

MINISTRY OF HIGHER EDUCATION, SCIENCE AND
INNOVATIONS OF THE REPUBLIC OF UZBEKISTAN

NATIONAL UNIVERSITY OF UZBEKISTAN NAMED AFTER
MIRZO ULUGBEK



RAQAMLI TEXNOLOGIYALAR VA
SUN'YI INTELLEKTNI RIVOJLANTIRISH
ILMIY-TADQIQOT INSTITUTI



INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE

PROCEEDINGS
of materials on topic:

«CONTEMPORARY TECHNOLOGIES OF
COMPUTATIONAL LINGUISTICS - CTCL.2025»



Crossref
we use DOIs



Google
scholar

OPEN ACCESS



TASHKENT - 2025

**“KOMPYUTER LINGVISTIKASINING ZAMONAVIY TEXNOLOGIYALARI – CTCL.2025”
NOMLI XALQARO ILMIY-AMALIY KONFERENSIYASINING ILMIY
MAQOLALARI VA TEZISLAR TO‘PLAMI. T.: O‘zMU, 2025.
2025-yil 25-aprel**

TASHKILIY QO‘MITA TARKIBI:

N.Z.Abduraxmonova.	O‘zMU Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası mudiri, f.f.d., professor.
G.B.Madiyeva.	Al-Farobiy nomidagi Qozoq Milliy universiteti professori, f.f.d.
Z. T.Taxirov.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası professori, f.f.n.
A.R.Iskandarova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası dotsenti, PhD.
D.A.Fattaxova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası katta o‘qituvchisi, PhD.
U.R.Xalmuxamedova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası katta o‘qituvchisi.
N.Sh.Inogamova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası katta o‘qituvchisi.
F.R.Topildiyeva.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası katta o‘qituvchisi.
D.K.Xudoyberganova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası katta o‘qituvchisi, PhD.
Z.M.Kasimova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası katta o‘qituvchisi, PhD.
J.N.O‘rozov.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası katta o‘qituvchisi, PhD.
G.G‘.Xoliyorova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
F.M.Aliyev.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
M.S.Jo‘raqulova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
Z.G.Qodirova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
M.F.Abdunazarova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
D.Xolbekova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
D.A.Mirvaliyeva.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
N.K.Tursunova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
U.M.Askarova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
S.A.Abduraxmonova.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası o‘qituvchisi.
R.T.O‘rinboyeva.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası magistranti.
L.A.Boboxo‘jayeva.	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası magistranti.
M.Sh.Xudoyberganova	Kompyuter lingvistikasi va amaliy tilshunoslik kafedrası magistranti.

Mazkur xalqaro ilmiy-amaliy konferensiya to'plami kompyuter lingvistikasining ayrim yo'nalishlari, jumladan, korpus lingvistikasi, til ta'limida kompyuter texnologiyalari, NLPda til va nutq tahlili (lingvistik analizator: morfologik, sintaktik va semantik tahlil, nutq analizi va sintezi), mashina tarjimasining lingvistik va dasturiy ta'minoti, kompyuter leksikografiyasi: ontologiya, tezaurus, shuningdek, zamonaviy tilshunoslik masalalari kabi mavzular tahlilga tortilgan maqola va tezislardan tashkil topgan.

Ushbu ilmiy maqola va tezislar to'plami kompyuter lingvistikasi, kompyuter texnologiyalari, amaliy filologiya kabi yo'nalishlarda ilmiy izlanishlar olib borayotgan professor-o'qituvchilar, doktorantlar, mustaqil izlanuvchilar, magistr va talabalar uchun mo'ljallangan.

To'plamga kiritilgan maqolalar va tezislar mazmuni hamda sifatiga mualliflar mas'uldir.

MAS'UL MUHARRIR:

prof., f.f.d. N.Z.Abduraxmonova

TAHRIR HAY'ATI:

f.f.n., prof. Z. T.Taxirov

f.f.n., dots. A.R.Iskandarova

ped.f.n., dots. D.A.Fattaxova

PhD D.K.Xudoyberganova

PhD Z.M.Kasimova

**СБОРНИК НАУЧНЫХ СТАТЕЙ И ТЕЗИСОВ
МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ
“СОВРЕМЕННЫЕ ТЕХНОЛОГИИ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ – CTCL.2025”.**
Т.: НУУз, 2025.
25 апреля 2025 г.

