O'ZBEKISTON MILLIY UNIVERSITETI

EURO
THE ASSOCIATION OF EUROPEAN OPERATIONAL RESEARCH SOCIETIES

EURO WORKING GROUP
ON OR AND ETHICS

LitORS
LITHUANIAN OPERATIONAL RESEARCH SOCIETY

AL-FARABI
KAZAKH NATIONAL UNIVERSITY

NARXOZ UNIVERSITY

TURAN UNIVERSITY

HAW HAMBURG
Hamburg University of Applied Sciences

UALR UNIVERSITY OF ARKANSAS AT LITTLE ROCK

KLAIPĖDA UNIVERSITY

VYTAUTAS MAGNUS UNIVERSITY

System Economics Association

ScienceInnovation
INTERNATIONAL SCIENTIFIC JOURNAL

VA BOSHQALAR...



IJTIMOIY TIZIMLARNI MODELLASHTIRISH VA RAQAMLI TRANSFORMATSIYA: ILMIY TAJRIBALAR, TEXNOLOGIYALAR VA RIVOJLANISH ISTIQBOLLARI (MSBC–2026)

4-XALQARO ILMIY-AMALIY KONFERENSIYA

2026-yil
23–25 sentyabr

Toshkent shahri,
O'zbekiston



Mazkur konferensiya ijtimoiy-gumanitar fanlar va kompyuter fanlari integratsiyasi, ijtimoiy-xulqiy hodisalarni modellashtirish, sun'iy intellekt, agentga asoslangan simulyatsiya tizimlari hamda raqamli jamiyatdagi kommunikativ jarayonlarni tadqiq etishga bag'ishlanadi.

- XALQARO ILMIY HAMKORLIK
- SUN'IY INTELLEKT VA RAQAMLI TEXNOLOGIYALAR
- IJTIMOIY-XULQIY HODISALARNI MODELLASHTIRISH
- ILMIY TAJRIBA VA DALLARGA ASOSLANGAN QARORLAR
- YOSH OLIMLARNI QI'LLAB-QUVVATLASH

Website:
https://msbc.tech/

E-mail:
jumagulsulaymonova0003@gmail.com

Tel:
+998 93 106 96 67

Tashkiliy qo'mita manzili:
100174, Toshkent sh.,
Universitet ko'chasi 2-uy,
O'zbekiston Milliy universiteti

O'ZBEKISTON RESPUBLIKASI OLIY TA'LIM, FAN VA INNOVATSIYA-LAR VAZIRLIGI

MIRZO ULUG'BEK NOMIDAGI O'ZBEKISTON MILLIY UNIVERSITETI

EUROPEAN ASSOCIATION OF OR SOCIETIES (EURO)

"SCIENCE AND INNOVATION" XALQARO ILMIY JURNALI

"IJTIMOIY TIZIMLARNI MODELLASHTIRISH VA RAQAMLI TRANSFORMATSIYA: ILMIY TAJRIBALAR, TEXNOLOGIYALAR VA RIVOJLANISH ISTIQBOLLARI (MSBC–2026)" MAVZUSIDAGI 4-XALQARO ILMIY-AMALIY KONFERENSIYASI

MATERIALLAR TO'PLAMI

2026-YIL 23-25-SENTYABR

COLLECTION OF MATERIALS

OF THE 4TH INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE ON THE TOPIC "MODELLING AND SIMULATION OF SOCIAL-BEHAVIOURAL PHENOMENA IN CREATIVE SOCIETIES (MSBC-2026)"

September 23-25, 2026

Toshkent – 2026

"Ijtimoiy tizimlarni modellashtirish va raqamli transformatsiya: ilmiy tajribalar, texnologiyalar va rivojlanish istiqbollari (MSBC–2026)" mavzusidagi 4-xalqaro ilmiy-amaliy konferensiyasi materiallar to'plami, (2026-yil 23-25-sentyabr) – T.: Science and Innovation, 2026.

Nashr etishda qatnashganlar:

Baxtiyor Rustamov, Gulira'no Rakhmatullayeva, Kozim Ergashov, Mansur Suyunkulov, Feruza Ollamurodova, Dilnoza Narzullayeva, Sevara O'rinboyeva, Asila Khaydaraliyeva, Anvar Ergashev

Tahrir hay'ati a'zolari:

Abduraxmonova Nilufar Zaynobiddin qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi mudiri, filologiya fanlari doktori, professor*
Taxirov Zair Taxirovich	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi filologiya fanlari doktori, professor*
Iskandarova Aybibi Raximbayevna	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi filologiya fanlari nomzodi, dotsent*
Kasimova Ziyoda Movlon qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi dots.v.b., PhD*
Qodirova Zebo Gulboyevna	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi katta o'qituvchisi, PhD*
Askarova Umidaxon Maxamadjon qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi o'qituvchisi*
Mirvaliyeva Dildora Anvarovna	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi o'qituvchisi*
Ibodullayeva Kunduz Rustam qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi o'qituvchisi*
Kuvondikova Gavhar Isomiddinovna	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi o'qituvchisi*
Abdunazarova Maftuna Faxriddinovna	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi tayanch doktoranti*
Abduraxmonova Sohiba Aktam qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi tayanch doktoranti*
Anvarova Durdona Akram qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi tayanch doktoranti*
O'rinboyeva Roza Toliboy qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi tayanch doktoranti*
Suyunova Mohinur Ilxom qizi	*O'zMU Jurnalistika va o'zbek filologiyasi fakulteti Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi tayanch doktoranti*

© Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti
© "Science and Innovation" xalqaro ilmiy jurnali
© Mualliflar

KIRISH SO‘ZI

Hurmatli hamkasblar, aziz o‘quvchilar!

Qo‘lingizdagi to‘plam O‘zbekiston Milliy universitetida 2026-yil 23–25-sentabr kunlari o‘tkazilgan “Ijtimoiy tizimlarni modellashtirish va raqamli transformatsiya: ilmiy tajribalar, texnologiyalar va rivojlanish istiqbollari (MSBC-2026)” mavzusidagi IV xalqaro ilmiy-amaliy anjuman materiallari asosida tayyorlandi. Anjuman Operatsion Tadqiqotlar Jamiyatlari Xalqaro Federatsiyasi (IFORS) va Yevropa Operatsion Tadqiqotlar Jamiyatlari Assotsiatsiyasi (EURO) qo‘llab-quvvatlashidagi MSBC turkumining to‘rtinchi anjumani bo‘lib, 2019-yilda Vilnyusda boshlangan, 2022-yilda yana Vilnyusda, 2024-yilda esa Almatida davom etgan an’anani ilk bor O‘zbekistonga olib keldi. Bu nafaqat universitetimiz, balki butun respublika ilm-fani uchun xalqaro ilmiy hamjamiyat bilan hamkorlikning yangi bosqichidir.

MSBC anjumanlarining o‘ziga xosligi shundaki, u ijtimoiy-gumanitar fanlar bilan kompyuter fanlari, operatsion tadqiqotlar va raqamli intellekt tutashgan nuqtada fikr almashish maydonini yaratadi. Bu yilgi anjuman dasturida 3 ta plenar majlisda 7 ta plenar ma’ruza va 14 ta sho‘ba majlisida 56 ta ma’ruza, jami 63 ta ma’ruza taqdim etildi. Anjuman gibrid shaklda o‘tkazilib, unda 12 ta davlatdan 57 nafar xorijlik olim ishtirok etdi; masofaviy ishtirokchilar uchun majlislar jonli efirga uzatildi. Sun’iy intellekt, katta til modellari, agentga asoslangan simulyatsiya, o‘yinlar nazariyasi, ijtimoiy tarmoqlar kontentini tahlil qilish, tabiiy tilni qayta ishlash kabi dolzarb yo‘nalishlardagi muhokamalar yosh tadqiqotchilar va magistrantlarga tajribali xorijlik olimlar bilan bevosita muloqot qilish imkonini berdi.

Ushbu to‘plam anjumanning 8-sho‘basi — Tabiiy tilni qayta ishlash (NLP) va amaliy tilshunoslik masalalari yo‘nalishiga bag‘ishlangan. Maqolalar anjuman saytiga topshirilgan materiallardan alohida, “Science and Innovation” xalqaro ilmiy jurnali bilan hamkorlikda qabul qilindi. To‘plamdan O‘zbekistonning turli oliy ta’lim muassasalari professor-o‘qituvchilari, tayanch doktorantlari, mustaqil izlanuvchilari va magistrantlarining 60 dan ortiq maqolasi o‘rin olgan. Ularda til texnologiyalari, kompyuter lingvistikasi va amaliy tilshunoslikning nazariy hamda amaliy masalalari yoritilgan.

To‘plamga qabul qilingan har bir maqola Index Copernicus, Zenodo va OpenAIRE xalqaro bazalarida hamda Google Scholar qidiruv tizimida indekslanadi, alohida DOI raqamiga ega bo‘ladi va ilmiyanjumanlar.uz platformasiga joylashtiriladi. Bu mualliflarning ilmiy natijalari xalqaro ilmiy muhitda ochiq va ishonchli tarzda tarqalishiga xizmat qiladi.

Anjumanni tashkil etishda hissa qo‘shgan IFORS va EURO, hamkor xorijiy universitetlar va tashkilotlar, “Science and Innovation” jurnali tahririyati, Dastur va Tashkiliy qo‘mitalar a’zolari hamda maqolalari bilan ishtirok etgan barcha mualliflarga chuqur minnatdorchilik bildiramiz. Umid qilamizki, ushbu to‘plam tilshunoslik va raqamli texnologiyalar sohasida izlanayotgan tadqiqotchilar uchun foydali manba bo‘ladi, kelgusi MSBC-2028 anjumani oldidan esa hamkorlikni yanada kengaytirishga turtki beradi.

Tashkiliy qo‘mita

Toshkent, 2026-yil oktyabr

THE IMPORTANCE OF THE FAMILY INSTITUTION IN PROMOTING HEALTHY LONGEVITY: MODERN SOCIOLOGICAL APPROACHES

Abdiraimova Shaxnoza Abdugappar qizi
Junior Researcher, Independent Scholar Institute of Family and Gender Studies
https://doi.org/10.5281/zenodo.23135839

Annotation. *This article examines the importance of the family institution in promoting healthy longevity from the perspective of modern sociological approaches. The study analyzes the role of family relationships, social support, intergenerational solidarity, and social capital as key sociological factors influencing healthy ageing and quality of life. Particular attention is paid to the contribution of the family institution to the psychological well-being, social inclusion, and active participation of older adults in society. The article also discusses the interaction between family-based support, healthcare services, and social policies in creating favorable conditions for healthy longevity. Based on contemporary sociological theories and the example of Uzbekistan, the study highlights the significance of strengthening the family institution as an important social resource for promoting healthy longevity and sustainable social development.*

Key words. *Healthy longevity, family institution, sociological approaches, healthy ageing, social capital, intergenerational solidarity, social support, active ageing.*

Annotatsiya. *Mazkur maqolada sog‘lom uzoq umr ko‘rishni ta'minlashda oila institutining ahamiyati zamonaviy sotsiologik yondashuvlar asosida tahlil qilingan. Tadqiqotda oilaviy munosabatlar, ijtimoiy qo‘llab-quvvatlash, avlodlararo birdamlik hamda ijtimoiy kapitalning sog‘lom qarish va uzoq umr ko‘rishga ta'siri yoritilgan. Shuningdek, oila institutining keksalarning psixologik farovonligi, ijtimoiy faolligi va hayot sifatini oshirishdagi o‘rni ilmiy jihatdan asoslab berilgan. Maqolada oila tomonidan ko‘rsatiladigan qo‘llab-quvvatlash, sog‘liqni saqlash tizimi va davlatning ijtimoiy siyosati o‘rtasidagi o‘zaro bog‘liqlik tahlil qilinib, ularning sog‘lom uzoq umr ko‘rishni ta'minlashdagi ahamiyati ochib berilgan. Zamonaviy sotsiologik nazariyalar hamda O‘zbekiston tajribasi asosida oila institutini mustahkamlash sog‘lom uzoq umr ko‘rishni ta'minlovchi muhim ijtimoiy resurslardan biri ekanligi asoslangan.*

Kalit so‘zlar: sog‘lom uzoq umr ko‘rish, oila instituti, sotsiologik yondashuvlar, sog‘lom qarish, ijtimoiy kapital, avlodlararo birdamlik, ijtimoiy qo‘llab-quvvatlash, faol qarish.

Аннотация. *В данной статье рассматривается значение института семьи в обеспечении здорового долголетия на основе современных социологических подходов. Исследование посвящено анализу влияния семейных отношений, социальной поддержки, межпоколенческой солидарности и социального капитала на здоровое старение и долголетие. Особое внимание уделено роли института семьи в укреплении психологического благополучия, социальной активности и повышении качества жизни пожилых людей. В статье также анализируется взаимосвязь семейной поддержки, системы здравоохранения и государственной социальной политики в создании благоприятных условий для здорового долголетия. На основе современных социологических теорий и опыта Узбекистана обосновывается, что укрепление*

института семьи является одним из важнейших социальных ресурсов, способствующих обеспечению здорового долголетия и устойчивого развития общества.

Ключевые слова. Здоровое долголетие, институт семьи, социологические подходы, здоровое старение, социальный капитал, межпоколенческая солидарность, социальная поддержка, активное старение.

INTRODUCTION

Population ageing has become one of the most significant demographic trends of the twenty-first century, making healthy longevity an important priority for both scientific research and public policy. In recent decades, increasing life expectancy has shifted attention from simply extending the length of life to ensuring that people remain healthy, socially active, and independent in old age. Consequently, healthy longevity is increasingly viewed as a multidimensional concept that reflects the combined influence of biological, medical, social, economic, and cultural factors.

Contemporary sociological studies indicate that the quality of family relationships, social support, social capital, and intergenerational solidarity significantly influence healthy ageing and the well-being of older adults. Among these factors, the family institution occupies a central position because it provides emotional support, strengthens social integration, facilitates the transmission of social values, and promotes active participation throughout the life course. Therefore, the family should be considered not only as a traditional social institution but also as an important sociological resource for promoting healthy longevity.

Theoretical contributions by Talcott Parsons, Pierre Bourdieu, and Robert Putnam provide a valuable framework for understanding the relationship between family, social capital, and healthy longevity. Their approaches emphasize that strong social ties, mutual trust, and family solidarity create favorable conditions for improving quality of life, reducing social isolation, and supporting active ageing.

Despite the growing body of international research on healthy longevity, the sociological role of the family institution in promoting healthy longevity has received limited attention, particularly in the context of Uzbekistan. Most studies have focused primarily on medical and demographic aspects, while the influence of family relationships and intergenerational solidarity remains insufficiently explored. This research seeks to address this gap by examining the family institution from a sociological perspective.

The aim of this study is to analyze the importance of the family institution in promoting healthy longevity through modern sociological approaches and to identify its contribution to improving the well-being, social participation, and quality of life of older adults in Uzbekistan.

Literature Review. The existing sociological literature indicates that healthy longevity is determined not only by biological and medical factors but also by social, cultural, and institutional influences. Particular attention has been given to the family institution as one of the key social environments that shapes individuals' well-being throughout the life course. Previous studies emphasize that family support, social capital, intergenerational solidarity, and strong social relationships contribute to healthy ageing, improve quality of life, and reduce the risks of social isolation among older adults. Contemporary sociological theories further suggest that strengthening the family institution and promoting social cohesion are essential components of healthy longevity. Based on these theoretical and empirical studies, the present research

examines the importance of the family institution in promoting healthy longevity through modern sociological approaches, with particular attention to the context of Uzbekistan.

Methods. This study is based on qualitative sociological research methods. The research applies a systematic review of scientific literature, comparative analysis, and sociological interpretation of theoretical approaches to examine the role of the family institution in promoting healthy longevity. The theories of Talcott Parsons, Pierre Bourdieu, and Robert Putnam were used as the main theoretical framework. In addition, national and international scientific publications, policy documents, and reports related to healthy ageing and family studies were analyzed. The collected information was interpreted through analysis, synthesis, and comparative methods to identify the sociological significance of the family institution in ensuring healthy longevity.

Results. The analysis indicates that the family institution plays a significant role in promoting healthy longevity. The findings suggest that strong family relationships, social support, intergenerational solidarity, and social capital positively influence the psychological well-being, social participation, and quality of life of older adults.

Table 1 summarizes the main sociological factors through which the family institution contributes to healthy longevity.

Table 1

Modern Sociological approaches to promoting healthy longevity through the family institution

Sociological factors	**Description**	**Contribution to healthy longevity**
Family institution	Emotional support, care, family responsibility	Improves psychological well-being and quality of life
Intergenerational solidarity	Mutual support between older and younger generations	Reduces loneliness and strengthens social integration
Social capital	Trust, cooperation, family and community networks	Enhances resilience to social risks and promotes active ageing
Social support	Emotional, material and informational assistance	Reduces stress and improves social participation
Healthy lifestyle within the family	Balanced nutrition, physical activity, preventive healthcare	Supports healthy ageing and prevents chronic diseases
Community and social	Participation in family and	Encourages active ageing and

participation	community activities	social inclusion

Source: Compiled by the author based on the analysis of sociological and interdisciplinary literature.

The analysis presented in Table 1 demonstrates that the family institution constitutes an important social environment in which various factors influencing healthy longevity are formed and maintained. Family relations, intergenerational solidarity, social capital, social support, healthy behavioural practices and participation in community life should not be considered as isolated factors. Rather, they form an interconnected system of social resources that can influence the psychological well-being, social activity and quality of life of older adults. From this perspective, healthy longevity should be understood not only as an increase in life expectancy, but also as the preservation of social connectedness, functional independence and opportunities for active participation in society.

Family institution. The family represents one of the fundamental social institutions through which individuals receive emotional, social and practical support throughout the life course. Its significance becomes particularly evident in later life, when changes in employment status, social roles, health conditions and patterns of everyday interaction may increase the need for assistance and social support. Stable family relationships can provide older adults with a sense of security, belonging and emotional stability. Regular communication with family members, participation in household and family activities, and the opportunity to receive assistance in everyday matters can contribute to maintaining social activity and psychological resilience.

The family also plays an important role in the formation of health-related behavioural patterns. Dietary habits, physical activity, attitudes towards preventive healthcare and approaches to managing everyday stress are often influenced by the family environment. Therefore, the family can function simultaneously as a source of emotional support and as a social mechanism through which health-promoting behaviours are reproduced across generations. Intergenerational solidarity. An important sociological dimension of healthy longevity is the quality of relationships between different generations. Intergenerational solidarity encompasses mutual assistance, emotional closeness, respect, communication and the exchange of knowledge and experience between younger and older family members. Such relationships allow older adults to maintain their social roles and feel that their experience and contribution remain valuable to the family and society.

In the context of Uzbekistan, intergenerational relations have particular social significance. The continued interaction between older parents, adult children and grandchildren may provide older people with opportunities for regular communication and practical assistance. At the same time, older generations contribute to the transmission of cultural traditions, family values and life experience. Consequently, intergenerational solidarity can serve as an important mechanism for reducing social isolation and strengthening social cohesion. Social capital. From a sociological perspective, the family is one of the primary environments in which social capital is accumulated and reproduced. Trust, reciprocity, cooperation and established social networks create resources that individuals can use when facing different social and economic challenges.

For older adults, these resources may include emotional assistance, information, practical help and access to wider community networks.

The concept of social capital developed in the works of P. Bourdieu and R. Putnam provides an important theoretical basis for understanding these processes. Strong interpersonal networks and reciprocal relations can increase an individual's ability to obtain support and adapt to changes occurring during later life. Thus, family-based social capital can indirectly contribute to healthy longevity by reducing social vulnerability and strengthening an individual's connection with the surrounding social environment.

Social support. Social support is another important factor linking family relations with healthy ageing. It may take emotional, material, informational and practical forms. Emotional support includes attention, understanding and encouragement, while practical support may involve assistance with household activities, transportation, access to services or other everyday needs.

The availability of social support can be particularly significant when older adults experience changes in their social roles or reduced opportunities for independent participation in certain activities. A supportive family environment may reduce feelings of loneliness and insecurity and contribute to greater psychological stability. At the same time, excessive dependence should be avoided; effective support should encourage the autonomy and active participation of older adults rather than replacing their social agency.

Healthy lifestyle within the family. The family environment has a considerable influence on the formation and maintenance of health-related behaviours. Eating patterns, physical activity, daily routines and attitudes towards preventive medical examinations are frequently shaped within the household. When healthy behaviours are supported by family members, their continuation may become part of everyday social practice rather than an individual effort. Balanced nutrition, regular physical activity, adequate rest and timely preventive healthcare are important components of healthy ageing. From a sociological perspective, the importance of these practices lies not only in their direct relationship with health but also in the social mechanisms through which they are encouraged, maintained and transmitted between generations. In this regard, the family can contribute to the development of a supportive environment that facilitates healthier patterns of everyday behaviour.

Community and social participation. The relationship between family and the wider community creates additional opportunities for maintaining social activity in later life. Participation in neighbourhood events, family gatherings, community initiatives and other forms of collective activity can expand the social networks of older adults and strengthen their sense of social belonging.

Social participation is closely associated with the concept of active ageing because it enables older people to remain involved in social life rather than becoming socially isolated. For this reason, the role of the family should not be limited to providing care. A family that encourages older members to participate in community activities can contribute to maintaining their social identity, independence and sense of usefulness.

Overall, the analysis indicates that the family institution should be considered one of the important sociological determinants of healthy longevity. Its influence is realised through several interconnected mechanisms, including emotional support, intergenerational solidarity, social capital, healthy behavioural practices and participation in family and community life. The effectiveness of these mechanisms depends not only on the internal stability of the family but

also on the broader social environment and the availability of institutional support. Therefore, promoting healthy longevity in Uzbekistan requires a comprehensive approach that recognises the family as an important social resource while simultaneously supporting the autonomy and active participation of older adults. Strengthening intergenerational relations, developing community-based support mechanisms and creating opportunities for social participation can contribute to improving the quality of life of older people. In this sense, the family institution should be viewed not merely as a source of care, but as an important social mechanism for supporting healthy, active and socially integrated ageing.

Conclusion

The findings of this study indicate that the family institution is one of the key sociological factors contributing to healthy longevity. Healthy ageing depends not only on biological and medical conditions but also on the quality of family relationships, social support, social capital, and intergenerational solidarity. These factors create favorable conditions for maintaining psychological well-being, social participation, and a high quality of life among older adults.

The analysis demonstrates that strong family ties reduce social isolation, encourage healthy lifestyles, and enhance the resilience of older people to social and psychological challenges. In the context of Uzbekistan, where the family remains a fundamental social institution, preserving family values and strengthening intergenerational relationships can significantly contribute to promoting healthy longevity.

The study also suggests that family-based support should be complemented by effective social policies, accessible healthcare services, and community participation. Integrating these sociological mechanisms will create more favorable conditions for active and healthy ageing.

In conclusion, strengthening the family institution should be regarded as an important priority in promoting healthy longevity. The findings of this study may serve as a scientific basis for developing social policies and programs aimed at improving the well-being, social inclusion, and quality of life of older adults in Uzbekistan.

REFERENCES

1. Abdiraimova Sh.A. Insonning uzoq umr ko‘rishini o‘rganishda aksiologik yondashuv // “Ilm-fan muammolari tadqiqotchilar talqinida” mavzusidagi respublika ilmiy konferensiyasi. 2024. – 5-b.
2. Abdiraimova Sh.A. Faol uzoq umr ko‘rishning ijtimoiy-demografik tamoyillari // Ta'limning zamonaviy transformatsiyasi. 2024. – 5-b.
3. Pescatello, L. S., Arena, R., Riebe, D., & Thompson, P. D. (2014). Physical Activity and Health: A Report of the American College of Sports Medicine. American Journal of Lifestyle Medicine, 8(2), 146–165.
4. World Health Organization (WHO). (2021). Healthy life expectancy and life expectancy at birth. WHO Global Health Observatory.
5. Sachs, J. D. (2015). The Age of Sustainable Development. Columbia University Press.

DIALEKTAL KORPUS YARATISHNING ZAMONAVIY TAJRIBALARI ASOSIDA QASHQADARYO SHEVASI LINGVISTIK BAZASINI SHAKLLANTIRISH

Abdullayeva Nasima Murodullayevna
Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti
Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi magistranti
https://doi.org/10.5281/zenodo.23135849

Annotatsiya. *Mazkur maqolada zamonaviy dialektal korpuslar yaratish tajribasi asosida Qashqadaryo shevasi, xususan, Qarshi shahri qarluq lahjasi uchun lingvistik baza shakllantirishning nazariy va amaliy masalalari tadqiq etiladi. Tadqiqotning maqsadi jahon dialektal korpusshunosligida shakllangan zamonaviy yondashuvlarni tahlil qilish hamda ularni Qarshi shevasi materiallarini raqamlashtirish va korpusga joylashtirish jarayoniga moslashtirishdan iborat. Tadqiqotda tavsifiy, qiyosiy, korpus-lingvistik va strukturaviy modellashtirish metodlaridan foydalanildi. Jahon tajribasi bilan qiyoslash natijasida Qarshi shevasi bazasining asosiy ilmiy qiymati audio, transkripsiya, metadata va lingvistik teglashni yagona raqamli muhitda birlashtirish imkoniyatida ekani asoslandi.*

Kalit so'zlar: *dialektal korpus, Qashqadaryo shevasi, Qarshi shevasi, lingvistik baza, korpus lingvistikasi, transkripsiya, metadata, annotatsiya, teglash, audio korpus.*

Abstract. *This article examines the theoretical and practical issues of developing a linguistic database for the Kashkadarya dialect, particularly the Karluk dialect of Qarshi city, based on modern practices in dialect corpus construction. The study aims to analyse contemporary approaches developed in international dialect corpus linguistics and adapt them to the digitisation and corpus-based organisation of Qarshi dialect materials. Descriptive, comparative, corpus-linguistic and structural modelling methods were employed.Comparison with international practices demonstrates that the principal scientific value of the Qarshi dialect database lies in integrating audio, transcription, metadata and linguistic annotation within a unified digital environment.*

Keywords: *dialect corpus, Kashkadarya dialect, Qarshi dialect, linguistic database, corpus linguistics, transcription, metadata, annotation, linguistic tagging, speech corpus.*

Аннотация. *В статье рассматриваются теоретические и практические вопросы формирования лингвистической базы кашкадарьинского диалекта, в частности карлукского говора города Карши, на основе современных подходов к созданию диалектных корпусов. Цель исследования заключается в анализе современных методов, сформировавшихся в мировой диалектной корпусной лингвистике, и их адаптации к процессу цифровой организации материалов каршинского диалекта. В исследовании использованы описательный, сравнительный, корпусно-лингвистический методы и структурное моделирование. Сопоставление с международным опытом показывает, что научная значимость базы каршинского диалекта заключается прежде всего в интеграции аудио, транскрипции, метаданных и лингвистической аннотации в единой цифровой среде.*

Ключевые слова: *диалектный корпус, кашкадарьинский диалект, каршинский диалект, лингвистическая база, корпусная лингвистика, транскрипция, метаданные, аннотация, лингвистическая разметка, речевой корпус.*

Kirish. Raqamli texnologiyalarning tilshunoslikka jadal kirib kelishi natijasida dialektologik tadqiqotlarning metodologik bazasi ham sezilarli darajada o‘zgarmoqda. An’anaviy dialektologiyada hududiy nutq xususiyatlari, asosan, ekspeditsion materiallar, so‘rovnomalar, lug‘atlar va tavsifiy tadqiqotlar orqali qayd etilgan bo‘lsa, zamonaviy korpus lingvistikasi sheva materiallarini audio, transkripsiya, metadata va lingvistik annotatsiyani birlashtirgan raqamli resurs sifatida tashkil etishni talab qilmoqda. “Bugungi kunda korpusning imkoniyatlari shu darajaga yetdiki, ushbu sohada erishilgan yutuqlardan nafaqat NLP sohasi yoki kompyuter lingvistikasi, balki tilga o‘qitishning metodika va pedagogika sohalari, diskurs tahlil, mashina tarjimasi va lingvistikaning o‘nlab sohalari (sotsiolingvistika, gendershunoslik, psixolingvistika va h.k.) unumli foydalanib, sohaga doir ijobiy natijalarga erishib kelinmoqda” [Abduraxmonova N.Z., 2023:29].

Dialektal korpusning ilmiy qiymati uning hajmi bilangina belgilanmaydi. Korpusning reprezentativligi, audio va matn o‘rtasidagi bog‘lanish, transkripsiya qatlamlarining aniqligi, informant haqidagi metadata, qidiruv imkoniyatlari hamda lingvistik teglash tizimi uning keyingi ilmiy tadqiqotlarda foydalanish imkoniyatini belgilaydi. Shu sababli zamonaviy dialekt korpuslarini yaratish jarayonida dala materiallarini yig‘ishdan boshlab ularni raqamli bazaga joylashtirishgacha bo‘lgan barcha bosqichlar o‘zaro bog‘langan tizim sifatida qaraladi.

So‘nggi xalqaro tadqiqotlarda ham aynan shu tendensiya kuzatilmoqda. Masalan, Bentum va hammualliflar Niderland dialektal nutqining tarixiy yozuvlari asosida audio, transkripsiya va metadata o‘rtasidagi avtomatik moslashtirish masalasini o‘rgangan. Tadqiqotda qariyb 300 soatlik audio materialning qo‘lda tayyorlangan transkripsiyalari bilan bog‘lanishi va ularning metadata bilan integratsiyasi ko‘rib chiqilgan [Bentum M. va boshq., 2024]. Tadqiqotchilar Kuparinen va Scherrer esa 2024-yilda dialektometriyada topic modeling metodidan foydalanish imkoniyatlarini tadqiq qilib, fin, norveg va shveytsar nemis tillari materiallari asosida oldindan qo‘lda annotatsiya qilinmagan transkripsiyalardan ham hududiy dialekt farqlarini aniqlash mumkinligini ko‘rsatgan [Kuparinen O. va boshq., 2024:1-12]. Shuningdek, 2024–2025-yillardagi multimodal dialekt korpusi tadqiqotlarida raqamli materiallarni yig‘ish, annotatsiyalash, segmentatsiya qilish, transkripsiyalash va sinxronlashtirish yagona texnologik jarayon sifatida qaralmoqda [Zhao Y. va boshq., 2025].

O‘zbek tilshunosligi uchun ayniqsa 2025-yilda Xorazm shevalarining o‘g‘uz shakliga oid leksik dataset va qoidaga asoslangan algoritm ishlab chiqilgani muhimdir. Unda dialektal so‘zlarning kirill va lotin shakllari, inglizcha tarjimasi, adabiy tildagi muqobili va qo‘llanish hududi kabi parametrlar jamlangan [Abdurakhmonova N. va boshq., 2025].

O‘zbek dialektologiyasida shevalarni tavsiflash bo‘yicha katta ilmiy tajriba shakllangan. S. Ashirboyevning o‘zbek dialektologiyasiga oid ishlari [Ashirboyev S., 2016] va transkripsiya tizimi Qarshi shevasi materiallarini raqamlashtirish uchun nazariy va amaliy asoslardan biri bo‘lib xizmat qiladi. Shu nuqtayi nazardan, Qashqadaryo shevasini zamonaviy korpus texnologiyalari asosida raqamlashtirish nafaqat mavjud dialektologik materiallarni saqlash, balki ularni qidiruv, statistik tahlil, avtomatik annotatsiya va tabiiy tilni qayta ishlash jarayonlariga jalb etish imkonini beradi.

Tadqiqot materiali va metodlari. Tadqiqotning asosiy obyektini Qarshi shahrining qarluq lahjasi shevalari, sheva vakillaridan yozib olingan audio ma'lumotlar va ularning transkripsiyalangan matnlari tashkil etadi. Tadqiqotning maqsadi Qarshi shahridagi qarluq lahjasining fonetik va morfologik xususiyatlarini o'rganish, ular asosida lingvistik baza yaratish hamda natijalarni o'zbek tili korpusining dialektal subkorpusiga joylashtirishdan iborat.

Tadqiqot materiallari Qarshi shahrining nisbatan qadimgi mahallalaridan yig'ildi. Jumladan, Navro'z, Bog'iston, Xudoyzod, Batosh, Chorgumbaz, Gulshan, Nasaf, Zog'za, Roguzar, Naxshab, Xontepa va Ko'chkak mahallalaridan audio nutqlar to'plandi. Informantlar tanlashda yosh, jins, ta'lim darajasi, yashash muhiti va nutq vaziyati kabi reprezentativlik mezonlariga e'tibor qaratildi.

Natijada 30 nafar informant nutqidan 5243 ta qarluq lahjasiga oid sheva so'zi qayd etildi. Ushbu materiallar asosida 121 ta audio yozuv va ularning transkripsiyalangan matnlari tayyorlandi.

Tadqiqotda quyidagi metodlar kompleks tarzda qo'llandi:

- ***tavsifiy metod*** – Qarshi shevasi fonetik, leksik va morfologik birliklarini tavsiflash;
- ***qiyosiy metod*** – dialektal birliklarni adabiy til shakllari bilan solishtirish;
- ***korpus-lingvistik metod*** – audio va matn materiallarini raqamli resurs sifatida tashkil etish;
- ***strukturaviy modellashtirish*** – lingvistik baza, metadata va teglash tizimini ishlab chiqish;
- ***raqamli qayta ishlash*** – audio segmentatsiyasi, transkripsiya va ma'lumotlarni strukturaviy bazaga joylashtirish.

Audio materiallar dastlab mobil qurilmalar yordamida yozib olindi, keyinchalik kompyuterga ko'chirilib, Audacity dasturida taxminan 30 soniyalik segmentlarga ajratildi. Segmentlar MP3 formatida arxivlashtirildi. Tadqiqotning asosiy maqsadi akustik-fonetik parametrlarni o'lchashdan ko'ra dialektal leksik, morfologik va sintaktik materiallarni tizimli jamlash bo'lgani sababli ushbu format tadqiqot ehtiyojlari uchun maqbul deb belgilandi. Segmentlangan audio materiallar ortiqcha pauza va texnik shovqinlardan tozalandi. Har bir audio faylga yagona identifikatsion kod berilib, kod tarkibida hudud, mahalla, jins, yosh va urug'ga oid ma'lumotlarni ifodalovchi parametrlar aks ettirildi. Transkripsiyalangan matnlar TXT formatida alohida saqlandi va audio materiallar bilan bog'landi.

