{uzb} = \oplus \downarrow P_x \oplus R_{i_1} \oplus \downarrow V_{z_1} \oplus \downarrow F_{y_1} \oplus \downarrow I_{f_1} \quad (1)$$

Здесь,

$W_{uzb}$ - слово,  $P_{x1}$ - префикс,  
 $V_{z1}$ - словообразовательный суффикс,  
 $F_{y1}$  – формаобразовательный суффикс,  
 $I_{f1}$ - словоизменятельный суффикс.

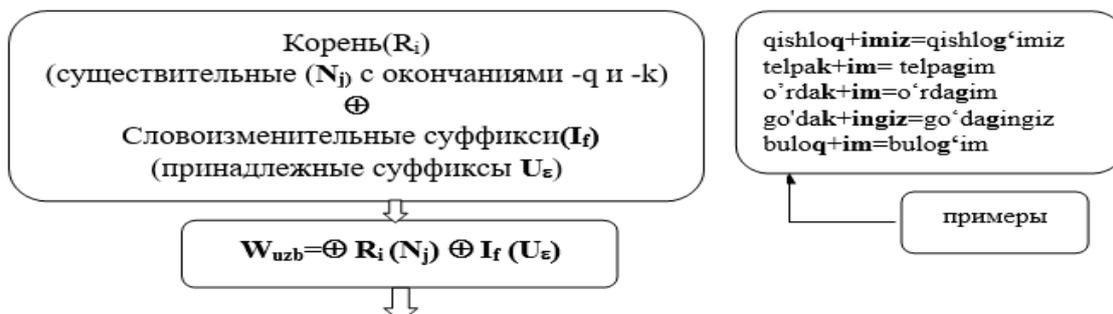
Если развернуть выражение (1), то получим следующее:

$$W_{uzb} = \oplus \downarrow \sum_{x_1=1}^m P_{x_1} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{z_1=1}^d V_{z_1} \oplus \downarrow \sum_{y_1=1}^v F_{y_1} \oplus \downarrow \sum_{f_1=1}^e I_{f_1} \quad (2)$$

Здесь:

$W_{uzb}$  — сложное слово на узбекском языке,  $P_{x1}$  — приставка,  $R_{i1}$  — корень слова,  $V_{z1}$  — словообразующий суффикс,  $F_{1y}$  — аффикс, формирующий форму,  $I_{f1}$  — аффикс, изменяющий слово,  $i, m, z, d, y, v, f, e$  — длина корня и аффиксов,  $\oplus$  — символ соединения,  $\downarrow$  — символ, который может быть или не быть. Изучив процесс словообразования в узбекском языке, мы пришли к выводу, что это сложный процесс, и для качественного электронного перевода необходимо глубоко изучить этот процесс. На этапе фонетического анализа мы отдельно формализовали его. В процессе словообразования в узбекском языке фонетический анализ основан в основном на 5 правилах, которые рассмотрены с примерами:

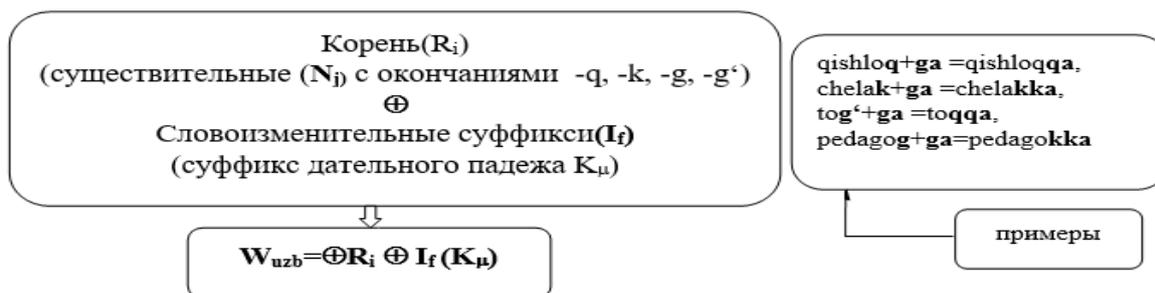
**1-правило фонетического анализа слов:**



$$W_{uzb} = \bigoplus_{i=1}^n R_i(N_j) \bigoplus_{f=1}^e I_f(U_\varepsilon) \quad (3)$$

Здесь (3):  
 $W_{uzb}$  – слово,  $R_i$ -корень,  
 $I_f$  – словоизменительный суффикс,  
 $U_\varepsilon$ -суффикс принадлежности,  
 $N_j$ -существительное

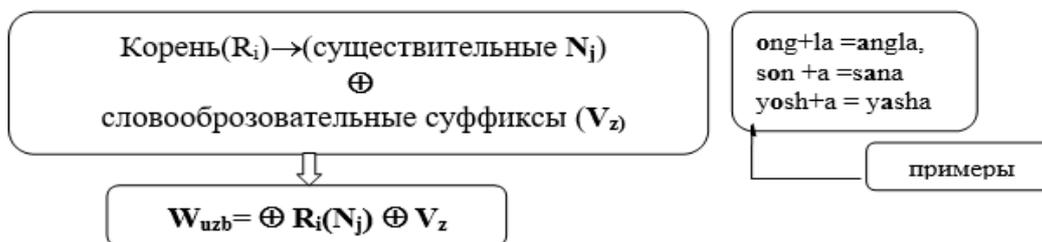
**2-правило фонетического анализа слов:**



$$W_{uzb} = \bigoplus_{i=1}^n R_i(N_j) \bigoplus_{f=1}^e I_f(K_\mu) \quad (4)$$

Здесь (2)  
 $W_{uzb}$  – слово,  $R_i$ - корень,  
 $I_f$  - словоизменительный суффикс,  
 $K_\mu$ - суффикс дательного падежа.

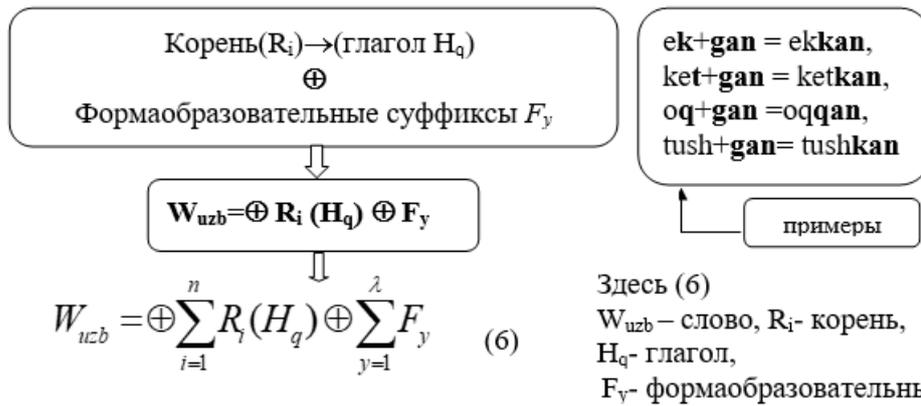
**3-правило фонетического анализа слов:**