СОСТАВ ОРГАНИЗАЦИОННОГО КОМИТЕТА:

<i>Н.З. Абдурахмонова.</i>	Заведующий кафедрой компьютерной и прикладной лингвистики НУУз, д.ф.н., профессор.
<i>Г.Б.Мадиева.</i>	Профессор Казахского Национального университета имени аль-Фараби, д.ф.н.
<i>З. Т. Тахиров.</i>	Профессор кафедры компьютерной и прикладной лингвистики, к.ф.н.
<i>А.Р.Искандарова.</i>	Доцент кафедры компьютерной и прикладной лингвистики, к.ф.н.
<i>Д.А. Фаттахова.</i>	Старшая преподавательница кафедры компьютерной и прикладной лингвистики, к.пед.н.
<i>У.Р. Халмухамедова.</i>	Старшая преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Н. Ш. Иногамова.</i>	Старшая преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Ф.Р. Топилдиева.</i>	Старшая преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Д.К.Худойберганава.</i>	Старшая преподавательница кафедры компьютерной и прикладной лингвистики, к.ф.н.
<i>З.М. Касимова.</i>	Старшая преподавательница кафедры компьютерной и прикладной лингвистики, к.ф.н.
<i>Ж.Н.Урозов.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики, к.ф.н.
<i>Г.Г. Холёрова.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Ф.М. Алиев.</i>	Преподаватель кафедры компьютерной и прикладной лингвистики.
<i>М.С. Джуракулова.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>З.Г. Кадирова.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>М.Ф. Абдуназарова.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Д. Холбекова.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Д.А. Мирвалиева.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>К.Р. Ибодуллаева.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Н.К. Турсунова.</i>	Преподаватель кафедры компьютерной и прикладной лингвистики.
<i>У.М. Аскарова.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>С.А. Абдурахмонова.</i>	Преподавательница кафедры компьютерной и прикладной лингвистики.
<i>Р.Т.Уринбоева.</i>	Магистрант кафедрой компьютерной и прикладной лингвистики.
<i>Л.А.Бобоходжаева.</i>	Магистрант кафедрой компьютерной и прикладной лингвистики.
<i>М.Ш.Худойберганава</i>	Магистрант кафедрой компьютерной и прикладной лингвистики.

Материалы международной научно-практической конференции, включающие статьи и тезисы, посвящены некоторым направлениям компьютерной лингвистики, в том числе корпусной лингвистике, компьютерным технологиям в языковом образовании, анализу языка и речи в NLP (Лингвистический анализатор: морфологический, синтаксический и семантический анализ, анализ речи и синтез), Лингвистические базы данных и программное обеспечение машинного перевода, Компьютерная лексикография: онтология, тезаурус, а также проблемы современной лингвистики.

Сборник научных статей и диссертаций предназначен для профессоров и преподавателей, докторантов, независимых исследователей, магистров и студентов, ведущих научные исследования в таких областях, как компьютерная лингвистика, компьютерные технологии и прикладная лингвистика.

Авторы несут ответственность за содержание и качество статей и тезисов, включенных в книгу.

ОТВЕТСТВЕННЫЙ РЕДАКТОР:

д.ф.н., проф. Н.Абдурахмонова

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

PhD, проф. З.Тахиров

PhD, доц. А.Искандарова

PhD З.Касимова

PhD Д.Фаттахова

PhD Д.Худойбергана

PROCEEDINGS OF ARTICLES AND THESIS
OF THE INTERNATIONAL SCIENTIFIC -PRACTICAL CONFERENCE TITLED
«CONTEMPORARY TECHNOLOGIES OF COMPUTATIONAL LINGUISTICS-CTCL.2025»
T.: NUUz, 2025.
April 25, 2025.

ORGANIZING COMMITTEE:

N. Z. Abdurakhmonova.	Head of Department of Computational and Applied Linguistics of UzMU, DSc, Professor.
G.B.Madiev.	Professor of the Al-Farabi Kazakh National University, Doctor of Philological Sciences
Z. T. Takhirov.	Professor of the Department of Computational and Applied Linguistics, PhD.
A.R. Iskandarov.	Associate Professor of the Department of Computational and Applied Linguistics, PhD.
D.A. Fattakhova.	Senior teacher of the Department of Computational and Applied Linguistics, PhD.
U.R. Khalmukhamedova.	Senior teacher of the Department of Computational and Applied Linguistics.
N. Sh. Inogamova.	Senior teacher of the Department of Computational and Applied Linguistics.
F.R. Topildieva.	Senior teacher of the Department of Computational and Applied Linguistics.
D.K. Khudoyberganova.	Senior teacher of the Department of Computational and Applied Linguistics, PhD.
Z.M. Kasimova.	Senior teacher of the Department of Computational and Applied Linguistics, PhD.
J.N.Urazov.	Senior teacher of the Department of Computational and Applied Linguistics, PhD.
G.G. Kholiyorova	Senior teacher of the Department of Computational and Applied Linguistics.
F.M. Aliyev.	Teacher of the Department of Computational and Applied Linguistics.
M.S. Jurakulova.	Teacher of the Department of Computational and Applied Linguistics.
Z.G. Kadirova.	Teacher of the Department of Computational and Applied Linguistics.
M.F. Abdunazarova.	Teacher of the Department of Computational and Applied Linguistics.
D. Kholbekova.	Teacher of the Department of Computational and Applied Linguistics.
D.A. Mirvaliyeva.	Teacher of the Department of Computational and Applied Linguistics.
N.K.Tursunova.	Teacher of the Department of Computational and Applied Linguistics.
U.M. Askarova.	Teacher of the Department of Computational and Applied Linguistics.
S.A. Abdurakhmonova.	Teacher of the Department of Computational and Applied Linguistics
R.T.Urinboeva.	Master student Departmentof Computational and Applied Linguistics.
L.A.Bobokhodzhaeva.	Master student Departmentof Computational and Applied Linguistics.
M.Sh.Khudoyberganova	Master student Departmentof Computational and Applied Linguistics