Jahon dialektal korpuslari tajribasi. Zamonaviy dialektal korpuslarni yaratish tajribasi bir nechta muhim yo'nalishlarning shakllanganini ko'rsatadi:

1. ***Audio va transkripsiyani integratsiyalash***. Niderland dialektal nutqi korpusini yaratish bo'yicha Bentum va hammualliflarning tadqiqotida tarixiy audio yozuvlar va mavjud transkripsiyalarni avtomatik ravishda moslashtirish masalasi markazga qo'yilgan. Mualliflar dialektal nutq uchun standart tilga asoslangan zamonaviy nutqni tanish texnologiyalaridan foydalanish imkoniyatini ham ko'rsatadi [Johannessen J.B., 2017].

Bu tajriba Qarshi shevasi uchun muhim metodologik xulosa – **audio va uning transkripsiyasi alohida fayllar sifatida emas, yagona identifikator orqali bog'langan birlik sifatida saqlanishi kerakligini berdi**.

2. ***Ko'p qatlamli transkripsiya***. Dialektal nutqni standart orfografiyaga to'liq moslashtirish uning fonetik o'ziga xosligini yo'qotishi mumkin. Aksincha, faqat fonetik transkripsiyadan foydalanish umumiy qidiruv va avtomatik qayta ishlashni murakkablashtiradi. Shu sababli zamonaviy dialekt korpuslarida fonetik/dialektologik va standartlashtirilgan

qatlamlarni birgalikda saqlash amaliyoti muhim ahamiyat kasb etmoqda [Bentum M. va boshq., 2024; Johannessen J.B., 2017].

Qarshi shevasi lingvistik bazasida ham ikki darajali – **orfografik (soddalashtirilgan) va fonetik (aniqlashtirilgan)** transkripsiya tizimi qo‘llangan. Bu yondashuv real talaffuzni saqlash va qidiruv uchun qulay shaklni ta’minlash o‘rtasidagi muvozanatni yuzaga keltiradi.

3. ***Dialektal ma’lumotlarni ko‘p parametrli dataset sifatida tashkil etish***. Xorazm shevalariga oid 2025-yilgi datasetda dialektal birliklarning turli yozuv shakllari, adabiy ekvivalentlari va hududiy qo‘llanishi bir ma’lumotlar tizimida berilgan [Abdurakhmonova N. va boshq., 2025]. Qarshi lingvistik bazasida esa ushbu tamoyil yanada kengroq ko‘rinishda namoyon bo‘lib, dialektal birlik bilan birga uning transkripsiyasi, TXT fayldagi identifikatori va informantga oid metama’lumotlar ham bir tizimga kiritilgan.

4. ***Avtomatik va statistik tahlilga tayyor korpus yaratish***. Kuparinen va Scherrerning tadqiqoti korpus transkripsiyalaridan oldindan belgilangan lingvistik xususiyatlarsiz ham topic modeling yordamida dialektal farqlarni aniqlash mumkinligini ko‘rsatadi [Kuparinen O. va boshq., 2024:1-12].

Demak, Qarshi lingvistik bazasining keyingi bosqichlarda kengaytirilishi uning nafaqat tavsifiy dialektologiya, balki avtomatik dialekt identifikatsiyasi, dialektometriya va NLP tadqiqotlari uchun ham ishlatilishiga imkon beradi.

Qarshi shevasi lingvistik bazasining shakllantirilishi. Tadqiqotning amaliy bosqichida Qarshi shahri shevasining audio va matn korpusi ikki asosiy komponentdan – **audio korpus va matn korpusidan** tashkil etildi. Korpus keyinchalik yangi materiallar bilan to‘ldirish imkoniyatini hisobga olgan holda loyihalashtirildi.

Lingvistik baza .xlsx formatida shakllantirildi. Unda sheva, lahja, so‘zning adabiy tildagi shakli, adabiy orfografik shakli, transkripsiyasi, TXT fayldagi belgilangan nomi, informantning manzili, yoshi,

	A	B	C	D	E	F	G	H	I	J
1										
2										
3	t/r	Adabiy tilda	Shevada (talaffuzi)	Transkripsiyasi	Lahjasi	Metama'lumot			Urug'i	Fayl (txt)
4						hudud	yoshi va jinsi	ma'lumoti va kasbi		
5234		kecha	kecha	kečä	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5235		pishirgan	pishirgan	piširgän	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5236		chuchvaramdan	chuvaramdan	čuvärämdän	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5237		olib chiqib beray	obchig'iberey	äbčiyïberäj	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5238		ozib ham	arrig'labam	ärrïylăbäm	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5239		ketibdi	ketipti	ketipti	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5240		tuzuk	tuziy	tuzij	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5241		narsa	narsa	närsä	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5242		yemasa ham	yemasayam	jemäsäjäm	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5243		kerak-da	kere-da	kere-dä	qarluq	Qarshi shahri Xontepa mfy	45 yosh, ayol	o'rta, uy bekasi		121-txt
5244										
5245										
5246										

Активация Windows
Чтобы активировать Windows, перейдите в раздел "Параметры".

Sheet1 Лист2 +

ma’lumoti, kasbi va urug‘iga oid ma’lumotlar aks ettirildi.

1-jadval. Qarshi shevasi lingvistik bazasi tuzilishidan namuna

Jadvaldan ko‘rinadiki, bazaning asosiy afzalligi dialektal birlikni faqat lug‘aviy birlik sifatida emas, balki uning **lingvistik shakli + transkripsiyasi + manbasi + informanti + hududi** bilan birgalikda saqlash imkoniyatidir. Bunday tashkil etish keyinchalik dialektal birlikni turli mezonlar asosida qidirish va guruhlash imkonini beradi.

Qarshi lingvistik bazasi uchun ishlab chiqilgan qidiruv tizimida so‘z yoki lemma bo‘yicha qidirish, grammatik teg bo‘yicha filtrlash, informant xususiyatlariga ko‘ra tanlash va audio bilan sinxron ko‘rish imkoniyatlari ko‘zda tutilgan. Shu jihatdan lingvistik baza oddiy elektron jadvaldan farqli ravishda korpusning ma’lumotlar qatlamini tashkil etadi. Excel formati

tadqiqotning dastlabki bosqichida ma'lumotlarni tezkor strukturaviy tashkil etish uchun qulay bo'lsa, keyingi bosqichda ushbu ma'lumotlarni TEI/XML yoki boshqa standart korpus formatlariga konvertatsiya qilish mumkin.

Transkripsiya va lingvistik teglash. Dialektal korpus yaratishda transkripsiya eng muhim bosqichlardan biridir. Chunki shevaning fonetik o'ziga xosligi standart imloga to'liq moslashtirilganda yo'qolib ketishi mumkin. Shu sababli Qarshi shevasi materiallarida ikki darajali transkripsiya tizimi tanlandi: birinchi daraja soddalashtirilgan orfografik qatlam, ikkinchi daraja esa fonetik jihatdan aniqlashtirilgan qatlamdan iborat. Ushbu yondashuv xalqaro dialekt korpuslarida uchraydigan ikki qatlamli transkripsiya prinsipiga mos keladi. Birinchi qatlam foydalanuvchi uchun qidiruv va matnni umumiy tahlil qilishni osonlashtirsa, ikkinchi qatlam dialektologik va fonetik tadqiqotlar uchun zarur bo'lgan tafsilotlarni saqlaydi.

Qarshi shevasi uchun transkripsiyada S. Ashurboyev tomonidan taklif etilgan o'zbek shevalarining fonetik xususiyatlarini izchil qayd etishga yo'naltirilgan transkripsiya tizimidan foydalanildi. Lingvistik teglash jarayonida esa har bir birlikni leksik, morfologik va fonetik xususiyatlariga ko'ra maxsus teglar bilan belgilash ko'zda tutilgan. Mazkur jarayonda Universal Dependencies standartlariga yaqin yondashuvdan foydalanilgan. Bunday teglash tizimi kelgusida korpusdan avtomatik qidiruv va statistik tahlil uchun foydalanish imkoniyatini oshiradi. Xususan, muayyan grammatik shaklning qaysi hududda, qaysi yosh guruhida yoki qaysi informantlar nutqida uchrashini aniqlash mumkin bo'ladi.

Annotatsiyaning yarim-avtomatik tashkil etilishi ham muhim metodologik jihatdir. Dissertatsiyada qayd etilganidek, annotatsiya mavjud standart til vositalari bilan moslashtirilgan. Bunday yondashuv qo'lda annotatsiya qilish xarajatlarini kamaytirish bilan birga, dialektal birliklarning standart til bilan o'zaro munosabatini ham hisobga olish imkonini beradi.

Asosiy natijalar. Tadqiqot natijasida Qarshi shahri qarluq lahjasiga mansub nutq namunalaridan quyidagi ko'rsatkich va natijalarga ega bo'lindi.

Ko'rsatkich	Natija
Informantlar	30 nafar
Qayd etilgan dialektal birliklar	5243 ta
Audio yozuvlar	121 ta
Transkripsiya	Orfografik + fonetik
Audio segmentatsiyasi	Taxminan 30 soniyalik segmentlar
Audio formati	MP3
Matn formati	TXT
Lingvistik baza	XLSX
Korpus komponentlari	Audio korpus + matn korpusi
Metadata	Hudud, manzil, yosh, ma'lumot, kasb, urug' va boshqa belgilar
Teglash	Leksik, morfologik va fonetik
Annotatsiya	Yarim-avtomatik
Qidiruv	So'z/lemma, teg, informant xususiyati, audio bilan bog'lash

2-jadval. Tadqiqotning asosiy miqdoriy va texnologik natijalari

Olingan natijalar bir necha jihatdan ilmiy ahamiyatga ega bo‘lib, quyidagi xulosalarni beradi:

1. 30 nafar informant nutqidan yig‘ilgan 5243 ta sheva birligi Qarshi shahri qarluq lahjasining real nutq materiallarini raqamli shaklda jamlash imkonini berdi. Materiallarning shaharning turli mahallalaridan olinishi hududiy qamrovni ta’minlashga xizmat qildi.

2. 121 ta audio va ularga mos transkripsion matnning mavjudligi korpusni faqat yozma dialekt materiallari majmuasi emas, balki **nutq korpusi** sifatida shakllantirish imkonini berdi. Audio va matnning yagona identifikatsion kod orqali bog‘lanishi keyingi bosqichlarda avtomatik alignment uchun metodologik asos yaratadi.

3. Ikki darajali transkripsiya Qarshi shevasining fonetik xususiyatlarini saqlash va qidiruvni soddalashtirish o‘rtasida muvozanat yaratadi. Bu jihat xalqaro dialekt korpuslari tajribasida ham muhim metodologik tamoyil sifatida namoyon bo‘lmoqda [Johannessen J.B., 2017; Bentum M., va boshq, 2024].

4. XLSX bazasida lingvistik birlik va informant metama’lumotlarining birgalikda saqlanishi kelgusida sotsiolingvistik tadqiqotlar uchun yangi imkoniyatlar yaratadi. Masalan, muayyan dialektal birlikning yosh, kasb, yashash hududi yoki boshqa ijtimoiy parametrlar bilan bog‘liqligini tekshirish mumkin.

5. Qidiruv tizimining lemma, grammatik teg va informant xususiyatlari bo‘yicha ishlashi korpusning amaliy qiymatini oshiradi. Bu imkoniyatlar kelgusida dialektal birliklarning chastotasi, hududiy tarqalishi va grammatik xususiyatlarini avtomatik hisoblash uchun kengaytirilishi mumkin.

Xalqaro tajribalar shuni ko‘rsatadiki, Qarshi shevasi lingvistik bazasining shakllanishi zamonaviy dialekt korpuslari rivojlanishidagi umumiy metodologik tendensiyalarga mos keladi. Shu bilan birga, uning asosiy o‘ziga xosligi o‘zbek tilining Qarshi guruhi qarluq shevasiga moslashtirilgan transkripsiya va metadata modelining ishlab chiqilganidadir.

Tadqiqotning ilmiy ahamiyati. Qarshi shahri qarluq lahjasi materiallarini korpus uchun tizimli ravishda yig‘ish, transkripsiyalash va lingvistik baza shaklida tashkil etish tamoyillari uning ilmiy ahamiyatini belgilaydi. Tadqiqot doirasida Qarshi shevasi ilk bor korpus uchun areologik nuqtayi nazardan o‘rganilib, hududiy o‘ziga xosliklar aniqlangan, fonetik va morfologik xususiyatlar adabiy til bilan qiyoslangan hamda to‘plangan faktlarni korpusga joylashtirish tamoyillari ishlab chiqilgan. Amaliy jihatdan esa yaratilgan baza:

- dialektologik tadqiqotlar;
- elektron dialekt lug‘atlari;
- o‘zbek tili korpusining dialektal subkorpusi;
- dialektal nutqni avtomatik tanish;
- nutq sintezatori;
- tabiiy tilni qayta ishlash;
- dialektlarni avtomatik tasniflash;
- sotsiolingvistik tadqiqotlar uchun ma’lumotlar manbai bo‘lishi mumkin.

Dialektal nutqning audio va transkripsiya shakllarining parallel saqlanishi kelgusida o‘zbek tilining kam resursli dialektlari uchun nutq texnologiyalarini ishlab chiqishda muhim ahamiyat kasb etadi.[Mengliev, D., Abdurakhmonova 2025]

Xulosa. Zamonaviy dialektal korpuslar tajribasini tahlil qilish shuni ko‘rsatadiki, bugungi kunda dialektologik materialni oddiy matn yoki lug‘at ko‘rinishida saqlash yetarli emas.

Dialektal nutqning tabiiy shaklini audio orqali saqlash, uni bir yoki bir nechta transkripsiya qatlamlari bilan bog‘lash, informant va hudud haqidagi metadata bilan ta’minlash hamda lingvistik jihatdan annotatsiyalash korpus yaratishning muhim shartlariga aylanmoqda. Tadqiqot natijasida shakllangan baza o‘zbek tili dialektal resurslarini raqamlashtirish uchun amaliy model vazifasini bajarishi mumkin. Ayniqsa, audio, matn, transkripsiya, metadata va teglashni bir tizimda birlashtirish kelgusida Qashqadaryoning boshqa hududlari materiallarini ham yagona korpus arxitekturasiga kiritish imkonini beradi. Kelgusidagi tadqiqotlarda bazani kengaytirish, audio-matn alignmentini avtomatlashtirish, morfologik va sintaktik annotatsiyani takomillashtirish, dialektal birliklarning chastota va hududiy tarqalishini statistik modellashtirish hamda mashinali o‘qitish asosida Qarshi shevasini avtomatik aniqlash masalalarini tadqiq etish maqsadga muvofiqdir.

FOYDALANILGAN ADABIYOTLAR

1. Abduraxmonova N.Z. Korpus lingvistikasi, (darslik), Toshkent : GlobeEdit, 2023. – B.29
2. McEnery T., Hardie A. Corpus Linguistics: Method, Theory and Practice. – Cambridge: Cambridge University Press, 2011. – 294 p.
3. Ashurboyev S. O‘zbek shevalarining fonetik tizimi va transkripsiya masalalari. – Toshkent: Fan, 1990. – 180 b.
4. Шерматов А. Узбекские народные говоры Кашкадарьинской области. – Ташкент: Фан, 1978. – 164 с.
5. Johannessen J.B. Annotations in the Nordic Dialect Corpus // Proceedings of the 21st Nordic Conference on Computational Linguistics (NoDaLiDa 2017). – Gothenburg, 2017. – P. 251–254.
6. Bentum M., Sanders E., van den Bosch A.P.J., Zeldenrust D., van den Heuvel H. Corpus Creation and Automatic Alignment of Historical Dutch Dialect Speech // Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024). – Torino, 2024. – P. 4021–4029.
7. Kuparinen O., Scherrer Y. Corpus-based dialectometry with topic models // Journal of Linguistic Geography. – 2024. – Vol. 12, No. 1. – P. 1–12. – DOI: 10.1017/jlg.2024.6.
8. Abdurakhmonova N., Astanova G., Akhmedov A., Mengliev D., Ibragimov B., Abdullayev A. Development of a lexical dataset and a rule-based algorithm for the analysis of Khorezm dialects of the Uzbek language // Data in Brief. – 2025. – Vol. 63. – Art. 112216. – DOI: 10.1016/j.dib.2025.112216.
9. Zhao Y., Ding Y., Min X. Construction of a multimodal dialect corpus based on deep learning and digital twin technology: A case study on the Hangzhou dialect // Language and Linguistics Compass. – 2025. – DOI: 10.1177/14727978241299701.
10. Abdullayeva N.M. **O‘zbek tili korpusi uchun Qashqadaryo shevasi lingvistik bazasini yaratish**: magistrlik dissertatsiyasi. – Toshkent: Mirzo Ulug‘bek nomidagi O‘zbekiston Milliy universiteti, 2026.
11. Mengliev, D., Abdurakhmonova, N., Kholmurodova, I., Ibragimov, B., Latipova, G., & Kadirova, Z. (2025, October). Bridging dialectal variations in Uzbek texts: A comparative evaluation of modern approaches. In *AIP Conference Proceedings* (Vol. 3377, No. 1, p. 070001). AIP Publishing LLC.

AXBOROT TEXNOLOGIYALARI TERMINLARINI ONTOLOGIK TASNIFLASHDAGI MUAMMOLAR VA ULARNING YECHIMLARI

Abdullayeva Shaxnoza Raximjon qizi
Tayanch doktorant O'zbekiston davlat jahon tillari universitet
Orcid: 0009-0000-8979-8087
shakhnozaa40@gmail.com (88)680 19 11
https://doi.org/10.5281/zenodo.23135867

Annotatsiya. *Mazkur tezisda axborot texnologiyalari terminlarini ontologik tasniflash jarayonida uchraydigan asosiy lingvistik va leksikografik muammolar tahlil qilinadi. Tadqiqotning maqsadi axborot texnologiyalari terminlarini ontologik modellashtirishda yuzaga keladigan qiyinchiliklarni aniqlash hamda ularni bartaraf etishning samarali usullarini ishlab chiqishdan iborat. Tadqiqot davomida tavsifiy, qiyosiy, komponent tahlil, terminologik tahlil hamda ontologik modellashtirish metodlaridan foydalanildi. Tahlil natijasida terminlarning polisemantikligi, sinonimiyasi, normativ ekvivalentlarning yetishmasligi, klassifikatsiya tizimlari o'rtasidagi tafovutlar va semantik munosabatlarni aniqlashdagi murakkabliklar asosiy muammolar sifatida belgilandi. Ushbu muammolarni bartaraf etish uchun terminologik standartlashtirish, ontologik xaritalash, korpus texnologiyalaridan foydalanish va ekspert verifikatsiyasiga asoslangan yondashuv taklif etildi. Tadqiqot natijalari axborot texnologiyalari terminlarining o'zbek tilidagi ontologik lug'atini yaratishda metodologik asos bo'lib xizmat qilishi mumkin.*

Kalit so'zlar: *ontologiya, ontologik lug'at, axborot texnologiyalari, terminologiya, tasnif, semantik munosabatlar*

Аннотация. *В тезисе анализируются основные лингвистические и лексикографические проблемы онтологической классификации терминов информационных технологий. Цель исследования заключается в выявлении трудностей, возникающих при онтологическом моделировании ИТ-терминов, и разработке эффективных способов их решения. В работе использованы описательный, сравнительный, компонентный, терминологический методы анализа, а также методы онтологического моделирования. Результаты исследования показали, что основными проблемами являются полисемия терминов, наличие синонимии, отсутствие нормативных эквивалентов, различия между классификационными системами и сложность выявления семантических отношений. Для их решения предложен подход, основанный на стандартизации терминологии, онтологическом картировании, использовании корпусных технологий и экспертной верификации. Полученные результаты могут служить методологической основой для создания онтологического словаря ИТ-терминов на узбекском языке.*

Ключевые слова: *онтология, онтологический словарь, информационные технологии, терминология, классификация, семантические отношения*

Abstract. *This thesis analyzes the major linguistic and lexicographic challenges involved in the ontological classification of Information Technology terms. The aim of the study is to identify the difficulties arising in ontology-based modeling of IT terminology and to develop effective approaches for addressing them. Descriptive, comparative, componential,*

terminological analysis and ontology modeling methods were employed. The findings indicate that polysemy, synonymy, lack of standardized equivalents, inconsistencies among classification systems, and difficulties in identifying semantic relations constitute the main challenges of ontological classification. To address these issues, an approach based on terminology standardization, ontology mapping, corpus technologies, and expert verification is proposed. The results may serve as a methodological foundation for developing an ontological dictionary of Information Technology terms in Uzbek.

Keywords: *ontology, ontological dictionary, information technology, terminology, classification, semantic relations*

Kirish

Axborot texnologiyalari sohasi raqamli transformatsiya jarayonlarining jadallashuvi natijasida eng tez rivojlanayotgan ilmiy-amaliy yo‘nalishlardan biriga aylandi. Mazkur sohada yuzaga kelayotgan yangi tushunchalar va terminlar terminologik tizimning uzluksiz kengayishiga olib kelmoqda. Shu sababli terminlarni tizimlashtirish, ularning konseptual aloqalarini aniqlash va ontologik modelga joylashtirish masalasi dolzarb ahamiyat kasb etmoqda. Ontologik tasniflash terminlar o‘rtasidagi giperonim-giponim, meronim-holonim hamda assotsiativ munosabatlarni aniqlash imkonini beradi.

Tadqiqot metodlari

Tadqiqotda tavsifiy, qiyosiy, komponent tahlil, terminologik tahlil va ontologik modellashtirish metodlaridan foydalanildi. Shuningdek, ACM Computing Classification System (CCS), ISO 704 va zamonaviy ontologiya nazariyasining tamoyillari metodologik asos sifatida qo‘llanildi.

Natijalar va muhokama

Tahlillar axborot texnologiyalari terminlarini ontologik tasniflashda bir qator tizimli muammolar mavjudligini ko‘rsatdi. Birinchi muammo terminlarning polisemantikligi bilan bog‘liq. Masalan, “port”, “thread”, “cache” kabi terminlar turli kontekstlarda turlicha ma’nolarni ifodalaydi. Bu esa ontologik modelda konseptlarni aniq chegaralashni qiyinlashtiradi.

Ikkinchi muammo sinonim terminlarning mavjudligidir. Bir tushunchaning bir nechta terminlar orqali ifodalanishi ontologik bazalarda takroriy yozuvlar paydo bo‘lishiga olib keladi. Uchinchi muammo esa o‘zbek tilida normativ ekvivalentlarning yetarli darajada shakllanmaganligi bilan bog‘liq. To‘rtinchi muammo turli klassifikatsiya tizimlari o‘rtasidagi tafovutlar bilan izohlanadi. ACM CCS, Computer Science Ontology va boshqa bilimlar bazalarida terminlarning joylashuvi hamda ierarxiyasi farqlanadi. Bu holat interoperabellikni pasaytiradi.

Beshinchi muammo semantik munosabatlarni aniqlash murakkabligidir. Ontologik lug‘atda har bir terminning boshqa terminlar bilan konseptual aloqalari ko‘rsatilishi lozim. Buning uchun katta hajmdagi korpus ma’lumotlari va ekspert baholash jarayonlari talab etiladi.

Muallif tomonidan terminologik standartlashtirish, ontologik xaritalash (mapping), korpus lingvistikasi vositalaridan foydalanish hamda ekspert verifikatsiyasiga asoslangan model taklif etildi. Mazkur model turli manbalardagi terminlarni yagona ontologik tuzilma asosida birlashtirish imkonini beradi.

Tahlillar axborot texnologiyalari terminlarini ontologik tasniflashda bir qator tizimli muammolar mavjudligini ko‘rsatdi. Birinchi muammo terminlarning polisemantikligi bilan

bog‘liq. Masalan, “port”, “thread”, “cache” kabi terminlar turli kontekstlarda turlicha ma’nolarni ifodalaydi. Bu esa ontologik modelda konseptlarni aniq chegaralashni qiyinlashtiradi.

Ikkinchi muammo sinonim terminlarning mavjudligidir. Bir tushunchaning bir nechta terminlar orqali ifodalanishi ontologik bazalarda takroriy yozuvlar paydo bo‘lishiga olib keladi. Uchinchi muammo esa o‘zbek tilida normativ ekvivalentlarning yetarli darajada shakllanmaganligi bilan bog‘liq.

To‘rtinchi muammo turli klassifikatsiya tizimlari o‘rtasidagi tafovutlar bilan izohlanadi. ACM CCS, Computer Science Ontology va boshqa bilimlar bazalarida terminlarning joylashuvi hamda ierarxiyasi farqlanadi. Bu holat interoperabellikni pasaytiradi.

Beshinchi muammo semantik munosabatlarni aniqlash murakkabligidir. Ontologik lug‘atda har bir terminning boshqa terminlar bilan konseptual aloqalari ko‘rsatilishi lozim. Buning uchun katta hajmdagi korpus ma’lumotlari va ekspert baholash jarayonlari talab etiladi.

Muallif tomonidan terminologik standartlashtirish, ontologik xaritalash (mapping), korpus lingvistikasi vositalaridan foydalanish hamda ekspert verifikatsiyasiga asoslangan model taklif etildi. Mazkur model turli manbalardagi terminlarni yagona ontologik tuzilma asosida birlashtirish imkonini beradi.

Tahlillar axborot texnologiyalari terminlarini ontologik tasniflashda bir qator tizimli muammolar mavjudligini ko‘rsatdi. Birinchi muammo terminlarning polisemantikligi bilan bog‘liq. Masalan, “port”, “thread”, “cache” kabi terminlar turli kontekstlarda turlicha ma’nolarni ifodalaydi. Bu esa ontologik modelda konseptlarni aniq chegaralashni qiyinlashtiradi.

Ikkinchi muammo sinonim terminlarning mavjudligidir. Bir tushunchaning bir nechta terminlar orqali ifodalanishi ontologik bazalarda takroriy yozuvlar paydo bo‘lishiga olib keladi. Uchinchi muammo esa o‘zbek tilida normativ ekvivalentlarning yetarli darajada shakllanmaganligi bilan bog‘liq.

To‘rtinchi muammo turli klassifikatsiya tizimlari o‘rtasidagi tafovutlar bilan izohlanadi. ACM CCS, Computer Science Ontology va boshqa bilimlar bazalarida terminlarning joylashuvi hamda ierarxiyasi farqlanadi. Bu holat interoperabellikni pasaytiradi.

Beshinchi muammo semantik munosabatlarni aniqlash murakkabligidir. Ontologik lug‘atda har bir terminning boshqa terminlar bilan konseptual aloqalari ko‘rsatilishi lozim. Buning uchun katta hajmdagi korpus ma’lumotlari va ekspert baholash jarayonlari talab etiladi.

Muallif tomonidan terminologik standartlashtirish, ontologik xaritalash (mapping), korpus lingvistikasi vositalaridan foydalanish hamda ekspert verifikatsiyasiga asoslangan model taklif etildi. Mazkur model turli manbalardagi terminlarni yagona ontologik tuzilma asosida birlashtirish imkonini beradi.

Xulosa

Tadqiqot natijalari axborot texnologiyalari terminlarini ontologik tasniflashda polisemiya, sinonimiya, normativ ekvivalentlarning yetishmasligi, klassifikatsion tafovutlar va semantik aloqalarni aniqlashdagi qiyinchiliklar asosiy muammolar ekanligini ko‘rsatdi. Taklif etilgan yondashuvlar ontologik lug‘atlar yaratish jarayonini takomillashtirishga, terminologik standartlashtirishni kuchaytirishga hamda o‘zbek tilidagi IT terminologiyasining ilmiy bazasini shakllantirishga xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR

1. Abduraxmonova N. Kompyuter lingvistikasi. – T.: Nodirabegim, 2021.

2. Gruber T. R. A Translation Approach to Portable Ontology Specifications // *Knowledge Acquisition*. – 1993. – Vol. 5. – № 2. – P. 199–220.
3. Cimiano P. *Ontology Learning and Population from Text: Algorithms, Evaluation and Applications*. – New York: Springer, 2006.
4. Staab S., Studer R. *Handbook on Ontologies*. – Berlin: Springer, 2009.
5. ISO 704:2022. *Terminology Work – Principles and Methods*. – Geneva: International Organization for Standardization, 2022.
6. ACM Computing Classification System (CCS). Association for Computing Machinery, 2012.
7. Buitelaar P., Cimiano P., Magnini B. *Ontology Learning from Text: Methods, Evaluation and Applications*. – Amsterdam: IOS Press, 2005.

BADIIY MULOQOTDA IMPLIKATURA INTERPRETATSIYASI VA PRAGMATIK TO'SIQ MUAMMOSI (ERKIN A'ZAM ASARLARI MISOLIDA)

Abdunazarova Maftuna Faxriddinovna
Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti,
Jurnalistika va o'zbek filologiyasi fakulteti, tayanch doktoranti
E-mail: abdunazarova_m@nuu.uz
ORCID ID: 009-0000-4417-1512
https://doi.org/10.5281/zenodo.23135887

Annotatsiya. *Ushbu maqolada zamonaviy antropotsentrik tilshunoslikning dolzarb yo'nalishlaridan biri — badiiy muloqotda implikatura (tagma'no) interpretatsiyasi va pragmatik to'siq fenomenining namoyon bo'lish xususiyatlari tadqiq etilgan. Tadqiqot obyekti sifatida taniqli o'zbek adibi Erkin A'zam asarlaridagi dialogik diskurslar tanlab olingan. Maqolada personajlar nutqidagi yashirin ma'nolarning tinglovchi tomonidan noto'g'ri talqin qilinishi, muloqot qoidalarining buzilishi hamda sotsiopragmatik omillar ta'sirida yuzaga keladigan kommunikativ muvaffaqiyatsizliklar tahlil qilingan. Erkin A'zam prozasida pragmatik to'siq shunchaki lisoniy xato emas, balki qahramonlar xarakterini, ularning dunyoqarashidagi ziddiyatlarni va ijtimoiy-psixologik muammolarni ochib beruvchi yetakchi badiiy-estetik vosita ekanligi isbotlangan. Tahlillar natijasida dialogik muloqotda lingvistik va ekstralingvistik omillarning o'zaro munosabati tizimlashtirilgan.*

Kalit so'zlar: *pragmalingvistika, dialogik diskurs, implikatura, interpretatsiya, pragmatik to'siq, kommunikativ muvaffaqiyatsizlik, nutqiy akt, sotsiopragmatika, kolliziya.*

Abstract. *This article investigates one of the current directions of modern anthropocentric linguistics — the features of implicature (hidden meaning) interpretation and the phenomenon of pragmatic failure in artistic communication. The dialogic discourse in the works of the famous Uzbek writer Erkin A'zam is selected as the object of research. The article analyzes communicative failures resulting from the listener's misinterpretation of implicit meanings in the characters' speech, violation of conversational maxims, and sociopragmatic factors. It is proven that in Erkin A'zam's prose, pragmatic failure is not merely a linguistic error, but a leading artistic and aesthetic tool that reveals the characters' personalities, conflicts in their worldviews, and socio-psychological issues. As a result of the analysis, the interaction of linguistic and extralinguistic factors in dialogic communication is systematized.*

Key words: *pragmalinguistics, dialogic discourse, implicature, interpretation, pragmatic failure, communicative failure, speech act, sociopragmatics, collision.*

Аннотация. *В данной статье исследуется одно из актуальных направлений современной антропоцентрической лингвистики — особенности интерпретации импликатуры (подтекста) и проявление феномена прагматической ошибки (барьера) в художественной коммуникации. В качестве объекта исследования выбран диалогический дискурс в произведениях известного узбекского писателя Эркина Аъзама. В статье анализируются коммуникативные неудачи, возникающие в результате неверного истолкования слушателем скрытых смыслов в речи персонажей, нарушения принципов общения и социопрагматических факторов. Доказано, что в прозе Эркина Аъзама*

прагматическая ошибка является не просто лингвистической погрешностью, а ведущим художественно-эстетическим средством, раскрывающим характеры героев, противоречия в их мировоззрении и социально-психологические проблемы. В результате анализа систематизировано взаимодействие лингвистических и экстралингвистических факторов в диалогическом общении.

***Ключевые слова**: прагмалингвистика, диалогический дискурс, импликатура, интерпретация, прагматическая ошибка, коммуникативная неудача, речевой акт, социопрагматика, коллизия.*

1. KIRISH

Badiiy muloqotda ko‘pincha kommunikativ mazmun bevosita ifodalanmaydi. Personaj nutqining yuzaki ma’nosi ortida kommunikativ vaziyat, suhbatdoshlar o‘rtasidagi munosabat, umumiy bilim va kommunikativ niyat asosida anglashiladigan qo‘shimcha ma’no yotadi. Shuning uchun dialogning pragmatik talqinida ifodalanayotgan va nazarda tutilayotgan ma’no munosabati muhim ahamiyat kasb etadi.

Implikatura — bu so‘zlovchi tomonidan to‘g‘ridan-to‘g‘ri aytilmagan, lekin adresat tomonidan ma’lum bir kontekst va kommunikativ intensiya asosida tushuniladigan mazmunni anglatadi. Ushbu yashirin ma’no muvaffaqiyatli tushunilgandagina kommunikativ jarayon davom etadi. Aksincha, so‘zlovchi nazarda tutgan ma’no bilan adresat nazarda tutgan ma’no o‘rtasida nomuvofiqlik bo‘lsa, pragmatik to‘siq yuzaga kelishi mumkin.

Tadqiqot maqsadi — Erkin A’zam asarlaridagi dialogik vaziyatlarda implikaturaning yuzaga kelishi va interpretatsiyasini tahlil qilish hamda pragmatik to‘siqning asosiy ko‘rinishlarini tasniflashdan iborat. Obyekti — dialogik diskurs; predmeti — dialoglarda implikatura, inferensiya va pragmatik to‘siqning o‘zaro munosabati.

2. NAZARIY ASOS VA ADABIYOTLAR SHARHI

Implikatura nazariyasining klassik asosini H. P. Gricening suhbat implikaturasi konsepsiyasi tashkil etadi: hamkorlik tamoyiliga ko‘ra so‘zlovchi ba’zan bevosita aytilmagan qo‘shimcha ma’noni tinglovchi tomonidan tiklanadigan tarzda yuzaga keltiradi [Grice, 1975]. Terkourafi implicit ma’noni kontekst va adresat inferensiyasi bilan bog‘laydi [Terkourafi, 2021], Sbisà esa presuppozitsiya va implikatura farqini ko‘rsatadi [Sbisà, 2021]; Tauchid va boshqalar pragmatik muvaffaqiyatsizlikni dialogik nomuvofiqlik sifatida tahlil qiladi [Tauchid et al., 2022].