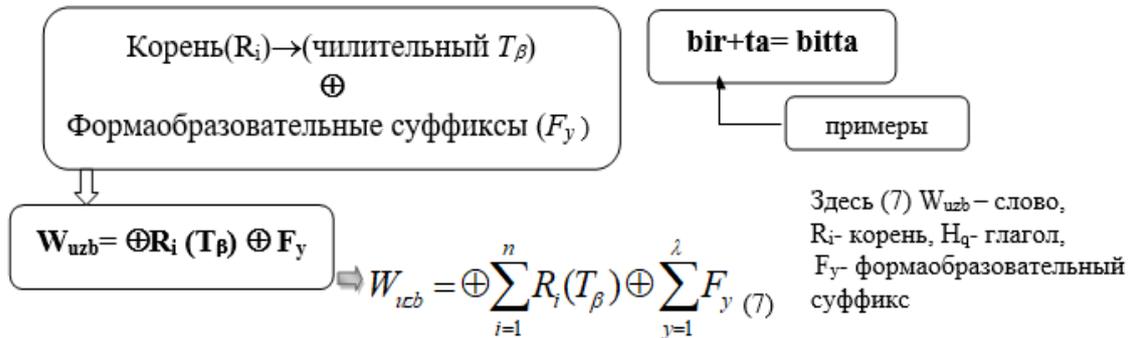
$$W_{uzb} = \bigoplus_{i=1}^n R_i(N_j) \bigoplus_{z=1}^d V_z \quad (5)$$

Здесь (5)  
 $W_{uzb}$  – слово,  $R_i$ - корень,  
 $N_j$ - существительное  
 $V_z$ - словообразовательные суффиксы

**4-правило фонетического анализа слов:**



**5-правило фонетического анализа слов:**



В узбекском языке лингво-математические модели фонетического анализа многосложных слов отражены в выражениях (1), (2), (3), (4) и (5). Обобщив разработанные лингво-математические модели, выражаем их в следующей блок-схеме:

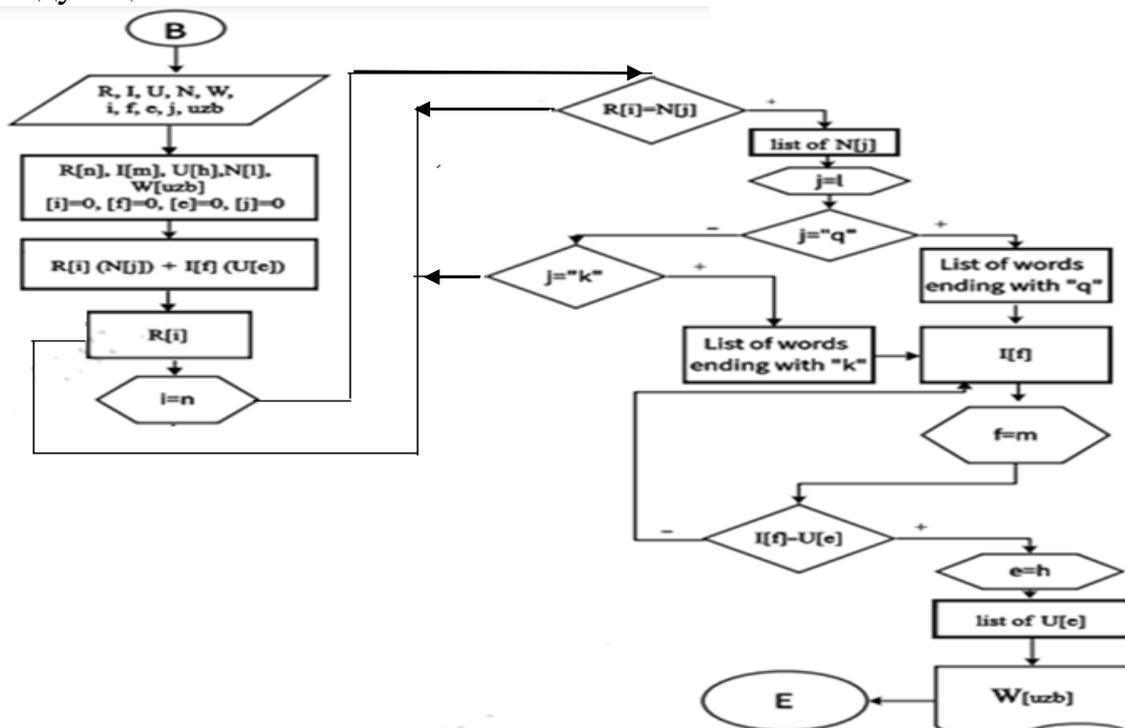


Рисунок 3. Блок-схема фонетического анализа при словообразовании в узбекском языке

Учитывая вышеизложенное, при осуществлении электронного перевода можно достичь точных и качественных результатов, которые покажут свои результаты. В нашей исследовательской работе также разработаны лингво-математические модели словообразования для каракалпакского и киргизского языков, которые принадлежат тюркской языковой семье. При этом для каракалпакского языка разработаны (2) и (3) в соответствии с выражением (1).

$$W_{kk} = \oplus \downarrow P_{x_2} \oplus R_{i_2} \oplus \downarrow V_{z_2} \oplus \downarrow F_{y_2} \oplus \downarrow I_{f_2} \quad (8)$$

$$W_{kk} = \oplus \downarrow \sum_{x_2=1}^j P_{x_2} \oplus \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{z_2=1}^d V_{z_2} \oplus \downarrow \sum_{y_2=1}^v F_{y_2} \oplus \downarrow \sum_{f_2=1}^e I_{f_2} \quad (9)$$

Здесь уточним следующие моменты:  $W_{kk}$  — каракалпакское слово,  $R_{i_2}$  — корень слова,  $V_{z_2}$  — словообразующий аффикс,  $F_{y_2}$  — аффикс, формирующий форму,  $I_{f_2}$  — аффикс, изменяющий слово,  $\oplus$  — символ соединения,  $\downarrow$  — символ обозначающий может быть или может не быть;  $x, i, z, y, f$  — длина морфем. Здесь индекс 2 обозначает второй язык. Для киргизского языка аналогичное выражение (1) разрабатывается следующим образом:

$$W_{kg} = \oplus \downarrow P_{x_3} \oplus R_{i_3} \oplus \downarrow V_{z_3} \oplus \downarrow F_{y_3} \oplus \downarrow I_{f_3} \quad (10)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=0}^j P_{x_3} \oplus \sum_{i_3=0}^n R_{i_3} \oplus \downarrow \sum_{z_3=0}^d V_{z_3} \oplus \downarrow \sum_{y_3=0}^v F_{y_3} \oplus \downarrow \sum_{f_3=0}^e I_{f_3} \quad (11)$$

Здесь уточним следующие моменты:

$W_{kg}$  — киргизское слово,  $R_{i_3}$  — корень слова,  $V_{z_3}$  — словообразующий аффикс,  $F_{y_3}$  — аффикс, формирующий форму,  $I_{f_3}$  — аффикс, изменяющий слово. Одной из особенностей структуры словообразования сложных слов в киргизском языке является то, что при переводе аффиксы притяжательности для 1-го и 2-го лица оказываются перед словом. Для уточнения обратимся к определению:

**Определение:** Когда сложное слово на киргизском языке переводится (например, с узбекского на киргизский ( $W_{uzb} \rightarrow W_{kg}$ )), если в слове встречается словообразовательный суффикс ( $I_f$ ), который является суффиксом притяжательности (то есть, суффиксы притяжательности для I и II лица: *-im, -ing, -imiz, -ingiz*), то в киргизском переводе аффиксы притяжательности превращаются в притяжательные местоимения такие как *menin, senin, bizdin, siznin* ( $G_t$ ), и они становятся отдельными словосочетаниями расположенными в начале слова.

$$\tilde{W}_{kg} = \oplus G_t \oplus R_i \oplus \downarrow V_z \oplus \downarrow F_y \oplus \downarrow \tilde{I}_{f1} \quad (12)$$

Здесь,

$\tilde{W}_{kg}$  - после перевода сложного слова на киргизский язык,  $G_t$  – притяжательное местоимение.