This proceedings of international scientific-practical conference including articles and theses deals with some directions of Computational Linguistics, including Corpus Linguistics, Computer Technologies in language education, Language and speech analysis in NLP (Linguistic analyzer: morphological, syntactic and semantic analysis, speech analysis and synthesis), Linguistic database and software of Machine Translation, Computer Lexicography: ontology, thesaurus, as well as Modern Linguistic Issues.

The proceedings of scientific articles and theses is intended for professors and teachers, doctoral students, independent researchers, masters and students who are conducting scientific research in such areas as Computational Linguistics, Computer Technologies, and Applied Linguistics.

The authors are responsible for the content and quality of the articles and theses included in the book.

RESPONSIBLE EDITOR:

DSc, prof. N.Abdurakhmonova

EDITORIAL BOARD:

PhD, prof. Z.Taxirov

PhD, dots. A.Iskandarova

PhD Z.Kasimova

PhD D.Fattaxova

PhD D.Xudoyberganova

MUNDARIJA

TABLE OF CONTENTS

N. ABDURAXMONOVA, U. TULIYEV, G. SHAMIYEVA. PARATRANSLATOR – PARALLEL KORPUSGA ASOSLANGAN ELEKTRON KONTEKSTOLOGIK TARJIMA LUG‘AT PLATFORMASINI YARATISHNING AYRIM MASALALARI XUSUSIDA	15
D. ZHUPARBEEK, S.S. NURKENOVA. THE USE OF WEBQUESTS IN FOREIGN LANGUAGE TEACHING TO FOSTER STUDENTS’ COMMUNICATIVE COMPETENCE	21
M. AMIRSHEIBANI, O.A. YO‘LDOSHEV. A CORPUS-BASED DESCRIPTION OF GRAMMATICAL ERRORS IN UZBEK EFL LEARNER WRITING	26
Н.А. ИСМАЙЛОВА. КОМПЬЮТЕРЛІК ЛИНГВИСТИКА ЖӘНЕ ЖАСАНДЫ ИНТЕЛЛЕКТ: ТІЛДІК КЕДЕРГІЛЕРДІ ЖЕҢУДІҢ ЗАМАНАУИ ТӘСІЛДЕРІ	33
Д. ТОҚМЫРЗАЕВ, А. ПИРМАНОВА. СОЗДАНИЕ ПОЭТИЧЕСКОГО ЧАСТОТНОГО СЛОВАРЯ СЛОВ В СТИХАХ АБАЯ В СООТВЕТСТВИИ С КОРПУСОМ, АВТОМАТИЗИРОВАННЫЙ ЧАСТОТНЫЙ СЛОВАРЬ (1-Й ЭТАП)	35
Г. МАДИЕВА, К. ПИРМАНОВА. СОЗДАНИЕ И АНАЛИЗ ЛИТЕРАТУРНО-ЯЗЫКОВОГО КОРПУСА СТИХОТВОРЕНИЙ АБАЯ.....	40
А. ОҢАЛБЕК, А. УӘЛИХАН. ПОТЕНЦИАЛ ИСПОЛЬЗОВАНИЯ ПОЭТИЧЕСКОГО КОРПУСА АБАЯ В ИССЛЕДОВАТЕЛЬСКИХ, ОБРАЗОВАТЕЛЬНЫХ И МЕТОДИЧЕСКИХ ЦЕЛЯХ	44
M.I. SUYUNOVA. SEMANTIC AMBIGUITY PROCESSING IN HUMAN COGNITION: A COMPARATIVE ANALYSIS OF NLP MODELS AND PSYCHOLINGUISTIC APPROACHES	55
M.YA. URAZALIYEVA. MACHINE LEARNING MODELLARI ASOSIDA O‘ZBEK TILIDAGI AUDIOMATNLARNI TAHLIL QILISH: KUN.UZ MISOLIDA.....	58
S. TURSUNBOYEVA. MILLIY KONTENT YARATISHDA ADABIY TIL ME‘YORLARIGA RIOYA QILISH.....	62
N.Z. ABDURAXMONOVA, K.CH. RAJABOVA. MULTIMEDIA KORPUSLARINING TIL O‘QITISHDAGI O‘RNI VA AHAMIYATI.....	66
R.S. SHAMSIYEVA. CHET ELLIK TALABALARNI O‘QITISHDA KORPUSGA ASOSLANGAN YONDASHUV	71
A.Q. ABDULLAYEV. SENTIMENT TAHLILDA TASNIFLASH SXEMASI	75
N.Z. ABDURAXMONOVA, D.G. ALLABERDIYEVA. QORAQALPOQ TILI ELEKTRON KORPUSINI YARATISH: NAZARIY, AMALIY VA STRATEGIK ASOSLAR	79
U.T. TURAYEVA. O‘ZBEK VA TURK TILLARIDA JOY NOMLARINING IFODALANISHI.....	83
G.SH. SAIDOVA, D.A. ABDUSATTAROVA. O‘ZBEK TILIDAGI METAFORALARNING INGLIZ TILIGA TARJIMASIDA TARJIMA MUAMMOLARI.....	89
M.S. SHARIPOV, L.U. KURBANOVA, M. ATANAZAROVA. XORAZM SHEVALARI SINONIMIK LUG‘ATINING MA‘LUMOTLAR BAZASINI ISHLAB CHIQISH.....	93
S.A. TURSUNOVA. JAHON TILSHUNOSLIGIDA PRAGMALINGVISTIKA VA PRAGMATIK TAHLIL	96