O‘zbek tilshunosligida badiiy va rasmiy matnlarning pragmalingvistik, sotsiolingvistik, lingvomadaniy xususiyatlari, shuningdek, raqamli texnologiyalar va matn semantikasi masalalari prof. D. Lutfullayeva, prof. M. Hakimov, prof. Sh. Safarov singari yetakchi olimlarimiz tomonidan chuqur tadqiq etilgan hamda o‘zbek tili materiali asosida o‘ziga xos ilmiy maktablar yaratilgan. Shunday qilib, mavjud tadqiqotlarda implikatura, pragmatik muvaffaqiyatsizlik va kommunikativ interpretatsiya alohida yo‘nalishlarda keng o‘rganilgan bo‘lsa-da, ularning badiiy dialogdagi o‘zaro munosabati, ayniqsa, o‘zbek nasri materialida pragmatik to‘siqning yuzaga kelish nuqtasi sifatida kompleks tahlil etilishi yetarlicha yoritilmagan. Mazkur tadqiqotning asosiy ilmiy yo‘nalishi aynan shu bo‘shliqni to‘ldirishga qaratilgan.

3. TADQIQOT METODOLOGIYASI

Tadqiqot sifat tahliliga asoslangan pragmalingvistik-diskursiv yondashuvda olib borildi; material sifatida “Guli-guli”dan uchta va “Shoirning to‘yi”dan ikkita dialogik vaziyat tanlandi — mezon: so‘zlovchi niyati bilan adresat interpretatsiyasi o‘rtasidagi nomuvofiqlik. Tahlil besh

bosqichda (lokutiv mazmun, kommunikativ niyat, kontekstual inferensiya, adresat interpretatsiyasi, to‘siq turi) amalga oshirildi.

4. TAHLIL VA NATIJALAR

4.1. Majoziy ma’noni literal qabul qilish: interpretativ pragmatik to‘siq

“Guli-guli” qissasidagi O‘rmon kalta va Soli baqiroq muloqoti majoziy ifodaning adresat tomonidan literal qabul qilinishi bilan bog‘liq pragmatik to‘siqni ko‘rsatadi. O‘rmon kalta o‘zini Jiydalining “haqiqiy patrioti” sifatida ko‘rsatib, qishloqning hozirgi holatini kuchli metaforik ifoda bilan baholaydi [E. A’zam, “Guli-guli”, 201–202-betlar].

— Номи достонларда зикр этилган, ўзи қўшиқ-яллаларда мадҳ этилган шундоқ сўлим, шундоқ гўзал маскан ташландиқ бир бойўғлилар маконига айланиб қолди-я! Мен бунга асло чидаёлмайман! Шахс сифатида! Мустақил шахс сифатида! – дея тантанавор тарзда эълон қилади Ўрмон калта.

— Оғзингизга қараб гапиринг, мухбир! – дейди умрида секин сўзламоқни билмаган уккикўз Соли бақироқ одатдагидек овозининг борича. – Бойўғли деганингиз нимаси? Ким ташландиқ, ким бойўғли? Қани!

— Авом! Қип-қизил омисиз, Солибой! – Ўрмон калта бошини орқага ташлаб қаҳ-қаҳ уради. – Ахир, бу муболаға-ку! Мажоз дейдилар буни! [E. A’zam, “Guli-guli”, 202-bet].

O‘rmon kaltaning kommunikativ niyati qishloqning tashlandiq holatini kuchaytirilgan obraz orqali baholashdir: “бойўғлилар макони” bu yerda denotativ emas, balki salbiy ijtimoiy-estetik bahoni ifodalovchi metafora. Soli baqiroq esa uni bevosita ma’noda qabul qilib, himoyalanuvchi savol beradi — bu esa O‘rmon kaltani “Мажоз дейдилар буни!” degan metapragmatik izoh bilan tuzatishga majbur qiladi. Demak, pragmatik to‘siqning sababi lisoniy shaklning o‘zi emas, balki adresatning majoziy signalni qayta ishlash uchun zarur kontekstual inferensiyani amalga oshirmaganidadir.

Mazkur misolda to‘siqni “literalizatsiya” modeli bilan ifodalash mumkin: majoziy signal → literal talqin → himoyalanuvchi reaksiya → metapragmatik tuzatish. Badiiy jihatdan bu O‘rmon kaltaning balandparvoz nutqi bilan Soli baqiroqning sodda reaksiyasi o‘rtasidagi kontrastni kuchaytiradi.

4.2. Umumiy kommunikativ kodning yetishmasligi: sotsiokognitiv to‘siq

“Guli-guli” qissasida “Гули-гули” birligi atrofida yuzaga kelgan munozara pragmatik to‘siqning boshqa shaklini ko‘rsatadi. Bu yerda muammo bitta so‘zning lug‘aviy ma’nosini bilmaslikdan ko‘ra, uning vaziyatdagi kommunikativ qiymatini birgalikda aniqlay olmaslik bilan bog‘liq.

— Оёғидаги пойабзали ғиж-ғиж этганга ўхшади, тамом.

— Қизиқ, бизникига келганда худди гапиргандай, “Гули-гули” дегандай бўлди-я! Билмадим, бу нима дегани экан.

— Қулоғингизга шундай чалингандир-да, Исматбой?

— Йўқ, аниқ-тиниқ эшитдим: гули-гули! [...]

— Гули-гули? Ўзбекча сўзга ўхшайди-ку, – деди Аҳад махсум. – Разведкачилар ўзбек бўларканми! [...]

— Бўлди, бўлди, укам, билдик. Биз ҳам сиз кийган этикдан кийганмиз! – деди Соли бақироқ. [E. A’zam, “Guli-guli”, 216-bet].

Bir xil signal — “Гули-гули” — atrofida bir nechta alternativ interpretatsiya paydo bo‘ladi: Ismatboy uni aniq eshitilgan birlik sifatida himoya qiladi, boshqa qahramon ayol ismi

bilan bog‘laydi, Ahad maxsum o‘zbekcha so‘zga o‘xshashini ta’kidlaydi, harbiy xizmatdan qaytgan yigit esa ruscha variantni taxmin qiladi. Pragmatik to‘siqning manbai — umumiy kommunikativ kodning yetishmasligi: qahramonlar bir xil tovushni eshitgan bo‘lsa-da, uning referensial-situativ qiymatini birgalikda aniqlay olmaydi, natijada kommunikativ jarayon bir ma’noni yetkazishdan ko‘ra ehtimoliy ma’nolar ishlab chiqarishga aylanadi.

Mazkur misol sotsiopragmatik jihatdan ham muhim: qahramonlarning ijtimoiy tajribasi va bilimlari turlicha bo‘lgani uchun ular bitta signalni turli “freymlar” orqali qayta ishlaydi — demak, pragmatik to‘siqda til kodidan tashqari umumiy tajriba va vaziyat bilimi ham hal qiluvchi rol o‘ynaydi.

4.3. Kommunikativ niyatlar to‘qnashuvi: shaxslararo pragmatik to‘siq

“Guli-guli” qissasidagi Olma kelin va hikoyachi o‘rtasidagi vaziyat pragmatik to‘siqning yana bir ko‘rinishini namoyon etadi: hikoyachi Olma kelindan tortinib, vaziyatga ehtiyotkorlik bilan kirib boradi, biroq kechasi kommunikativ vaziyat birdan o‘zgaradi.

— Э-э, мушуги бор! – дея баҳона қилмоқчи бўлдим.

— Келинг, укажон, келаверинг. Мушукни оғилхонага қамаб даф қиламан! Қарасам – девор кемтигидан Олма келиннинг ўзи илжайибгина турибди. Бошқа баҳона йўқ! [...]

Бир маҳал юмшоққина нимадир юзимга суйкалаётгандек бўлдию чўчиб уйғониб кетдим. Мушукми?! Биров тез-тез нафас олаётир. Атир ҳиди.

— Ҳеч кимга айтманг, ўлайжон, сизни би-ир яхши кўриб кетдим-да.

— Э, боринг-э! – дедим дағ-дағ титраб. [...]

Қўрққанимдан тонггача мижжа қоқмай ётдим. [E. A’zam, “Guli-guli”, 219-bet].

Olma kelinning replikasi ochiq mazmunda sevgi izhorini bildiradi, biroq kontekstda shaxslararo yaqinlikni kuchaytiruvchi signal sifatida ishlaydi; hikoyachining “Э, боринг-э!” javobi va “дағ-дағ титраб” holati esa bu yaqinlikni qabul qilmaslik va kommunikativ masofani tiklash funksiyasini bajaradi. Demak, to‘siqning manbai so‘zning tushunilmasligi emas, balki bir xil vaziyatga berilgan ikki xil baho: so‘zlovchi yaqinlashuvni ko‘zlasa, adresat uni noqulay va xavfli deb qabul qiladi, qo‘rquvning kuchayishi esa muloqotning muvaffaqiyatsizlik bilan yakunlanganini tasdiqlaydi.

Bu misol pragmatik to‘siqning psixologik va interaksion xarakterini ko‘rsatadi: kommunikativ signal mavjud, signal anglangan, biroq unga berilgan interpretativ va emotsional javob kutilgan yo‘nalishga mos kelmaydi [Mengliyev B. va boshq].

4.4. Ritorik so‘roq orqali yashirin ayblov: implikativ to‘siq

“Shoirning to‘yi” qissasida Oqsoqol shoir marhum do‘sti haqida rasmiy-hamdardlik nutqini davom ettirayotgan vaziyatda kutilmagan savol bilan to‘xtatiladi. Savolning qisqaligi uning pragmatik kuchini kamaytirmaydi, aksincha, kontekst sababli yashirin ayblovni kuchaytiradi.

— Биз у билан ҳар вақт, ҳар жойда бирга эдик... Афсус ва надоматлар бўлсинким, шундоқ ноёб истеъдод соҳиби, ажойиб дўстим ноҳақ туҳматларга дучор бўлиб, халқ душмани сифатида олис совуқ ўлкаларга сургун этилди.

Ён тарафдан кимдир қалтис, омонсиз луқма ташлади:

— Сиз-чи? Дўстингизни ҳимоя қилмадингизми?

Ҳайъатда бесаранжомлик бошланди. [E. A’zam, “Shoirning to‘yi”, 5-bet].

Savolning lokutiv mazmuni oddiy axborot talab qilsa-da, kontekst uning illokutiv kuchini o‘zgartirib, Oqsoqolning o‘tmishdagi pozitsiyasini shubha ostiga qo‘yadi: tagma’no darajasida

"siz do'stingizni himoya qilmadingiz" degan xulosa yuzaga keladi. Yashirin ayblov ochiq hukm emas, so'roq shaklida berilgani uchun adresat bevosita javob bermay, rasmiy nutqini davom ettiradi — "Ҳайъатда бесаранжомлик бошланди" tasviri savolning kuchli ta'sirini tasdiqlaydi. Demak, bu ritorik so'roq ham axborot so'rovchi nutqiy akt, ham implikatura orqali yashirin ayblovni faollashtiruvchi vosita bo'lib xizmat qiladi.

4.5. Noqulay savolga javob bermaslik: strategik pragmatik to'siq

"Shoirning to'yi" qissasining 22-betidagi vaziyatda yosh qahramon tarixiy va axloqiy javobgarlikka oid savollarni ketma-ket beradi. Savollarning yakuniy qismi Ajoyib domlani aniq javob berishga majbur qiluvchi kommunikativ vazifani bajaradi.

> *— Заҳматкаш домла, кўзингиз тушгандир, ўша жиноятномада яна икки кишининг кўргазмаси бор. Бир аёл билан бир эркакнинг. ... Хўш, ўша «олижаноб» эркак киму «жабрдийда» хотин ким?.. ... Ўша йиллари ким бу ташкилотнинг раҳбари эди? Мана шу бир тўп саволга сиз жавоб берсангиз, ҳурматли Оқсоқол!*
>
> ...
>
> *— Демагогияни бас қилинг, ҳўв ука, — деб қўйди Ажойиб домла ҳамон ишшайганича.* [E. A'zam, "Shoirning to'yi", 22-bet].

So'zlovchi savollar orqali faktik axborot olish bilan birga, adresatni ijtimoiy-axloqiy javobgarlik pozitsiyasiga olib keladi; Ajoyib domlaning "Демагогияни бас қилинг" javobi esa mazmuniy savolga javob bermay, kommunikativ predmetni "kim?" savolidan "savol beruvchining nutqi qanday?" degan baholovchi yo'nalishga buradi. Shu tariqa strategik tuzilish o'zgaradi: faktik savol → javob berish ehtiyoji → savolni diskreditatsiya qilish → predmetni almashtirish — bu yerda to'siq adresatning bilmasligidan emas, javob berishdan qochish ehtiyojidan kelib chiqadi.

Asar kontekstida bu replikani yanada kengroq kommunikativ tizimda ko'rish mumkin. Ajoyib domla "Демагогияни бас қилинг" degach, Jasur shoir "Демагогия? Сизлар яхши кўрган сўз бу! Иложсиз, чорасиз қолганда шуни қурол қилиб одамни айбламоқчи бўласизлар" deya javob qaytaradi. Demak, strategik to'siqning o'zi ham yangi bahs predmetiga aylanadi [A'zam, "Shoirning to'yi", 22-bet].

5. MUHOKAMA

Tahlil natijalari Erkin A'zam dialoglarida pragmatik to'siqning kamida uch darajasini ko'rsatadi: interpretativ (majoziy signalning literal qabul qilinishi — "бойўғлилар макони" misolida), sotsiokognitiv (umumiy kod yetishmasligi — "Гули-гули" va Olma kelin misollarida) va strategik ("Shoirning to'yi"dagi yashirin ayblov va javobdan qochish). Bu shuni ko'rsatadiki, pragmatik muvaffaqiyatsizlik faqat "tushunmovchilik" emas: ma'no to'g'ri anglanib, unga boshqacha kommunikativ javob berilishi ham mumkin. Ilmiy yangilik — pragmatik to'siq ma'no dekodlanishi, sotsiokognitiv moslashuv va kommunikativ strategiya sathlarida tizimlashtirilib, adib uslubida o'quvchi interpretatsiyasini faollashtiruvchi poetik vositaga aylanadi.

6. XULOSA

Olib borilgan tahlil shuni ko'rsatdiki, Erkin A'zam badiiy dialoglarida implikatura va pragmatik to'siq ajralmas hodisalar: to'siq so'zlovchi yaratgan signal bilan adresat tiklagan ma'no o'rtasidagi nomuvofiqlikdan yuzaga keladi — majoziy ifodaning literal talqini, umumiy kod yetishmasligi, kommunikativ niyatlar to'qnashuvi yoki strategik javobsizlik shaklida.

Shunday qilib, pragmatik to'siqni oddiy "tushunmovchilik" sifatida baholash yetarli emas — u qahramonlarning kognitiv tajribasi va ijtimoiy pozitsiyasini ochuvchi diskursiv hodisa

bo‘lib, konfliktni kuchaytirish, satirik baho yaratish va kitobxon interpretatsiyasini rag‘batlantirish funksiyalarini bajaradi.

FOYDALANILGAN ADABIYOTLAR

1. A’zam, E. “Guli-guli” (qissa). Ziyouz elektron kutubxonasi nusxasi.
2. A’zam, E. “Shoirning to‘yi” (qissa). Ziyouz elektron kutubxonasi nusxasi.
3. Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and Semantics 3: Speech Acts (pp. 41–58). Academic Press.
4. Dahlman, R. C. (2022). The Non-saying of What Should Have Been Said. Acta Analytica, 37, 441–451. https://doi.org/10.1007/s12136-021-00481-1
5. Marty, P., & Romoli, J. (2021). Presuppositions, implicatures, and contextual equivalence. Natural Language Semantics, 29, 229–280. https://doi.org/10.1007/s11050-021-09176-0
6. Permatasari, K. (2022). Aspek-aspek Pragmatik pada Percakapan dalam Dialog pada Novel Jepang dan Novel Terjemahannya. Sintaks: Jurnal Bahasa & Sastra Indonesia, 2(1), 10–20. https://doi.org/10.57251/sin.v2i1.189
7. Ronai, E., & Xiang, M. (2021). Pragmatic inferences are QUD-sensitive: an experimental study. Journal of Linguistics. https://doi.org/10.1017/S0022226720000250
8. Sbisà, M. (2021). Presupposition and implicature: Varieties of implicit meaning in explicitation practices. Journal of Pragmatics, 182, 176–188. https://doi.org/10.1016/j.pragma.2021.05.027
9. Tauchid, A., Kumala, E. T. N., & Rizkiana, A. N. (2022). Pragmatic Failure of Students Conversation in Speaking Class. IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, 10(1), 975–983. https://doi.org/10.24256/ideas.v10i1.2785
10. Terkourafi, M. (2021). Pragmatics as an interdisciplinary field. Journal of Pragmatics, 179, 77–84. https://doi.org/10.1016/j.pragma.2021.04.015
11. Mengliyev, B., Xoliyorov, O., & Abdurahmonova, N. (2014). O‘zbek tilidan Universal qo ‘llanma (qayta ishlangan 3-nashri). Toshkent: akademnashr, 389.

SUN'IY INTELEKT TEXNOLOGIYALARI YORDAMIDA ERGANOMIKANI TAHLIL QILISH ORQALI KORXONALARDA MEHNAT XAVFSIZLIGINI TA'MINLASH

Abduraxmanova Azoda Djo'rayevna[1], Solixov Murod Botir o'g'li[2]
[1]Islom Karimov nomidagi Toshkent davlat texnika universiteti "Hayot faoliyati xavfsizligi" kafedrasi t.f.d.professor
[2]"Crafers" M.Ch.J fabrikasi Mehnat muhofazasi va texnika xavfsizligi muhandisi
https://doi.org/10.5281/zenodo.23135913

Annotatsiya*: Mazkur tezisda zamonaviy korxonalarda ergonomik xavflarni aniqlash va mehnat xavfsizligini ta'minlashda sun'iy intellekt texnologiyalaridan foydalanish imkoniyatlari yoritilgan. Texnologik taraqqiyot va ishlab chiqarish jarayonlarining raqamlashtirilishi natijasida ayrim ish faoliyatlarida xodimlarning uzoq vaqt davomida bir xil yoki noqulay tana holatida ishlashi, kamharakatlilik va takroriy harakatlarning ortishi ergonomik xavflarning yuzaga kelishiga sabab bo'lmoqda. Bunday xavflarni o'z vaqtida aniqlash va baholash xodimlarning sog'lig'ini saqlash hamda mehnat xavfsizligini oshirishda muhim ahamiyatga ega.*

Tadqiqotda ergonomik xavflarni baholashning RULA (Rapid Upper Limb Assessment) va REBA (Rapid Entire Body Assessment) metodlarining mazmuni va ularni sun'iy intellekt texnologiyalari bilan integratsiya qilish imkoniyatlari ko'rib chiqiladi.

Shuningdek, tadqiqotda sun'iy intellekt yordamida real ish sharoitlariga yaqin video va tasvirli o'quv materiallarini yaratish, xodimlarga to'g'ri va noto'g'ri ish holatlarini ko'rgazmali tarzda namoyish etish hamda ergonomik xavfsizlik bo'yicha bilim va amaliy ko'nikmalarni oshirish imkoniyatlari taklif etiladi. Natijada sun'iy intellekt texnologiyalarini ergonomik baholash va xodimlarni o'qitish jarayonlariga joriy etish orqali korxonalarda ergonomik xavflarni muntazam monitoring qilish, xavfli ish joylarini aniqlash va mehnat xavfsizligini takomillashtirishga xizmat qiluvchi integratsiyalashgan yondashuvni shakllantirish mumkin.

Kalit so'zlar: *ergonomika, sun'iy intellekt, mehnat xavfsizligi, mehnat muhofazasi, ergonomik xavf, Computer Vision, Human Pose Estimation, RULA, REBA, ergonomik baholash, ish joyi, mushak-skelet tizimi, xavflarni monitoring qilish, raqamli texnologiyalar, xodimlarni o'qitish.*

Bugungi kunda raqamli texnologiyalar va avtomatlashtirilgan tizimlarning jadal rivojlanishi natijasida mehnat jarayonlarining mazmuni va xodimlarning ish faoliyati sezilarli darajada o'zgarib bormoqda. Zamonaviy korxonalarda ko'plab ishlab chiqarish, operatorlik, boshqaruv va ofis faoliyatlari texnologik vositalar, kompyuterlar va avtomatlashtirilgan uskunalar bilan bog'liq bo'lib, ayrim ish jarayonlarida xodimlarning uzoq vaqt davomida o'tirgan yoki bir xil tana holatida ishlashi kuzatilmoqda. Bunday holatlarda ish joyining ergonomik jihatdan to'g'ri tashkil etilmasligi, tana holatining uzoq vaqt davomida noto'g'ri yoki noqulay pozitsiyada saqlanishi, takroriy harakatlar va statik jismoniy zo'riqish xodimlarning sog'lig'i hamda mehnat qobiliyatiga salbiy ta'sir ko'rsatishi mumkin.

Ergonomik xavf omillari, xususan, noqulay tana holatlari, takroriy harakatlar, uzoq vaqt davomida bir xil pozitsiyada ishlash va ish joyining noto'g'ri tashkil etilishi mushak-skelet tizimi bilan bog'liq muammolarning yuzaga kelishiga sabab bo'luvchi muhim omillardan hisoblanadi.

Jahon sog‘liqni saqlash tashkiloti ma’lumotlariga ko‘ra, mushak-skelet tizimi bilan bog‘liq holatlar insonning harakatlanishi va mehnat faoliyatini cheklashi mumkin.

Shu sababli korxonalarda mavjud ergonomik xavflarni faqat xodimning shikoyati yoki davriy tekshiruv natijasida aniqlash bilan cheklanib qolmasdan, ularni muntazam, obyektiv va imkon qadar avtomatlashtirilgan tarzda tahlil qilish muhim ahamiyat kasb etmoqda.

Bugungi kunda sun’iy intellekt, kompyuter ko‘rishi (Computer Vision) va inson tana holatini aniqlash (Human Pose Estimation) texnologiyalarining rivojlanishi ergonomik xavflarni avtomatik tahlil qilish uchun yangi imkoniyatlarni yaratmoqda. Ushbu texnologiyalar yordamida videotasvir yoki fotosuratlardan inson tanasining asosiy nuqtalarini aniqlash va xodimning ish jarayonidagi tana holatini baholash mumkin. Shuningdek, 2026-yilda e’lon qilingan tadqiqotlarda sun’iy intellekt va inson tana holatini aniqlash texnologiyalaridan foydalanib, ergonomik xavflarni avtomatlashtirilgan tarzda baholash yo‘nalishida 35 ta tadqiqot tahlil qilingan. Shu nuqtayi nazardan, sun’iy intellekt texnologiyalari yordamida ergonomik xavflarni aniqlash, baholash va ular asosida mehnat xavfsizligini takomillashtirish bugungi kundagi dolzarb ilmiy-amaliy masalalardan biri hisoblanadi.

Korxonalarda ergonomik xavflarni aniqlashda amaliyotda mutaxassis tomonidan xodimning ish jarayonini kuzatish, fotosurat yoki videotasvirlarni tahlil qilish hamda maxsus ergonomik baholash metodlaridan foydalanish keng qo‘llaniladi. Biroq ish jarayonlarining uzluksizligi, xodimlar sonining ko‘pligi va turli ish joylarining mavjudligi sababli barcha ishchilarning tana holatini doimiy ravishda vizual nazorat qilish katta vaqt va inson resurslarini talab qiladi.

Tadqiqotning asosiy maqsadi — sun’iy intellekt texnologiyalari yordamida korxonalardagi xodimlarning ish jarayonidagi tana holati va ergonomik xavf omillarini tahlil qilish, RULA va REBA metodlari asosida baholash hamda aniqlangan xavflarni kamaytirish orqali mehnat xavfsizligini ta’minlashga qaratilgan yondashuvni ishlab chiqishdan iborat.

Ushbu maqsadga erishish uchun xodimlarning ish jarayonini foto va videotasvirlar orqali tahlil qilish, inson tana qismlarining holatini aniqlash, tegishli ergonomik ko‘rsatkichlarni hisoblash va aniqlangan natijalar asosida ish joylarini takomillashtirish bo‘yicha tavsiyalar ishlab chiqish ko‘zda tutiladi.

RULA — ishchining asosan yuqori tana qismlariga tushayotgan ergonomik yuklamani tezkor aniqlash imkonini beruvchi baholash metodidir. Ushbu metod yordamida ishchining yuqori qo‘li, pastki qo‘li, bilagi, bo‘yni, gavdasi va oyoqlarining holati, shuningdek mushaklardan foydalanish chastotasi va qo‘llanilayotgan kuch yoki yuklama hisobga olinadi. RULA baholash jarayonida tana qismlarining holati maxsus jadvallar asosida ballarga aylantiriladi va yakunda umumiy 1 dan 7 gacha bo‘lgan RULA balli olinadi.

RULA yakuniy ballining ortishi ergonomik xavf darajasining oshishini anglatadi. Umuman olganda, past ball ish holatining nisbatan maqbul ekanligini, yuqori ball esa ish joyida ergonomik o‘zgarishlar va tezkor choralar zarurligini bildiradi. RULA yakuniy natijasi 1–7 oralig‘ida bo‘lib, natijaga qarab keyingi ergonomik tadbirlar belgilanadi.

Masalan, xodim kompyuter oldida uzoq vaqt davomida bo‘ynini oldinga egib, yelkasini ko‘tarib va bilagini noqulay holatda ushlab ishlasa, RULA yordamida ushbu holatning ergonomik xavfi aniqlanishi mumkin. Xuddi shuningdek, ishlab chiqarish liniyasida xodimning qo‘lini takroran yuqoriga ko‘tarib ishlashi yoki uzoq vaqt davomida gavdasini oldinga egib turishi ham RULA orqali baholanishi mumkin.

REBA — Rapid Entire Body Assessment, ya’ni butun tana holatini tezkor ergonomik baholash metodidir. RULAga nisbatan REBA butun tanaga kengroq e’tibor qaratadi va ayniqsa egilish, burilish, yuk ko‘tarish, itarish, tortish, qo‘l bilan bajariladigan jismoniy ishlar va noqulay tana holatlari mavjud bo‘lgan faoliyatlarni baholashda qo‘llaniladi.

Mazkur yondashuvni korxonalarda qo‘llash natijasida ergonomik xavflarni aniqlash jarayonini tizimlashtirish va qisman avtomatlashtirish imkoniyati yaratiladi. Xususan, xodimning ish jarayonidan olingan video yoki fotosurat AI tizimi orqali tahlil qilinib, uning tana holatidagi asosiy ergonomik xavf omillari aniqlanishi mumkin.

Bundan tashqari, tadqiqot doirasida sun’iy intellekt yordamida real ishlab chiqarish holatlariga yaqin video va tasvirli o‘quv materiallarini yaratish, xodimlarga noto‘g‘ri va to‘g‘ri ish holatlarini vizual tarzda namoyish etish hamda ergonomik xavflarni amaliy misollar orqali tushuntirish imkoniyatlarini o‘rganish nazarda tutiladi. Bunday yondashuv ergonomik talablarni xodimlarga oddiy nazariy ma’lumot sifatida emas, balki real ish jarayoniga yaqin ko‘rgazmali materiallar orqali yetkazish imkonini beradi.

Tahlil natijasida:

• xodimning noto‘g‘ri tana holati aniqlanadi;
• bo‘yin, bel, yelka, tirsak, bilak va tizza holatlari baholanadi;
• uzoq vaqt davom etuvchi statik holatlar aniqlanadi;
• takroriy harakatlar va noqulay pozitsiyalar qayd etiladi;
• RULA yoki REBA bo‘yicha xavf darajasi aniqlanadi;
• yuqori xavfli ish joylari ustuvor ravishda belgilanadi;
• ish joyini qayta tashkil etish bo‘yicha tavsiyalar ishlab chiqiladi;
• xodim uchun to‘g‘ri ish holati va xavfsiz ish usuli ko‘rsatiladi.

Tadqiqot natijasida korxonalarda ergonomik xavflarni baholashning AI asosidagi integratsiyalashgan modelini shakllantirish ko‘zda tutiladi. Ushbu modelda sun’iy intellekt xodimning tana holatini aniqlaydi, RULA va REBA esa aniqlangan holatning ergonomik xavf darajasini baholash uchun metodik asos sifatida xizmat qiladi.

Natijada korxonadagi har bir ish joyi bo‘yicha ergonomik xavf xaritasini shakllantirish imkoniyati paydo bo‘lishi mumkin. Masalan, ish joylari: past xavf → o‘rta xavf → yuqori xavf → juda yuqori xavf darajalari bo‘yicha tasniflanadi.

Bunday tasnif mehnat muhofazasi mutaxassisiga qaysi ish joylariga birinchi navbatda e’tibor qaratish kerakligini aniqlash imkonini beradi. Yuqori xavfli ish joylarida texnologik jarayonni o‘zgartirish, ish jihozlarini ergonomik jihatdan moslashtirish, ish va dam olish rejimini qayta ko‘rib chiqish, mexanizatsiya vositalaridan foydalanish yoki xodimlarni qo‘shimcha o‘qitish kabi choralar ishlab chiqilishi mumkin.

Sun’iy intellekt texnologiyalaridan foydalanish ergonomik xavflarni baholash jarayonini zamonaviylashtirish va mehnat xavfsizligini boshqarish samaradorligini oshirishning istiqbolli yo‘nalishlaridan biridir. RULA va REBA metodlari esa ushbu tizim uchun ilmiy va metodik asos vazifasini bajarishi mumkin. Bundan tashqari, sun’iy intellekt yordamida yaratilgan realistik video va tasvirlardan foydalanish ergonomik xavfsizlik bo‘yicha o‘quv mashg‘ulotlarining ko‘rgazmaliligini oshirishga imkon beradi. Natijada xodimlarga nazariy ma’lumotni yetkazish bilan bir qatorda, noto‘g‘ri va to‘g‘ri ish holatlarini vizual tarzda namoyish etish orqali ularning amaliy ko‘nikmalarini shakllantirish mumkin.

FOYDALANILGAN ADABIYOTLAR

1. McAtamney L., Corlett E.N. RULA: a survey method for the investigation of work-related upper limb disorders // Applied Ergonomics. – 1993. – Vol. 24, No. 2. – P. 91–99.
2. Hignett S., McAtamney L. Rapid Entire Body Assessment (REBA) // Applied Ergonomics. – 2000. – Vol. 31, No. 2. – P. 201–205.
3. Kim W., Sung J., Saakes D., Huang C., Xiong S. Ergonomic postural assessment using a new open-source human pose estimation technology (OpenPose) // International Journal of Industrial Ergonomics. – 2021. – Vol. 84.
4. Development of a fully automated RULA assessment system based on computer vision // International Journal of Industrial Ergonomics. – 2021. – Vol. 86.
5. Iyer H., Jeong H. A systematic review of computer vision-based human pose estimation for occupational safety // International Journal of Industrial Ergonomics. – 2026. – Vol. 113.
6. World Health Organization (WHO). Musculoskeletal health. – Geneva: World Health Organization, 2022. Ushbu manbada mushak-skelet tizimi kasalliklarining global tarqalishi va ularning mehnat faoliyatiga ta'siri yoritilgan.
7. Page M.J., McKenzie J.E., Bossuyt P.M., et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews // BMJ. – 2021.
8. Güler R.A., Neverova N., Kokkinos I. DensePose: Dense Human Pose Estimation In The Wild // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2018.

DEVELOPING WRITTEN TRANSLATION COMPETENCE THROUGH AI-ASSISTED LEARNER PROFILING AND DIFFERENTIATED INSTRUCTION

Akhmedova Anorkhon Nasibali kizi[1], Bokijonova Sayyorakhon Tukhtamurod kizi[2]
[1]PhD in Philological Sciences, Associate Professor
Uzbek State University of World Languages
[2]Teacher Uzbek State University of World Languages
https://doi.org/10.5281/zenodo.23136063

Abstract. *The rapid integration of generative artificial intelligence (AI) into translation education has created new opportunities for providing timely feedback to student translators. However, AI-generated feedback is often provided in a relatively uniform form and may not adequately address differences in learners' translation competence and developmental needs. This conceptual study proposes an AI-assisted differentiated feedback model for developing written translation competence among EFL translation students. The framework integrates three theoretical perspectives: translation competence, Differentiated Instruction, and human-mediated AI-supported learning. It conceptualizes a five-stage pedagogical cycle consisting of translation task, AI-assisted diagnostic analysis, learner needs profiling, differentiated feedback, and revision and reflection. Within this model, AI is used primarily to identify potential recurring patterns in students' translation performance, while teachers interpret AI-generated information and determine appropriate feedback strategies. Differentiation is achieved by adapting the focus, explicitness, complexity, and level of scaffolding according to individual learner needs. The proposed framework shifts the role of AI from automated correction toward diagnostic and pedagogical support. The study contributes a theoretically grounded approach to integrating generative AI and differentiated instruction into written translation education and provides a basis for future empirical investigation of its effectiveness.*

Keywords: *artificial intelligence; differentiated instruction; translation competence; AI-assisted feedback; written translation; translation education; EFL learners; learner-centred learning.*

Introduction.

The rapid development of generative artificial intelligence (Gen AI) is transforming the practices through which translation is produced, evaluated, and taught. Artificial intelligence tools can now generate translations, identify linguistic problems, suggest alternative formulations, provide explanations, and support revision. These developments have created new opportunities for translator education, but they have also challenged conventional approaches to teaching and assessing written translation. Rather than simply asking whether artificial intelligence should be used in translation classrooms, an increasingly important question concerns how AI can be integrated pedagogically in ways that promote the development of independent translation competence.

Written translation competence is a multidimensional construct that involves substantially more than bilingual language proficiency. Research conducted by the PACTE group has demonstrated that translation competence involves interconnected dimensions of knowledge and

skills and develops progressively through translation training. Their research on written translation has examined competence acquisition from both process- and product-oriented perspectives and has emphasized the role of translation-related knowledge, strategic behavior, instrumental competence, and problem-solving. Research into competence levels in written translation has further identified language competence, cultural and thematic competence, instrumental competence, translation service provision competence, and translation problem-solving competence as relevant dimensions for describing developing translators.

The multidimensional nature of translation competence has direct implications for feedback. A single global evaluation such as "good translation" or "needs improvement" provides insufficient information for learners who may experience very different translation problems. One student may struggle primarily with lexical selection and grammatical accuracy, whereas another may demonstrate adequate linguistic competence but have difficulty with terminology, cultural adaptation, textual cohesion, or the justification of translation decisions. Consequently, effective feedback in translation education needs to provide information that is sufficiently specific to the learner's current competence profile.