Например, на узбекском: *bog' + bon + lar + imiz + ning* (наши садовники)

на киргизском : *bizdin bag + ban + lar + ning* (наши садовники)

В предыдущих работах, где анализировались двухязычные электронные переводчики, были приведены последовательности словообразования для узбекского и каракалпакского языков, поэтому в нашей работе мы приведем только последовательности словообразования для киргизского языка.

При разборе сложных слов на части для киргизского языка сталкиваемся со следующими случаями *в виде численных методов*:

$$\begin{aligned}
 W_{kg1} &= \sum_{t_3=1}^{\zeta} G_{t_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y_3=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f_3=1}^e I_{f_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg2} &= \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y_3=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg3} &= \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f_3=1}^e I_{f_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg4} &= \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z_3=1}^d V_{z_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y_3=1}^v F_{y_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f_3=1}^e I_{f_3}(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
 W_{kg5} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.
 \end{aligned}$$

$$\begin{aligned}
W_{kg6} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
W_{kg7} &= \sum_{i_3=1}^n R_{i_3} \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_{z_3}(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_{y_3}(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_{f_3}(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_{y_3}(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
W_{kg8} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
W_{kg9} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
W_{kg10} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \quad \text{здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.
\end{aligned}$$

$$\begin{aligned}
W_{kg11} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
W_{kg12} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right. \\
W_{kg13} &= \sum_{i=1}^n R_i \oplus \left\{ \begin{array}{l} \sum_{z=1}^d V_z(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{f=1}^e I_f(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \\ \sum_{y=1}^v F_y(l_k) \sim N(\omega) \text{ здесь } \omega = l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1} \end{array} \right.
\end{aligned}$$

Здесь,  $W_{kg1}, W_{kg2}, W_{kg3}, W_{kg4}, W_{kg5}, W_{kg6}, W_{kg7}, W_{kg8}, W_{kg9}, W_{kg10}, W_{kg11}, W_{kg12}, W_{kg13}$  – слова на киргизском языке,  $N(\omega)$  – один предикат, каждое слово  $\omega$  выражается как комбинация двух или более слов  $l_0 \wedge l_1 \wedge l_2 \wedge \dots \wedge l_{k-1}$ , здесь  $k$  – длина суффикса.

В узбекском, каракалпакском и киргизском языках, принадлежащих к тюркской языковой семье, лингво-математические модели образования предложений разрабатывались в частном порядке, в этих случаях лингво-математические модели разрабатывались и для части времен глаголов:

Лингво-математические модели словообразования во временах глаголов в узбекском, каракалпакском и киргизском языках (для прошедшего времени)

$$W_{uzb} = \oplus \downarrow \sum_{x_1=1}^m P_{x_1} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{w_1=1}^{\tau} B_{w_1} \oplus \downarrow \sum_{c_1=1}^l L_{c_1} \oplus \downarrow \sum_{e_1=1}^o K_{e_1} \oplus \downarrow \sum_{d_1=1}^p O_{d_1} \quad (13)$$

$$W_{qq} = \oplus \downarrow \sum_{x_2=1}^m P_{x_2} \oplus \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{w_2=1}^{\tau} B_{w_2} \oplus \downarrow \sum_{c_2=1}^l L_{c_2} \oplus \downarrow \sum_{e_2=1}^o K_{e_2} \oplus \downarrow \sum_{d_2=1}^p O_{d_2} \quad (14)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=1}^m P_{x_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \downarrow \sum_{w_3=1}^{\tau} B_{w_3} \oplus \downarrow \sum_{c_3=1}^l L_{c_3} \oplus \downarrow \sum_{e_3=1}^o K_{e_3} \oplus \downarrow \sum_{d_3=1}^p O_{d_3} \quad (15)$$

Здесь  $P_x$ -префикс,  $R_i$ -корень,  $V_w$ - суффикс обозначающий прошедшее время,  $L_c$ -1-лично-числовые суффиксы,  $K_e$ -2-лично-числовые суффиксы,  $O_d$ -3- лично-числовые суффиксы

Лингво-математические модели словообразования во временах глаголов в узбекском, каракалпакском и киргизском языках (для настоящего времени)

$$W_{uzb} = \oplus \downarrow \sum_{x_1=1}^m P_{x_1} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{g_1=1}^b Q_{g_1} \oplus \downarrow \sum_{c_1=1}^l L_{c_1} \oplus \downarrow \sum_{e_1=1}^o K_{e_1} \oplus \downarrow \sum_{d_1=1}^p O_{d_1} \quad (16)$$

$$W_{qq} = \oplus \downarrow \sum_{x_2=1}^m P_{x_2} \oplus \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{g_2=1}^b Q_{g_2} \oplus \downarrow \sum_{c_2=1}^l L_{c_2} \oplus \downarrow \sum_{e_2=1}^o K_{e_2} \oplus \downarrow \sum_{d_2=1}^p O_{d_2} \quad (17)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=1}^m P_{x_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \downarrow \sum_{g_3=1}^b Q_{g_3} \oplus \downarrow \sum_{c_3=1}^l L_{c_3} \oplus \downarrow \sum_{e_3=1}^o K_{e_3} \oplus \downarrow \sum_{d_3=1}^p O_{d_3} \quad (18)$$

Здесь  $Q_g$ - суффикс обозначающий настоящего времени

Лингво-математические модели словообразования во временах глаголов в узбекском, каракалпакском и киргизском языках (настоящее время)

$$W_{uzb} = \oplus \downarrow \sum_{x_2=1}^m P_{x_2} \oplus \sum_{i_1=1}^n R_{i_1} \oplus \downarrow \sum_{u_1=1}^s T_{u_1} \oplus \downarrow \sum_{c_1=1}^l L_{c_1} \oplus \downarrow \sum_{e_1=1}^o K_{e_1} \oplus \downarrow \sum_{d_1=1}^p O_{d_1} \quad (19)$$

$$W_{qq} = \oplus \downarrow \sum_{x_2=1}^m P_{x_2} \oplus \sum_{i_2=1}^n R_{i_2} \oplus \downarrow \sum_{u_2=1}^s T_{u_2} \oplus \downarrow \sum_{c_2=1}^l L_{c_2} \oplus \downarrow \sum_{e_2=1}^o K_{e_2} \oplus \downarrow \sum_{d_2=1}^p O_{d_2} \quad (20)$$

$$W_{kg} = \oplus \downarrow \sum_{x_3=1}^m P_{x_3} \oplus \sum_{i_3=1}^n R_{i_3} \oplus \downarrow \sum_{u_3=1}^s T_{u_3} \oplus \downarrow \sum_{c_3=1}^l L_{c_3} \oplus \downarrow \sum_{e_3=1}^o K_{e_3} \oplus \downarrow \sum_{d_3=1}^p O_{d_3} \quad (21)$$

Здесь  $T_u$ - суффикс обозначающий будущее время

На основе вышеуказанного анализа словообразования времён глаголов в тюркских языках — узбекском, каракалпакском и киргизском — были разработаны лингвоматематические модели.

В третьей главе диссертационной работы: «Лингво-математические модели и алгоритмы для построения предложений в трёхязычном электронном переводе, относящемся к тюркской языковой семье» приведена лингвистическая построения предложений, формальные модели и лингво-математические модели на узбекском, каракалпакском и киргизском языках.