K. IBODULLAYEVA. DISKURS TUSHUNCHASI VA UNING INTERTEKSTUALLIK HODISASI BILAN BOG‘LANISHI	100
X.U. MAXSUDOVA. MASHINASOZLIK TERMINOLOGIYASINING SHAKLLANISH BOSQICHLARI	104
M. SHAYDULLAYEVA, M. FAZLIDDINOVA, X. ISMANQULOVA, A. MAMARAYIMOVA. O‘ZBEK TILINING MORFOLOGIK ANALIZATORINI TADQIQ ETISH VA YARATISH	107
G.A. SHAMSIYEVA. MASHINA TARJIMASI UCHUN TARJIMA XOTIRASINI YARATISH TAMOYILLARI	110
У.С. ЙИГИТАЛИЕВ, Ш.У. УМИДЖОНОВ. ВЫРАЖЕНИЕ ЯЗЫКОВОГО АТОМА «РОД» НА ГРАММАТИЧЕСКОМ УРОВНЕ УЗБЕКСКОГО ЯЗЫКА.....	114
M.F. ABDUNAZAROVA. PRAGMALINGVISTIKA VA UNDA DISKURS TUSHUNCHASI	117
A. ISKANDAROVA. “QIYOMAT” ROMANIDAGI FRAZEMALARNI PARALLEL KORPUSDA LINGVISTIK TEGGLASH MASALASI	121
Z.M. KASIMOVA. SEMANTIC-STRUCTURAL ANALYSIS OF EXPRESSIONS OF THE BEGINNING OF THE XX CENTURY	125
D.D. FATTAKHOVA, D.A.MIRVALIYEVA. THE FORMATION AND DEVELOPMENT OF LINGVODIDACTICS	129
И. ДЖУМАНИЯЗОВА. ОСОБЕННОСТИ ПОСТРОЕНИЯ И ФУНКЦИОНИРОВАНИЯ АВТОРСКИХ КОПИУСОВ	137
G.G‘. XOLIYOROVA. O‘ZBEK TILI TEZAURUSI SEMANTIK MAYDONLARI VA STRUKTURASI.....	141
S.K. SA‘DULLAYEV. O‘ZBEK TILIDAGI IJTIMOIIY TARMOQ MATNLARI SENTIMENT TAHLILI	145
M.Q. JALG‘ASBEKOVA. O‘ZBEK ADABIY TILINING QOZOQ TILIGA TA‘SIRI.....	149
A.A. MOHICHEHRA, N.Z. ABDURAXMONOVA. O‘ZBEK TILINING SIFAT LEKSEMALARIDA KONTEKSTOLOGIK SINONIMLAR BAZASINI YARATISH MASALALARI.....	152
S. OTAMURODOVA. POLYSEMOUS NATURE OF HIGH-FREQUENCY PHRASAL VERBS	160
G.D. YO‘LDOSHBOYEVA. O‘ZBEK TILIDAGI ZAMONAVIY TARJIMA TIZIMLARINI YARATISHNING DOLZARB MASALALARI	164
Г. ХИДИРОВА. ПРИМЕНЕНИЕ КОРПУСНОГО ПОДХОДА В ИССЛЕДОВАНИЯХ ЯЗЫКА	169
M.J. JO‘RABOYEVA, O. HAZRATQULOV. ZAMONAVIY TILSHUNOSLIK VA TARJIMASHUNOSLIKNING DOLZARB MUAMMOLARI.....	173
M.B. MUXAMADKULOVA, A.R. ISKANDAROVA. O‘ZBEK VA TATAR MILLIY KORPUSLARI TOKSONOMIYASI.....	176
R.T. O‘RINBOYEVA, U.T. TO‘RAYEVA. NUTQ SINTEZATORI UCHUN MO‘LJALLANGAN NLP PAKETLARI	181
N.R. ALOYEV. KORPUS MATNLARINI TEMATIK MODELLASHTIRISHDAGI ZAMONAVIY YONDASHUVLAR	185