Traditional teacher feedback remains an essential component of translation education, particularly because teachers can interpret translation solutions in relation to communicative purpose, genre, cultural context, and professional conventions. However, providing detailed individualized feedback on multiple translation tasks is time-consuming and cognitively demanding for instructors. This challenge becomes particularly significant in higher education settings where teachers work with heterogeneous groups of students who differ in linguistic proficiency, translation experience, technological competence, and learning needs.

Generative AI has introduced a potential means of supplementing conventional feedback practices. Large language models can analyse learner-produced texts rapidly, generate explanations, suggest alternative formulations, and provide immediate responses to students. A recent systematic review of AI-assisted translation teaching and learning in higher education found that AI-supported approaches can contribute to translation quality, learner engagement, classroom innovation, personalization, and reflective learning. At the same time, the review identified substantial challenges related to AI literacy, ethical issues, teacher preparation, and the possibility of using AI as a shortcut rather than as a learning resource.

Recent empirical research specifically examining AI-generated feedback in translation education provides an important basis for the present study. Su, Xu, and Liu (2025) compared ChatGPT-generated feedback with teacher feedback across four translation tasks using a mixed-methods crossover experimental design. Their findings indicated that ChatGPT tended to provide direct reference translations and relatively general feedback, with particular attention to lexical and morphosyntactic issues, whereas teacher feedback tended to be more indirect and elaborated. Students also demonstrated greater acceptance of teacher feedback, while their perceptions of ChatGPT feedback were mixed in terms of reliability, comprehensibility, and affective connection.

Similarly, research on students' engagement with AI-driven feedback during translation revision has demonstrated that learners do not simply accept AI-generated feedback automatically. A 2025 mixed-methods study involving Master's students in translation examined their engagement with ChatGPT-generated feedback during revision and found complex relationships between cognitive, affective, and behavioral engagement. More recent research involving undergraduate and postgraduate translation students has also highlighted limitations of

large-language-model feedback, including cultural blind spots, stylistic flattening, and insufficient sensitivity to contextual factors. These findings suggest that the effectiveness of AI feedback depends substantially on how learners interpret, evaluate, and use the feedback rather than merely on the availability of AI-generated comments.

These findings point to an important pedagogical problem. AI-generated feedback is typically produced in response to a particular translation, but learners do not have identical feedback needs. If two students make different types of translation errors, providing them with essentially the same feedback logic may not support their development equally effectively. A learner who repeatedly produces literal translations may require explicit guidance on restructuring and translation strategies, whereas a learner with strong basic translation skills may benefit more from indirect feedback that encourages independent problem-solving. Similarly, an advanced learner may require feedback on stylistic, pragmatic, or genre-related decisions rather than on basic grammatical accuracy.

Differentiated Instruction (DI) provides a theoretical basis for addressing this problem. DI is grounded in the recognition that learners differ in readiness, needs, interests, and learning characteristics and that instructional processes can be adapted accordingly. Rather than providing identical learning experiences to all students, differentiated approaches seek to establish appropriate pathways toward shared learning objectives. In language education, differentiation can involve variation in content, learning processes, products, assessment, and the amount or type of instructional support provided to learners.

The application of differentiated instruction to translation education is particularly relevant because translation classrooms are inherently heterogeneous. Students may possess different levels of source-language comprehension, target-language proficiency, terminology knowledge, cultural awareness, translation strategy use, and revision ability. A uniform feedback model may therefore fail to address the specific competence dimensions that individual learners need to develop.

The combination of AI and differentiated feedback provides a potentially valuable solution. AI can assist in identifying recurring patterns in student translations, while differentiated instruction can provide the pedagogical principle for determining what type of feedback should be offered to different learners. In this model, AI does not simply generate more feedback; rather, it contributes to the profiling and personalization of feedback.

The proposed approach can be conceptualized as a five-stage pedagogical cycle: translation task→AI-assisted diagnostic analysis→learner profile → differentiated feedback → revision and reflection.

At the first stage, students produce a written translation. At the second stage, AI is used to identify potential patterns in the translation, such as lexical, grammatical, terminological, textual, pragmatic, or strategic problems. At the third stage, the teacher interprets these observations and establishes a learner profile. At the fourth stage, feedback is differentiated according to the learner's specific needs. At the final stage, students revise their translations and reflect on the feedback received. Subsequent translation tasks can then provide new evidence for updating the learner profile.

This approach also responds to an important limitation in current discussions of AI in translator education. Recent research has increasingly examined whether AI can produce useful translation feedback, but less attention has been given to the pedagogical question of how AI-generated feedback should be differentiated according to individual learner needs. The 2026

systematic review of empirical studies in higher-education translation education explicitly identifies the need for more focused pedagogical approaches to AI integration rather than treating AI primarily as a technological tool. Thus, the central issue is not simply whether AI can provide feedback, but how AI-supported feedback can be incorporated into a learner-centred instructional design.

The present article addresses this issue by showcasing an AI-assisted differentiated feedback approach for developing written translation competence among EFL translation students. The approach positions AI as a diagnostic and feedback-support tool while retaining teacher mediation as an essential component of pedagogical decision-making. The teacher is responsible for validating AI-generated observations, identifying the learner's priority development areas, and selecting an appropriate level and type of feedback.

The article is guided by the following research questions:

RQ1. What recurring categories of problems can be identified in EFL students' written translations through AI-assisted analysis?

RQ2. How can AI-generated diagnostic information be used to differentiate feedback according to individual translation-learning needs?

RQ3. How can differentiated AI-assisted feedback support the development of written translation competence through revision and reflection?

The article contributes to translation pedagogy by shifting the focus from AI-generated feedback as a uniform technological product toward AI-assisted feedback as a differentiated pedagogical process. By connecting translation competence, learner differences, AI-assisted analysis, and differentiated feedback, the proposed approach seeks to provide a more learner-centred model for integrating generative AI into written translation education.

Literature Review

Written TranslationCompetence

Written translation competence has been widely conceptualized as a complex and multidimensional construct that cannot be reduced to bilingual language proficiency. The PACTE research group conceptualizes translation competence as a system of interrelated sub-competences that includes bilingual, extra-linguistic, translation knowledge, instrumental, and strategic components (PACTE, 2003). From this perspective, competent translation requires translators to mobilize different types of knowledge and skills and to regulate the translation process strategically.

The empirical research conducted by PACTE has provided substantial evidence for the progressive development of translation competence through translation training. Their experimental research has examined how students acquire different components of translation competence and how this development can be observed through both translation products and processes (Hurtado Albir, 2017). In particular, the acquisition of instrumental competence has been shown to constitute an important dimension of translator development, as students progressively learn to use documentary and technological resources more effectively during translation (Kuznik & Olalla-Soler, 2018).

A competence-oriented perspective is also reflected in the European Master's in Translation (EMT) competence framework. The EMT framework identifies several competence areas required of professional translators, including language and culture, translation, technology, personal and interpersonal, and service provision competences (European Commission, 2022). The framework is particularly relevant in the context of artificial

intelligence because contemporary translators increasingly work in technologically mediated environments and are expected to combine linguistic expertise with technological and strategic skills.

Hurtado Albir et al. (2018) further emphasize the importance of describing translation competence in terms of progressive levels of development. Their competence framework identifies areas such as language competence, cultural and thematic competence, instrumental competence, translation service provision, and translation problem-solving competence. Such a multidimensional understanding has important implications for translation teaching because students at the same educational level may demonstrate different strengths and weaknesses across these dimensions.

For written translation education, this means that assessment and feedback should move beyond identifying whether a translation is simply "correct" or "incorrect." A student may produce linguistically accurate sentences while experiencing difficulties with terminology, genre conventions, pragmatic adaptation, cohesion, or strategic decision-making. Conversely, another student may have strong translation strategies but make recurrent grammatical or lexical errors. Consequently, the development of written translation competence requires feedback that identifies specific areas for improvement and provides learners with opportunities to address their individual weaknesses.

Feedback in Translation Education

Feedback plays a central role in developing translation competence because translationlearning depends heavily on students' ability to recognize problems, evaluate possible solutions, and revise their own work. Unlike many language-learning tasksin which a single correct answer may be expected, translation frequently involves several acceptable solutions. Therefore, effective feedback should not only identify errors but also help students understand why a particular translation choice may be inappropriate in a given communicative, textual, or cultural context.

Traditional teacher feedback has several advantages in this respect. Teachers can interpret translation decisions in relation to the source text, target readership, genre, communicative purpose, and translation strategy. They can also adjust explanations according to students' levels of knowledge and previous performance. At the same time, individualized feedback on written translations is demanding for teachers because it requires considerable time and cognitive effort, particularly when students produce lengthy texts.

The issue becomes more significant in heterogeneous university classrooms. Students entering translation programmes may differ considerably in English proficiency, target-language competence, previous translation experience, terminology knowledge, digital literacy, and familiarity with translation strategies. A uniform feedback approach may therefore provide the same amount and type of information to learners whose developmental needs are substantially different.

Research on translation competence acquisition supports the importance of learners' active engagement with translation problems rather than merely receiving corrected versions of their texts. PACTE's research emphasizes the interaction between declarative and procedural knowledge and the development of strategic competence through translation practice (Beeby et al., 2015). This suggests that feedback should ideally encourage learners to analyse translation problems and make informed decisions rather than simply replace their solutions with teacher-provided answers.

Feedback can therefore be understood not merely as an assessment mechanism but as part of the learning process. In a competence-oriented translation classroom, feedback should help students identify recurring problems, understand the underlying causes, evaluate alternatives, revise their translations, and gradually develop greater independence.

Artificial Intelligence-Assisted Feedback in Translation Education

The rapid development of generative artificial lintelligence has created new possibilities for providing feedback on student translations. Large language models can analyse learner-produced texts, identify potential linguistic problems, generate alternative formulations, and provide explanations almost immediately. These capabilities make AI potentially useful as a supplementary resource for translation teaching, particularly when teachers need to provide feedback to relatively large groups of students.

Recent research has increasingly examined the effectiveness and limitations of AI-generated translation feedback. Su et al. (2025), for example, compared ChatGPT-generated feedback with teacher feedback across four translation tasks using a mixed-methods crossover experimental design. Their findings showed that ChatGPT frequently relied on direct reference translations and relatively general forms of feedback, while teacher feedback tended to be more indirect and elaborated. The study also found that students generally demonstrated greater acceptance of teacher feedback, whereas their perceptions of ChatGPT feedback were mixed with respect to reliability, comprehensibility, and affective connection (Su et al., 2025).

These findings indicate that AI-generated feedback should not automatically be considered equivalent to expert teacher feedback. The ability of a large language model to generate fluent explanations does not guarantee that its evaluation of a translation is appropriate. Translation quality is strongly dependent on contextual, cultural, textual, and pragmatic considerations that may not always be adequately recognized by AI systems.

This issue is further demonstrated by research into students' engagement with AI-driven feedback during translation revision. A mixed-methods study of Master's students in translation found that learners' engagement with ChatGPT-generated feedback involved cognitive, affective, and behavioural dimensions rather than simple acceptance or rejection of AI suggestions (Investigating student engagement with AI-driven feedback in translation revision, 2025). This suggests that students need to evaluate AI feedback critically and learn how to distinguish useful suggestions from inappropriate or contextually weak recommendations.

More recent research has also highlighted the influence of text type, proficiency, and learner attitudes on students' responses to AI feedback. Research involving undergraduate and postgraduate translation students found that AI-generated feedback could be affected by limitations such as cultural blind spots, stylistic flattening, and insufficient sensitivity to contextual factors ("Navigating AI feedback in translation training," 2026). Such findings reinforce the importance of maintaining human oversight when AI is incorporated into translator education.

Systematic reviews provide a broader picture of the potential and limitations of AI in translation education. Kan et al. (2026), reviewing empirical research on AI-assisted translation teaching and learning in higher education, identified benefits associated with translation quality, learner engagement, classroom innovation, personalization, and reflective learning. At the same time, the authors identified challenges involving AI literacy, ethics, teacher preparation, and the risk that students may use AI as a shortcut rather than as a learning resource.

Similarly, Nguyen et al. (2025) found that AI translation tools can support efficiency, vocabulary and grammar development, post-editing skills, motivation, professional preparation, accessibility, reflective learning, and teacher support. However, their systematic review also identified risks including overreliance on AI, contextual inaccuracies, insufficient training, pedagogical gaps, ethical concerns, and unequal access to digital technologies.

Taken together, this body of research suggests that the pedagogical value of AI feedback depends not simply on whether AI can generate comments, but on how those comments are incorporated into the learning process. A human-in-the-loop approach is therefore particularly relevant. The teacher can validate AI-generated observations, reject inappropriate suggestions, identify pedagogically significant patterns, and guide students in interpreting the feedback.

Differentiated Instruction and Individualized Feedback

Differentiated Instruction (DI)provides a useful pedagogical framework for addressing differences between learners. Tomlinson (2014) defines differentiated instruction as an approach inwhich teachers proactively adapt aspects of teaching in response to learner differences. Differentiation may involve content, process, product, and the learning environment, while maintaining common educational objectives.

The relevance of differentiation to higher education has received increasing attention. Tajik et al. (2024), for instance, investigated differentiated instruction in an EFL university context and examined how teachers could adapt content, process, product, learning environment, and assessment to address students' different needs. Their findings demonstrate the practical relevance of differentiated approaches in heterogeneous language-learning settings.

Research in higher education also indicates that implementing differentiated instruction can be challenging. Teachers may experience difficulties related to time, workload, classroom organization, institutional expectations, and the need to design multiple learning pathways (Educator Perceptions of Differentiated Instruction in the Higher Education Sector, 2025). These challenges are particularly relevant to translation education, where teachers already need to evaluate complex learner-produced texts (Tohirov Z. et. al., 2025).

Recent evidence nevertheless suggests that differentiated instruction can have meaningful positive effects on diverse language learners. A systematic review and meta-analysis by Sedanza and Abatayo (2026), based on 45 controlled studies and 56 effect sizes involving 16,773 learners across 18 countries, reported a substantial overall effect of differentiated instruction on English learners' outcomes. The authors also emphasized considerable contextual variation, indicating that the effectiveness of differentiation depends on factors such as educational level, learning context, and learner population.

In translation education, differentiated feedback can be understood as an application of this broader principle to the feedback stage of learning. Rather than giving all students identical comments, the teacher can prioritize feedback according to each learner's most important developmental needs. For example, a student with recurrent lexical and grammatical problems may receive more explicit corrective feedback and targeted language support, while a student with stronger linguistic competence but weak strategic decision-making may receive prompts that encourage analysis and justification of translation choices.

This distinction is particularly important because effective differentiation does not necessarily mean giving weaker students more feedback and stronger students less feedback. Instead, the type, focus, explicitness, and level of support can be adjusted according to the

learner's current competence. Such an approach is compatible with the developmental and multidimensional nature of translation competence.

Toward AI-Assisted Differentiated Feedback

The literature reviewed above reveals an important intersection between three research areas: translation competence, AI-assisted feedback, and differentiated instruction. Translation competence research demonstrates that learners develop multiple interconnected competences and may experience different developmental trajectories (PACTE, 2003; Hurtado Albir et al., 2018). Research on AI feedback shows that generative AI can provide rapid and potentially useful feedback but also has limitations related to contextual understanding, reliability, and learner acceptance (Su etal., 2025; Kan et al., 2026). Research on differentiated instruction, meanwhile, emphasizes the need to adapt instructional support to learner differences (Tomlinson, 2014; Tajik et al., 2024).

However, these strands of research have not yet been sufficiently integrated into a pedagogical model specifically designed for written translation competence. Existing studies of AI in translation education have predominantly focused on questions such as whether AI improves translation performance, how students perceive AI tools, how learners engage with AI-generated feedback, and what risks accompany AI-supported translation learning. Although personalization has been identified as a potential benefit of AI-assisted education, comparatively less attention has been devoted to the systematic differentiation of feedback according to learners' specific translation competence profiles.

This gap is particularly important because AI systems can potentially analyse large amounts of learner-produced text and identify recurring patterns across multiple translation tasks. Such information could support teachers in determining which problems are isolated errors and which represent persistent areas of developmental need. However, AI identification alone does not constitute differentiated instruction. Pedagogical differentiation requires the teacher to interpret the diagnostic information and determine the appropriate feedback focus, degree of explicitness, and level of learner support.

Accordingly, the present study conceptualizes AI-assisted differentiated feedback as a human-mediated pedagogical cycle rather than as automated correction. The cycle consists of five interconnected stages:

translation task → AI-assisted diagnostic analysis → learner needs identification → differentiated feedback → revision and reflection.

At the first stage, students complete a written translation task. At the second stage, an AI tool assists in identifying potential problems and recurring patterns. At the third stage, the teacher interprets these observations and determines the learner's priority developmental needs. At the fourth stage, feedback is differentiated according to those needs. Finally, students revise their translations and reflect on the feedback received. Evidence from subsequent tasks can then be used to refine the learner's developmental profile.

The proposed approach thus combines the efficiency of AI-assisted analysis with the pedagogical judgement of the teacher and the learner-centred principles of differentiated instruction. Its primary purpose is not to automate translation assessment but to make feedback more targeted, developmental, and responsive to individual learner needs.

This conceptual integration provides the basis for examining how AI can be used not simply to produce feedback faster, but to support teachers in making feedback more differentiated and pedagogically meaningful for EFL translation students.

Conclusion

This conceptual study proposed an AI-assisted differentiated feedback model for developing written translation competence among EFL translation students. The model integrates three perspectives: the multidimensional nature of translation competence, the principles of Differentiated Instruction, and the emerging role of generative AI in translation education.

The proposed five-stage cycle – translation task → AI-assisted diagnostic analysis → learner needs profiling → differentiated feedback → revision and reflection—positions AI as a diagnostic and feedback-support tool rather than an autonomous evaluator. AI can assist in identifying recurring translation problems, while teachers remain responsible for interpreting these problems and determining the appropriate type and level of feedback.

The main contribution of the framework is the conceptualization of feedback itself as a dimension of differentiation. Instead of providing uniform AI-generated corrections, feedback can be adapted to learners' specific developmental needs, including linguistic accuracy, terminology, translation strategies, textual cohesion, and pragmatic or cultural competence.

Overall, the framework suggests that the most productive role of AI in translation education is not to replace teacher feedback but to support more targeted, flexible, and learner-centred feedback practices.

REFERENCES

1. Abdurakhmonova, N. (2016). The bases of automatic morphological analysis for machine translation. *Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta*, *2*(38), 12-7.
2. Abduraxmonova, N. Z. (2018). Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) il dis. aftoref. *2018). Abdurakhmonova N. The bases of automatic morphological analysis for machine translation. Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta. 2016; 2 (38): 12-014.-264 c*.
3. Abdurakhmonova, N. Z., Ismailov, A. S., & Mengliev, D. (2022, November). Developing NLP tool for linguistic analysis of turkic languages. In *2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)* (pp. 1790-1793). IEEE.
4. Abdurakhmonova, N., Shamieva, N., & Adali, E. (2024, October). Exploring the Semantic Complexity of Adjective-Noun Collocations between Uzbek and English for Improved Machine Translation. In *2024 9th International Conference on Computer Science and Engineering (UBMK)* (pp. 1-4). IEEE.
5. Axmedova A.N. Talabalarning yozma tarjima kompetensiyasini korpus vositasi yordamida takomillashtirish bo‘yicha mashq va topshiriqlar tizimi // Filologiya va pedagogika, 2025. e-ISSN 3060-4885 №12. (13.00.00. № 16) B. 95–97. https://tilvaadabiyot.uz/f/12-son_filologiya_va_pedagogika2025.pdf
6. Axmedova A.N., Tuxtasinov I.M. Yozma tarjimani o‘qitishda zamonaviy metod va texnologiyalar integratsiyasi // Filologiya masalalari, 2025 ISSN: 1994 – 4233 Vol. 55, №4. B. 3-21 https://philmatt.uzswlu.uz/article/download/14
7. Beeby, A., Castillo, L., Fox, O., Galán-Mañas, A., Hurtado Albir, A., Kuznik, A., Massana, G., Neunzig, W., Olalla-Soler, C., Rodríguez-Inés, P., & Romero, L. (2015). Results of PACTE’s experimental research on the acquisition of translation competence: The

acquisition of declarative and procedural knowledge in translation. The Dynamic Translation Index. *Translation Spaces, 4*(1), 29–53. https://doi.org/10.1075/ts.4.1.02bee

8. Hurtado Albir, A. (Ed.). (2017). *Researching translation competence by PACTE Group*. John Benjamins Publishing Company. https://doi.org/10.1075/btl.127
9. Hurtado Albir, A., Galán-Mañas, A., Kuznik, A., Olalla-Soler, C., Rodríguez-Inés, P., & Romero, L. (2018). Competence levels in translation: Working towards a European framework. *The Interpreter and Translator Trainer, 12*(2), 111–131. https://doi.org/10.1080/1750399X.2018.1466093
10. Kan, L., Kamal, S. S. L. A., Osman, S. Z. b. M., Yan, Y., & Liu, J. (2026). AI-assisted translation teaching and learning in higher education: A systematic review of empirical studies. *Arab World English Journal, 17*(1), 139–160. https://doi.org/10.24093/awej/vol17no1.9
11. Kuznik, A., & Olalla-Soler, C. (2018). Results of PACTE group's experimental research on translation competence acquisition: The acquisition of the instrumental sub-competence. *Across Languages and Cultures, 19*(1), 19–51. https://doi.org/10.1556/084.2018.19.1.2
12. Mengliev, D. B., Abdurakhmonova, N. Z., Barakhnin, V. B., Iskandarova, A. R., Topildiyeva, F. R., & Akhmedov, E. Y. (2024, November). Development of an algorithm for automatic analysis of sentiment in school essays of the Uzbek language. In *2024 IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE)* (pp. 1570-1573). IEEE.
13. Nguyen, T. N. N., Tran, T. T., Nguyen, N. H. A., Lam, H. P., Nguyen, H. M. S., & Tran, N. A. T. (2025). The benefits and challenges of AI translation tools in translation education at the tertiary level: A systematic review. *International Journal of TESOL & Education, 5*(2), 132–148. https://doi.org/10.54855/ijte.25527
14. PACTE Group. (2003). Building a translation competence model. In F. Alves (Ed.), *Triangulating translation: Perspectives in process oriented research* (pp. 43–66). John Benjamins Publishing Company. https://doi.org/10.1075/btl.45.06pac
15. Sedanza, M. A. C., & Abatayo, J. A. (2026). Beyond the average: A systematic review and meta-analysis of differentiated instruction for diverse English learners. *Frontiers in Education, 11*, 1906974. https://doi.org/10.3389/feduc.2026.1906974
16. Subban, P., Suprayogi, M. N., Preston, M., Liyani, A. N., & Ratri, A. P. P. (2025). "Differentiation is sometimes a hit and miss": Educator perceptions of differentiated instruction in the higher education sector. *The Asia-Pacific Education Researcher, 34*, 873–884. https://doi.org/10.1007/s40299-024-00904-8
17. Su, Y., Xu, S., & Liu, K. (2025). Adapt or adopt? Examining the efficacy of ChatGPT in providing translation feedback. *The Interpreter and Translator Trainer, 19*(3–4), 296–316. https://doi.org/10.1080/1750399X.2025.2541486
18. Tajik, O., Noor, S., & Golzar, J. (2024). Investigating differentiated instruction and the contributing factors to cater EFL students' needs at the collegial level. *Asian-Pacific Journal of Second and Foreign Language Education, 9*, 74. https://doi.org/10.1186/s40862-024-00299-5
19. Tohirov Z., Tukhtasinov I., Akhmedova A., Babajanova L., Egamberganova A., Yakubov M. Development and Evaluation of a Combined Model for Genre Classification of Uzbek Media Texts. Scopus. // 2025 IEEE XVII International Scientific and Technical Conference

on Actual Problems of Electronic Instrument Engineering (APEIE). 2025. https://ieeexplore.ieee.org/document/11289361

20. Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.
21. Xu, S., Su, Y., & Liu, K. (2025). Investigating student engagement with AI-driven feedback in translation revision: A mixed-methods study. *Education and Information Technologies, 30*(12), 16969–16995. https://doi.org/10.1007/s10639-025-13457-0
22. Yang, C., Hou, S., Chen, J., & Wider, W. (2026). Prioritizing translator competencies in the era of artificial intelligence: A Delphi method study. *Discover Education, 5*, 605. https://doi.org/10.1007/s44217-026-01642-3
23. Chen, S., & Lou, J. (2026). Navigating AI feedback in translation training: How text type, proficiency, and attitude shape students' acceptance behaviors. *Frontiers in Artificial Intelligence, 9*. https://doi.org/10.3389/frai.2026.1727544

O‘ZBEK XALQ DOSTONLARINING PARALLEL KORPUSINI YARATISH VA UNING LINGVISTIK AHAMIYATI

Nafisa Amanlikova
TDTU o‘qituvchisi
https://doi.org/10.5281/zenodo.23136085

Annotatsiya: *Mazkur maqolada o‘zbek xalq dostonlari asosida parallel korpus yaratishning lingvistik va tarjimashunoslikdagi ahamiyati yoritiladi. Dostonlarning original va tarjima variantlarini yagona elektron resursga birlashtirish, ularni qiyosiy o‘rganish hamda tarjima ekvivalentlarini aniqlash imkoniyatlari ko‘rib chiqiladi. Shuningdek, bunday korpusning o‘zbek tili tadqiqotlari, folklor va mashina tarjimasi uchun ahamiyati asoslanadi.*

Kalit so‘zlar: *o‘zbek xalq dostonlari, parallel korpus, korpus lingvistikasi, tarjimashunoslik, mashina tarjimasi, "Alpomish", tarjima ekvivalenti.*

Abstract: *This article examines the importance of creating a parallel corpus of Uzbek folk epics for linguistics and translation studies. It explores the integration of original and translated texts into a unified electronic resource, their comparative analysis, and the identification of translation equivalents. The study also highlights the potential of such a corpus for Uzbek language research, folklore studies, and machine translation.*

Keywords: *Uzbek folk epics, parallel corpus, corpus linguistics, translation studies, machine translation, "Alpamysh", translation equivalent.*

Аннотация. *В статье рассматривается значение создания параллельного корпуса узбекских народных дастанов для лингвистики и переводоведения. Анализируются возможности объединения оригинальных и переводных текстов в единый электронный ресурс, их сопоставительного изучения и выявления переводческих эквивалентов. Обосновывается значение такого корпуса для исследования узбекского языка, фольклора и развития машинного перевода.*

Ключевые слова: *узбекские народные дастаны, параллельный корпус, корпусная лингвистика, переводоведение, машинный перевод, «Алпамыш», переводческий эквивалент.*

Korpus lingvistikasining muhim yo‘nalishlaridan biri parallel korpuslarni yaratish va ulardan foydalanish hisoblanadi. Parallel korpuslar bir tildagi original matn va uning boshqa tildagi tarjimasini o‘zaro bog‘langan holda saqlaydi. Bunday korpuslar tarjima jarayonini qiyosiy o‘rganish, tarjima ekvivalentlarini aniqlash, ikki tilli lug‘atlar yaratish va mashina tarjimasi tizimlarini rivojlantirishda muhim vosita hisoblanadi. Zanettin parallel korpuslarning tarjimashunoslikdagi imkoniyatlarini alohida ta’kidlab, bunday resurslar tarjima matnlarini katta hajmda va tizimli ravishda tahlil qilish imkonini berishini ko‘rsatadi [1].

O‘zbek tili uchun ham elektron korpuslar va parallel matnlar bazasini shakllantirish masalasi so‘nggi yillarda dolzarb yo‘nalishlardan biriga aylandi. Xususan, O‘zbek milliy korpusi tarkibida parallel korpus yo‘nalishining shakllantirilayotgani o‘zbek tilidagi qiyosiy va tarjimashunoslik tadqiqotlari uchun muhim imkoniyat yaratmoqda [2].

Parallel korpus — ikki yoki undan ortiq tillardagi bir-biriga mazmunan mos matnlar majmuasidir. Odatda parallel korpusda bir tildagi original matn va uning tarjimasi birgalikda

joylashtiriladi. Masalan, o‘zbek tilidagi "Alpomish" dostoni va uning ingliz tilidagi tarjimasi bitta parallel korpus tarkibiga kiritilishi mumkin.

Parallel korpuslar dastlab tarjimashunoslik va qiyosiy tilshunoslikda keng qo‘llanilgan bo‘lsa, bugungi kunda ulardan kompyuter lingvistikasi va tabiiy tilni qayta ishlash texnologiyalarida ham foydalanilmoqda. Xususan, mashina tarjimasi tizimlarini yaratishda parallel ma'lumotlar asosiy o‘quv resurslaridan biri hisoblanadi.

Parallel korpuslardan quyidagi maqsadlarda foydalanish mumkin:

- tarjima ekvivalentlarini aniqlash;
- ikki tilning grammatik va leksik xususiyatlarini qiyoslash;
- tarjimon tomonidan qo‘llangan usullarni o‘rganish;
- milliy-madaniy birliklarning tarjimasini tahlil qilish;
- ikki tilli lug‘atlar yaratish;
- til o‘rgatish uchun material tayyorlash;
- mashina tarjimasi modellarini yaratish va takomillashtirish.

O‘zbek Milliy korpusi doirasida parallel korpusni shakllantirish bo‘yicha ishlar olib borilmoqda. Mazkur resurs orqali o‘zbek tilidagi matnlarning boshqa tillardagi variantlarini qiyoslash imkoniyati yaratilmoqda [2].

O‘zbek-rus parallel korpuslarini yaratishga bag‘ishlangan tadqiqotlarda badiiy adabiyot namunalaridan foydalanish ham muhim yo‘nalish sifatida ko‘rsatilgan. Jumladan, A. Qodiriyning "O‘tkan kunlar" romani va uning rus tilidagi tarjimasi asosida parallel korpus yaratish bo‘yicha tadqiqot amalga oshirilgan. Ushbu ishda original va tarjima matnlarini elektron shaklda birlashtirish, ularni qidirish va qiyoslash imkoniyatlari ko‘rib chiqilgan [3].

Bundan tashqari, o‘zbek va qozoq tillari o‘rtasidagi parallel ma'lumotlar bazasini yaratishga bag‘ishlangan tadqiqotlar ham mavjud. O‘zbek-qozoq parallel datasetini shakllantirishga bag‘ishlangan tadqiqotda parallel ma'lumotlardan mashina tarjimasi tizimlarini yaratish va sinovdan o‘tkazishda foydalanish mumkinligi ko‘rsatilgan [4].

"Alpomish" dostoni tarjimalarini o‘rgangan tadqiqotlarda ham milliy-madaniy va maishiy leksikaning boshqa tillarda berilishi masalalariga alohida e'tibor qaratilgan [5]. Shuningdek, dostonning ingliz va nemis tillaridagi tarjimalarida epik klishelarning ifodalanishi ham tadqiq qilingan [6].

Demak, dostonlarning parallel korpusini yaratish orqali bunday birliklarning originaldagi shakli, konteksti va tarjimadagi variantlarini birgalikda o‘rganish mumkin. Masalan, "Alpomish"ning ingliz tilidagi tarjimalaridan biri Qosimboy Ma'murov tarjimasi asosida nashr qilingan bo‘lib, u Toshkent davlat sharqshunoslik universitetining raqamli kutubxonasida ham mavjud [7].

Dostonning o‘zbekcha original matni va inglizcha tarjimasini parallel shaklda joylashtirish orqali tadqiqotchi quyidagi masalalarni o‘rganishi mumkin:

Tadqiqot obyekti	**O‘rganish imkoniyati**
So‘zlar	Tarjima ekvivalentlarini aniqlash
Frazeologizmlar	Iboralarning tarjimada berilishini o‘rganish
Milliy realiyalar	Madaniy birliklarning tarjimasini tahlil qilish
Antroponimlar	Ismlarning tarjima va transliteratsiyasini o‘rganish
Toponimlar	Joy nomlarining berilishini qiyoslash
Epik klishelar	Takroriy poetik birliklarning tarjimasini o‘rganish
Badiiy vositalar	Tashbeh, mubolag‘a va metaforalarning tarjimasini qiyoslash

Masalan, o‘zbekcha matnda uchragan muayyan birlikni korpusdan izlash orqali uning inglizcha tarjimada qanday berilgani aniqlanadi. Agar bir birlik bir necha marta turlicha tarjima qilingan bo‘lsa, tarjimonning tanlovlari ham qiyosiy jihatdan o‘rganilishi mumkin.

O‘zbek xalq dostonlari parallel korpusini yaratishda bir necha bosqichga amal qilish lozim. *Birinchi bosqichda* dostonlarning elektron yoki raqamlashtirilishi mumkin bo‘lgan original matnlari yig‘iladi. Keyingi bosqichda ularning mavjud xorijiy tillardagi tarjimalari aniqlanadi. *Ikkinchi bosqichda* matnlar elektron shaklga keltiriladi va texnik jihatdan qayta ishlanadi. Agar matn skaner ko‘rinishida bo‘lsa, OCR texnologiyasi yordamida matnga aylantiriladi va xatolar tekshiriladi. *Uchinchi bosqichda* original va tarjima matnlari gap yoki boshqa birliklar bo‘yicha o‘zaro moslashtiriladi. Ushbu jarayon “alignment” deb ataladi. *To‘rtinchi bosqichda* korpusga qo‘shimcha lingvistik ma’lumotlar kiritilishi mumkin. Masalan, so‘z turkumi, lemma, frazeologik birlik, milliy-madaniy birlik yoki atoqli ot kabi belgilar bilan annotatsiya qilish mumkin. *Yakuniy bosqichda* foydalanuvchi uchun qidiruv tizimi yaratiladi. Bunda tadqiqotchi muayyan so‘zni kiritib, uning original matndagi konteksti va tarjimadagi variantlarini ko‘rishi mumkin.Xulosa qilib aytganda, o‘zbek xalq dostonlarining parallel korpusini yaratish zamonaviy korpus lingvistikasi, tarjimashunoslik va kompyuter lingvistikasi uchun istiqbolli yo‘nalishlardan biridir. Dostonlar xalqning nafaqat badiiy merosi, balki uning tili, madaniyati, urf-odatlari va dunyoqarashini aks ettiruvchi muhim manba hisoblanadi.