В узбекском языке после этапа словообразования наступает этап построения предложения. На этапе построения предложения в первую очередь необходимо изучить компоненты, требуемые для построения предложения. Основу построения предложения составляют подлежащее и сказуемое. Подлежащее представляет собой субъект действия, а сказуемое — компонент, выражающий действие в предложении. С лингвистической точки зрения подлежащее и сказуемое называются главными членами предложения. Кроме того, в предложении также существуют второстепенные члены.

На схеме, приведённой на рисунке 4 можно увидеть визуальное представление этой информации:



Рисунок 4. Лингвистическая классификация членов, составляющих основу предложения в узбекском языке

Если после этапа словообразования перейти к этапу построения предложения в узбекском языке, можно убедиться, насколько богата грамматика узбекского языка.

В узбекском языке предложения по своей структуре делятся на простые и сложные. Простые предложения по своей структуре делятся на:

1. Простые нераспространённые предложения,
2. Простые распространённые предложения (см. рисунок 5).

При такой классификации учитывается участие или неучастие второстепенных членов предложения.



Рисунок-5. Лингвистические виды простых предложений в узбекском языке

Предложение, состоящее только из главных членов (подлежащего и сказуемого), называется простым нераспространённым предложением. Примеры простых нераспространённых предложений можно представить следующим образом:

*Axborot keltirildi, Kompyuter ishlayapti. ....*

*подлежащее сказуемое подлежащее сказуемое*

Если прокомментировать этот пример, то в структуре построения предложений в узбекском языке подлежащее всегда стоит на первом месте, а сказуемое — в конце предложения, после подлежащего. Этот порядок считается строгим.

Предложение, построенное с участием главных и второстепенных членов, называется простым распространённым предложением. Например:

*Bu yilgi seminar yanvar oyining boshida o'tdi*  
*определение подлежащее обстоятельство сказуемое*

*Jismoniy mashqlar organizmni har xil kasalliklardan saqlaydi*  
 определение подлежащее дополнение определение дополнение сказуемое

Схема структуры предложения в узбекском языке с участием главных и второстепенных членов.

На основе представленных выше схем были разработаны лингво-математические модели построения простых предложений в узбекском языке.

$$S_{uzbsimple1} = E_1(x_i, \dots, x_n) \cup K_1(x_i, \dots, x_n) \quad (22)$$

*Вu yerda, S<sub>uzbsimple1</sub>- простое нераспространённое предложение, E<sub>1</sub>-сказуемое, подлежащее, K<sub>1</sub>-, i и n – шаги изменения каждого предиката в пределах x множеств.*

*Izoh, E<sub>1</sub> ∈ x, x={ изменяется с i до n} i ≥ 1, n ≥ 1*

*K<sub>1</sub> ∈ x, x={ изменяется с i до n} i ≥ 1, n ≥ 1*

Например:

*Axborot keltirildi (kel+tir+il+di). Информация была представлена.*

В процессе анализа данного предложения сначала оно разбивается на части (токены), и определяется их тип. Каждый токен далее разделяется на морфемы на основе морфологического анализа — то есть на корень (R<sub>i</sub>) и аффиксы. В данном случае подлежащее является простой (непроизводной) формой слова и не делится на морфемы, тогда как сказуемое разбивается на морфемы. На этой основе был представлен следующий пример **в численном виде**.

$$S_{uzbsimple1} = \left\{ \begin{array}{l} E_1(x_i, \dots, x_n) \cup \\ K_1(x_i, \dots, x_n) \rightarrow R_{i1} \oplus F_{y1} \oplus F_{y1} \oplus F_{y1} \end{array} \right.$$

Здесь,

*S<sub>uzbsimple1</sub>- простое нераспространённое предложение, E<sub>1</sub>- подлежащее, K<sub>1</sub>-сказуемое, i и n – шаги изменения каждого предиката в пределах x множеств, R<sub>i1</sub>-корень, F<sub>y1</sub>-формаобразующие суффиксы, ∪ - соединительный знак.*

*Примечание, E<sub>1</sub> ∈ x, x={изменяется с i до n} i ≥ 1, n ≥ 1*

*K<sub>1</sub> ∈ x, x={изменяется с i до n} i ≥ 1, n ≥ 1*

Как видно из выражения (2) (стр 33), в простых нераспространённых предложениях участвуют только два члена — подлежащее и сказуемое.

На следующем этапе была построена лингвоматематическая модель для простых распространённых предложений, которая была выражена следующим образом:

$$S_{uzbsimple2} = E_1(x_i, \dots, x_n) \cup \downarrow O_1(x_i, \dots, x_n) \cup \downarrow H_1(x_i, \dots, x_n) \cup \downarrow D_1(x_i, \dots, x_n) \cup K_1(x_i, \dots, x_n) \quad (23)$$

*Здесь, S<sub>uzbsimple2</sub>- простое нераспространённое предложение, E<sub>1</sub>-сказуемое, K<sub>1</sub>-подлежащее, O<sub>1</sub>-дополнительное, D<sub>1</sub>-определение, H<sub>1</sub>-обстоятельство, i и n – шаги изменения каждого предиката в пределах x множеств, ∪ - соединительный знак.*

Примечание,  $E1 \in x, x = \{\text{изменяется с } i \text{ до } n\}$

$K1 \in x, x = \{\text{изменяется с } i \text{ до } n\}$

Например: В этом году семинар состоялся в начале января.

Bu yilgi seminar yanvar oyining boshida o'tdi,

определение подлежащее обстоятельство сказуемое

$Bu\ yil+gi\ seminar\ yanvar\ oy+i+ning\ bosh+i+da\ o\ 't+di.$

В этом предложении есть главные и второстепенные члены предложения. Ниже приведено образование предложения **в виде численного метода** на узбекском языке:

$$S_{uzbsimple2} = \begin{cases} D1(x_i, \dots, x_n) \rightarrow R_{i1} \text{ } \_R_{i1} \oplus F_{y1} \cup \\ \cup E1(x_i, \dots, x_n) \cup \\ H1(x_i, \dots, x_n) \rightarrow R_{i1} \text{ } \_R_{i1} \oplus F_{y1} \oplus I_{f1} \text{ } \_R_{i1} \oplus F_{y1} \oplus I_{f1} \cup \\ K1(x_i, \dots, x_n) \rightarrow R_{i1} \oplus F_{y1} \end{cases}$$

Здесь,

$S_{uzbsimple2}$ -,  $E1$ - подлежащее,  $K1$ -сказуемое,  $R_{i1}$ -корень,  $F_{y1}$ -формаобразовательный суффикс,  $I_{f1}$  – словоизменятельный суффикс,  $i$  и  $n$  – шаги изменения каждого предиката в пределах  $x$  множеств,  $\cup$  - соединительный знак.

Поскольку языки тюркской языковой семьи относятся к классу агглютинативных языков, то приведенные выше лингво-математические модели применимы также к каракалпакскому и киргизскому языкам соответственно.

Математические модели узбекского языка представлены также в трудах М. Хакимова, в которых были разработаны математические модели для целевых типов предложений в узбекском языке (такие как повествовательные, вопросительные, восклицательные предложения).

В соответствии с (7) и (8) построения предложений в узбекском языке, представляем следующие (9, 10) для каракалпакского и киргизского языков:

$$S_{kksimple1} = E_2(x_i, \dots, x_n) \cup K_2(x_i, \dots, x_n) \quad (24)$$

$$S_{kksimple2} = \downarrow E_2(x_i, \dots, x_n) \cup O_2(x_i, \dots, x_n) \cup \downarrow H_2(x_i, \dots, x_n) \cup \downarrow \\ \cup \downarrow D_2(x_i, \dots, x_n) \cup \downarrow K_2(x_i, \dots, x_n) \quad (25)$$

Здесь (24 и 25),

$S_{kksimple1}$ - простое распространенное предложение на каракалпакском языке,

$S_{kksimple2}$ - и простое нераспространенное предложения в каракалпакском языках.