N.R. ALOYEV. ROBERTA TIL MODELI VA NMF TEMATIK MODELLASHTIRISH USULI	192
M. BOTIROVA, Z.G. QODIROVA. KORPUS UCHUN LUG‘AT TO‘PLASH VA ULARNI TASNIFLASH.....	203
N. ASROROVA. COMBINING CORPUS LINGUISTICS WITH CRITICAL DISCOURSE ANALYSIS	209
J.N. O‘ROZOV. AMALIY TILSHUNOSLIKDA LINGVISTIK EKSPERTIZA VA UNING AHAMIYATI	213
D.M. JAMOLITDINOVA, M.SH. AKBAROVA. BADIY MATN LINGVOPOETIKASIDA G‘AYRIODATIY BIRIKMALAR.....	217
K.K. MUSULMONOVA. LINGVISTIK EKSPERTIZANING KRIMINAL VAZIYATLARGA TATBIQI MASALASI.....	221
I.A. XOLMURODOVA. TEZAUROSLARDA TUR-JINS MUNOSABATI.....	224
D. XUDAYBERGANOVA. DOSTONLAR LEKSIKASIDA UCHRAYDIGAN O‘XSHATISHLI BIRLIKLARNING TIPLARI.....	229
SH. ABDULLAYEVA. O‘ZBEK VA NEMIS TILLARIDA FRAZEOLOGIK BIRLIKLARNING LINGVOKULTUROLOGIK XUSUSIYATLARI.....	233
N. SHAMIYEVA, M. JUMAYEVA, O.U. KHAZRATKULOV. SYNONYMOUS AND ANTONYMOUS SERIES OF ADJECTIVES IN THE ENGLISH THESAURUS	236
R.T. O‘RINBOYEVA, I.J. ASQAROV. SAMARQAND SHEVASI DIALEKTINI TRANSKRIPSIYA QILISH TAMOYILLARI.....	239
M. ASQAROVA. KO‘P KOMPLEMENTLI TERMINLARNI PARALLEL KORPUSDA MOSLASHTIRISH MUAMMOLARI INGLIZ VA RUS TILLARI MISOLIDA	244
B.B. ELOV, M.T. SAMATBOYEVA. NER: TEGGLASH MUOMMOLARI	247
G.B. ERKINJONOVA. SHAXS-SON QO‘SHIMCHALARINI AVTOMATIK POS-TEGLASH.....	251
SH.M. HAMROYEVA, M.X. ALIMBEKOV. “ABDURAUF FITRAT KORPUSI” MATNLARINI LEKSIK QATLAMGA KO‘RA TEGGLASH.....	255
S. NURMONOVA, N. TO‘XTABOYEVA. AVTOMATIK TARJIMADAGI MUAMMOLAR	258
R. YULDASHEVA. O‘ZBEK TILIDA KOLLOKATSIYALARGA DOIR YONDASHUVLAR	261
R. YANGIBAYEVA. DIAKRON KORPUS YARATISHNING NAZARIY VA AMALIY AHAMIYATI	266
Z. XUSAINOVA, S. YANGIBAYEVA. MUSTAQILLIK DAVRI NASHRLARIGA ASOSLANGAN DIAKRON KORPUS YARATISHNING LINGVISTIK TA‘MINOTI.....	270
M.S. SHARIPOV, O.O. SOBIROV. SAVOL-JAVOB TIZIMLARIDA LEMMALASH ALGORITMINING O‘RNI.....	274
K.T. ODILOVA. COGNITIVE FEATURES OF ADVERTISING TEXTS WITH VERBAL AND NONVERBAL COMPONENTS.....	278
L. BABAJANOVA. SEMANTIK VALENTLIK TUSHUNCHASINING YUZAGA KELISH USULLARI	294

MUSTAQILLIK DAVRI NASHRLARIGA ASOSLANGAN DIAKRON KORPUS YARATISHNING LINGVISTIK TA'MINOTI

Xusainova Zilola

Filologiya fanlari falsafa doktori (PhD).
Alisher Navoiy nomidagi Toshkent davlat
o'zbek tili va adabiyoti universiteti.
E-mail: xusainovazilola@navoiy-uni.uz
ORCID: 0000-0003-4357-7515,

Yangibayeva Surayyo

ToshDO'TAU, Kompyuter lingvistikasi mutaxassisligi
I bosqich magistranti.
E-mail: surayyoyangibayeva4@gmail.com

Annotatsiya. Ushbu maqolada mustaqillik davri nashrlariga asoslangan diaxron korpus yaratish va uning lingvistik ta'minoti masalalari yoritiladi. Tadqiqotda mustaqillikdan keyingi davrda o'zbek tilida yuz bergan o'zgarishlarni o'rganish uchun diaxron korpusning ahamiyati tahlil qilinadi. Shuningdek, tarixiy matnlarni raqamlashtirish, morfologik va sintaktik belgilash texnologiyalari, leksik o'zgarishlarni aniqlash metodlari hamda lingvistik resurslarni yaratish jarayonlari ko'rib chiqiladi. Tadqiqotning dolzarbligi mustaqillik davri nashrlarining tilshunoslik nuqtayi nazaridan tahlil qilinmaganligi bilan bog'liq.