O‘zbek tilida parallel korpus yaratish bo‘yicha olib borilayotgan tadqiqotlar, jumladan, “O‘tkan kunlar” romani asosida o‘zbek-rus parallel korpusi va o‘zbek-qozoq parallel datasetini yaratishga qaratilgan ishlar ushbu sohada muayyan ilmiy va amaliy tajriba shakllanganini ko‘rsatmoqda [3; 4].

Shu asosda o‘zbek xalq dostonlari, xususan, “Alpomish” dostoni original va tarjima variantlarini o‘z ichiga oluvchi maxsus parallel korpus yaratish mumkin. Bunday resurs yordamida tarjima ekvivalentlari, milliy-madaniy birliklar, frazeologizmlar, epik klishelar va badiiy tasvir vositalarining tarjimada berilishi tizimli ravishda tadqiq etiladi.

Bundan tashqari, yaratiladigan korpus o‘zbek tilining elektron lingvistik resurslarini boyitadi hamda kelajakda mashina tarjimasi, elektron lug‘atlar, tarjima dasturlari va turli NLP tizimlarini yaratish uchun ma’lumotlar bazasi bo‘lib xizmat qilishi mumkin. Shunday qilib, o‘zbek xalq dostonlarining parallel korpusini yaratish milliy adabiy merosni raqamlashtirish bilan birga, uni zamonaviy ilmiy tadqiqotlar va sun’iy intellekt texnologiyalari bilan bog‘lash imkonini beradi.

FOYDALANILGAN ADABIYOTLAR

1. Zanettin, F. (2015). Parallel corpora in translation studies: Issues in corpus design and analysis. In *The Routledge Handbook of Translation Studies and Linguistics*. Routledge.
2. Uzbek National Corpus. *Parallel corpus*. O‘zbek Milliy korpusi. https://uzbekcorpus.uz/enParallel
3. Nigmatova, L. K., Saidova, M. R., Djuraeva, Z. R., Sharipov, S. S., & Avezov, S. S. (2024). An online platform for Uzbek-Russian and Russian-Uzbek parallel corpora: Linguistic challenges and prospects exemplified by A. Kadyri's novel “Bygone Days”. *2024 9th International Conference on Computer Science and Engineering (UBMK)*. https://doi.org/10.1109/UBMK63289.2024.10773409
4. Akhmedova, S., et al. (2024). Parallel texts dataset for Uzbek-Kazakh machine translation. *Data in Brief, 53*, 110194. https://doi.org/10.1016/j.dib.2024.110194

5. Shomuratova, I. B. (2026). Household lexicon in English translations of the Alpomish epic. *Komparativistika, 11*(11).
6. Asqarova, Sh. (2024). Epik klishelar hamda ularning ingliz va nemis tillaridagi talqini ("Alpomish" dostoni misolida). *Ijtimoiy-gumanitar fanlarning dolzarb muammolari, 4*(8).
7. Ma'murov, Q. (Trans.). (2023). *Alpomish: Uzbek heroic epos*. Tashkent State University of Oriental Studies Digital Library.

SINONIMLAR O‘QUV LUG‘ATLARIDA IZOHLASH MUAMMOSI VA MULTIMEDIALI YONDASHUV

Sabina Amonova
O‘zMU 2-bosqich magistranti
https://doi.org/10.5281/zenodo.23136106

Annotatsiya. *Maqolada sinonimlarni o‘quv lug‘atlarida izohlash muammosi multimediali yondashuv nuqtayi nazaridan tahlil qilinadi. Asosiy e’tibor xorijiy elektron lug‘atlar tajribasini o‘rganish, ularni didaktik va texnologik belgilariga ko‘ra tasniflash hamda bu tajribani o‘zbek tili o‘quv lug‘atchiligiga tatbiq etish imkoniyatlariga qaratildi. Tadqiqotda Oxford Learner’s Dictionaries, Cambridge Dictionary, Longman va Merriam-Webster tipidagi platformalarda sinonimlar faqat ma’nodosh so‘zlar ro‘yxati sifatida emas, balki talaffuz, misol, kollokatsiya, uslubiy belgi, kontekst va vizual birliklar orqali tushuntirilishi aniqlanadi. Natijada o‘zbek tili uchun sinonim izohining ko‘p qatlamli modeli: semantik farq, uslubiy bo‘yoq, qo‘llanish doirasi, audio talaffuz, kontekstli misol, interaktiv mashq va korpusga tayangan qidiruv kabi komponentlar asosida taklif etiladi.*

Kalit so‘zlar: *sinonim, o‘quv lug‘ati, elektron leksikografiya, multimediali yondashuv, xorijiy tajriba, izoh, semantik farq, audio, vizualizatsiya, interaktivlik.*

Abstract. *This article analyzes the problem of explaining synonyms in learner’s dictionaries from the perspective of a multimedia approach. Particular attention is paid to the study of the experience of foreign electronic dictionaries, their classification according to didactic and technological features, and the possibilities of applying this experience to Uzbek learner lexicography. The study demonstrates that on platforms such as Oxford Learner’s Dictionaries, Cambridge Dictionary, Longman, and Merriam-Webster, synonyms are presented not merely as lists of semantically related words, but are explained through pronunciation, examples, collocations, stylistic labels, contextual information, and visual elements. As a result, a multilayered model for explaining synonyms in Uzbek is proposed, incorporating such components as semantic distinction, stylistic connotation, domain of use, audio pronunciation, contextual examples, interactive exercises, and corpus-based search.*

Keywords: *synonym, learner’s dictionary, electronic lexicography, multimedia approach, foreign experience, explanation, semantic distinction, audio, visualization, interactivity.*

Аннотация. *В статье с точки зрения мультимедийного подхода анализируется проблема толкования синонимов в учебных словарях. Основное внимание уделяется изучению опыта зарубежных электронных словарей, их классификации по дидактическим и технологическим признакам, а также возможностям применения данного опыта в узбекской учебной лексикографии. В ходе исследования установлено, что на таких платформах, как Oxford Learner’s Dictionaries, Cambridge Dictionary, Longman и Merriam-Webster, синонимы представлены не только в виде списка слов, близких по значению, но и объясняются посредством произношения, примеров, коллокаций, стилистических помет, контекстуальной информации и визуальных элементов. В результате предлагается многоуровневая модель толкования синонимов для узбекского языка, включающая такие компоненты, как семантическое различие, стилистическая окраска, сфера употребления,*

аудиопроизношение, контекстуальные примеры, интерактивные упражнения и корпусный поиск.

Ключевые слова: *синоним, учебный словарь, электронная лексикография, мультимедийный подход, зарубежный опы.*

Kirish. Til ta'limida o'quv lug'ati oddiy ma'lumotnoma emas, balki o'quvchining so'z tanlash, ma'no farqlash, og'zaki va yozma nutq qurish ko'nikmasini shakllantiradigan didaktik vositadir. Ayniqsa sinonimlar lug'ati ta'lim jarayonida muhim o'rin tutadi, chunki bir-biriga yaqin ma'noli birliklar nutqda aynan bir xil vazifani bajarmaydi. Masalan, *chiroyli, go'zal, ko'rkam, dilbar* so'zlari umumiy "estetik yoqimlilik" ma'nosiga ega bo'lsa-da, ularning uslubiy bo'yoqi, qo'llanish muhiti, emotsional darajasi va birikish imkoniyatlari farqlanadi. An'anaviy qog'oz lug'atlarda bunday nozik farqlar ko'pincha qisqa izoh yoki ro'yxat orqali beriladi. O'quvchi esa qaysi sinonimni qaysi vaziyatda ishlatish kerakligini mustaqil aniqlashga majbur bo'ladi.

Zamonaviy elektron va multimediali lug'atlar bu muammoni birmuncha yengillashtiradi. Chet el tajribasida sinonimlar audio talaffuz, real misollar, kollokatsiyalar, rasm, animatsion ko'rsatkich, rangli belgilash, interaktiv test va shaxsiy lug'at papkasi bilan birgalikda taqdim etiladi. Oxford Learner's Dictionaries platformasi ta'rif, rasm, misol gap va sinonimlarni bir sahifada uyg'unlashtiradi; Cambridge Dictionary esa o'rganuvchilar uchun ma'no, talaffuz, tarjima va thesaurus imkoniyatlarini birlashtirgan elektron muhit sifatida ishlaydi. Bu holat sinonim izohining faqat matniy emas, balki ko'p kanalli idrok orqali berilishi lozimligini ko'rsatadi.

Maqolaning maqsadi – sinonimlarni o'quv lug'atlarida izohlashda uchraydigan muammolarni aniqlash, xorijiy multimediali lug'atlar tajribasini tasniflash va o'zbek tili o'quv lug'atchiligi uchun amaliy model taklif qilishdan iborat. Tadqiqotning nazariy asosini elektron leksikografiya, o'quv leksikografiyasi, multimediali ta'lim va kompyuter lingvistikasi bo'yicha ilmiy qarashlar tashkil etadi.

Sinonimlarni izohlashdagi asosiy muammolar. Sinonimlar o'quv lug'atida izohlanayotganda uchta asosiy muammo yuzaga chiqadi. *Birinchisi* – semantik yaqinlik va semantik farqni ajratish muammosi. O'quvchi ko'pincha sinonimlarni "bir xil ma'noli so'zlar" deb qabul qiladi. Aslida sinonimlar to'liq almashina olmaydi; ular ma'no qirrasi, baholash darajasi, emotsional bo'yoq, uslubiy mansublik va kontekstga moslashish jihatidan farqlanadi. Shuning uchun o'quv lug'atida sinonim qator berilganda har bir birlikning markaziy semasi va farqlovchi belgisi aniq ko'rsatilishi zarur.

Ikkinchi muammo – uslubiy va pragmatik izohning yetishmasligi. Masalan, yuz, bet, chehra, jamol kabi birliklar ma'nodosh bo'lsa-da, neytral, so'zlashuv, badiiy yoki ko'tarinki uslubga mansubligi bilan farqlanadi. Agar lug'atda faqat sinonimlar ro'yxati berilsa, o'quvchi bet so'zini rasmiy matnda, jamol so'zini oddiy maishiy suhbatda noo'rin qo'llashi mumkin. Shuning uchun sinonim izohi "ma'nosi yaqin" degan umumiy belgi bilan cheklanmasligi, balki nutq vaziyati bilan bog'lanishi lozim.

Uchinchi muammo – o'quvchining qabul qilish usuli bilan bog'liq. Boshlang'ich yoki o'rta darajadagi o'quvchiga uzun nazariy ta'rifdan ko'ra jadval, rang, belgi, real gap, tovushli namuna va qisqa mashq ko'proq yordam beradi. Mayerning multimediali ta'lim nazariyasiga ko'ra, o'quvchi ma'lumotni so'z va tasvir kanallari orqali uyg'un qabul qilganda o'zlashtirish

samaraliroq bo‘ladi; u multimediali prinsipni aynan so‘z va rasmni birgalikda berish zarurati bilan izohlaydi [Mayer., 2009: 223–241]. Demak, sinonim izohi ham ko‘p kanalli bo‘lishi kerak.

Xorijiy tajriba: multimediali yondashuvning tasnifi. Chet el elektron lug‘atlari tajribasini o‘rganish shuni ko‘rsatadiki, multimediali yondashuv bir xil shaklda emas, balki bir necha qatlamda namoyon bo‘ladi. Uni quyidagicha tasniflash mumkin:

Tajriba turi	Xorijiy namuna	Sinonim izohiga ta’siri	O‘zbek o‘quv lug‘ati uchun xulosa
Audio-fonetik yondashuv	Cambridge Dictionary, Merriam-Webster	So‘zning UK/US talaffuzi, urg‘usi va eshitish orqali eslab qolish imkonini beradi.	Sinonim qatoridagi har bir so‘z uchun adabiy talaffuz audio yozuvi berilishi kerak.
Vizual-semantik yondashuv	Oxford Learner’s Dictionaries	Rasm va kontekst orqali ma’no farqi tezroq anglashiladi.	Aniq predmet, belgi, harakat bildiruvchi sinonimlar rasm yoki piktogramma bilan izohlanadi.
Kontekst-korpus yondashuvi	Longman, COBUILD an’anasi	Real misol gaplar va kollokatsiyalar sinonimning qo‘llanish chegarasini ko‘rsatadi.	O‘zbek milliy korpusidan olingan sodda misollar qo‘shilishi zarur.
Thesaurus-integrativ yondashuv	Oxford Learner’s Thesaurus, Merriam-Webster Thesaurus	Sinonimlar darajalanadi, antonim va yaqin birliklar bilan bog‘lanadi.	Sinonimlar faqat ro‘yxat emas, semantik guruh va uslubiy belgi bilan beriladi.
Interaktiv-o‘quv yondashuv	Mobil lug‘at ilovalari	Test, tanlash, saqlash, ovozli qidirish orqali o‘quvchi faol ishtirok etadi.	Har bir sinonim qatoridan keyin mini-mashq va avtomatik tekshiruv bo‘lishi kerak.

Jadvaldan ko‘rinadiki, xorijiy tajriba sinonimlarni izohlashda bir vaqtning o‘zida fonetik, semantik, uslubiy, kontekstual va interaktiv komponentlardan foydalanadi. Lew va de Schryver raqamli inqilob sharoitida lug‘at foydalanuvchilarining xatti-harakati o‘zgargani, internet qidiruv tizimlari an’anaviy lug‘atga raqobatchi bo‘lib qolganini ta’kidlaydi; ular kelajak lug‘atlari foydalanuvchi ehtiyojiga moslashadigan moslashuvchan vositalar bo‘lishi kerakligini qayd etadi [Lew, de Schryver., 2014: 341–359]. Bu fikr sinonimlar lug‘ati uchun ham muhim: elektron muhitda foydalanuvchi faqat “ma’nosi nima?” deb so‘ramaydi, balki “qaysi so‘z qaysi vaziyatda to‘g‘ri?” degan javobni ham kutadi.

Xorijiy olimlar fikrlari va ularning amaliy ahamiyati. Atkins va Rundell amaliy leksikografiyani lug‘at tuzish jarayonining ehtiyoj, ma’lumot tanlash, ma’no tuzilishi va foydalanuvchiga qulay taqdimot bilan bog‘liq tizimli faoliyat sifatida ko‘rsatadi [Atkins, Rundell., 2008: 20]. Bu yondashuv sinonimlar o‘quv lug‘ati uchun shuni anglatadiki, har bir izoh o‘quvchining yoshiga, til bilish darajasiga va kommunikativ ehtiyojiga mos bo‘lishi kerak. Chet el o‘quv lug‘atlarida aynan shu prinsip asosida murakkab ilmiy ta’riflar soddalashtiriladi, ko‘p ma’noli so‘zlar alohida bloklarga ajratiladi va har bir ma’no uchun misol beriladi.

Levy va Steel elektron til lug‘atlaridan foydalanuvchi talabalar tajribasini tahlil qilib, o‘quvchilar elektron resurslardan so‘z tarjimasi, ma’no aniqlash, talaffuz va qo‘llanishni

tekshirish uchun mustaqil foydalanishini ko‘rsatadi [Levy, Steel., 2015: 177–196]. Bu kuzatuv sinonimlar izohida mustaqil o‘rganishga yo‘naltirilgan dizayn zarurligini anglatadi: o‘quvchi bir so‘zdan ikkinchisiga o‘tganda, tizim unga ma’no farqi, uslubiy belgi va kontekstni o‘zi solishtirib ko‘rish imkonini berishi kerak.

Dziemianko elektron va qog‘oz lug‘atlar o‘rtasidagi farqni faqat tashqi shakl bilan emas, balki qidiruv, gipermatn, tovush, rasm va foydalanuvchi harakatini boshqaruvchi funksiyalar bilan bog‘laydi [Dziemianko., 2012: 319–341]. Shundan kelib chiqib, multimediali sinonimlar lug‘ati oddiy PDF yoki qog‘oz nusxaning elektron ko‘rinishi bo‘lib qolmasligi zarur. U dinamik qidiruv, sinonimlar tarmog‘i, talaffuzni eshitish, misollarni daraja bo‘yicha saralash, so‘zlarni shaxsiy ro‘yxatga qo‘shish va mashq natijasini ko‘rish imkoniyatiga ega bo‘lishi lozim.

Mayerning multimediali ta’lim prinsiplaridan biri – ortiqcha axborotni kamaytirish va kerakli belgilarni ajratib ko‘rsatishdir. Sinonim izohida bu prinsip juda muhim: bitta sahifaga ortiqcha uzun matn, keraksiz rasm yoki ko‘p rangli bezak joylashtirilsa, o‘quvchining e’tibori asosiy farqqa emas, tashqi ko‘rinishga ketadi. Demak, multimediali yondashuv “ko‘proq material qo‘shish” emas, balki ma’no farqini aniqroq ko‘rsatadigan vositani tanlashdir.

O‘zbek tili sinonimlari uchun multimediali izoh modeli. O‘zbek tili o‘quv lug‘atlarida sinonimlarni multimediali izohlash uchun quyidagi model taklif etiladi. Birinchi bosqichda sinonim qatorning bosh so‘zi tanlanadi va barcha birliklar semantik markaz atrofida guruhlanadi. Masalan: chiroyli – go‘zal – ko‘rkam – dilbar. Ikkinchi bosqichda har bir sinonimning farqlovchi belgisi beriladi: chiroyli – neytral, umumiy; go‘zal – badiiy va ko‘tarinki; ko‘rkam – tashqi ko‘rinish va salobatni bildiradi; dilbar – emotsional, inson yoki obrazga nisbatan qo‘llanishi kuchli. Uchinchi bosqichda misol gaplar qo‘shiladi: “chiroyli bino”, “go‘zal manzara”, “ko‘rkam yigit”, “dilbar qiz”. To‘rtinchi bosqichda multimedia elementlari ulanadi: talaffuz audioyozuvi, rasm yoki piktogramma, urg‘u belgisi, jadval va interaktiv savol.

Bunday modelda kompyuter lingvistikasi yutuqlaridan ham foydalanish mumkin. Tukeyev va Abdurakhmonova hammuallifligidagi tadqiqotda o‘zbek so‘zlarini morfologik tahlil qilish, qo‘shimchalar to‘plamini tizimlashtirish va stemming algoritmlarini qurish masalasi yoritilgan [Tukeyev va boshq., 2024: 1760–1764]. Sinonimlar lug‘ati uchun bunday yondashuv foydalanuvchi so‘zning turli shakllarini yozganda ham asosiy lemma va sinonim qatoriga olib borish imkonini beradi. Masalan, chiroyliroq, chiroylilik, go‘zallik kabi shakllar avtomatik ravishda tegishli asosiy birlikka bog‘lanishi mumkin.

Multimediali izoh modeli ayniqsa maktab o‘quvchilari va til o‘rganuvchilar uchun qulay. Chunki ular bir vaqtning o‘zida ko‘rish, eshitish, o‘qish va bajarish kanallari orqali bilim oladi. Masalan, “tez” sinonim qatori uchun tez, chaqqon, ildam, shoshilinch birliklari ajratilsa, tizim ularning farqini harakat tezligi, xulq-atvor, rasmiylik va vaziyat bo‘yicha ko‘rsatadi. O‘quvchi audio orqali talaffuzni eshitadi, animatsiya orqali harakat darajasini ko‘radi, misol gap orqali kontekstni tushunadi va test orqali o‘zlashtirishini tekshiradi.

Amaliy tavsiyalar. *Birinchi tavsiya* – sinonim izohini qatlamli tuzish. Har bir maqola boshida qisqa umumiy ma’no, keyin farqlovchi belgi, undan so‘ng misol, uslubiy izoh, kollokatsiya va mashq berilishi kerak. *Ikkinchi tavsiya* – audio va talaffuzni majburiy komponentga aylantirish. Cambridge Dictionary va Merriam-Webster tajribasi shuni ko‘rsatadiki, audio talaffuz o‘quvchi uchun so‘zning fonetik shaklini mustahkamlaydi. *Uchinchisi* – sinonimlarni rang yoki piktogramma orqali belgilash: neytral so‘zlar bir belgi, badiiy so‘zlar ikkinchi belgi, so‘zlashuv birliklari uchinchi belgi bilan ajratiladi. *To‘rtinchi tavsiya* – korpusga asoslangan misollar bazasini yaratish. O‘zbek tilidagi real matnlardan

olingan, lekin o‘quvchi darajasiga moslashtirilgan gaplar sinonimlarning tabiiy qo‘llanishini ko‘rsatadi. *Beshinchi tavsiya* – interaktiv mashqlarni lug‘at maqolasining ajralmas qismi sifatida kiritish. Masalan, “Berilgan gapda qaysi sinonim mos?” yoki “Rasmga mos so‘zni tanlang” kabi topshiriqlar o‘quvchini passiv o‘qishdan faol tanlovga olib keladi. *Oltinchi tavsiya* – o‘zbek tili sinonimlari uchun mobil ilova va veb-platforma yaratishda foydalanuvchi tiplari ajratilishi lozim: maktab o‘quvchisi, abituriyent, chet ellik o‘rganuvchi, tarjimon va jurnalist. Har bir foydalanuvchi uchun izoh chuqurligi va multimedia darajasi farq qilishi mumkin. Bu jihat xorijiy elektron lug‘atlarning foydalanuvchiga moslashuvchanlik tajribasi bilan uyg‘un keladi.

Xulosa. Sinonimlarni o‘quv lug‘atlarida izohlash muammosi faqat leksik ma’no masalasi emas, balki semantik, uslubiy, pragmatik va didaktik muammodir. An’anaviy ro‘yxatli izoh o‘quvchiga sinonimlar orasidagi nozik farqlarni yetarlicha anglatmaydi. Chet el tajribasi esa sinonimlarni audio, rasm, misol, kollokatsiya, thesaurus, interaktiv mashq va foydalanuvchiga mos interfeys orqali tushuntirish samarali ekanini ko‘rsatadi. Shunday qilib, o‘zbek tili o‘quv lug‘atchiligida multimediali yondashuv qo‘shimcha bezak emas, balki izohning sifatini oshiruvchi asosiy metodik vosita sifatida qaralishi kerak.

Taklif etilgan model sinonim qatorni semantik markaz, farqlovchi belgi, uslubiy belgi, kontekst, audio, vizual element va interaktiv topshiriq asosida izohlashni nazarda tutadi. Bunday yondashuv o‘quvchilarning so‘z boyligini oshirish, nutqiy aniqlikni shakllantirish va o‘zbek tilining elektron ta’lim resurslarini rivojlantirishga xizmat qiladi. Kelgusida bu model asosida maktab o‘quvchilari uchun sinov elektron sinonimlar lug‘atini yaratish va uning samaradorligini tajriba-sinov orqali tekshirish maqsadga muvofiqdir.

FOYDALANILGAN ADABIYOTLAR:

1. Abdurakhmonova N. Kompyuter lingvistikasi (darslik), Toshkent: Nodirabegim, 2021. – B. 400.
2. Atkins B. T. S., Rundell M. The Oxford Guide to Practical Lexicography. – Oxford: Oxford University Press, 2008. – 552 p.
3. Crystal D. A Dictionary of Linguistics and Phonetics. 6th ed. – Oxford: Blackwell Publishing, 2008. – 529 p.
4. Dziemianko A. On the use(fulness) of paper and electronic dictionaries // Electronic Lexicography / ed. S. Granger, M. Paquot. – Oxford: Oxford University Press, 2012. – pp. 319–341.
5. Landau S. I. Dictionaries: The Art and Craft of Lexicography. 2nd ed. – Cambridge: Cambridge University Press, 2001. – 477 p.
6. Levy M., Steel C. Language learner perspectives on the functionality and use of electronic language dictionaries // ReCALL. – 2015. – Vol. 27, Issue 2. – pp. 177–196. DOI: 10.1017/S095834401400038X.
7. Lew R., de Schryver G.-M. Dictionary users in the digital revolution // International Journal of Lexicography. – 2014. – Vol. 27, Issue 4. – pp. 341–359. DOI: 10.1093/ijl/ecu011.
8. Mayer R. E. Multimedia Learning. 2nd ed. – Cambridge: Cambridge University Press, 2009. – pp. 223–241.
9. Oxford Learner’s Dictionaries. Online learner dictionary platform: definitions, pictures, examples, synonyms, pronunciation audio. – Accessed: 24.05.2026.
10. Cambridge Dictionary. Online dictionary and thesaurus for learners of English: meanings, definitions, British/American audio pronunciation and translations. – Accessed: 24.05.2026.

KORPUSGA ASOSLANGAN OMONIMLAR LUG'ATINI YARATISHDA OMONIMLARNI TASNIFLASH VA LOYIHALASHTIRISH MASALALARI

Durdona Anvarova
O'zMU 2-kurs tayanch doktoranti
https://doi.org/10.5281/zenodo.23136121

Annotatsiya. *Mazkur maqolada korpusga asoslangan omonimlar lug'atini yaratish jarayonining ikki o'zaro bog'liq jihati – omonimiya hodisasini lingvistik tasniflash tamoyillari va lug'atning makro- hamda mikrostrukturaviy loyihasi – tahlil etiladi. Omonimiya va polisemiyani ajratishning an'anaviy (etimologik, semantik) mezonlari izchil ko'rib chiqiladi. Ushbu mezonlar zamonaviy korpus lingvistikasi, xususan, so'z ma'nosi induksiyasi (word sense induction) usullari nuqtai nazaridan qayta ko'rib chiqiladi. Shuningdek, o'zbek tilshunosligidagi omonimshunoslik an'analari bilan xalqaro leksikografik tajriba o'rtasidagi uyg'unlik masalalari o'rganiladi. O'zbek tilining aglyutinativ morfologiyasi korpusga asoslangan lug'at loyihalashga qo'yadigan o'ziga xos talablari ko'rib chiqiladi.*

Kalit so'zlar: *omonimiya, polisemiya, korpus lingvistikasi, leksikografiya, so'z ma'nosi induksiyasi (WSI), makrostruktura, mikrostruktura, o'zbek tili milliy korpusi.*

Аннотация. *В данной статье анализируются два взаимосвязанных аспекта процесса создания корпусного словаря омонимов – принципы лингвистической классификации явления омонимии и макро- и микроструктурное проектирование словаря. Последовательно рассматриваются традиционные (этимологические, семантические) критерии разграничения омонимии и полисемии. Данные критерии пересматриваются с точки зрения современной корпусной лингвистики, в частности, методов индукции значения слова (word sense induction). Также изучаются вопросы сочетания традиций изучения омонимов в узбекском языкознании с международным лексикографическим опытом. Рассматриваются специфические требования, предъявляемые агглютинативной морфологией узбекского языка к проектированию корпусного словаря.*

Ключевые слова: *омонимия, полисемия, корпусная лингвистика, лексикография, индукция значения слова (WSI), макроструктура, микроструктура, Национальный корпус узбекского языка.*

Abstract. *This article analyzes two interrelated aspects of the process of creating a corpus-based dictionary of homonyms: the principles of linguistic classification of the phenomenon of homonymy and the macro- and microstructural design of the dictionary. Traditional (etymological, semantic) criteria for distinguishing homonymy and polysemy are considered consistently. These criteria are reviewed from the perspective of modern corpus linguistics, specifically word sense induction methods. Also, issues of harmony between the traditions of homonymy in Uzbek linguistics and international lexicographic experience are studied. The specific requirements of the agglutinative morphology of the Uzbek language for the design of a corpus-based dictionary are considered.*

Keywords: *homonymy, polysemy, corpus linguistics, lexicography, word meaning induction (WSI), macrostructure, microstructure, national corpus of the Uzbek language.*

So‘nggi yillarda O‘zbekistonda raqamli lingvistik resurslar yaratishga davlat darajasida katta e’tibor qaratilmoqda. Xususan, sun’iy intellekt tizimlarida o‘zbek tilini samarali joriy etish maqsadida milliy til korpusini shakllantirish bo‘yicha keng ko‘lamli ish boshlangan: 2025 yil oktyabr oyida Raqamli texnologiyalar vazirligi rahbariyati tomonidan 2026 yil oxiriga qadar o‘n milliard so‘z hajmidagi o‘zbek tili korpusini tuzish va bu asosda kamida beshta ochiq manbali katta til modelini yaratish rejalari e’lon qilindi [Shermatov, 2025]. Bunday miqyosdagi korpus nafaqat sun’iy intellekt modellarini o‘qitish, balki an’anaviy leksikografik amaliyotni, jumladan, omonimlar lug‘atini tuzish ishini ham tubdan yangilash imkonini beradi.

Ma’lumki, o‘zbek tilshunosligida omonimlar masalasi uzoq tarixga ega. Sh. Rahmatullaevning “O‘zbek tili omonimlarining izohli lug‘ati” hanuzgacha sohaning asosiy manbalaridan biri hisoblanadi [Rahmatullaev Sh., 1984], shuningdek, M. Mirtojievning semasiologiya va leksik semantikaga oid tadqiqotlari omonimiya va polisemiya chegaralarini nazariy jihatdan asoslashga muhim hissa qo‘shgan [Mirtojiev M., 2013; Mahmudov, Nurmonov, 2007]. Biroq bu tadqiqotlarning aksariyati asosan leksikograf-tadqiqotchining til haqidagi intuisiyasi va cheklangan misollar bazasiga tayangan holda amalga oshirilgan edi. Bu esa XX asr leksikografiyasining umumjahon miqyosidagi tabiiy xususiyati edi. Zamonaviy korpus lingvistikasi esa lug‘at tuzish jarayonini empirik dalillarga, ya’ni katta hajmdagi tabiiy matnlar majmuasidagi real qo‘llanish holatlariga asoslashga imkon beradi.

Shu munosabat bilan mazkur maqolaning maqsadi korpusga asoslangan omonimlar lug‘atini yaratishda. Birinchidan, omonimlarni ilmiy jihatdan to‘g‘ri tasniflash mezonlarini, ikkinchidan, bunday lug‘atning tarkibiy (makro- va mikrostrukturaviy) loyihasini belgilovchi asosiy masalalarni xalqaro leksikografik va kompyuter lingvistikasi tajribasi asosida tahlil etishdan iborat.

Tilshunoslik an’anasida omonimiya va polisemiya o‘rtasidagi chegara doimo bahsli masala bo‘lib kelgan. Ko‘pchilik tadqiqotchilar ikki asosiy mezonga tayanadi: birinchisi so‘zning etimologik kelib chiqishi, ikkinchisi esa ma’nolar o‘rtasidagi semantik bog‘liqlik darajasi. Agar bir xil shaklga ega ikki yoki undan ortiq ma’no tarixan turli manbalardan kelib chiqqan bo‘lsa va ular orasida sinxron nuqtai nazardan hech qanday mazmuniy aloqa sezilmasa, bunday hodisa omonimiya deb tan olinadi; agar ma’nolar bir o‘zakdan metaforik yoki metonimik ko‘chish yo‘li bilan rivojlangan bo‘lsa va til egalari ularni intuitiv ravishda bog‘liq deb qabul qilsa, bu polisemiya hisoblanadi [Vicente, Falkum, 2017; Cruse, 1986: 80].

Shunga qaramay, bu mezonlarning har ikkisi ham mutlaq emas. Xususan, etimologik mezon ba’zan aldamchi bo‘lishi mumkin. Tarixan bir o‘zakdan tarqalgan so‘zlar zamonaviy til ongida butunlay alohida leksik birliklar sifatida idrok etilishi mumkin. Shu bois ko‘plab zamonaviy tadqiqotlarda sinxron yondashuv ustuvor deb hisoblanadi: bunda ikki ma’no faqat ularning hozirgi tilda bir-biri bilan qanday aloqada idrok etilishiga qarab omonim yoki polisemantik deb baholanadi, ularning tarixiy kelib chiqishidan qat’i nazar. Semantik yaqinlik mezonining o‘zi ham doimiy keskin chegara bermaydi: masalan, ingliz tilidagi “bank” (daryo sohili) va “bank” (moliya muassasasi) so‘zlari ko‘pchilik lug‘atlarda omonim sifatida berilsa-da, ayrim tilshunoslar ularni tarixan bir etimologik o‘zakka borib taqaluvchi, juda uzoqlashgan polisemiya hodisasi deb ham tahlil qiladilar. Demak, omonimiya va polisemiya o‘rtasidagi chegara uzluksiz ko‘rinishida namoyon bo‘ladi. Bu esa amaliy leksikografiyada, xususan, lug‘at maqolasini loyihalashda doimiy metodologik qiyinchilik tug‘diradi.

Psixolingvistik tadqiqotlar shuni ko‘rsatadiki, omonimlarning turli ma’nolari inson ongida alohida-alohida mustaqil shaklda saqlanadi. Polisemantik so‘zlarning ma’nolari esa

umumiy "o‘zak ma’no" atrofida birlashgan tarmoq sifatida idrok etiladi. Bu farq nutqni qayta ishlash tezligidagi farqlarda ham o‘z aksini topadi: polisemantik so‘zlar odatda omonimlarga nisbatan tezroq anglab yetiladi, chunki ularning ma’nolari o‘rtasida umumiy kognitiv "ko‘prik" mavjud. Bunday psixolingvistik dalillar omonimiya va polisemiyaning faqat til tarixi yoki lug‘aviy an’ana masalasi emas, balki til egasining ongida ob’yektiv asosga ega bo‘lgan hodisa ekanini tasdiqlaydi. Bu esa korpusga asoslangan lug‘at uchun ma’nolar klassifikasiyasini faqat statistik mezonlargagina emas, balki kognitiv asoslanganlikka ham tayanishini zarurligini ko‘rsatadi.

Jahon tilshunosligida omonimlarni tasniflashning bir necha asosiy yo‘nalishi shakllangan. Eng keng tarqalgan tasnif to‘liq va qisman omonimlarga bo‘lish hisoblanadi. To‘liq omonimlar so‘zning barcha grammatik shakllarida tovush va yozilish jihatidan bir xillik saqlanishini nazarda tutadi. Qisman omonimlarda esa bu moslik faqat ayrim shakllarda, masalan, faqat bosh kelishikda yoki faqat birlik sonda kuzatiladi. Ikkinchi muhim tasnif esa omonimlarning leksik, grammatik va leksik-grammatik turlarga ajratilishidir. Leksik omonimlar bir so‘z turkumiga mansub bo‘lib, faqat ma’nosi bilan farqlanuvchi so‘zlardir. Grammatik omonimiya esa, aksincha, turli so‘z turkumlariga mansub so‘z shakllarining tovush jihatidan mos kelishida namoyon bo‘ladi. Leksik-grammatik omonimiya esa bu ikkisining oraliq holatini tashkil etadi.