$E_2$ - подлежащее,

$K_2$ - сказуемое,

$O_2$ -дополнительное,

$D_2$ -определение,

$H_2$ -обстоятельство,  $i$  и  $n$  – шаги изменения каждого предиката  
внутри множества  $x$ ,

$\cup$  - знак соединения,

$\downarrow$  - знак, который может прийти, а может и не прийти.

Примечание,  $E1 \in x, x = \{\text{изменяется с } i \text{ до } n\} i \geq 1, n \geq 1$

$K1 \in x, x = \{\text{изменяется с } i \text{ до } n\} i \geq 1, n \geq 1$

$$S_{kgsimple1} = E_3(x_i, \dots, x_n) \cup K_3(x_i, \dots, x_n) \quad (26)$$

$$S_{kgsimple2} = \downarrow E_3(x_i, \dots, x_n) \cup \downarrow O_3(x_i, \dots, x_n) \cup \downarrow H_3(x_i, \dots, x_n) \cup \downarrow \\ \cup \downarrow D_3(x_i, \dots, x_n) \cup \downarrow K_3(x_i, \dots, x_n) \quad (27)$$

Здесь (26 и 27),

$S_{kgsimple1}$ - простое распространенное предложение на киргизском языке,

$S_{kgsimple2}$ - простое нераспространенное предложения на киргизском языке.

$E_3$ - подлежащее,

$K_3$ - сказуемое,

$O_3$ -дополнительное,

$D_3$ -определение,

$H_3$ -обстоятельство,  $i$  и  $n$  – шаги изменения каждого предиката  
внутри множества  $x$ ,

$\cup$  - знак соединения,

$\downarrow$  - знак, который может прийти, а может и не прийти.

Примечание,  $E1 \in x, x = \{\text{изменяется с } i \text{ до } n\} i \geq 1, n \geq 1$

$K1 \in x, x = \{\text{изменяется с } i \text{ до } n\} i \geq 1, n \geq 1$

На основе разработанных выше лингвистико-математических моделей был также разработан алгоритм выполнения трехязычного электронного перевода (рисунок б):

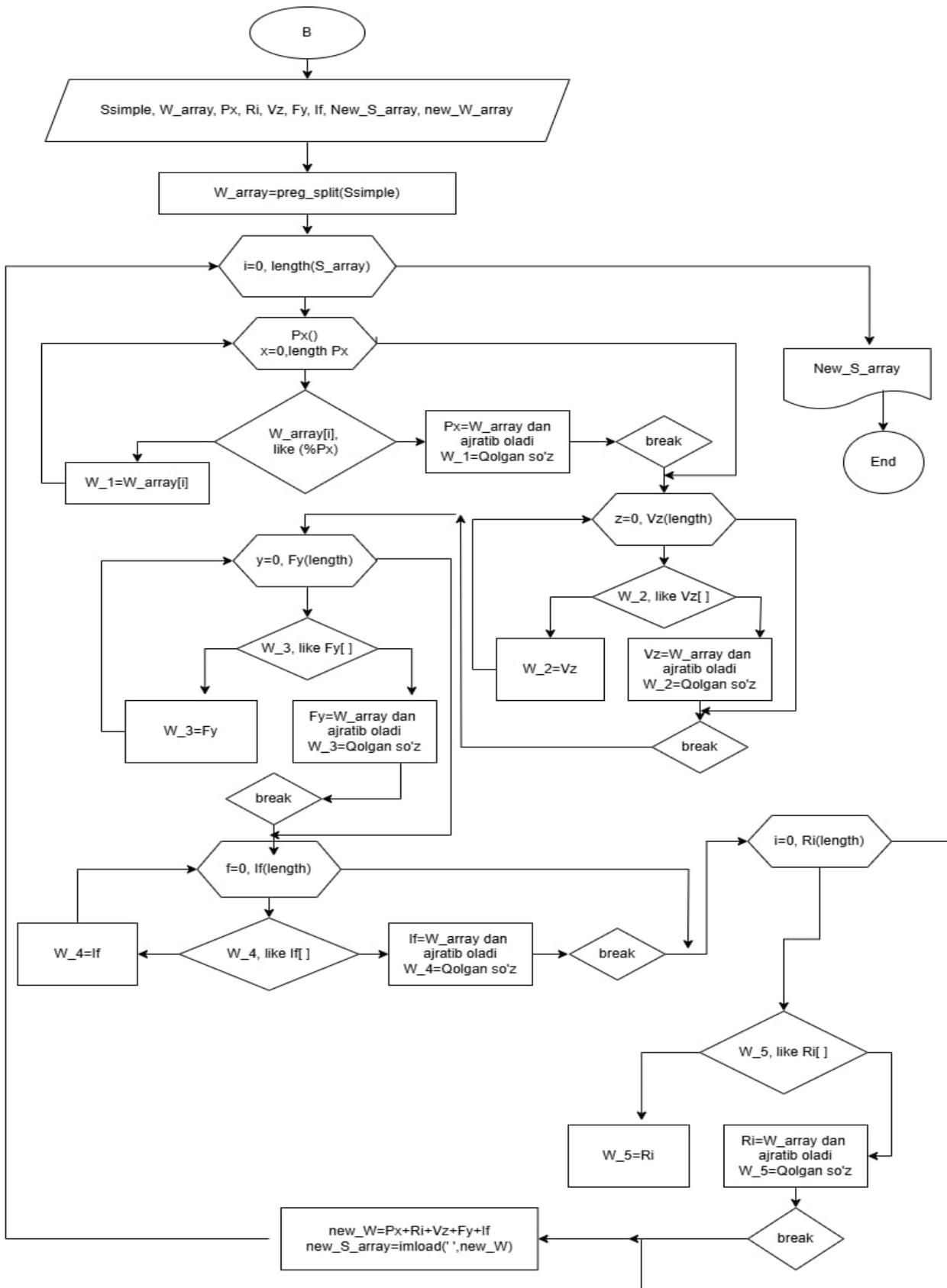


Рисунок-6. Алгоритм для трёхязычного электронного перевода в виде блок-схемы (А-алгоритм)

Если представить математическую модель и алгоритм расчета комбинированной метрики, оценивающей качество трёхязычного электронного перевода, то формула расчета METEORE будет следующей:

$$\text{METEOR} = \text{GM} \times (1 - \alpha) \quad (28)$$

$$\text{GM} = \frac{P \cdot R}{\alpha \cdot P + (1 - \alpha) \cdot R}$$

Формула для ROUGE-1 среднее гармоник GM (Harmonic Mean) – точность (P) и полнота (R);

$\alpha$  (альфа) — весовой коэффициент, регулирующий баланс между точностью и полнотой, обычно равный 0,9.

$$\text{ROUGE-1} = 2 \cdot \frac{\text{ROUGE-1} \cdot \text{Precision} \cdot \text{ROUGE-1} \cdot \text{Recall}}{\text{ROUGE-1} \cdot \text{Precision} + \text{ROUGE-1} \cdot \text{Recall}} \quad (29)$$

Объединив эти метрики, получаем следующую комбинированную метрику:

$$\text{MRS}_{\text{hb}} = \text{METEORE} + \text{ROUGE-1} + \text{Suffix}_{\text{PR}}$$

$$\text{MRS}_{\text{hb}} = \frac{P \cdot R}{\alpha \cdot P + (1 - \alpha) \cdot R} + \frac{P \cdot R}{P + R} + \text{Suffix}_{\text{PR}} \quad (30)$$

На (24) существуют параметры точности и полноты, на основе исследовательской работы вводим суффиксный параметр.