Kalit so'zlar: diaxronik korpus, lingvistik ta'minot, tarixiy matnlar, morfologik tahlil, lemmatizatsiya, normalizatsiya.

Annotation. This article explores the creation of a diachronic corpus based on publications from the independence period and its linguistic support. The study analyzes the significance of a diachronic corpus in examining linguistic changes in the Uzbek language after independence. Additionally, it discusses the digitization of historical texts, morphological and syntactic annotation technologies, methods for identifying lexical changes, and the process of developing linguistic resources. The relevance of the study lies in the fact that publications from the independence period have not yet been analyzed from a linguistic perspective.

Keywords: diachronic corpus, linguistic support, historical texts, morphological analysis, lemmatization, normalization.

Аннотация. В данной статье рассматриваются вопросы создания диахронного корпуса, основанного на изданиях периода независимости, и его лингвистического обеспечения. В исследовании анализируется значение диахронного корпуса для изучения изменений, произошедших в узбекском языке после обретения независимости. Также рассматриваются процессы оцифровки исторических текстов, технологии морфологической и синтаксической разметки, методы выявления лексических изменений и создание лингвистических ресурсов. Актуальность исследования обусловлена тем, что издания периода независимости ранее не подвергались лингвистическому анализу.

Ключевые слова: диахронный корпус, лингвистическое обеспечение, исторические тексты, морфологический анализ, лемматизация, нормализация.

Kirish

Mustaqillikdan keyin o'zbek tilida sezilarli o'zgarishlar yuz berdi. Jumladan, kirill yozuvidan lotin yozuviga o'tish, yangi atamalar paydo bo'lishi, leksik va grammatik tuzilmalarning boyib borishi kuzatildi. Ushbu jarayonlarni tizimli o'rganish va tahlil qilish uchun mustaqillik davri nashrlari asosida diaxronik korpus yaratish dolzarb masala hisoblanadi. Diaxron

korpuslar til o'zgarishlarining dinamikasini kuzatish, matnlarning turli davrlardagi leksik va grammatik o'zgarishlarini aniqlash hamda lingvistik modellash jarayonida muhim vosita bo'lib xizmat qiladi [Rissanen M., 2000]. Ushbu tadqiqot diaxron korpus yaratish jarayonida lingvistik ta'minotning asosiy jihatlarini tahlil qilishga qaratilgan.

Asosiy qism

1. Mustaqillik davri nashrlari va diaxron korpus

Mustaqillik davri matbuot va adabiyotlarining lingvistik o'zgarishlarni aks ettirishdagi ahamiyati katta. Ushbu davrda chop etilgan rasmiy hujjatlar, ilmiy maqolalar, ommaviy axborot vositalari materiallari va adabiy asarlar til o'zgarishlarini kuzatish uchun muhim manba hisoblanadi [Biber D., 1993]. Mana shu ahamiyatlarni inobatga olgan holda o'zbek tilining 33 yillik ya'ni, 1991-yildan 2024-yilgacha bo'lgan *mustaqillik davriga oid* nashrlarini (asosan badiiy asarlar) o'zida jamlagan yana bir diaxron korpus yaratish ishlari olib borilmoqda. Bu kabi diaxron korpuslarda leksik-semantik o'zgarishlar, terminologik yangilanishlar va stilistik transformatsiyalar yaqqol namoyon bo'ladi. Masalan, mustaqillikdan keyin siyosiy va iqtisodiy sohalarida jadal rivojlanish natijasida yangi atamalar va iboralar paydo bo'ldi. Shuningdek, adabiy matnlarda ham milliy o'ziga xoslikni aks ettirish tendensiyalari kuzatilmoqda. Ushbu o'zgarishlarni aniq tahlil qilish uchun diaxron korpusning qo'llanilishi muhim ahamiyat kasb etadi. Bu jarayonda lingvistik modellash, leksik-statistik tahlil va sun'iy intellekt asosida ishlov berish texnologiyalaridan foydalanish diaxron tadqiqotlarni yanada samarali olib borishga yordam beradi. Mustaqillik davriga oid matnlar diaxronik korpusga kiritilsa, o'zbek tilining rivojlanish dinamikasini yanada aniqroq tahlil qilish imkoniyati kengayadi.