Bundan tashqari, "omonim" tushunchasining o‘zi ham keng va tor ma’nolarda qo‘llaniladi. Keng ma’noda omonimlarga yozilishi bir xil, lekin talaffuzi har xil bo‘lgan geteronimlar (so‘zning urg‘usiga qarab ma’nosi o‘zgaruvchi shakllar), talaffuzi bir xil, lekin yozilishi har xil bo‘lgan omofonlar, hamda yozilishi bir xil bo‘lgan omograflar ham kiritiladi. Tor ma’noda esa omonim deb faqat ham tovush, ham yozilish jihatidan bir xil bo‘lgan, lekin ma’nosi mutlaqo bog‘liq bo‘lmagan so‘zlar tushuniladi. Lug‘at tuzish amaliyotida aynan shu tor ma’no – ya’ni leksik omonimiya – asosiy obyekt bo‘lib xizmat qiladi, chunki omofoniya va omografiya ko‘proq yozuv tizimi va fonetika sohasiga oid masala hisoblanadi [Vinogradov, 1978:299].

O‘zbek tilshunosligida omonimiya masalasi asosan leksik-semantik yo‘nalishda o‘rganilgan. Sh. Rahmatullaevning izohli lug‘ati amaliy leksikografik manba sifatida omonimlarni alohida lug‘at maqolalari shaklida – raqamlangan bosh so‘zlar ko‘rinishida – tartibga solgan edi. Bu esa keyingi tadqiqotlar uchun tayanch sifatida xizmat qilmoqda. M. Mirtojievning ilmiy maktabi esa omonimiya va sinonimiya hodisalarini semiologik birlik sifatida izohlashga harakat qilgan [Mirtojiev, 2013]. So‘nggi yillarda esa o‘zbek tilshunosligida omonimiyani lingvokognitiv aspektda o‘rganish yo‘nalishi rivojlanmoqda. Bunda omonimiya hodisasi nafaqat tizimli-struktur, balki til egasining dunyoni idrok etish tarzi bilan bog‘liq hodisa sifatida ham tahlil qilinadi. Ayniqsa e’tiborliki, so‘nggi tadqiqotlarda omonimiya va polisemiya masalasi endi faqat badiiy matn yoki an’anaviy lug‘atshunoslik doirasida emas, balki korpus lingvistikasi nuqtai nazaridan tadqiq etila boshlandi. Xususan, frazeologik omonimiya va polisemiyaning korpus materialida namoyon bo‘lishi misolida ham o‘rganila boshlandi [Nurboeva, 2023:114–116]. Bu esa o‘zbek omonimshunosligining korpusga asoslangan leksikografiya bilan tabiiy ravishda integrasiyalashuv jarayonida ekanini ko‘rsatadi.

Korpusga asoslangan leksikografiyada ikki asosiy metodologik yo‘nalish farqlanadi: "korpusga asoslangan" (corpus-based) va "korpus tomonidan boshqariladigan" (corpus-driven) yondashuvlar. Birinchisida leksikograf avvaldan so‘zning ma’nosi haqida intuitiv tasavvurga ega bo‘ladi va korpusdan faqat bu tasavvurni tasdiqlash yoki misollar bilan boyitish uchun foydalanadi. Ikkinchisida esa jarayon teskari yo‘nalishda kechadi: tadqiqotchi so‘zning

korpusdagi barcha qo‘llanish holatlarini to‘playdi, ularni tahlil qilib, guruhlarga (ya’ni alohida ma’nolar yoki, omonimik birliklarga) ajratadi va faqat shundan keyin lug‘at maqolasini tuzadi. Amaliyotda, ayniqsa katta va xilma-xil milliy korpuslar bilan ishlashda, bu ikki yondashuvning uyg‘unlashuvi, ya’ni leksikografning nazariy tayyorgarligi bilan korpus beradigan statistik dalillarning birgalikda qo‘llanilishi eng samarali usul hisoblanadi.

Kompyuter lingvistikasida so‘z ma’nolarini korpus materiali asosida avtomatik aniqlash vazifasi “so‘z ma’nosi induksiyasi” (word sense induction, WSI) deb ataladi. Bu vazifa “so‘z ma’nosini aniqlashtirish” (word sense disambiguation, WSD) dan farq qiladi. WSD avvaldan berilgan, tayyor ma’nolar ro‘yxatiga tayanib, matndagi konkret qo‘llanishni shu ro‘yxatdagi biror ma’noga bog‘laydi. WSI esa aksincha – ma’nolar ro‘yxatining o‘zini ham korpus materialidan keltirib chiqaradi. Bunda “so‘z ma’nosi – bu korpusdagi qo‘llanish holatlari to‘plamidan ajratilgan abstraksiyadir” degan tamoyilga tayanadi [Kilgarriff; Ustalov et., 2018]. WSI usullari asosan quyidagi yo‘nalishga bo‘linadi:

	Yo‘nalish	**Vazifasi**
1	kontekstni klasterlash	so‘zning har bir qo‘llanishi atrofidagi matn birligi vektor ko‘rinishida ifodalanib, o‘xshash vektorlar bir guruhga birlashtiriladi
2	so‘zlarni klasterlash	so‘z bilan sintaktik yoki statistik bog‘liq boshqa so‘zlar guruhlarga ajratiladi
3	birga qo‘llanish graflari	co-occurrence graphs usuli orqali amalga oshiriladi

Shunday qilib, zamonaviy korpus lingvistikasi omonimlarni aniqlash jarayonini butunlay avtomatlashtirmaydi, balki leksikografning qaror qabul qilish jarayonini empirik dalillar bilan boyituvchi qudratli vosita sifatida xizmat qiladi. Aynan shu nuqtai nazar, ya’ni inson eksperti bilan hisoblash algoritmlarining hamkorligi korpusga asoslangan omonimlar lug‘ati loyihasining metodologik poydevorini tashkil etishi lozim.

Lug‘atshunoslik nazariyasida har qanday lug‘atning tashkiliy tuzilishi ikkita asosiy darajaga bo‘linadi: makrostruktura va mikrostruktura. Makrostruktura – bu lug‘atdagi bosh so‘zlar (lemma) ro‘yxatining umumiy tartibi va prinsipi, ya’ni lug‘at maqolalari bir-biriga nisbatan qanday joylashtirilgani (alifbo tartibida, chastota – qo‘llanish darajasi bo‘yicha yoki boshqa mezonlar asosida). Mikrostruktura esa alohida olingan har bir lug‘at maqolasining ichki tarkibi – lemma, talaffuz, grammatik belgilar, etimologik ma’lumot, ta’rif, misollar va boshqa komponentlarning qanday tartibda berilishidir [Lehmann, n.d.; Holova, Nashirova, 2024:41-42].

Shuningdek, omonimlar lug‘atining mikrostrukturasida omonimlarning turi (to‘liq/qisman, leksik/grammatik/leksik-grammatik) ham alohida belgi sifatida ko‘rsatilishi maqsadga muvofiq, chunki bu ma’lumot foydalanuvchiga, ayniqsa, tilni o‘rganuvchilar va tarjimonlarga so‘zning qaysi grammatik shakllarda omonimlik xususiyatini saqlab qolishi yoki yo‘qotishi haqida aniq tasavvur beradi. Umuman olganda, zamonaviy korpusga asoslangan lug‘at loyihasi an’anaviy leksikografik tamoyillarni (makro/mikrostruktura nazariyasi) inkor etmaydi, balki ularni empirik dalillar bilan boyitib, yanada moslashuvchan va foydalanuvchi ehtiyojlariga yo‘naltirilgan shaklga keltiradi.

Xulosa shuki, korpusga asoslangan omonimlar lug‘atini yaratish ikki o‘zaro chambarchas bog‘liq masalani – nazariy-tasnifiy va loyihaviy-tarkibiy masalalarni – birgalikda hal etishni talab etadi. Omonimiya va polisemiya o‘rtasidagi chegara mutlaq emas, u etimologik, semantik va psixolingvistik mezonlarning o‘zaro ta’sirida belgilanadigan tushuncha sifatida namoyon

bo‘ladi. Shu bois har qanday lug‘at loyihasi bu chegarani belgilashda aniq metodologik pozisiyani tanlashi shart. Zamonaviy korpus lingvistikasi, xususan, so‘z ma’nosi induksiyasi usullari, bu jarayonni butunlay avtomatlashtirmasa-da, leksikografning qarorlarini empirik dalillar bilan asoslashga kuchli vosita beradi. Lug‘atning makro- va mikrostrukturasi esa an’anaviy leksikografik tamoyillarni korpus beradigan chastota va kontekst ma’lumotlari bilan boyitish orqali loyihalashtirilishi lozim. O‘zbek tili uchun bunday lug‘at yaratish jarayonida milliy tilshunoslik maktabining boy tajribasini (Rahmatullaev, Mirtojiev va boshqalar) zamonaviy milliy korpus resurslari va kompyuter lingvistikasi usullari bilan uyg‘unlashtirish – sohaning eng dolzarb va istiqbolli yo‘nalishlaridan biri bo‘lib qolmoqda.

FOYDALANILGAN ADABIYOTLAR

1. Anvarovna, H. A., & Burievna, S. N. (2024). The representation of macro and microstructure of bilingual dictionaries. *Science and Innovation*, 3(Special Issue 4), 41–45.
2. Atkins, B. T. S., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford University Press.
3. Čermák, F. (2010). Notes on compiling a corpus-based dictionary. *Lexikos*, 20, 559–579. https://doi.org/10.5788/20-0-156
4. Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.
5. Exploring the representation of word meanings in context: A case study on homonymy and synonymy. (2021). *arXiv*. https://arxiv.org/abs/2106.13553
6. Homonyms in modern German: Ways of occurrence and typological characteristics. (2018). *Philology Journal*. philology-journal.ru
7. Introduction to corpus-based lexicographic practice. (2022). DARIAH-Campus.
8. Kilgarriff, A. (2018). Russian word sense induction by clustering averaged word embeddings. *arXiv*. https://arxiv.org/abs/1805.02258
9. Krovetz, R. (1997). Homonymy and polysemy in information retrieval. *Proceedings of ACL-EACL'97*, 72–79.
10. Lehmann, C. (n.d.). Functions and types of dictionaries. *Linguistic Description*. https://www.christianlehmann.eu/ling/ling_meth/ling_description/lexicography/functions_and_kinds.html
11. Mahmudov, N., & Nurmonov, A. (2007). Miraziz Mirtojiev. *O‘zbek tili va adabiyoti*, (4).
12. Mirtojiev, M. (2013). *O‘zbek tili fonetikasi*. Fan.
13. Navruzova, M. G‘., & Muxamatova, D. X. (2025). O‘zbek tili korpusi uchun tibbiy tezaurus lug‘atlar yaratish zarurati xususida. *Oriental Renaissance*, 5(9), 438–443.
14. Nurboeva, M. (2023). Korpus lingvistikasida frazeologik omonimiya va polisemiya tadqiqi. *ZDIT*, (2(20)), 114–116.
15. Opposing effects of semantic diversity in lexical and semantic relatedness decisions. (2015). *PMC (NCBI)*.
16. O‘zbek tili milliy korpusida so‘zshaklni leksikografik baza asosida qidiruv imkoniyatlari: Monografiya. (2021). ResearchGate.
17. O‘zbek tilining milliy korpusi: Muammo va vazifalar: Xalqaro ilmiy-amaliy konferensiya materiallari. (2022). ToshDO‘TAU.
18. Rahmatullaev, Sh. (1984). *O‘zbek tili omonimlarining izohli lug‘ati*. O‘qituvchi.
19. The role of corpora in dictionary making. (2022). FutureLearn.

20. Rossini Favretti, R., Tamburini, F., & Zaninello, A. (2011). Exploiting corpus evidence for automatic sense induction. *Proceedings of CICLing*, University of Bologna.
21. Shermatov, Sh. (2025, 8-oktabr). *O‘zbek tili korpusi va milliy sun'iy intellekt modeli yaratish bo‘yicha ma'ruza* [Ma'ruza]. O‘zbekiston Respublikasi Oliy Majlisi Qonunchilik palatasi (spot.uz nashri bo‘yicha).
22. Sketch Engine. (2024). *Automatic identification of word senses* [Texnik qo‘llanma]. sketchengine.eu
23. Vicente, A., & Falkum, I. L. (2017). Polysemy. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press.
24. Vinogradov, V. V. (1978). *Tanlangan asarlar: Rus grammatikasi bo‘yicha tadqiqotlar*. Nauka.
25. Word-sense induction. (n.d.). In *Wikipedia*.

KORPUS TAHLILI ASOSIDA NLP TEXNOLOGIYALARINING RIVOJLANISHI VA LINGVOSTATISTIK TADQIQOTLARDAGI AHAMIYATI

Umida Askarova
Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi o'qituvchisi O'zbekiston Milliy universiteti
https://doi.org/10.5281/zenodo.23136134

Annotatsiya. *Mazkur maqolada korpus lingvistikasi, lingvostatistika va tabiiy tilni qayta ishlash (NLP) texnologiyalarining o'zaro bog'liqligi yoritiladi. Korpus real nutqni aks ettiruvchi elektron matnlar majmuasi sifatida zamonaviy tilshunoslikning muhim empirik manbasi hisoblanadi. Maqolada miqdoriy va sifat tahlili, chastota ko'rsatkichlari, statistik modellar hamda ko'p o'lchovli yondashuvlarning o'rni tahlil qilinadi. Shuningdek, BNC, Brown va COCA kabi yirik korpuslarning amaliy ahamiyati hamda J. Zipf, G.U. Yule, W. Herdan, D. Biber, M. Stubbs, R.H. Baayen va J. Sinkler kabi olimlarning ilmiy hissasi ko'rib chiqiladi. Shuningdek, O'zbek tili milliy korpusining NLP, mashina tarjimasi, avtomatik matn tahlili va sun'iy intellekt texnologiyalaridagi imkoniyatlariga ham to'xtalib o'tilgan.*

Kalit so'zlar: *korpus lingvistikasi, lingvostatistika, tabiiy tilni qayta ishlash, o'zbek tili milliy korpusi, chastota tahlili, konkordans, mashina tarjimasi, sun'iy intellekt.*

Abstract. *This article examines the interrelationship between corpus linguistics, linguistic statistics, and natural language processing (NLP) technologies. A corpus, defined as a collection of electronic texts reflecting actual speech, serves as a vital empirical source for modern linguistics. The article analyzes the roles of quantitative and qualitative analysis, frequency metrics, statistical models, and multidimensional approaches. It also discusses the practical significance of major corpora—such as the BNC, Brown, and COCA—and the scholarly contributions of researchers including G.K. Zipf, G.U. Yule, W. Herdan, D. Biber, M. Stubbs, R.H. Baayen, and J. Sinclair. The capabilities of the National Corpus of the Uzbek Language in the fields of NLP, machine translation, automatic text analysis, and artificial intelligence technologies were also highlighted.*

Keywords: *corpus linguistics, linguistic statistics, natural language processing, National Corpus of the Uzbek Language, frequency analysis, concordance, machine translation, artificial intelligence.*

Аннотация. *В статье рассматривается взаимосвязь между корпусной лингвистикой, лингвистической статистикой и технологиями обработки естественного языка (NLP). Корпус, определяемый как совокупность электронных текстов, отражающих реальную речевую практику, служит важнейшим эмпирическим источником для современной лингвистики. В работе анализируются роль количественного и качественного анализа, частотные показатели, статистические модели и многомерные подходы. Также обсуждается практическая значимость крупных корпусов (таких как BNC, Brown и COCA) и научный вклад таких исследователей, как Г.К. Ципф, Дж.У. Юл, Г. Хердан, Д. Бибер, М. Стаббс, Р.Х. Баайен и Дж. Синклер. Также были отмечены возможности Национального корпуса узбекского языка в сферах обработки*

естественного языка (NLP), машинного перевода, автоматического анализа текста и технологий искусственного интеллекта.

Ключевые слова: *корпусная лингвистика, лингвистическая статистика, обработка естественного языка, Национальный корпус узбекского языка, частотный анализ, конкорданс, машинный перевод, искусственный интеллект.*

Zamonaviy tilshunoslikning muhim yo‘nalishi bo‘lgan Korpus lingvistikasi real nutqni aks ettiruvchi, kompyuterda saqlanadigan va tizimlashtirilgan elektron matnlar majmuasini tez hamda aniq tahlil qilishga asoslanadi. Bu soha tilni o‘rganishda empirik ma’lumotlarga tayanish va nazariyalarni amalda sinash imkonini beradi. O‘zbek tili milliy korpusi tilimizning o‘ziga xos xususiyatlarini saqlash, raqamlashtirish va NLP hamda sun’iy intellekt (SI) texnologiyalari uchun asosiy ma’lumotlar bazasi sifatida xizmat qiladi [Elov B. va boshqalar, 2025: 1].

NLP sohasi katta hajmdagi til ma’lumotlariga tayanadi. O‘zbek tili korpusi mashina tarjimasi (parallel korpuslar orqali), matnlarni avtomatik tasniflash, so‘z tanib olish va matn tahlili kabi amaliy masalalarni hal etishda asosiy infratuzilma vazifasini o‘taydi. Til birliklarini miqdoriy o‘rganish ehtiyoji natijasida lingvostatistika shakllandi. U intuitiv xulosalardan voz kechib, empirik dalillar beradi. Korpuslar real nutqning eng ishonchli manbai, lingvostatistika esa til tizimining ichki qonuniyatlarini ochib beruvchi vositadir. Matematik va kompyuter lingvistikasi, amaliy tilshunoslik hamda stilometriya kabi sohalar aynan korpuslardan olingan statistik natijalarga tayanadi.

Til korpusi ikki xil tahlil qilinadi:

Miqdoriy tahlil. Lingvistik xususiyatlarni tasniflaydi va statistik modellar ishlab chiqadi. Bu modellar qaysi hodisalar tilning haqiqiy qonuniyati va qaysilari tasodif ekanini aniqlashga xizmat qiladi.

Sifat tahlili miqdoriy tahlil orqali korpusda kuzatilgan hodisalarning to‘liq va batafsil ta’rif berishga qaratilgan. Korpusning sifat tahlili aniq xulosalarni chiqarishga imkon beradi. Chunki bu holda ma’lumotlarni cheklangan tasniflar to‘plamiga kiritish shart emas. Ham miqdoriy, ham sifat tahlillarini amalga oshirish natijasida korpusni tahlil qilish samaradorligi oshadi. Biroq miqdoriy tahlil evaziga olingan natijalar sifat tahlilidan olingan natijalarga qaraganda samaradorligi pastroq hisoblanadi” [Yuldashev A., 2024: 2]

Bugungi kunda korpus asosida miqdoriy tahlilni amalga oshirishda bir qator statistik yondashuvlar qo‘llaniladi:

1. Tavsiflovchi statistik yondashuv kuzatilgan ma’lumotlarning asosiy xususiyatlarini umumlashtirib, ularni sodda va tushunarli shaklda ifodalash imkonini beradi.
2. Inferensial statistik yondashuv esa tavsiflovchi tahlil natijalariga tayanib, muayyan savollarga javob topish yoki ilmiy gipotezalarni ilgari surishga xizmat qiladi.
3. Baholovchi statistik yondashuv gipotezaning ma’lumotlar bilan qay darajada tasdiqlanishini aniqlash, shuningdek, matematik model yoki nazariy taqsimotning real holatga mosligini tekshirishda muhim ahamiyat kasb etadi.

Korpus ustida amalga oshirilgan miqdoriy tahlil natijalarini taqqoslash va yashirin shablonlarni aniqlash uchun quyida keltirilgan ko‘p o‘lchovli statistik usullardan foydalaniladi:

- *omillar tahlili (Factor Analysis);*
- *ko‘p o‘lchovli masshtablash (Multidimensional Scaling);*

- *klaster tahlili (Cluster Analysis);*
- *log-linear modellari (Log-linear Models)* [Yuldashev A., 2024: 2]

Korpusda jamlangan so‘zlarni odatda ikki xil usulda tartiblash mumkin. Shulardan biri — raqamli saralash bo‘lib, u miqdoriy ma’lumotlar bilan ishlashning eng sodda va qulay usullaridan biridir. Bu yondashuvda obyektlar ma’lum bir tizim asosida guruhlanadi va har bir guruhga tegishli matnlardagi elementlar soni arifmetik jihatdan hisoblab chiqiladi. Oddiy chastota hisoblash jarayonida mavjud ma’lumotlar foydalanuvchiga alifbo yoki sonlar tartibida taqdim etiladi, hamda hosil qilingan ro‘yxat ehtiyojga qarab o‘sish yoki kamayish tartibida joylashtirilishi mumkin. Matnni tahlil qilishda esa turli birliklar — so‘z, token, so‘z birikmasi yoki lemma — necha marta uchrashi aniqlanadi.

Natijada so‘zlarning chastota ro‘yxati shakllanadi. Bunday ro‘yxatni o‘rganish matn tuzilishi haqida muayyan tasavvur beradi va keyingi tahlil bosqichlarini rejalashtirishga yordam beradi. Shu bilan birga, alifbo tartibida tuzilgan so‘zlar ro‘yxati umumiy ma’lumotnoma sifatida ham qo‘llanishi mumkin. Alfavit tartibidagi chastota ro‘yxatlari ilmiy jihatdan ham muhim bo‘lib, ular ko‘pincha yangi gipotezalarni ilgari surish va asosli qarorlar qabul qilishga xizmat qiladi.

O‘zbek tili korpusida chastotani hisoblashdan oldin esa unda qo‘llanilgan belgilar, so‘zlar, iboralar, frazalar va gaplar bilan ishlash tamoyillarini puxta o‘ylab chiqish zarur. Bu yondashuv o‘zbek tilining turli lingvistik xususiyatlari yuzasidan noto‘g‘ri kuzatuv va xulosalarning oldini olishga yordam beradi.

Yuqorida keltirilgan nazariy va uslubiy yondashuvlar korpus lingvistikasining amaliy ahamiyatini yanada aniqroq tushunishga imkon beradi. Ayniqsa, real korpuslar misolida ushbu tamoyillarning qanday ishlashini ko‘rish ularning samaradorligini yaqqol namoyon etadi. Shu nuqtai nazardan, Britaniya milliy korpusi turli ijtimoiy guruhlar, yosh guruhlari, jinslar va turli hududlar vakillari tomonidan ishlatiladigan tillarni ham o‘z ichiga oladi. Lug‘at tuzish, yangi so‘zlar va iboralarni aniqlash, so‘zlarning ishlatilish chastotasini o‘rganish kabi jarayonlarda keng qo‘llanadi. Ushbu jihatlar Britaniya milliy korpusi (BNC)ning tilshunoslik, leksikografiya va boshqa turli ilmiy sohalarda muhim manba sifatida qiymatini ko‘rsatadi. BNCning “Umumiy” va “Tafsilot” (Overwiev, Details) qismida ishlatilgan so‘zlar turli janrlar kesimida ajratilgan. [Yulbarsov O., 2025: 3]

Korpusning foydalilik mohiyati shundan iboratki, u lug‘atlar uchun konkordanslar yaratish, statistik ma’lumotlarni shakllantirish va matn tarkibini tizimli ravishda tuzishda nihoyatda muhim vazifani bajaradi. Ayniqsa, ingliz tilining zamonaviy xususiyat va jihatlarini aniqlash jarayonida konkordans lug‘atini tuzish hamda N-gram modeli asosida so‘zlarni qidirish ishlari aynan shu turdagi korpuslarga tayanadi. Jumladan, Britaniya Milliy Korpusi, Brown korpusi va Corpus of Contemporary American English (COCA) kabi yirik til korpuslari bu borada keng qo‘llanilgan va dolzarb manbalar sifatida xizmat qilgan.

Hozirgi davrda ushbu korpuslar zamonaviy elektron til resurslari sifatida tan olinadi. Xususan, COCA nafaqat so‘z va iboralarning qo‘llanish chastotasini aniqlash, balki ularni chuqur tahlil qilish imkonini ham beradi. Bundan tashqari, u matnlarni turli janrlarga ajratgan holda, so‘zlarning real kontekstda qanday ishlatilishini aniq va tushunarli tarzda ko‘rsatib beradi.

Jahon tilshunosligida korpus matnlariga tayanib lingvostatistik tadqiqot olib borgan yetakchi olimlar va ularning ilmiy hissasi quyidagilardan iborat:

Jorj Kingsli Zipf (*George Kingsley Zipf*): XX asr boshlarida lingvostatistikaning poydevorini qo‘ydi. Uning *“Human Behavior and the Principle of Least Effort”* asarida til

birliklarining chastotaviy taqsimoti matematik qonuniyatlarga va insonning kommunikativ ehtiyojlariga bo‘ysunishi isbotlangan. Bu g‘oyalar kompyuter lingvistikasi va axborot nazariyasida keng qo‘llaniladi. [Zipf G.K., 1949: 4]

Jorj Udni Yul (*George Udny Yule*): *"The Statistical Study of Literary Vocabulary"* asarida statistik tahlilni adabiy matnlarga tatbiq etib, lug‘at boyligi va so‘z takrorlanishini o‘lchash usullarini ishlab chiqdi. [Yule G.U., 2014: 5]

Vilhelm Herdan (*Gustav Herdan*): *"Type-Token Mathematics"* asari orqali matndagi so‘z turlari va ularning takrorlanish nisbatini matematik asosda tahlil qilib, sohani qat’iy ilmiy yo‘nalishga aylantirdi. [Herden W., 1960: 6]

Daglas Biber (*Douglas Biber*): Korpus ma’lumotlari asosida og‘zaki va yozma nutq hamda turli janrlarni solishtirib, "Multi-dimensional analysis" (ko‘p o‘lchovli tahlil) usulini yaratdi. U grammatik va leksik birliklar chastotasini hisoblash orqali matnlarni uslubiy o‘lchamlarga (*dimension*) ajratgan. [Biber D., 1998: 7]

Maykl Stabbs (*Michael Stubbs*): Korpus doirasida kollokatsiya tahlilini olib borib, so‘zlarning birga kelish chastotasini o‘rgandi. Shuningdek, semantic prosody yondashuvi orqali so‘zlarning kontekstdagi yashirin baholovchi (ijobiy yoki salbiy) ma’nolari hamda matn naqshlarini (*pattern*) aniqladi.

R. Harald Baayen (*R. Harald Baayen*): *"Word Frequency Distributions"* asarida korpus ma’lumotlariga tayanib leksik va morfologik jarayonlarning chastota taqsimotini yoritgan. [Baayen R.H., 2001: 8]

Jon Sinkler (*John Sinclair*): *"Prospects of automatic lexicography"* tadqiqotida korpus ma’lumotlari yordamida lug‘at tuzishni avtomatlashtirish, yangi so‘zlarni aniqlash va lug‘atlarni doimiy yangilab borish imkoniyatlarini asoslab bergan. [Sinkler J., 1996: 9]

S. Rzayeva: Yangi so‘zlar tahlilida tarjima matnini asl nusxa bilan solishtirish metodiga tayanadi.

Xulosa. Ushbu tadqiqot tahlillari ko‘rsatadiki, korpus lingvistikasi va lingvostatistika bir-birini to‘ldiruvchi hamda milliy tilshunoslikni raqamli bosqichga olib chiquvchi ajralmas ilmiy sohalardir. Keng hajmli matnlar korpusi asosida olib boriladigan miqdoriy va sifat tahlillari tilning real qo‘llanish holati hamda ichki qonuniyatlarini xolis va empirik dalillar yordamida o‘rganish imkonini beradi. J.K. Zipf, G.U. Yule, W. Herdan, D. Biber, M. Stubbs, R.H. Baayen va J. Sinkler kabi jahon tilshunoslarining ilmiy qarashlari matnlardagi leksik-grammatik taqsimot muayyan matematik va statistik modellarga bo‘ysunishini to‘liq isbotladi.

O‘zbek tili milliy korpusining izchil rivojlantirilishi nafaqat nazariy tilshunoslikni boyitadi, balki tabiiy tilni qayta ishlash (NLP) infratuzilmasining asosi sifatida xizmat qiladi. Korpus ma’lumotlari va lingvostatistik usullarning tatbiq etilishi o‘zbek tili uchun mukammal mashina tarjimasi tizimlarini yaratish, matnlarni avtomatik klassifikatsiya qilish, leksikografik jarayonlarni avtomatlashtirish hamda milliy tilimizning sun’iy intellekt texnologiyalarida munosib o‘rin egallashiga mustahkam poydevor bo‘lib xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR

1. Elov B. va boshqalar. O‘zbek tili korpusi: tarixiy rivojlanishi, maqsad va vazifalari. *"Kompyuter lingvistikasi: muammolar, yechim, istiqbollar"* V Xalqaro ilmiy-amaliy konferensiya. Vol. 1 №. 01 (2025). – B. 168–184.
2. Yuldashev A. Til korpusi matnlari chastotasi, konkordans va KWIC. Xalqaro ilmiy-amaliy konferensiya. 2024. – B. 143–144.

3. Yulbarsov O. O'zbek tili veb korpusi yordamida sohaviy neologizmlar bazasini yaratish. F.f.f.d. (PhD) diss. – B. 27.
4. Zipf G.K. *Human Behavior and the Principle of Least Effort.* – Cambridge, 1949.
5. Yule G.U. *The Statistical Study of Literary Vocabulary.* – Cambridge, 1944 (reprinted 2014).
6. Herdan G. *Type-Token Mathematics.* – The Hague: Mouton & Co., 1960.
7. Biber D. et al. *Corpus Linguistics.* – Cambridge University Press, 1998.
8. Baayen R.H. *Word Frequency Distributions.* – Dordrecht: Kluwer Academic Publishers, 2001.
9. Sinclair J. *Prospects for automatic lexicography.* // Lexicographica. Proceedings of the Seventh International Symposium on Lexicography VII, 1996. – Pp. 1–11.

O‘ZBEK TILIDA OLMOSHLARNING REFERENSIAL XUSUSIYATLARI VA DISKURSDAGI KOMMUNIKATIV VAZIFALARI

Saodat Boysariyeva
TerDU doktoranti
https://doi.org/10.5281/zenodo.23136215

Annotatsiya. *Maqola o‘zbek tilidagi olmoshlarning referensial xususiyatlari va diskursdagi kommunikativ vazifalari tahlil qilinadi. Olmoshlarning matn birliklari o‘rtasidagi semantik bog‘lanishni ta’minlash, diskurs mavzusini davom ettirish va kommunikativ tejamkorlikni yuzaga keltirishdagi roli yoritiladi. Shuningdek, kishilik, ko‘rsatish va o‘zlik olmoshlarining anaforik xususiyatlari hamda o‘zbek tiliga xos pro-drop hodisasining referensni shakllantirishdagi o‘rni ko‘rib chiqiladi. Olmoshlarning referensial tabiati ularni matn va diskurs yaxlitligini ta’minlovchi muhim vosita sifatida baholash imkonini beradi.*

Kalit so‘zlar: *o‘zbek tili, olmosh, referens, anafora, koreferens, diskurs, pro-drop.*

Аннотация. *В тезисе анализируются референциальные свойства местоимений узбекского языка и их коммуникативные функции в дискурсе. Рассматривается роль местоимений в обеспечении смысловой связи между текстовыми единицами, продолжении дискурсивной темы и коммуникативной экономии. Особое внимание уделяется анафорическим свойствам личных, указательных и возвратного местоимений, а также роли характерного для узбекского языка явления pro-drop в формировании референции. Референциальная природа местоимений позволяет рассматривать их как важное средство обеспечения связности текста и дискурса.*

Ключевые слова: *узбекский язык, местоимение, референция, анафора, кореференция, дискурс, pro-drop.*

Abstract. *This thesis analyzes the referential properties of pronouns in the Uzbek language and their communicative functions in discourse. It examines the role of pronouns in establishing semantic links between textual units, maintaining discourse topics, and ensuring communicative economy. Particular attention is paid to the anaphoric properties of personal, demonstrative, and reflexive pronouns, as well as the role of the pro-drop phenomenon characteristic of Uzbek in forming reference. The referential nature of pronouns allows them to be viewed as an important means of maintaining textual and discourse coherence.*

Keywords: *Uzbek language, pronoun, reference, anaphora, coreference, discourse, pro-drop.*

O‘zbek tilshunosligida olmoshlar uzoq yillar davomida asosan grammatik birlik sifatida o‘rganilib kelindi. An’anaviy grammatik tavsiflarda olmoshlar mustaqil ma’noga ega bo‘lmagan, ot, sifat, son kabi so‘zlarning o‘rnida qo‘llanuvchi hamda predmet, shaxs, belgi yoki miqdorga ishora qiluvchi so‘z turkumi sifatida izohlanadi. Bunday yondashuvda olmoshlarning asosiy vazifasi boshqa nominativ birliklarni takror qo‘llashning oldini olish va gap tarkibida muayyan sintaktik vazifani bajarish bilan chegaralanadi. Olmoshlarning grammatik xususiyatlarini o‘rganish ularning til tizimidagi o‘rnini belgilashda muhim ahamiyatga ega. Biroq zamonaviy tilshunoslikda olmoshlarga faqat grammatik birlik sifatida qarash yetarli emasligi

ta'kidlanmoqda. Chunki olmoshlarning haqiqiy mazmuni ko'pincha uning shakliy xususiyatlari bilan emas, balki nutq vaziyati, matn tarkibi va kommunikativ jarayondagi boshqa birliklar bilan munosabati orqali namoyon bo'ladi.

Tilshunoslikda referens nazariyasining rivojlanishi natijasida olmoshlar yangi nuqtayi nazardan tadqiq etila boshlandi. Xususan, olmoshlar muayyan obyekt yoki shaxsning nomini takrorlamaydigan, balki unga ishora qiluvchi, ya'ni referens munosabatini yuzaga keltiruvchi birliklar sifatida talqin qilinadi. Shu jihatdan olmoshlar nafaqat gap darajasidagi grammatik vosita, balki matn va diskurs yaxlitligini ta'minlovchi muhim kommunikativ vosita hisoblanadi.

Halliday va Hasan referensial bog'lanishni matn yaxlitligini ta'minlovchi vositalaridan biri sifatida talqin qiladi. Kishilik va ko'rsatish olmoshlarining matn davomida referensial aloqalarni yuzaga keltirishini qayd etadilar. Ularning fikricha, ayrim til birliklari o'z ma'nosini mustaqil ravishda emas, balki boshqa birliklarga murojaat qilish orqali namoyon qiladi [Halliday M., Hasan R., 1976: 31-87]. **Bu** xususiyat ayniqsa olmoshlarda yaqqol ko'rinadi. Masalan, *Alisher Navoiy o'zbek adabiyotining buyuk namoyandalaridan biridir. U ko'plab ilmiy va badiiy asarlar yaratgan.* **U** olmoshi oddiygina **Alisher Navoiy** o'rnida kelgan grammatik birlik emas. Olmosh oldingi gapdagi referentni qayta faollashtirib, ikki gap o'rtasida semantik bog'lanish hosil qilmoqda. Natijada olmosh matn yaxlitligini ta'minlovchi vositaga aylanadi.