Здесь,

$\text{MRS}_{\text{hb}} - \text{METEOR}, \text{ROUGE-1}, \text{Suffix};$

*hb-hybrid (гибридная или комбинированная метрика)*

*P-точность, R-полнота и Suffix-оценивают качество перевода с учетом совпадения морфологических элементов, т.е. суффиксов. В диссертации представлены конкретные примеры оценки перевода на трёх языках.*

Вычислительный алгоритм для оценки электронного перевода:

1-шаг. Начало программы;

2-шаг. Ввод текста ( $T \downarrow$ );

3-шаг. Реализуется алгоритм А;

4-шаг. Полученный текст проверяется на основе METEOR, т.е.:

сравнивается с эталонным текстом через  $\text{METEOR} = \text{GM} \times (1 - \alpha)$

( $T \downarrow = T \uparrow$ ), результаты подсчитываются и делается следующий шаг;

5-шаг.  $T \downarrow$  рассчитывается на основе ROUGE-1, т.е. вычисляется 1-грамм и делается следующий шаг;

6-шаг.  $T \downarrow = T \uparrow$  в котором сравниваются суффиксы и вычисляется среднее арифметическое;

7-шаг. Среднее арифметическое значений METEOR+ROUGE+SUFFIXPR вычисляется и ассимилируется в  $\text{MRS}_{\text{hb}}$ , после чего выполняется переход к следующему шагу.

Шаг 8. Выводится  $\text{MRS}_{\text{PR}}$ .

Шаг 9. Программа завершается.

*Примечание: алгоритм А -приведено на 6 рисунке);  $T \downarrow$ - входящий текст;  $T \uparrow$ - исходящий текст;*

$MRS_{hb}$ -METEOR, ROUGE, SUFFIX<sub>Precision, Recall</sub>, *hb*- аббревиатура слова «гибрид».

Этот алгоритм вычисляет процент контекстной, лексической и грамматической точности полученного перевода.

Четвертая глава диссертации: «Функциональный модуль и программный комплекс трехязычного электронного перевода, относящийся к тюркской языковой семье» разработаны инфологическая модель базы данных электронного переводчика, электронный словарь и функциональный модуль переводчика, а также инструкция по использованию программы.

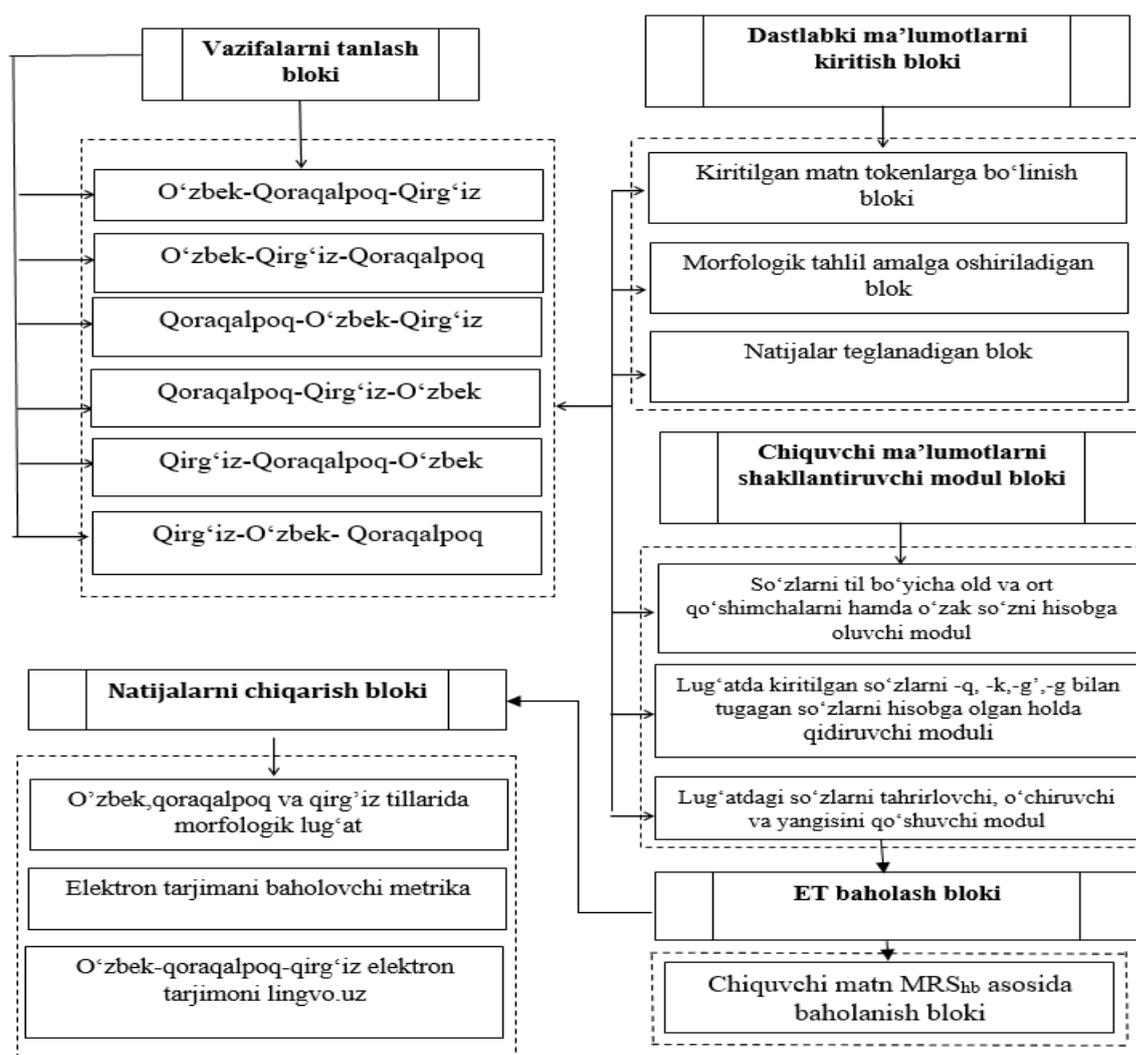


Рисунок-7. Функциональная схема программного обеспечения

В исследовательской работе электронный перевод состоит из следующих блоков (рис.-6):

1.Блок выборки функций: в этом блоке выбирается задача, после префикс, аффикс и корень который учитывает все эти морфемы;

2.Блок формирующий исходящие данные: здесь осуществляется редактирование, удаление и добавление нового выбранного слова из электронного словаря;

3. Блок образующая исходных данных;

4. Блок сопоставления объектов и оценки электронного перевода;

5. Блок вывода результатов.

Разработан программное средство для электронного перевода на основе лингво-математических моделей трёхязычных языках (узбекский, каракалпакский и киргизский). Лингво-математические модели и программное средство могут быть использованы при разработке электронного перевода для других тюркских языков.