2. Lingvistik ta'minot tarkibiy qismlari

Har qanday korpusni yaratishda lingvistik jarayonlar muhim ahamiyat kasb etadi. Bu jarayonlar til o'zgarishlarini tizimli ravishda o'rganish, matnlarni sintaktik va semantik jihatdan chuqur tahlil qilish hamda lingvistik ma'lumotlarni aniq va ishonchli tarzda qayta ishlash imkonini beradi. Bundan tashqari, diaxron korpuslarga ham lingvistik tadqiqotlar uchun mustahkam ma'lumotlar bazasini shakllantirishda asosiy vositalardan biri sifatida qaraladi. Diaxron korpus yaratishda ham quyidagi lingvistik jarayonlar muhim ahamiyatga ega:

2.1. Matnlarni tanlash va tasniflash

Matnlarning stilistik va funksional turlari bo'yicha tasniflanishi, nashr yili va mualliflik omillari hisobga olinishi kerak. Shuningdek, matnlar yozuv tizimi (kirill yoki lotin) bo'yicha ajratilishi lozim. Masalan, Helsinkidagi diaxron korpus tadqiqotlarida matnlarni kategoriyalarga ajratishning samarali usullari qo'llanilgan [Rissanen M., 2000].

2.2. Normalizatsiya va transliteratsiya

Lotin va kirill yozuvidagi matnlarni bir-biriga moslashtirish, eski imlo qoidalarini yangi standartlarga mos ravishda tahlil qilish muhimdir [Jurafsky D. and Martin J.H., 2021]. Masalan, "*xalq xo'jaligi*" iborasi keyingi yillarda "*iqtisodiyot*" so'zi bilan almashtirilgan. Yangi imlo qoidalar bo'yicha eski imlo qoidalarida "*kechayu-kunduz*", "*ertalabki*" shaklida qo'llanilgan so'zlar keying yillarda yangi imlo qoidalariga ko'ra "*kecha-yu kunduz*", "*ertalabgi*" shaklida qo'llanilishi kabilar.

2.3. Morfologik va sintaktik belgilash

Morfologik analizatorlar va avtomatlashtirilgan tahlil usullari orqali so'z shakllari va ularning grammatik xususiyatlari belgilanishi zarur. TreeTagger va UDPipe kabi vositalardan foydalanish mumkin [Kuznetsov I., 2014].

2.4. Leksik o'zgarishlarni kuzatish

Mustaqillikdan keyingi davrda paydo bo'lgan yoki faol qo'llanila boshlagan so'z va iboralar o'rganiladi. Masalan, "*mustaqil davlat*", "*milliy g'oya*", "*ijtimoiy tarmoq*" kabi terminlarning tarixi tahlil qilinadi.

Mana shu ahamiyatlarni o'zida jamlagan holda mustaqillik davri nashrlariga asoslangan diaxron korpusning lingvistik ta'minoti bosqichma-bosqich yaratilmoqda. Lingvistik taminotni yaratishda dastlab 33 yil oralig'ida nashr qilingan badiiy asarlar to'plandi. Ushbu davrda nashr

qilingan matnlar ichidan biz aynan badiiy asarlarni tanlab olishimizdan maqsad badiiy asarlardagi til o'zgarishlarini o'rganish, badiiy asar tilida eskirgan yoki yangidan paydo bo'lgan so'zlarni aniqlash, ya'ni umuman olganda badiiy asar leksikasini o'rganishdir. Shu sababli to'plangan matnlar dastlab janrlarga, nashr qilingan yillarga ko'ra guruhlariga ajratildi. Hozirda to'plangan matnlarni korpus talablariga mos shaklda tozalash ishlari olib borilmoqda. Diaxron korpus yaratishda lingvistik ta'minot quyidagi qamrovlarni o'z ichiga olganini ko'rishimiz mumkin:

1-jadval. Mustaqillik davri nashrlari diaxron korpusi qamrovi

Davr	Yillar oralig'i	Janri	Til doirasi
Mustaqillik davri nashrlari	1991-2024	Badiiy asarlar	Bir tilli (o'zbek)

Ushbu qamrovga asoslangan holda quyidagi matnlar to'plandi (2-jadval). Matnlarni tozalash jarayonida bu raqamlar hali o'zgarish ehtimoli mavjud. Chunki izlanishlar hali davom etmoqda.

2-jadval. To'plangan matnlar

Mualliflar	Badiiy asar	Yillar kesimi	Janri	Formati	Yozuv shakli
110 ta	220 ta	1991-2024	Roman, qissa, hikoya, esse, drama	pdf	Krill, latin,

3. Duch kelinadigan muammolar:

Diaxron korpus yaratishda bir qator texnik, lingvistik va metodologik muammolarga duch kelinadi. Ushbu qiyinchiliklarni hal etish korpusning aniqligi va samaradorligini oshirishga xizmat qiladi.

1-jadval. Diaxron korpus yaratishda yuzaga keladigan muammolar

Matnlarning aksariyat qismi raqamlashtirilmagan:	Mustaqillik davrining dastlabki yillariga oid gazeta va jurnal maqolalarining aksariyati faqat bosma nashr shaklida mavjud.
Yozuv tizimining o'zgarishi:	Lotin-kirill o'zgarishlari lingvistik tahlil jarayonlarini murakkablashtiradi.
Leksik va grammatik o'zgarishlar:	Yangi atamalarining paydo bo'lishi va ba'zi eski iboralarning eskirishi diaxronik korpus tahlilini murakkablashtiradi.