Shuningdek, zamonaviy kognitiv va diskursiv yondashuvlarda olmoshlarning qo'llanishi so'zlovchining tinglovchiga ma'lum bo'lgan axborotni qanday boshqarishi bilan bog'lanadi. Ya'ni olmosh tanlovi nafaqat grammatik qoidalar, balki referentning diskursdagi faolligi, oldindan ma'lumligi va kommunikativ ahamiyati bilan ham belgilanadi. Shu nuqtayi nazardan o'zbek tilidagi olmoshlarning referensial xususiyatlarini o'rganish muhim ilmiy masalalardan biridir. Chunki o'zbek tilining agglutinativ tabiati, grammatik jins kategoriyasining mavjud emasligi, shuningdek, pro-drop hodisasi (eganing tushirilishi) olmoshlarning diskursdagi faoliyatini o'ziga xos tarzda namoyon qiladi.

Olmoshning tilshunoslikdagi talqini: an'anaviy grammatikadan diskursiv yondashuvgacha

Olmoshlar tilshunoslik tarixida eng qadimgi davrlardan boshlab o'ziga xos grammatik kategoriya sifatida o'rganib kelinadi. An'anaviy grammatikalarda olmoshlar, avvalo, *mustaqil atash ma'nosiga ega bo'lmagan, boshqa so'zlarning o'rnida qo'llanuvchi birliklar* sifatida tavsiflangan [Sayfullayeva R., Mengliyev B. 2010: 227-232]. Bu qarashlarga ko'ra olmoshlarning asosiy vazifasi nutqda takroriylikni kamaytirish va gap tarkibida ma'lum sintaktik vazifani bajarishdan iborat deb qaralgan.

O'zbek tilshunosligida ham olmoshlar so'z turkumlari tizimida o'rganilib, ularning grammatik xususiyatlari, ma'no turlari va sintaktik vazifalariga e'tibor qaratilgan. Xususan, olmoshlar shaxs, predmet, belgi, miqdor kabi tushunchalarni bevosita nomlamasdan, ularga ishora qiluvchi birliklar sifatida izohlanadi. Bunda olmoshlarning:

- kishilik (*men, sen, u, biz, siz, ular*);
- o'zlik (*o'z*);
- ko'rsatish (*bu, shu, o'sha*);
- so'roq (*kim, nima, qaysi*);
- belgilash (*hamma, barcha, har kim*);
- bo'lishsizlik (*hech kim, hech nima*);
- gumon (*kimdir, nimadir*)

kabi turlari ajratiladi. Biroq olmoshlarning bunday grammatik tavsifi ularning til tizimidagi shakliy va kategorial xususiyatlarini ochib bersa-da, ularning nutq jarayonidagi haqiqiy kommunikativ imkoniyatlarini to‘liq aks ettirmaydi. Chunki olmoshning asosiy o‘ziga xosligi uning nimani anglatishida emas, balki qanday sharoitda va qaysi referentga murojaat qilishida namoyon bo‘ladi. Hozirda olmoshlarga nisbatan bunday tor grammatik qarash kengaytirilib, ular **referensial birliklar** sifatida tadqiq qilina boshlandi. Referens nazariyasiga ko‘ra til birliklari voqelikdagi obyektlar bilan munosabatga kirishadi va ma’lum referentlarni ko‘rsatadi. Olmoshlar esa ushbu munosabatni o‘ziga xos tarzda amalga oshiradi: ular ko‘pincha obyektni nomlamaydi, balki unga ishora qiladi. Masalan, *Professor yangi maqola yozdi. U maqolada o‘zbek tilining korpus masalalarini yoritdi.* Mazkur misolda **u** olmoshi grammatik jihatdan **professor** o‘rnida kelgan birlik sifatida ko‘rilishi mumkin. Ammo diskurs nuqtayi nazaridan uning vazifasi bundan kengroqdir. U avval kiritilgan referentni saqlab turadi, uni keyingi axborot bilan bog‘laydi va matn davomiyligini ta’minlaydi. Demak, olmosh bu yerda nafaqat almashtirish, balki referensni boshqarish vositasi sifatida faoliyat ko‘rsatmoqda. Bu masalaga zamonaviy matn lingvistikasi vakillari ham alohida e’tibor qaratgan.

Halliday va Hasan referensni matn qismlari o‘rtasidagi semantik bog‘lanishni ta’minlovchi asosiy vositalardan biri sifatida ko‘rsatadi. Ularning fikricha, ayrim birliklarning ma’nosi matn ichidagi boshqa birliklarga bog‘liq holda shakllanadi [Halliday M., Hasan R., 1976: 31-87]. Olmoshlar ana shunday birliklarning eng faol ko‘rinishidir. Masalan: *Yangi ilmiy loyiha boshlandi. Bu loyiha o‘zbek tili korpusini rivojlantirishga qaratilgan.* Bu yerda **bu** ko‘rsatish olmoshi faqatgina **loyiha** so‘zining o‘rnini bosmaydi. U oldingi gapda ifodalangan butun vaziyatni diskursda qayta faollashtiradi. Shu sababli ko‘rsatish olmoshlarining vazifasi oddiy grammatik almashtirishdan ko‘ra kengroq bo‘lib, ular tinglovchi e’tiborini muayyan axborotga yo‘naltiradi.

Lyonsning referens va deiksis haqidagi qarashlarida ham olmoshlarning kontekstga bog‘liqligi asosiy xususiyat sifatida ko‘rsatiladi [Lyons J, 1977:174-229]. Ayniqsa, **men, sen, bu, u** kabi birliklarning ma’nosi nutq vaziyatisiz to‘liq aniqlanmaydi. Masalan, **men** olmoshi doimiy bitta shaxsni bildirmaydi; uning referenti har bir nutq vaziyatida o‘zgaradi. Demak, olmoshlarning ma’nosi lug‘aviy emas, balki kommunikativ jarayonda shakllanuvchi ma’nodir. Olmoshlarni faqat grammatik o‘rinbosar birlik sifatida talqin qilish ularning til tizimidagi barcha imkoniyatlarini qamrab olmaydi. Diskursiv yondashuv esa olmoshlarni:

- referentni ko‘rsatuvchi;
- matn birliklarini bog‘lovchi;
- axborot oqimini boshqaruvchi;
- kommunikativ maqsadni amalga oshiruvchi

muhim til birliklari sifatida baholash imkonini beradi.

O‘zbek tilidagi olmoshlarning referens turlari

Referens hodisasi til birliklarining muayyan obyekt, shaxs, hodisa yoki vaziyatga murojaat qilish munosabatini ifodalaydi. Olmoshlar referensni amalga oshiruvchi eng faol birliklardan biri bo‘lib, ularning ma’nosi ko‘pincha matn yoki nutq vaziyati orqali aniqlanadi. Shu sababli olmoshlarni o‘rganishda ularning faqat grammatik shakli emas, balki **referent bilan munosabati** ham muhim ahamiyat kasb etadi. O‘zbek tilidagi olmoshlarning referens xususiyatlarini bir necha jihatdan ko‘rib chiqish mumkin.

1. Anaforik referens

Anaforik referens olmoshning matnda avval qo‘llangan birlikka murojaat qilishidir. Bunda olmosh o‘zidan oldingi referentni qayta faollashtiradi va matn birliklari o‘rtasida bog‘lanish hosil qiladi.

Masalan, *Abdulla Qodiriy o‘zbek adabiyotining yirik namoyandalaridan biridir. U o‘z asarlarida xalq hayotini chuqur tasvirlagan.* Ushbu misolda **u** kishilik olmoshi oldingi gapdagi **Abdulla Qodiriy** antroponimiga murojaat qilmoqda. Tuzilish jihatidan:

- **Abdulla Qodiriy** –oldingi referens birlik (antecedent);
- **u** – anaforik birlik;
- ikkala birlik – bitta referentga tegishli.

An’anaviy grammatik nuqtayi nazardan bu holat “ot o‘rnida kelgan olmosh” sifatida izohlanishi mumkin. Biroq diskurs tahlili nuqtayi nazaridan **u** olmoshi matndagi mavzuni davom ettirish, axborot oqimini uzluksiz saqlash va referentni tinglovchi xotirasida faol holatda ushlab turish vazifasini bajaradi. O‘zbek tilida anaforik referens ayniqsa **u, ular, uning, ularning** kabi birliklarda faol namoyon bo‘ladi. Masalan: *O‘zbekiston Milliy universiteti yangi ilmiy loyihani amalga oshirmoqda. Uning doirasida o‘zbek tili korpusini rivojlantirish ishlari olib boriladi.* Bu yerda **uning** egalik olmoshi **O‘zbekiston Milliy universiteti** bilan bog‘lanib, keyingi axborotni oldingi referent bilan aloqador qiladi.

2. Ko‘rsatish olmoshlarining referens xususiyati

Ko‘rsatish olmoshlari (**bu, shu, o‘sha, mana bu, ana u**) referens munosabatlarini ifodalashda alohida o‘rin tutadi. Ular ko‘pincha matnda ma’lum obyektga e’tibor qaratish, avvalgi axborotni qayta faollashtirish yoki butun vaziyatga murojaat qilish vazifasini bajaradi. Masalan, *O‘zbek tili uchun elektron korpus yaratildi. Bu loyiha milliy til texnologiyalarini rivojlantirishda muhim ahamiyatga ega.* Bu yerda **bu** faqat **loyiha** so‘ziga emas, balki oldingi gapda ifodalangan **elektron korpus yaratish jarayoni**ga ishora qilmoqda. Shuning uchun ko‘rsatish olmoshlarining referensi ba’zan bitta so‘z bilan emas, balki butun gap yoki diskurs bo‘lagi bilan bog‘lanadi. Bu holat ayniqsa avtomatik koreferens tahlilida murakkablik tug‘diradi. Chunki tizim faqatgina yaqin joylashgan otni emas, balki mazmuniy jihatdan bog‘langan katta birlikni ham aniqlashi kerak bo‘ladi.

3. O‘zlik olmoshining referens xususiyati

O‘zbek tilidagi **o‘z** olmoshi referens jihatdan eng murakkab birliklardan biridir. Chunki uning asosiy xususiyati – gapdagi boshqa bir shaxs bilan bir referentga ega bo‘lishidir. Masalan, *Talaba o‘z maqolasini taqdim etdi.* Bu gapda: **talaba = o‘z** munosabati mavjud. Ya’ni **o‘z** olmoshi mustaqil referent bildirmaydi, balki gapdagi shaxsga qaytadi. Tilshunoslikda bunday hodisa **bog‘langan anafora** (bound anaphora) deb yuritiladi. Biroq murakkab sintaktik tuzilmalarda o‘zlik olmoshining referentini aniqlash qiyinlashadi: *Aziz Jamshidga o‘z kitobini berdi.*

Bu gapda **o‘z**:

- *Azizga;*
- *yoki Jamshidga*

tegishli bo‘lishi mumkin. Referentni aniqlash uchun faqat grammatik shakl emas, balki kontekst va kommunikativ maqsad ham hisobga olinadi. Shu sababli **o‘z** olmoshi o‘zbek tili uchun koreferens modellarini ishlab chiqishda alohida e’tibor talab qiladigan birliklardan biridir.

4. Yashirin referens va pro-drop hodisasi

O‘zbek tilining muhim xususiyatlaridan biri – ega olmoshining ko‘pincha tushirilishidir. Bunday holatda referent ochiq shaklda ifodalanmaydi, balki fe’lning shaxs-son shakli orqali anglashiladi. Masalan, *Men bugun maqolani yozdim* va *Bugun maqolani yozdim.* Ikkinchi gapda **men** olmoshi mavjud emas, ammo fe’l shakli orqali nutq egasi aniqlanadi. Bu hodisa o‘zbek tilida referens tushunchasining faqat mavjud so‘zlar bilan chegaralanmasligini ko‘rsatadi. Ya’ni diskursda ba’zan referent nol shaklda ifodalanadi. NLP tizimlari uchun bu alohida muammo hisoblanadi, chunki koreferensni aniqlash jarayonida nafaqat mavjud olmoshlar, balki yashirin referentlarni ham aniqlash talab qilinadi.

O‘zbek tilidagi olmoshlar referensni turli shakllarda amalga oshiradi. Kishilik olmoshlari asosan anaforik bog‘lanishni, ko‘rsatish olmoshlari diskursga murojaatni, o‘zlik olmoshi esa bog‘langan anaforani ifodalasa, pro-drop hodisasi yashirin referenslarni yuzaga keltiradi.

O‘zbek tilida olmoshlarning diskursdagi kommunikativ vazifalari

Olmoshlarning til tizimidagi o‘rni faqat muayyan so‘zning o‘rnini egallash bilan cheklanmaydi. Diskurs jarayonida ular axborotni tashkil etish, ishtirokchilar o‘rtasidagi munosabatlarni ifodalash, mavzuni davom ettirish va matn yaxlitligini ta’minlash kabi bir qator kommunikativ vazifalarni bajaradi. Shu jihatdan olmoshlar nafaqat grammatik, balki **diskursiv birliklar** sifatida ham namoyon bo‘ladi. Zamonaviy diskurs tahlilida til birliklarining ma’nosi ularning faqat lug‘aviy yoki grammatik xususiyatlari orqali emas, balki kommunikativ vaziyatdagi faoliyati orqali belgilanadi. Olmoshlar aynan shunday birliklardandir: ular o‘z referentini ko‘pincha matn ichidan yoki nutq vaziyatidan oladi.

1. Matn yaxlitligini ta’minlash vositasi sifatida olmoshlar

Matnning asosiy belgilaridan biri uning qismlari o‘rtasidagi bog‘lanishdir. Olmoshlar ana shu bog‘lanishni ta’minlovchi vositalardan biri hisoblanadi. Ayrim til birliklari o‘z ma’nosini boshqa birliklarga murojaat qilish orqali hosil qiladi. Olmoshlar aynan shunday referensial birliklar qatoriga kiradi. Masalan, *Tilshunos olim yangi tadqiqot natijalarini e’lon qildi. U mazkur masala bo‘yicha yangi qarashlarni ilgari surdi.* Agar ikkinchi gapda **tilshunos olim** yana takrorlansa, matnda ortiqcha takroriylik yuzaga keladi. **U** olmoshi esa bir tomondan takrorni kamaytiradi, ikkinchi tomondan esa ikki gap o‘rtasidagi semantik bog‘lanishni saqlaydi. Demak, olmoshning vazifasi oddiy almashtirish emas, balki **matn davomiyligini boshqarish**dir.

2. Diskurs mavzusini davom ettirish vositasi

Har qanday diskursda muayyan obyekt yoki shaxs asosiy mavzu sifatida shakllanadi. Olmoshlar ushbu mavzuni keyingi gaplarda davom ettirish imkonini beradi. Masalan, *O‘zbek tili korpusi zamonaviy tilshunoslik uchun muhim manbadir. U katta hajmdagi lingvistik ma’lumotlarni jamlaydi. Uning imkoniyatlari tadqiqotchilar uchun kengdir.* Ushbu misolda:

- **O‘zbek tili korpusi** – asosiy diskurs mavzusi;
- **u** – mavzuni davom ettiruvchi birlik;
- **uning** – shu referentga qayta murojaat qiluvchi egalik shakli.

Natijada matn bir butun axborot tizimi sifatida shakllanadi. Bu hodisa diskursda **referentning faolligi** bilan bog‘liq. Referent qanchalik faol bo‘lsa, unga murojaat qilish uchun to‘liq nomlash o‘rniga qisqa shakllar, ayniqsa olmoshlar qo‘llanadi.

3. Kommunikativ tejamkorlikni ta’minlash

Nutq jarayonida so‘zlovchi har bir ma’lum axborotni qayta-qayta to‘liq ifodalashga ehtiyoj sezmaydi. Tinglovchi uchun allaqachon ma’lum bo‘lgan referent olmosh yordamida ifodalanishi mumkin. Masalan, *Abu Ali ibn Sino tibbiyot tarixida katta iz qoldirgan buyuk*

olimdir. Abu Ali ibn Sino ko'plab ilmiy asarlar yaratgan. Ushbu shakl grammatik jihatdan to'g'ri bo'lsa-da, kommunikativ jihatdan og'ir.

Tabiiy diskursda: *Abu Ali ibn Sino tibbiyot tarixida katta iz qoldirgan buyuk olimdir. U ko'plab ilmiy asarlar yaratgan* shakli qo'llanadi. Bu holatda **u** olmoshi kommunikativ tejamkorlikni ta'minlaydi. Biroq bu tejamkorlik shunchaki qisqartirish emas, balki **tinglovchining oldingi bilimiga tayanadigan kommunikativ strategiyadir.**

4. Nutq ishtirokchilarini belgilash vazifasi

Kishilik olmoshlari diskursdagi kommunikativ ishtirokchilarni belgilashda alohida ahamiyatga ega. **Men**, **biz** olmoshlari nutq egasini; **sen**, **siz** olmoshlari tinglovchini;

u, **ular** esa nutq jarayonida bevosita ishtirok etmayotgan shaxs yoki predmetlarni ko'rsatadi. Masalan, *Biz o'zbek tili korpusini rivojlantirish bo'yicha yangi loyihani boshladik.* Bu gapdagi **biz** faqat ko'plik shaklidagi shaxsni bildirmaydi. U ma'lum kommunikativ guruhni – so'zlovchi va u mansub bo'lgan jamoani ifodalaydi. Shu sababli birinchi shaxs ko'plik olmoshi (**biz**) ba'zan diskursda ijtimoiy yaqinlik, hamkorlik yoki umumiylik ma'nolarini ham yuzaga keltiradi.

5. Axborotning ma'lum va yangi qismlarini boshqarish

Diskursda barcha axborotlar bir xil darajada yangi yoki ma'lum bo'lmaydi. Olmoshlar ko'pincha tinglovchi uchun allaqachon tanish bo'lgan referentlarni ifodalash orqali yangi axborotga e'tibor qaratishga imkon beradi. Masalan, *Yangi algoritm ishlab chiqildi. U o'zbek tilidagi matnlarda koreferensni aniqlash imkonini beradi.* Birinchi gapda **yangi algoritm** kiritiladi. Ikkinchi gapda esa **u** orqali ayni referent saqlanib, asosiy e'tibor algoritmning xususiyatiga qaratiladi. Demak, olmoshlar diskursda nafaqat bog'lovchi vosita, balki **axborot tuzilishini boshqaruvchi mexanizm** sifatida ham xizmat qiladi.

O'zbek tilidagi olmoshlarni referensial va diskursiv nuqtayi nazardan o'rganish ularning an'anaviy grammatik tavsifdan ko'ra kengroq funksional imkoniyatlarga ega ekanligini ko'rsatadi. Olmoshlar nafaqat boshqa nominativ birliklarning takrorlanishini kamaytiradi, balki matnda avval kiritilgan referentni qayta faollashtiradi, diskurs mavzusini davom ettiradi va matn birliklari o'rtasidagi semantik aloqani ta'minlaydi. Tahlil natijasida o'zbek tilidagi olmoshlarning referensial xususiyatlari bir necha ko'rinishda namoyon bo'lishi aniqlandi. Kishilik olmoshlari, xususan, *u, ular, uning, ularning* shakllari anaforik bog'lanishni yuzaga keltirsa, ko'rsatish olmoshlari (*bu, shu, o'sha*) ayrim hollarda alohida otga emas, balki butun gap yoki vaziyatga murojaat qilishi mumkin. O'zlik olmoshi (*o'z*) esa bog'langan anafora bilan bog'liq bo'lib, uning referentini aniqlashda sintaktik munosabatlar bilan bir qatorda kontekst va kommunikativ omillar ham muhim ahamiyat kasb etadi. Shuningdek, o'zbek tilidagi pro-drop hodisasi referensning har doim ham ochiq shaklda ifodalanmasligini ko'rsatadi. Eganing tushirilishi natijasida referent fe'lning shaxs-son shakli va kontekst asosida tiklanadi. Mazkur xususiyat o'zbek tilidagi koreferensni avtomatik aniqlashda nafaqat ifodalangan olmoshlarni, balki yashirin referentlarni ham hisobga olish zarurligini ko'rsatadi.

Demak, o'zbek tilida olmoshlarni tadqiq etishda ularni faqat so'z turkumi yoki grammatik o'rinbosar birlik sifatida emas, balki referensni boshqaruvchi, diskurs mavzusini davom ettiruvchi va matn yaxlitligini ta'minlovchi diskursiv birlik sifatida o'rganish muhimdir. Bunday yondashuv o'zbek tili korpusi uchun koreferens annotatsiyasi va avtomatik koreferens tahlili modellarini ishlab chiqish uchun ham nazariy asos bo'lib xizmat qilishi mumkin.

FOYDALANILGAN ADABIYOTLAR

1. Halliday M., Hasan R. *Cohesion in English*. London: Longman. 1976. – 374 p.
2. Lyons J. *Semantics. Volume 2*. Cambridge: Cambridge University Press. 1977. – 897 p.
3. Sayfullayeva R. R., Mengliyev B va boshq. *Hozirgi o'zbek adabiy tili*. Toshkent: Fan va texnologiya. 2010. – 404 b.

ЛИНГВИСТИЧЕСКИЕ ИНДИКАТОРЫ ДЛЯ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ И ПОИСКА АФОРИЗМОВ

Джуманиязова Интизор Давлатназар кизи
Старший преподаватель
университет Маъмун, Хива
https://doi.org/10.5281/zenodo.23136261

Аннотация. *В статье рассматривается проблема автоматического обнаружения, классификации и индексирования афоризмов в текстовых корпусах. Задача анализируется на основе моделей обнаружения фрагментов текста, классификации предложений и ранжирования в рамках НЛП и интерпретируется как сложная задача типа «needle-in-the-haystack». Она основана на том, что для обнаружения афоризмов следует использовать комбинацию графико-пунктуационных, количественных, лексико-морфологических, синтаксических, семантических, прагматических и стилистических индикаторов; процесс представлен в виде четырехэтапного конвейера, состоящего из поиска кандидатов, оценки, разметки и ранжирования. Также рассматривается необходимая для узбекского языка инфраструктура предварительной обработки (нормализация, токенизация, лемматизация, POS-тегирование, синтаксический анализ зависимостей) на основе существующих лингвистических ресурсов (BBPOS, узбекский UD-корпус).*

Ключевые слова: *Выявление афоризмов, span detection, sentence classification, ranking, quotability, узбекский язык, морфологический анализ, POS-тегирование, dependency parsing, sentence embedding, BBPOS, узбекский UD treebank, текстовый корпус, поисковая система, стилистико-риторические индикаторы, семантико-тематический анализ.*

Annotatsiya. *Maqolada matn korpuslarida aforizmlarni avtomatik aniqlash, tasniflash va qidiruv tizimiga indekslash masalasi ko‘rib chiqiladi. Vazifa NLPning span detection, sentence classification va ranking modellari asosida tahlil qilinib, "needle-in-the-haystack" tipidagi murakkab muammo sifatida talqin etiladi. Aforizmni aniqlash uchun grafik-punktuatsion, kvantitativ, leksik-morfologik, sintaktik, semantik, pragmatik va stilistik indikatorlar majmuiy qo‘llanilishi kerakligi asoslanadi; jarayon kandidat topish, baholash, teglash va reytinglashdan iborat to‘rt bosqichli pipeline sifatida taqdim etiladi. Shuningdek, o‘zbek tili uchun zarur bo‘lgan oldindan ishlov berish infratuzilmasi (normalizatsiya, tokenizatsiya, lemmatizatsiya, POS-teglash, dependency parsing) mavjud lingvistik resurslar (BBPOS, Uzbek UD treebank) asosida yoritiladi.*

Kalit so‘zlar: *aforizmni aniqlash, span detection, sentence classification, ranking, quotability, o‘zbek tili, morfologik tahlil, POS-tegirlash, dependency parsing, sentence embedding, BBPOS, Uzbek UD treebank, matn korpusi, qidiruv tizimi, stilistik-ritorik indikatorlar, semantik-tematik tahlil.*

Abstract. *The article considers the issue of automatic detection, classification and indexing of aphorisms in text corpora. The task is analyzed based on the NLP span detection, sentence classification and ranking models and interpreted as a complex "needle-in-the-*

haystack" type problem. It is based on the fact that a combination of graphic-punctuation, quantitative, lexical-morphological, syntactic, semantic, pragmatic and stylistic indicators should be used to detect aphorisms; the process is presented as a four-stage pipeline consisting of candidate finding, evaluation, tagging and ranking. Also, the pre-processing infrastructure necessary for the Uzbek language (normalization, tokenization, lemmatization, POS-tagging, dependency parsing) is covered based on existing linguistic resources (BBPOS, Uzbek UD treebank).

Keywords: *aphorism detection, span detection, sentence classification, ranking, quotability, Uzbek language, morphological analysis, POS-tagging, dependency parsing, sentence embedding, BBPOS, Uzbek UD treebank, text corpus, search engine, stylistic-rhetorical indicators, semantic-thematic analysis.*

Афоризм — одна из единиц, принадлежащих к компактному, но насыщенному содержанием слою языка. Он выражает обобщенный опыт, оценку, предостережение или моральную норму в краткой форме. Поэтому изучение афоризма не только как художественно-эстетического феномена, но и как прагматической, семантической и дискурсивно сложной языковой структуры, а также его автоматическое выделение в цифровом тексте стало одной из актуальных проблем современной вычислительной лингвистики.

Она включает в себя автоматическую или полуавтоматическую идентификацию, фильтрацию, классификацию и индексирование афоризмов в корпусе в формате, подходящем для поисковых систем. В современной практике обработки естественного языка значимые, но численно редкие единицы моделируются несколькими способами: span detection (определение границ), sentence classification (разделение предложений на теги) и ranking (выделение вероятных единиц). В частности, работы по quote detection рассматривали обнаружение цитаты как задачу span detection, тогда как quotability identification моделировала «вероятность превращения фрагмента текста в цитату» как задачу ранжирования. Исследования по quotable phrase extraction показали, что основными проблемами таких задач являются неопределённость границ, уникальность положительных примеров и синтаксико-семантическая сложность. Обнаружение афоризма объединяет в себе именно эти три трудности.

В этом смысле задача автоматического обнаружения афоризма представляет собой не чисто филологическое описание, а поиск уникальной, обладающей высокой семантической плотностью, стилистически маркированной и относительно независимой от контекста единицы на основе статистических и формальных признаков. Подобные задачи в NLP часто предстают как проблема типа «needle-in-the-haystack» («игла в стоге сена»), когда из множества похожих на правило кандидатов необходимо выделить именно коммуникативно и стилистически «ядерную» единицу. Риторика формирует структуру языка, а структура языка проявляется в риторике. Найти поверхностный паттерн легко, однако выделить случай, обладающий подлинной риторической функцией, значительно сложнее, и сочетание вручную построенных фильтров с машинным обучением считается одним из наиболее эффективных путей.

В данном разделе подчёркивается, что процесс обнаружения афоризмов не может ограничиваться единственным признаком. Напротив, совместно должны действовать графические, количественные, лексико-морфологические, синтаксические, семантические,

прагматические, стилистические показатели и метаданные. Афоризм часто краток, однако не всякое краткое предложение является афоризмом. В нём присутствует обобщающее суждение, но не всякое обобщение обладает афористической силой; в нём может присутствовать образность, однако не всякое образное предложение самостоятельно вне контекста. Модель автоматического обнаружения должна опираться не на «один сильный признак», а на совокупность признаков.

Автоматическое обнаружение афоризма распадается как минимум на четыре вычислительные задачи:

- нахождение кандидатов;
- оценка кандидата как афоризма или неафоризма;
- тегирование типа и атрибутов афоризма;
- корректное размещение и ранжирование в поисковой системе.

На практике эти этапы не изолированы друг от друга, а образуют последовательно работающий pipeline.

В исследованиях quote detection показано, что эффективными подходами являются обнаружение цитаты в виде span detection, а в исследованиях quotability — рассмотрение фрагмента текста как объекта ранжирования. Для афоризма также полезно объединить эти две модели: нахождение спана афоризма в полном тексте — задача детекции, а выдача из корпуса наиболее подходящего для пользователя афоризма — задача поиска и ранжирования.

Особенность афоризма заключается в том, что в практической корпусной системе большая его часть проявляется на уровне предложения. Однако отдельные афоризмы могут представлять собой двухпредложенческую или сложно-композиционную конструкцию. Поэтому для автоматической системы предлагается следующий рабочий принцип:

- *первичная единица детекции — предложение,*
- *вторичная единица детекции — пара соседних предложений или короткий спан.*

Такой подход учитывает и проблему неопределённости границ «quotable phrase», отмеченную Бендерски и соавторами. В отдельных случаях единица может быть фрагментом, иногда — полным предложением, а иногда — состоять из нескольких предложений.

Здесь необходимо чётко разграничить классификацию и ранжирование. В классификации модель отвечает на вопрос «Является ли это афоризмом или нет?». В ранжировании же отвечают на вопрос «Какие единицы в данном документе или во всём корпусе больше похожи на афоризм?». Ранжирование очень полезно для таких стилистико-семантических единиц, как афоризм, поскольку аннотатор или исследователь чаще всего хочет увидеть «наиболее сильных кандидатов». Исследования quotability, опираясь именно на метод passage ranking, тестировали модели на основе BERT и получили результаты по показателям nDCG и rank correlation.

Обнаружение афоризма тесно связано с качеством языково-специфичной предобработки. Для узбекского языка этот вопрос важен, поскольку язык, с одной стороны, агглютинативен и морфологически богат, а с другой — существует в среде параллельного использования латинской и кириллической графики. В исследовании BBPOS отмечено, что узбекский язык морфологически богат и является сильно агглютинативным, и хотя официальным письмом является латиница, кириллическое

письмо по историческим причинам по-прежнему активно используется. В этой работе POS-теггеры на основе BERT достигли средней точности 91%, превзойдя многоязычный BERT и теггер на основе правил. Наряду с этим в 2025 году был представлен первый узбекский UD treebank, в котором 500 предложений и 5850 токенов размечены по лемме, POS, морфологическим признакам и синтаксическим зависимостям.

Следовательно, для системы автоматического обнаружения афоризмов реальной и целесообразной считается следующая базовая инфраструктура:

- *нормализация латиницы/кириллицы;*
- *сегментация предложений;*
- *токенизация;*
- *лемматизация;*
- *UPOS/XPOS тегирование;*
- *выделение морфологических FEATS;*
- *dependency parsing;*
- *формирование sentence embedding.*

Первые пять из них необходимы для извлечения индикаторов, а последние два — для сложного синтаксического и семантического поиска.

Хотя афоризмы зачастую внешне выглядят как «обычное предложение», отличает их от прочих чаще всего портативность выражения, независимость от контекста, lexical distinctiveness и обобщение. Классические исследования memorable language показали, что запоминающиеся цитаты чаще используют необычный лексический выбор, но при этом строятся на сравнительно обычном синтаксическом каркасе; они также менее привязаны к конкретной ситуации, то есть более «free-standing» и пригодны для переноса в другие контексты. Эти выводы крайне важны для афоризма: при автоматическом обнаружении «портативный» и «многократно используемый» характер идеи может рассматриваться как отдельный признак.

Предлагаемая система индикаторов для обнаружения афоризма делится на семь основных групп:

- *графико-пунктуационные;*
- *количественно-композиционные;*
- *лексико-морфологические;*
- *синтаксико-конфигуративные;*
- *семантико-тематические;*
- *прагматико-дискурсивные;*
- *стилистико-риторические индикаторы.*

К ним добавляются отдельные индикаторы источника и метаданных. Именно эта многослойность помогает отличить афоризм от лозунга, заголовка, цитаты, определения или пословицы.

Наиболее яркой особенностью афоризма является его краткость. Однако краткость нельзя измерить абсолютной длиной. Например, заголовок из 3–4 слов короткий, но не всегда является афоризмом, в то время как сложное предложение из 20–25 слов иногда может быть явно афористичным. Поэтому уместно рассматривать длину как оптимальный диапазон, а не как «жесткий предел». В "Quotable phrase extraction" также было обнаружено, что слишком длинные варианты имеют меньшую запоминаемость, а слишком короткие — не обладают семантической полнотой.

В рамках данного исследования предлагаются следующие рабочие критерии:

1. Диапазон от 4 до 25 токенов является основной зоной, подходящей для афоризма;
2. 1 предложение — наиболее типичный случай;
3. 2 предложения — ограниченный, но допустимый случай;
4. 3 или более предложений — только в особых контекстах.

Кроме того, для формальной оценки предложения полезны такие показатели, как clause_count, punctuation_density, average_dependency_depth. Обычно синтаксическая глубина афоризма невелика, он, как правило, строится по модели «одно центральное утверждение + одно противопоставление или сравнение».

Существует три основных признака, позволяющих отличить афоризм в лексическом слое:

1. абстрактный тематический лексикон;
2. оценочный или нормативный лексикон;
3. лексическая отличительность.

Афоризмы часто формируются вокруг абстрактных семантических полей, таких как «жизнь», «время», «язык», «знание», «истина», «человек», «память», «совесть», «труд», «свобода». Такой лексикон увеличивает потенциал для обобщения. Кроме того, исследования запоминающегося языка показали, что запоминающиеся и часто повторяющиеся фразы часто имеют необычный выбор слов, но их синтаксическая структура более привычна. Этот вывод также согласуется с афоризмами. Модель может использовать редкие лексические сочетания, с одной стороны, и обычную структуру частей речи, с другой стороны, в качестве сигналов.

В заключение, предложенная в статье система многоуровневых индикаторов позволяет отличать афоризм от слогана, заголовка, цитаты, определения или пословицы и может служить теоретической и методологической основой для создания автоматических систем обнаружения и поиска афоризмов в узбекском языке. Важным направлением будущих исследований станет тестирование этих индикаторов на практических корпусных данных и эмпирическая оценка их эффективности с использованием моделей машинного обучения.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ:

1. Джуманиязова, И.Д. (2024). Прагматико-семантическая классификация афоризмов в корпусе узбекского языка. Международный журнал искусства слова, 7(2), 45–58.
2. MacLaughlin, A., & Smith, D. (2021). Quotation Recommendation and Interpretation with Structured Retrieval Models. Proceedings of the 2021 Conference on EMNLP.
3. Mamasaidov, M., & Shopulatov, A. (2023). TahrirchiBERT base. Hugging Face Model Hub.
4. Mansurov, B., & Mansurov, A. (2021). UzBERT: Pretraining a BERT Model for Uzbek. CoRR, abs/2108.09814.
5. Sharipov, M., Kuriyozov, E., Yuldashov, O., & Sobirov, O. (2023). UzbekTagger: The Rule-Based POS Tagger for Uzbek Language. Proceedings of SIGUL 2022.
6. Tekir, S., Güzel, A., Tenekeci, S., & Haman, B. (2023). Quote Detection: A New Task and Dataset for NLP. Proceedings of the 7th Joint SIGHUM Workshop, Dubrovnik, Croatia.