## ЗАКЛЮЧЕНИЕ

В результате исследования, проведенного по диссертационной работе на тему «Лингво-математические модели и вычислительные алгоритмы для трёхязычного электронного перевода относящихся к тюркским семействам», представлены следующие выводы:

Проведен лексико-морфологический анализ узбекского, каракалпакского и киргизского языков, разработана лингвоматематическая модель образования простых и сложных словосочетаний. В результате она стала основой для контекстно-целесообразного построения простых и сложных слов в узбекском, каракалпакском и киргизском языках;

Разработаны лингвоматематические модели процесса образования простых и сложных предложений для узбекского, каракалпакского и киргизского языков. В результате данные лингвоматематические модели послужили для оценки качества электронного перевода последовательности построения предложений на узбекском, каракалпакском и киргизском языках;

На основе лингвоматематических моделей простых и сложных слов и структуры предложения разработаны лингвоматематические модели и алгоритмы процессов образования простых и сложных слов и структуры предложения с учетом времен глаголов для узбекского, каракалпакского и киргизского языков. В результате проведенных работ удалось повысить контекстуальную корректность перевода текста при электронном переводе в 1-2 раза;

На основе базы данных трехязычного электронного словаря и электронного переводчика для узбекского, каракалпакского и киргизского языков разработана комбинированная метрика и алгоритм ее расчета для оценки электронного перевода на узбекский, каракалпакский и киргизский языки для поиска слов научно-технической тематики и определения взаимопереводимости текстов. В результате качество оценки трехязычного электронного перевода повысилось на 78,2%;

По результатам проведенных научно-исследовательских работ составлены акты в ООО «BMW», ООО «TCW» и АО «SHARK». На основании актов результаты исследований внедрены в практику на предприятиях сектора UzINFOCOM, UNICON и RTSIRITI, в результате чего точность трехязычного электронного перевода повышена на 86,7%.

**SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES  
DSc.13/30.12.2019.T.07.01 AT TASHKENT UNIVERSITY OF  
INFORMATION TECHNOLOGIES**

---

**TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES NAMED  
AFTER MUHAMMAD AL-KWAREZMI**

**UZAKOVA MAMURA ABDURAYIMOVNA**

**LINGUO-MATHEMATICAL MODELS AND COMPUTATIONAL  
ALGORITHMS FOR TRILINGUAL ELECTRONIC TRANSLATION OF  
TURKIC LANGUAGES FAMILY**

05.01.07 – Mathematical Modeling. Numerical methods and set of applications

**DISSERTATION ABSTRACT OF THE DOCTOR OF PHILOSOPHY (PHD)  
ON TECHNICAL SCIENCES**

The theme of doctor of philosophy (PhD) on technical sciences was registered at the Supreme  
ation Commission at the Ministry of Higher education, Science and Innovations of  
the republic of Uzbekistan under number № B2025.1.PhD/T5300

The dissertation has been prepared at Tashkent University of Information Technologies.

The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)) on  
the website (www.tuit.uz) and on the website of "Ziyonet" Information and educational portal  
(www.ziyonet.uz)

**Scientific adviser:** Nazirova Elmira Shodmonovna  
doctor of technical sciences, professor

**Official opponents:** **Muhammadiyeva Dildora Kabilovna**  
doctor of technical sciences, associate professor  
**Bakayev Ilhom Izatovich**  
doctor of philosophy technical sciences,  
associate professor

**Leading organization:** Karshi State university

The dissertation defense will be held at the meeting of the Academic Council of the Tashkent  
University of Information Technologies DSc.13/30.12.2019.T.07.01 on "30" 10 at 16<sup>00</sup> in  
2025. (Address: 100084, Tashkent, Amir Temur Street, 108.

Tel.: (99871) 238-64-43, e-mail: iktuit@tuit.uz).

The dissertation can be viewed at the Information Resource Center of the Tashkent University of  
Information Technologies named after Muhammad al-Khwarizmi (registered under number №372).

Address: 100084, Tashkent, Amir Temur Street, 108. Tel.: (99871) 238-64-70.

The dissertation abstract was distributed on "18" 10, 2025.

(Registration report number 31 dated "18" 10, 2025.)



*M.M. Musaev*

**M.M. Musaev**  
Chairman of the scientific council  
awarding scientific degrees,  
doctor of technical sciences, professor

*J.X. Djumanov*

**J.X. Djumanov**  
Scientific secretary of scientific council  
awarding scientific degrees,  
doctor of technical sciences, professor

*F.M. Nuraliyev*

**F.M. Nuraliyev**  
Chairman of the academic  
seminar under the scientific council  
awarding scientific degrees,  
doctor of technical sciences, professor

## INTRODUCTION (abstract of PhD dissertation)

**The purpose of the research** is to develop linguo-mathematical model and algorithms, software package for trilingual electron translation Uzbek, Karakalpak and Kyrgyz language.

**The object of the research** is the linguistic corpus of Uzbek, Karakalpak and Kyrgyz languages.

**The subject of the research** is the development of a linguo-mathematical model, algorithms and a software package for implementing trilingual electronic translation for Turkic languages as well as the development of a mathematical model and computational algorithm for a combined metric to evaluate trilingual electronic translation.

**The scientific novelty of the study is as follows:**

based on the lexical and morphological base of the Uzbek, Karakalpak and Kyrgyz languages, linguo-mathematical models and algorithms for the formation of complex words and sentence structures have been developed for the Uzbek, Karakalpak and Kyrgyz languages;

based on linguo-mathematical models of the formation of simple and complex words and sentence structures, linguomathematical models and algorithms for simple word formation processes have been developed for the phonetic structure of the Uzbek, Karakalpak and Kyrgyz languages;

based on linguo-mathematical models of simple and complex words and sentences, linguomathematical models and algorithms for simple and complex word formation processes and sentence structures have been developed for the Uzbek, Karakalpak and Kyrgyz languages, taking into account verb tenses;

a trilingual electronic dictionary for the Uzbek, Karakalpak and Kyrgyz languages and an electronic translator database for searching scientific and technical words and determining the mutual translation of texts were developed, as well as a combined metric and a computational algorithm for calculating it for evaluating electronic translation into the Uzbek, Karakalpak and Kyrgyz languages.

**Implementation of research results:** based on the linguo-mathematical models, algorithms and the created software package developed as a result of the research conducted within the framework of the dissertation:

the software package developed based on the algorithm and mathematical model for constructing complex words in the Uzbek language was introduced to the “The computer workshop” LLC publishing house (reference № 33-8/4723 of the Ministry Digital Technologies Republic of Uzbekistan dated July 7, 2025). The results obtained were used to analyze the Uzbek words in the “Uzmorphoanalyzer - a program for morphological analysis of Uzbek words” program.

the software package developed based on the algorithm and linguo-mathematical model for constructing simple sentences in the Uzbek, Karakalpak and Kyrgyz languages was introduced to the “Big Makro Word” LLC editorial office for translating scientific and technical texts (reference № 33-8/4723 of the Ministry Digital Technologies Republic of Uzbekistan dated July 7, 2025). As a result, the total amount of work was reduced by 1-2 times in terms of time;

the software developed on the basis of the algorithm for forming complex words in the Uzbek, Karakalpak and Kyrgyz languages and the linguistic-mathematical model was implemented in the joint-stock company “SHARQ” publishing house to eliminate shortcomings in the translation of texts (reference № 33-8/4723 of the Ministry Digital Technologies Republic of Uzbekistan dated July 7, 2025). As a result, an increase in the overall efficiency of work by 78.2% was achieved.

**Structure and volume of the dissertation.** The dissertation consists of an introduction, four chapters, a conclusion, a list of references, and appendices. The dissertation has a total of 110 pages.