Ushbu biz keltirib o'tgan muammolar diaxron korpus yaratishda bir qancha murakkabliklarni yaratishi mumkin. Xususan, mustaqillik davrining dastlabki yillariga oid gazeta va jurnal maqolalarining aksariyati faqat bosma nashr shaklida mavjudligi raqamlashtirish jarayonining murakkabligini oshiradi va matnlarni OCR (Optical Character Recognition) texnologiyalari orqali aniq qayta ishlash muammosini keltirib chiqaradi. Ayniqsa, eski matnlarda ishlatilgan yozuv shriftlari va bosma sifati OCR tizimlarining aniqligiga ta'sir qilishi mumkin. Bu muammoni bartaraf etish uchun yanada ilg'or sun'iy intellekt asosidagi matn tanib olish texnologiyalarini joriy qilish zarur. Shuningdek, raqamlashtirilgan matnlarni verifikatsiya qilish va tahrirlash jarayonlarini avtomatlashtirish uchun lingvistik modellar va neyron tarmoqlar asosida ishlaydigan vositalardan foydalanish maqsadga muvofiqdir. Yozuv tizimining o'zgarishi ham ayniqsa, ma'lumotlar bazasini shakllantirishda bir qancha imloviy va texnik muammolarni keltirib chiqaradi. Bu muammoni bartaraf etishda esa avtomatik transliteratsiya tizimidan foydalanishimiz lozim. Natijalarning aniq bo'lishi uchun esa ushbu tizimlarni yanada rivojlantirish talab qilinadi.

Bundan tashqari korpuslarni yaratishda tabiiy til masalalari bilan ishlashda mashinali o'rganish kabi ham to'siqqa duch kelindi – uning algoritmlari odatda raqamlar ustida amallarni bajaradi. Tabiiy tilda esa matnlar qayta ishlanadi. Shunday qilib, tabiiy tildagi matnni raqamlarga aylantirishga to'g'ri keladi. Ushbu jarayon NLPda *matn vektorizatsiyasi* deb nomlanadi. Ushbu bosqich matnli ma'lumotlarni tahlil qilish uchun mashinali o'rganish jarayonidagi asosiy qadamdir va turli vektorlashtirish algoritmlari yakuniy natijalarga keskin ta'sir qiladi [Elov B.B., Xusainova Z.Y., Xudayberganov N.U. 2023].

Xulosa

Mustaqillik davri nashrlari asosida diaxron korpus yaratish o'zbek tilidagi o'zgarishlarni tizimli o'rganish uchun muhim vositadir. Ushbu tadqiqotda diaxron korpus yaratish jarayonida lingvistik ta'minotning asosiy jihatlari ko'rib chiqildi va mavjud muammolar aniqlandi. Korpus lingvistikasi tamoyillariga asoslangan holda tarixiy matnlarni jamlash, ularni lingvistik jihatdan belgilash va analiz qilish orqali mustaqillik yillarida yuz bergan o'zgarishlarni ilmiy asosda o'rganish imkoniyati yaratildi. Bundan tashqari, diaxron korpusning qo'llanilishi o'zbek tilining rivojlanish tendensiyalarini aniqlashda muhim rol o'ynaydi. Korpus lingvistikasi yondashuvlari asosida leksik va grammatik o'zgarishlarni tizimli tahlil qilish, matnlarni avtomatik tarzda tasniflash va semantik jihatdan o'rganish imkoniyati paydo bo'ladi. Bu esa nafaqat tilshunoslik, balki sun'iy intellekt asosidagi tabiiy tilni qayta ishlash (NLP) sohasida ham katta ahamiyat kasb etadi.

Adabiyotlar:

1. Biber D. Representativeness in corpus design. Literary and Linguistic Computing, 1993. 8(4), 243-257.
2. Elov B.B., Xusainova Z.Y., Xudayberganov N.U. O'zbek tili korpusi matnlari uchun tf-idf statistik ko'rsatkichni hisoblash. Maqola: 2023: 1783-1784
3. Gries S.T. Quantitative Corpus Linguistics with R: A Practical Introduction. Routledge. 2009.
4. Jurafsky D. and Martin J.H. Speech and Language Processing (3rd ed.) Pearson. 2021.
5. Kilgariff A. and Grefenstette G. Introduction to the Special Issue on the Web as Corpus. Computational Linguistics, 2003. 333-347.
6. Kuznetsov I. Russian National Corpus: Structure and Applications. Computational Linguistics and Intelligent Text Processing, 2014. 781-795.
7. McEnery T. and Hardie, A. (2012). Corpus Linguistics: Method, Theory and Practice. Cambridge University Press.
8. Rissanen M. The Helsinki Corpus of English Texts: A Diachronic Perspective on English Syntax. Language and Computers, 2000. 30(1), 21-42.
9. Sinclair J. Corpus, Concordance, Collocation. Oxford University Press. 1991.
10. Stubbs M. Words and Phrases: Corpus Studies of Lexical Semantics. Blackwell Publishing. 2001.