**E'LON QILINGAN ISHLAR RO'YXATI**  
**СПИСОК ОПУБЛИКОВАННЫХ РАБОТ**  
**LIST OF PUBLISHED WORKS**

**I bo'lim (I часть, I part)**

1. Э. Назирова, Ш.Абидова, М.Узакова “Туркий тиллар учун икки тилли электрон таржиманинг модел ва алгоритмлари” Мухаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университети. Монография. ББК:22.18:81.2 Тошкент-2022 й. «Бизнес-полиграф».- 102 б.
2. Elmira Nazirova, Shakhnoza Abidova, Khurshida Karabayeva, Mamura Uzakova Mathematical Software For Creating A Trilingual Electronic Translator Between The Languages Of The Turkish-Language Family. Journal of Computational Analysis and Applications, Volume 33, ISSN 15211398, 15729206, SJR 0.19, United States-2024. -P.220-223 (05.00.00; №2)
3. Uzakova M.A. Turkiy tillar oilasiga mansub qirg'iz tilida so'z birikmalarini hosil qilishning lingvo-matematik modeli va algoritmi “TATU xabarlarlari” ilmiy-texnika va axborot-tahliliy jurnali. №2 (70), Toshkent-2024. – В 75-87. <http://jurnal.tuit.uz> (OAK rayosatining 05.00.00; №31).
4. Nazirova Elmira Shodmonovna, Abidova Shakhnoza Bakhodirovna, Uzakova Ma'mura Abdurayimovna, Boymurodov Farrukh Farhod Mathematical model for evaluation of trilingual electronic translation J.:“Bulletin of TUIT: Management and Communication Technologies”. <https://doi.org/10.61663/244tuitmct10> (05.00.00; OAK rayosat qarori 30.07.2020. № 4(40) 01-10/1103), Toshkent-2024. -P-7.
5. Nazirova Elmira, Abidova Shakhnoza, Usmanova Kamola Creating open-corres parallel corpus bases for machine translation J.:“Bulletin of TUIT: Management and Communication Technologies”. <https://doi.org/10.61663/251tuitmct5>. (05.00.00; OAK rayosat qarori 30.07.2020. № 01-10/1103) 2025. -P -7.

**II bo'lim(II часть, II part)**

6. E. Nazirova, Sh.Abidova, M. Uzakova “Mathematical models of complex word formation in uzbek, karakalpak and kyrgyz languages belonging to the turkic language family” // International conference “Modern Problems of Applied Science and Engineering” Samarkand– 2024, -P .030043-1-030043-7.
7. Elmira Nazirova, Shakhnoza Abidova, Yusuf Kholmuminov, Farrukh Boymurodov “Linguistic-mathematical model of kyrgyz and karakalpak languages implementing bilingual electronic translation with feedback” // International scientific conference “Modern Problems and Perspectives of Applied Mathematics”. AIP Conference Proceedings. (ISSN 0094-243X, E-ISSN 1551-7616) Karshi-2025. – P.050001-1 - 050001-6

8. E.Sh. Nazirova, M.A. Uzakova Efficient construction of suitable trees for syntactic analysis of Uzbek sentences. International Scientific Forum. Proceedings of the multi-disciplinary scientific - practical conference, Tashkent-2023. -P.13-15.

9. E.Sh. Nazirova, Sh.A. Abidova, M.A. Uzakova Turkiy tillari tizimi tabiiy tillarini qayta ishlash modeli o'zbek tili misolida // "Kompyuter lingvistikasining zamonaviy texnologiyalari CTCL – 2023" mavzusidagi vazirlik miqyosidagi ilmiy-amaliy anjuman, www.myscience.uz sertifikat №7606 Toshkent-2023. -B.153-157

10. Abidova Sh.B., Uzakova M.A Yangi avlod o'quv lug'atlari va ularning mobil ilovalarini yaratish O'zbekiston Respublikasi Raqamli Texnologiyalari Vazirligi Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti "Matematik modellashtirish algoritmlash va dasturlashning dolzarb muammolari mavzusidagi respublika ilmiy-texnik konferensiyasi" UO'K 510.5, 519.768.2. Tashkent- 2023 yil, -B.102-107

11. E. Nazirova, N. Abduraxmanova, Sh. Abidova, M.Uzakova "Morphological analysis of word formation in the Uzbek, Karakalpak and Kyrgyz languages belonging to the Turkic family of languages" 11th th International Conference on Computer Processing of Turkic Languages, "TurkLang-2023", ББК 81.1, ISBN 978-5-9690-1197 Bukhara-2023. -P.152-155.

12. E. Nazirova, F.Boymurodov, M. Uzakova, O.Tuyakov "O'zbek tili so'zlarini imloviy tahrirlash va qayta ishlash zamonaviy algoritmi" amaliy muammolar va istiqbollari Respublika ilmiy - amaliy konferensiya, 24-25 may. Qarshi-2024. -B 321-323.

13. E. Nazirova, Sh. Abidova, M. Uzakova, Y.Xolmo'minov "Teskari aloqaga ega bo'lgan ikki tili elektron tarjimoni amalga oshiruvchi qirg'iz vaqt qilmoq tillarning eng va matematik modeli" amaliy matematikaning zamonaviy muammolar va istiqbollari Respublika ilmiy - amaliy konferensiya, Qarshi-2024 yil, -B.323-325.

14. E. Nazirova, F.Boymurodov, M. Uzakova, Y.Xolmo'minov "Turkiy tillar oilasiga mansub o'zbek va qirg'izistonlarida elektron tarjimonning lima matematik modelini ishlab chiqish" amaliy matematikaning zamonaviy muammolar va istiqbollari Respublika ilmiy - amaliy konferensiya, Qarshi-2024 yil, -B. 326-328.

15. E.Sh. Nazirova, Sh.B. Abidova, M.A. Uzakova "O'zbek tilida fonetik tahlil jarayonining lingvo-matematik model va algoritmlari" // "Scientific foundations of raising the use of information technologies to a new level and modern problems of automation" III-international scientific conference Tashkent-2024. -P.486-493.

<https://doi.org/10.5281/zenodo.14214728>

16. E.Sh. Nazirova, M.A. Uzakova Linguo mathematical models of sentences by types of Uzbek language for electronic translation Scientific review of the problems and prospects of modern science and education. Volume №1. Germany-2025 -P.4-8.

<https://zenodo.org/badge/DOI/10.5281/zenodo.15049689.svg>

17. Nazirova Elmira Shodmonovna, Abidova Shakhnoza Baxadirovna, Uzakova Mamura Abduraimovna, Kholmuminov Yusuf Kholbuta o'g'li "Linguistic mathematical models and algorithms of the process phonetic analysis in the uzbek language" // Science and education: development vectors. UDK 378 (063) BBK 74.58 Astana, Kazakhstan-2025. – C. 11-14.
18. E. Nazirova M. Uzakova Turkiy tillar uchun uch tilli tarjimon dasturi. O'zbekiston Respublikasi Adliya vazirligining Elektron hisoblash mashinalari uchun yaratilgan dasturni rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma № DGU 41673. 30.07.2024
19. E. Nazirova M. Uzakova, Sh.Abidova, Yu.Xolmuminov Uch tilli elektron tarjimon. O'zbekiston Respublikasi Adliya vazirligining Elektron hisoblash mashinalari uchun yaratilgan dasturni rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma № DGU 46782. 08.01.2025
20. E. Nazirova M. Uzakova, Sh.Abidova, Yu.Xolmuminov Uch tilli elektron tarjimon uchun ma'lumotlar bazasi. O'zbekiston Respublikasi Adliya vazirligining Ma'lumotlar bazasining rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma № BGU 1810. 02.01.2025

Avtoreferat “Bulletin of TUIT: management and Communication Technologies: Managment and Communication” ilmiy jurnali tahririyatida tahrirdan o‘tkazildi va o‘zbek, rus, ingliz tillaridagi matnlarini mosligi tekshirildi.