K-MEANS KLASTERLASH ALGORITMIGA NEYRON TARMOQ ARXITEKTURASINI QO'LLASH

Ergashev Sirojiddin
Muassasa: Abdulla Qodiriy nomidagi Jizzax davlat pedagogika universiteti
Mail: s.b.ergashev@gmail.com
https://doi.org/10.5281/zenodo.23136270

Annotatsiya. *Tezisda nazoratsiz o'qitish usuliga kiruvchi K-means klasterlash algoritmini neyron tarmoq modeli sifatida qarash mumkinligiga asoslanadi. Ishning asosiy mazmunida K-means algoritmining sentroidlarni tayinlash mexanizmi hamda SOM (Self-Organizing Map) raqobatlashuvchi neyron qatlamining ekvivalentlik nazariyasini ko'rib chiqish, algoritmning formal matematik ta'rifi va uni tasvir ma'lumotlarida sinash uchun mo'ljallangan tadqiqot rejasi bayon qilinadi.*

Kalit so'zlar: *Klasterlash, xatolik funksiyasi, sun'iy intellekt, K-means, SOM, neyron tarmoq, faollashtirish funksiyasi, bias, chuqur o'qitish, vizualizatsiya, simulyator.*

Аннотация. *Тезис основан на том, что алгоритм кластеризации K-means, относящийся к методам обучения без учителя, может рассматриваться как модель нейронной сети. В основном содержании работы излагается рассмотрение теории эквивалентности между механизмом назначения центроидов алгоритма K-means и соревнующимся нейронным слоем SOM (Self-Organizing Map), формальное математическое определение алгоритма, а также план исследований, предназначенный для его тестирования на графических данных.*

Ключевые слова: *Кластеризация, функция потерь, искусственный интеллект, K-means, SOM, нейронная сеть, функция активации, баис, глубокое обучение, визуализация, симулятор.*

Abstract. *The abstract is based on the idea that the K-means clustering algorithm, an unsupervised learning method, can be viewed as a neural network model. The main content of the work outlines the consideration of the equivalence theory between the centroid assignment mechanism of the K-means algorithm and the competitive neural layer of SOM (Self-Organizing Map), the formal mathematical definition of the algorithm, and a research plan designed to test it on image data.*

Keywords: *Clustering, loss function (cost function), artificial intelligence, K-means, SOM, neural network, activation function, bias, deep learning, visualization, simulator.*

Kirish

Raqamli tasvir va video shaklidagi ma'lumotlar hajmining tez sur'atlarda o'sishi, ularni avtomatik tarzda mazmunli qismlarga ajratish vositalariga bo'lgan ehtiyojni kuchaytirmoqda. Ushbu vazifani hal qiluvchi klassik yechimlardan biri, belgi bo'yicha o'xshash obyektlarni guruhlarga birlashtiruvchi klasterlash bo'lib, uning eng ko'p tarqalgan ko'rinishi K-means hisoblanadi. Shu bilan birga, mavjud adabiyotda K-means bilan neyron tarmoqlar orasidagi bog'liqlik ko'pincha yuzaki tilga olinadi, aniq matematik moslik ko'rsatilmaydi. Ikkinchi tomondan, algoritmning tasvirlarni tahlil qilishdagi amaliy samaradorligi turli sinov shartlarida farqli natija berishi mumkin, bu esa aniq eksperimental tekshiruvni talab qiladi. Ushbu ikki

bo‘shliq nazariy moslikning to‘liq ko‘rsatilmaganligi va amaliy sinovning yetarli emasligi mazkur tadqiqotning boshlanish nuqtasini tashkil etadi.

Olib borilayotgan tadqiqotning asosiy maqsadi quyidagilardan iborat: K-means algoritmini raqobatlashuvchi neyron tarmoq sifatida formal tarzda ta’riflash va shu asosda real tasvir ustida o‘tkaziladigan eksperiment uchun aniq metodika ishlab chiqish.

Algoritmning formal ta’rifi

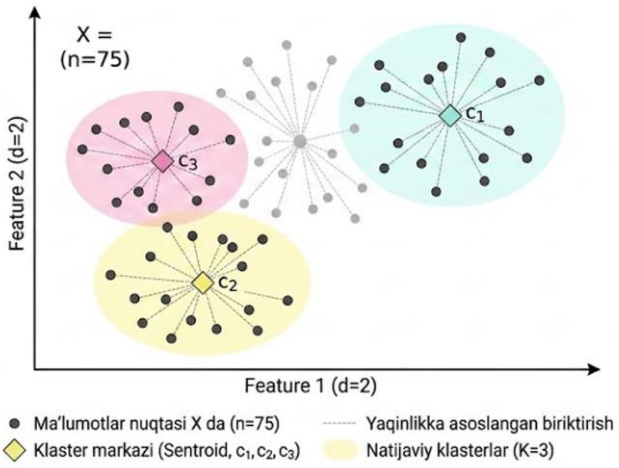
Faraz qilaylik, $\boldsymbol{d}$ o‘lchamli fazoda joylashgan $\boldsymbol{n}$ ta obyektdan iborat $\boldsymbol{X}$ to‘plam berilgan. Masala bu obyektlarni $\boldsymbol{K}$ ta guruhga shunday taqsimlashdan iboratki, har bir guruh ichidagi obyektlar o‘z markaziga (sentroidiga) nisbatan iloji boricha yaqin joylashsin (1-rasm). Bu talab bo‘yicha xatolik funksiyasining minimallashtirilishi quyidagicha bo‘ladi:

$$J(C,\mu) = \Sigma_{j=1}^{K} \Sigma_i \in C_j ||x_i - \mu_j||^2 \rightarrow min \quad (1)$$

Funksiyani minimallashtirish ikkita takrorlanuvchi qadam orqali amalga oshiriladi. Birinchi qadamda har bir $\boldsymbol{x}_i$ obyekt o‘ziga eng yaqin joylashgan $\boldsymbol{c}_i$ markazga biriktiriladi (1-rasm):

$$c_i = arg\ min_j ||x_i - \mu_j||^2 \quad (2)$$

Ikkinchi qadamda esa har bir $\boldsymbol{c}_i$ markaz o‘ziga tegishli $\boldsymbol{n}$ ta $\boldsymbol{x}_i$ obyekt joylashuvining umumiy o‘rtacha qiymati sifatida yangilanadi [1]. Asosiy muhim xususiyat shundaki, ikkala qadam ham xatolik funksiyasining qiymatini kamaytiradi yoki o‘zgarishsiz qoldiradi. Biroq bu faqat kichik hajmli ma’lumotlarda qiymat minimumga yetib borishini kafolatlaydi, ammo katta hajmli ma’lumotlarda har doim ham optimal yechimni bermaydi. Agar dastlabki tasodifiy $\boldsymbol{c}_i$ markazlar bir-biriga yaqin joylashib qolsa, algoritm samarasiz yechimda «qotib qolishi» mumkin. Bu $\boldsymbol{c}_i$ markaz qiymatlarini qayta hisoblashni talab qiladi.



1-rasm. $\boldsymbol{d}$ o‘lchamli fazoda joylashgan $\boldsymbol{n}$=75 ta obyektdan iborat $\boldsymbol{X}$=3 to‘plam namunasi.

Ushbu kamchilikni bartaraf etish uchun tadqiqotda (K-means++) metodikasidan foydalanish kerak. Har bir yangi **c** markaz mavjud markazlardan qanchalik uzoq bo‘lsa, aniqligi shuncha yuqori bo‘ladi:

$$P(x_i) \propto D(x_i)^2 \quad (3)$$

bu yerda $P(x_i)$ iborasi - $\boldsymbol{i}$ obyektning navbatdagi klaster markazi (sentroid) sifatida tanlanish ehtimolligini anglatadi. $D(x_i)$ - $\boldsymbol{i}$ obyektning eng yaqin $\boldsymbol{c}_i$ markazgacha bo‘lgan oraliq masofa.

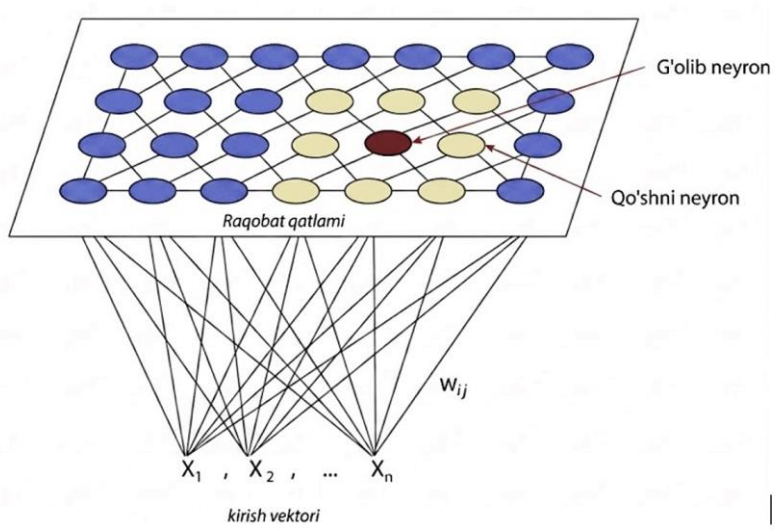
Ma'lumotlarni hisoblash nuqtai nazaridan algoritm chiziqli murakkablikka ega - vaqt bo'yicha $O(nKdt)$ va xotira bo'yicha $O(nd + Kd)$ lekin bu uni katta o'lchamli tasvir ma'lumotlariga qo'llashda samarali qiladi [2].

Neyron tarmoq sifatidagi talqini

Algoritmning arxitekturaviy qismini Kohonenning o'z-o'zini tashkillashtiruvchi xarita nazariyasida ko'rish mumkin. Bunda har bir $\boldsymbol{\mu}_j$ sentroid alohida neyron rolini bajaradi: kirish signali (obyekt vektori) barcha neyronlarga bir vaqtda ma'lumot uzatiladi, so'ng ular o'rtasida «raqobat» boshlanadi va ma'lumotlarga eng yaqin turgan yagona neyron «g'olib» deb e'lon qilinadi (2-rasm). Faqat shu g'olib neyronning og'irliklari yangilanadi:

$$\mu_j \leftarrow \mu_j + \eta(x_i - \mu_j) \quad (4)$$

bu yerda η - o'zgarish tezligini boshqaruvchi parametr. Ushbu mexanizm klassik SOM tarmog'idan bitta muhim jihati bilan farqlanadi: SOM da g'olib neyron bilan birga uning topologik qo'shnilari ham yangilanadi, K-means da esa faqat g'olib neyroning o'zi o'zgaradi. Shu bois K-means ni SOM nazariyasining qo'shnilik aloqasi olib tashlangan versiyasi sifatida qarash mumkin, bu talqin algoritmni chuqur o'qitish tizimlariga integratsiya qilish uchun nazariy ko'prik vazifasini bajaradi [3].



2-rasm. Kohonen o'z-o'zini tashkillashtiruvchi xarita arxitekturasi.

Tasvir tahliliga qo'llash rejasi

Ishlab chiqilgan nazariy modelni tekshirish uchun keyingi bosqichda quyidagi tadqiqot rejasi amalga oshirilishi ko'zda tutilgan. Kirish ma'lumoti sifatida rangli tasvir tanlanadi. Undagi har bir piksel x = (r, g, b) vektori ko'rinishida tasvirlanib, piksellar to'plami klasterlash uchun standart X to'plamiga aylantiriladi. Klasterlash sklearn dasturiy kutubxonasi yordamida amalga oshiriladi. Olingan natijalar ko'p qatlamli perseptron (MLP) orqali klaster chegaralarini baholashda qo'llanadi hamda neyron tarmog'ining ushbu chegaralarni qanchalik aniq baholashi tekshiriladi[4, 5].

Xulosa

Tezisda K-means klasterlash algoritmi raqobatlashuvchi neyron tarmoq sifatida ko'rib chiqildi. Kohonen nazariyasi bilan bog'lanish algoritmning arxitekturaviy mohiyatini ochib berdi, shu bilan birga uni real ma'lumotlarda sinash uchun aniq metodika shakllantirildi. Ishlab chiqilgan reja asosida olingan tadqiqotning amaliy natijalari maqolaning navbatdagi bosqichida taqdim etiladi.

FOYDALANILGAN ADABIYOTLAR

1. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1(14), 281–297.
2. Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, 78(9), 1464–1480. https://doi.org/10.1109/5.58325.
3. Ergashev S.B., & Yusupov R.M. (2023). Tyuring mashinasi va sun'iy neyron to'rlari. Илм-фан ва инновацион ривожланиш / Наука и инновационное развитие, 6(6), 27-38.
4. Ergashev S.B. (2023). Simulyatorlarda o'qitish imkoniyatlarining samaradorligini baholash. O'ZBEKISTON MILLIY UNIVERSITETI XABARLARI, 1(2181-7324), 221-223.
5. Ergashev S.B., Yusupov R.M., & Qarshiyev A.A. (2024). SUN'IY INTELLEKT FANINI O 'QITISHDA SIMULYATORNI QO'LLASH. SAMBHRAM XABARNOMASI, 1(1), 124-126.

FRAZEOLOGIK BIRLIKLARNI KORPUS UCHUN TEGLASH MASALASIDA ZAMONAVIY YONDASHUVLAR VA TADQIQOTLAR SHARHI

G'anijonova Shaxlo Dilmurod qizi
Kompyuter lingvistikasi magistratura mutaxassisligi 1- kurs magistranti
O'zbekiston Milliy universiteti, Toshkent, O'zbekiston.
https://doi.org/10.5281/zenodo.23136293

Annotatsiya. *Mazkur maqola o'zbek tili milliy korpusida frazeologik birliklarni teglash va ularni avtomatik lingvistik annotatsiya qilishning nazariy-metodik asoslarini ishlab chiqishga bag'ishlanadi. Frazeologik birliklarning semantik jihatdan yaxlitligi, shuningdek, tarkibining o'zgaruvchanligi sabab NLP tizimlarida oddiy so'z birikmalaridan ularni ajratish biroz murakkab. Tadqiqotda tizimli lisoniy tahlil, korpus metodlari va zamonaviy neyron tarmoq modellari imkoniyatlarini qiyoslash metodologiyasidan foydalanish orqali frazeologizmlarning elektron bazasini shakllantirishda xalqaro tajribalar va ularning o'zbek tili korpusiga tatbiq etilish imkoniyatlari ko'rib chiqiladi.*

Kalit so'zlar: *O'zbek tili milliy korpusi, frazeologik birliklar, tabiiy tilni qayta ishlash, lingvistik annotatsiya, avtomatik teglash, ko'p komponentli birikmalar, semantik yaxlitlik.*

Abstract. *This article is devoted to developing the theoretical and methodological foundations for tagging phraseological units in the Uzbek national language corpus and their automatic linguistic annotation. Due to the semantic integrity of phraseological units, as well as the variability of their composition, distinguishing them from ordinary word combinations in NLP systems presents considerable complexity. The study examines international practices in the development of electronic databases of phraseological units through the application of systematic linguistic analysis, corpus-based methods, and a comparative methodological framework for evaluating the capabilities of modern neural network models, as well as the possibilities of their implementation within the Uzbek language corpus.*

Keywords: *Uzbek national language corpus, phraseological units, natural language processing (NLP), linguistic annotation, automatic tagging, multi-word expressions (MWE), semantic integrity.*

Аннотация. *Данная статья посвящена разработке теоретико-методических основ тегирования фразеологических единиц в национальном корпусе узбекского языка и их автоматической лингвистической аннотации. Семантическая целостность фразеологических единиц, а также вариативность их состава обусловливают значительную сложность их разграничения с обычными словосочетаниями в системах обработки естественного языка (NLP). В исследовании рассматриваются международные практики формирования электронных баз данных фразеологических единиц посредством применения системного лингвистического анализа, корпусных методов и сравнительной методологии оценки возможностей современных моделей нейронных сетей, а также возможности их внедрения в корпус узбекского языка.*

Ключевые слова: *национальный корпус узбекского языка, фразеологические единицы, обработка естественного языка, лингвистическая аннотация, автоматическое тегирование, многословные выражения, семантическая целостность.*

Kirish.

Frazeologik birliklar tilning semantik va madaniy jihatdan eng murakkab qatlamlaridan biri hisoblanadi va ular o‘zining ko‘chma ma’nosi va barqaror tarkibi bilan boshqa leksik birliklardan ajralib turadi. Frazeologizmlar tilning “oltin fondi” hisoblanib, ularning ma’nosi ko‘pincha komponentlarning oddiy yig‘indisidan kelib chiqmaydi. Shuning uchun frazeologik birliklarni korpus lingvistikasi va NLP tizimlarida avtomatik aniqlash ham qiziqarli, ham murakkab vazifalardan biri hisoblanadi. O‘zbek tili frazeologiyasini raqamlashtirish, xususan, frazeologik birliklarni milliy korpusda avtomatik identifikatsiyalash masalasi hozirda yechimini kutayotgan dolzarb ilmiy muammolardandir.

Shavkat Rahmatullayevning tadqiqotlarida ta’kidlanganidek, frazeologik birliklar nutqimizning ko‘rki bo‘lish bilan birga, o‘zining barqarorligi va yaxlitligi bilan ajralib turadi. Biroq bu xususiyatlar raqamli muhitda algoritmlar uchun jiddiy qiyinchiliklarni tug‘dirishi mumkin. An’anaviy tilshunoslikda iboralarning semantik va uslubiy xususiyatlari atroflicha o‘rganilgan bo‘lsa ham, milliy korpusda avtomatik teglash va annotatsiya qilish uchun yangi texnologik yondashuvlar talab etiladi. So‘nggi yillarda ushbu muammo NLP tadqiqotlarida alohida e’tibor qaratilayotgan yo‘nalishlardan biriga aylandi.

Jahon miqyosida ko‘p so‘zli birikmalar (Multiword Expressions — MWE) va idiomalarni identifikatsiya qilish uchun zamonaviy transformator modellaridan foydalanish sezilarli darajada kengaydi. Jurafsky va Martin ta’kidlaganidek, til modellarining kontekstual semantikani anglash qobiliyati frazeologizmlarni matn tarkibida aniqlashda muhim o‘rin tutadi. O‘zbek tili agglyutinativ til sifatida o‘ziga xos murakkab morfologik tarkibga ega bo‘lib, frazeologik birliklarning tarkibiy qismlari turli grammatik shakllarda keladi yoki gap ichida bir-biridan uzoqda joylashadi, bu esa avtomatik tahlil jarayonini yanada murakkablashtiradi.

Maqolada frazeologik birliklarni avtomatik teglash va aniqlashga doir lingvistik hamda kompyuter lingvistikasi sohasidagi tadqiqotlar tahlil etiladi. Shuningdek, xalqaro tajribalar va ularning o‘zbek tili korpuslariga tatbiq etilish imkoniyatlari ko‘rib chiqiladi.

Mavzuga oid adabiyotlar tahlili.

Frazeologik birliklarni raqamli muhitda o‘rganishga oid tadqiqotlar dastlab an’anaviy leksikografiya doirasida shakllangan bo‘lsa, endilikda bu yo‘nalishda korpus lingvistikasi va NLP usullaridan foydalanish rivojlanib bormoqda.

O‘zbek frazeologiyasining dastlabki shakllanishida Shavkat Rahmatullayevning tadqiqotlari muhim o‘rin tutadi. Muallifning “O‘zbek tilining izohli frazeologik lug‘ati” va “Nutqimiz ko‘rki” asarlari bugungi kunda ham shu sohaning muhim manbalaridan biri hisoblanadi [15], [16]. Rahmatullayev frazeologizmlarni “ko‘chma ma’noli turg‘un bog‘lanmalar” deb ta’riflaydi va ularni erkin birikmalardan ajratuvchi asosiy mezon deb ma’no yaxlitligini ko‘rsatadi. Mazkur qarash kompyuter lingvistikasidagi non-compositionality tushunchasiga yaqin turadi. Ammo klassik tadqiqotlarda frazeologizmlarning korpusdagi xususiyatlari yetarlicha o‘rganilmaganligi sabab zamonaviy korpus lingvistikasida bu masala yetarli darajada o‘zlashtirilmagan yo‘nalishlardan biri bo‘lib qolmoqda.

O‘zbek olimlaridan N. Abduraxmonova korpus lingvistikasining zamonaviy yo‘nalishlari qatorida avtomatik annotatsiya va semantik teglash masalalariga alohida e’tibor qaratadi [1]. Mazkur yondashuvlar frazeologik birliklarni korpus tarkibida mustaqil semantik birlik sifatida ajratish masalasi bilan ham chambarchas bog‘liqdir. Agostini va hammualliflar tomonidan yaratilgan UZWORDNET tizimi esa o‘zbek tilining 28 000 dan ortiq synsetlarni qamrab olgan

yirik leksik-semantik bazasi sifatida tavsiflanadi [6]. Bu tadqiqot o‘zbek tili frazeologizmlari uchun semantik annotatsiya va NLP texnologiyalarini rivojlantirishda muhim bosqich hisoblanadi.

O‘zbek tili tabiiy tilni qayta ishlash sohasida morfologik analizator va annotatsiyalangan datasetlar yaratishga oid tadqiqotlar ham olib borilgan. Xususan, Abduraxmonova, Mengliyev va Baraxninlar yaratgan morfologik analizator va keyinchalik nashr etilgan annotatsiyalangan morfologik dataset o‘zbek tili tabiatini kompyuter modellariga o‘rgatish sohasini rivojlantirishga xizmat qildi [3], [4]. Iboralarni avtomatik teglash jarayonida morfologik segmentatsiya muhim ahamiyatga ega. MorphUz analizatori o‘zbek tilidagi o‘zak va qo‘shimchalarni aniqlash asosida lingvistik birliklarni avtomatik tahlil qilish imkonini beradi, bu esa frazeologik konstruksiyalarni korpus muhitida aniqlash samaradorligini oshiradi [5]. Yodgorovning “Formation of a database of phraseological units based on the corpus of the Uzbek language” tadqiqoti esa bu ikki yo‘nalish: lingvistik tavsif va korpus texnologiyasini frazeologiya sohasida birlashtirish uchun frazeologik birliklar bazasini shakllantirish masalalariga e’tibor qaratadi [19]. Shunday bo‘lsa-da, mavjud o‘zbek tili korpuslarida asosiy e’tibor morfologik va sintaktik teglashga qaratilgan bo‘lib, frazeologik birliklarni alohida semantik segment sifatida ajratish hamon hal etilmagan masala bo‘lib qolmoqda.

Xalqaro adabiyotlarda frazeologiya masalasi, asosan, “Multiword Expressions” (MWE) atamasi bilan o‘rganiladi. Kompyuter lingvistikasida o‘ziga xos o‘ringa ega bo‘lgan Jurafsky va Martinning “Speech and Language Processing” nomli darsligida ko‘p so‘zli birikmalarni aniqlash NLP tizimlari uchun murakkab vazifalardan biri sifatida ko‘rsatiladi [13]. Ide va boshqalar tomonidan olib borilgan tadqiqot esa bu sohadagi metodologik xilma-xillikni tizimlashtiradi va iboralar identifikatsiyasining ikki yo‘nalishda rivojlanayotganini ko‘rsatadi. Bunda statistik qarash so‘zlarning birga kelish ehtimolligi, o‘zaro axborot koeffitsientlariga tayanadi. Chuqur o‘rganish uslubi esa kontekstual embeddinglar orqali matn ichida idiomalarni topishga qaratiladi [12].

V. Nedumpozhimana va boshqalarning “Shapley Idioms” tadqiqoti ham transformator modellar sohasida alohida ahamiyatga ega bo‘lib, mualliflar diqqat mexanizmlarining idiomalarni aniqlashdagi rolini tahlil qiladi [14]. Transformator modellari nafaqat ko‘chma ma’noni anglashda yaqin atrofdagi so‘zlarga, balki uzoq masofadagi kontekstual bog‘liqliklarga ham tayanadi. Bu xususiyat o‘zbek tilidagi komponentlari gap ichida bir-biridan uzoqlashib kelgan iboralar, ya’ni distant frazeologizmlar uchun, ayniqsa, muhimdir.

Ko‘p tilli tahlil yo‘nalishida I. Zaitova va boshqalarning multilingual tahlili esa BERT asosidagi modellarning idiomalar va mikrosintaktik konstruksiyalarga nisbatan ishlash xususiyatini yanada chuqurroq ochib beradi [20]. Tadqiqot natijalariga ko‘ra, modelning idiomatik ma’noni to‘g‘ri talqin etishi ko‘p jihatdan ta’limiy korpusning hajmi va sifatiga bog‘liq. Kuzatuv o‘zbek tili uchun annotatsiyalangan frazeologik korpus yaratishning nafaqat lingvistik, balki texnologik jihatdan ham zarurligini yana bir bora tasdiqlaydi. Xuddi shu model asosida Yayavaram va boshqalar ham tadqiqot olib borib, BERT modelini idiomalarni aniqlashga moslashtirish bo‘yicha yangicha yechim taklif etishdi. Bunda so‘zlarning o‘zaro yaqinligi (word cohesion) va tarjima usullari yordamida model ko‘chma ma’noni literal ma’nodan ajrata olish qobiliyati sezilarli yaxshilanganligi kuzatildi [18]. Agglyutinativ tuzilishga ega bo‘lgan o‘zbek tili uchun bu tadqiqot, ayniqsa, perspektiv bo‘lib, morfologik tomondan murakkab komponentlardan tashkil topgan frazeologizmlarni aniqlashda qo‘shimcha samaradorlik berishi kutiladi.

Frazeologik birliklarning nutqdagi faolligi va zamonaviy qo'llanish darajasini aniqlash masalasi Budiltseva va Novikova tadqiqotida ham ko'rib chiqilgan. Mualliflar somatik komponentli frazeologizmlar misolida ayrim idiomalar faol iste'molda saqlanib qolishini, ayrimlari asta-asta passiv qatlamga o'tib qolishini qayd etadi. Bunda "aktual frazeologik korpus" modeli, ya'ni frazeologik birliklarni qo'llanilish chastotasiga ko'ra saralash usuli taklif etilgan [8]. Ularning ushbu yondashuvini o'zbek tili frazeologizmlari uchun ham amaliy jihatdan qo'llash mumkin, chunki milliy korpus asosida faol va passiv qatlamdagi iboralarni ajratish semantik teglash sifatini sezilarli miqdorda oshirishi mumkin.

Xorij tajribasi bilan bir qatorda bu sohada turkiy tillarning tadqiqotlarini ham yodga olish o'rinli bo'ladi. Sababi turkiy tillar doirasidagi so'nggi ishlar ham muhim hisoblanadi. Xususan, G.Aslantash va T.Gungo'rlarning "A Unified Turkic Idiom Understanding Benchmark" deb atalgan tadqiqoti turk, ozarbayjon, o'zbek kabi beshta turkiy tillar uchun umumiy idiomalar bazasini yaratdi va o'zaro transfer qobiliyatini (cross-lingual transfer) sinovdan o'tkazdi [7]. Sh.Hakimovning "TurkicNLP" loyihasi esa turkiy tillar uchun yagona API va CoNLL-U formatidagi annotatsiya standartlarini taklif etib, o'zbek tili NLP tizimini xalqaro platformaga ulash uchun imkon yaratdi [11]. D.O.Dobrovolskiyning frazeologiyada korpus metodologiyasini qo'llashga bag'ishlangan ishlari esa qiyosiy perspektiva beradi: muallif parallel korpuslar orqali ayrim frazeologik birliklarning faqat ma'lum grammatik shaklda yoki ma'lum uslubda ishlatilishini ta'kidlaydi — bu kuzatuv o'zbek tili uchun ham semantik teglash tizimini boyitish zarurligini ko'rsatadi [9], [10].

Tadqiqot metodologiyasi.

Frazeologik birliklarni avtomatik aniqlash va teglashga oid tadqiqotlar metodologik jihatdan qoidalarga asoslangan, statistik va zamonaviy neyron tarmoq (transformator) modellari imkoniyatlarini qiyoslashga tayanadi. Tadqiqot doirasida quyidagi yondashuvlar tahlil qilindi:

1. Qoidalarga asoslangan yondashuv: Sh.Rahmatullayevning frazeologik lug'atlaridagi birliklar asosida lug'at va lingvistik qoidalarga tayanib ishlaydi [15], [16].
2. Statistik yondashuv: Korpusdagi so'zlarning o'zaro bog'liqlik koeffitsientlari (Mutual Information, Log-likelihood) va birga qo'llanish chastotasini tahlil qiladi [11], [18].
3. Transformator modellar: BERT va mBERT kabi modellar yordamida frazeologizmlarning kontekstual semantikasini va diqqat mexanizmlari orqali idiomatiklik darajasini aniqlashga qaratiladi [7], [13], [17].

   Bu yondashuvlarning o'zbek va turkiy tillar korpuslari misolidagi qiyosiy samaradorlik ko'rsatkichlarini quyidagi jadval umumlashtiradi:

1-jadval. NLP modellari samaradorligining qiyosiy tahlili

modellash usuli	aniqlik (F1-score)	tadqiqot yo'nalishi
qoidalarga asoslangan (rule-based)	65 %	lug'at va morfologik qoidalar
statistik metodlar (MI, log-likelihood)	72 %	So'zlarning korpusda birga kelish ehtimoli

transformator modellari (BERT)	85 %	kontekstual semantika va diqqat mexanizmi
Ko'p tilli modellar (mBERT/Turkic)	88 %	turkiy tillar uchun transfer qobiliyati

Ko'rib chiqilgan tadqiqotlarda o'zbek tilining agglyutinativ tabiati va komponentlari gap ichida uzoq joylashgan frazeologizmlarni aniqlashda transformator modellarining kontekstual imkoniyatlari yuqori samaradorlik berishi kuzatildi.

Tahlil va natijalar.

Mavjud adabiyotlar tahlili shuni ko'rsatadiki, frazeologizmlarni teglash masalasi NLP tadqiqotlarida faol rivojlanayotgan yo'nalishlardan biriga aylangan. Xalqaro tadqiqotlarda, ayniqsa, ingliz va yevropa tillari korpuslarida transformator modellardan foydalanish keng tarqalib, ayrim ishlarda frazeologik birliklarni aniqlashda yuqori samaradorlik ko'rsatkichlari qayd etilgan [7], [13]. Qoidalarga asoslangan va statistik metodlar ham frazeologik birliklarni aniqlashda qo'llaniladi, biroq komponentlari gap ichida uzoq joylashgan frazeologizmlarni aniqlashda ularning imkoniyatlari cheklanishi mumkin.

O'zbek tili bilan bog'liq tadqiqotlarda esa asosiy e'tibor ko'proq morfologik va sintaktik annotatsiyaga qaratilgan. Abdurakhmonova va boshqalar tomonidan yaratilgan morfologik analizator va annotatsiyalangan datasetlar o'zbek tilini avtomatik tahlil qilish yo'nalishidagi muhim ishlardan hisoblanadi [3], [4]. Biroq frazeologik birliklarni alohida semantik birlik sifatida avtomatik aniqlash va teglash masalasi hali yetarli darajada ishlab chiqilmagan. Turkiy tillar doirasidagi ayrim tadqiqotlar esa umumiy annotatsiya standartlari va ko'p tilli modellarni qo'llash imkoniyatlariga e'tibor qaratadi [2], [20].

Xulosa va takliflar.

Ko'rib chiqilgan manbalar asosida shuni aytish mumkinki, xalqaro va mahalliy tadqiqotlarda transformator modellari, statistik metodlar va qoidalarga asoslangan usullardan keng foydalanilmoqda. O'zbek tilida esa frazeologik birliklarni avtomatik tahlil qilish bo'yicha tadqiqotlar hali yetarli darajada shakllanmagan bo'lib, mavjud ishlarda, asosan, morfologik va sintaktik annotatsiyaga e'tibor qaratilgan. Shu bilan birga, turkiy tillar doirasidagi zamonaviy tadqiqotlar hamda ko'p tilli NLP modellari o'zbek tili korpuslarida frazeologik birliklarni semantik va kontekstual jihatdan avtomatik aniqlash bo'yicha yangi imkoniyatlar mavjudligini ko'rsatadi.

FOYDALANILGAN ADABIYOTLAR

1. Abdurakhmonova N. Korpus lingvistikasi. – Toshkent, 2024. – 320 b.
2. Abdurakhmonova N., Ismailov A., Mengliev D. Developing NLP Tool for Linguistic Analysis of Turkic Languages // 2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). – 2022. – Pp. 1–6.
3. Abdurakhmonova N., Mengliev D., Barakhnin V. Development of Intellectual Web System for Morph Analyzing of Uzbek Words // Applied Sciences. – 2021. – Vol. 11. – No. 19. – Pp. 9117.
4. Abdurakhmonova N., Shirinova R., Sayfullayeva R., Mengliev D., Ibragimov B., Ernazarova M. An annotated morphological dataset for Uzbek word forms: Towards rule-based and machine learning approaches // Data in Brief. – 2025. – Vol. 61. – Pp. 111702.

5. Abdurakhmonova N., Alisher I., Sayfulleyeva R. Morphological analyzer for the Uzbek language // 2022 7th International Conference on Computer Science and Engineering (UBMK). – 2022. – Pp. 61–66.
6. Agostini A., Usmanov T., Khamdamov U., Abdurakhmonova N., Mamasaidov M. UzWordNet: A lexical-semantic database for the Uzbek language // Proceedings of the 11th Global Wordnet Conference. – 2021. – Pp. 8–19.
7. Aslantaş G., Güngör T. A Unified Turkic Idiom Understanding Benchmark: Idiom Detection and Semantic Retrieval Across Five Turkic Languages // Proceedings of SIGTURK. – 2026. – Pp. 12–24.
8. Budiltseva M.B., Novikova N.S. Life of an idiom – defining the current corpus of phraseological units: experimental research experience // RUDN Journal of Language Studies, Semiotics and Semantics. – 2023. – Vol. 14. – No. 3. – Pp. 931–945.
9. Dobrovol'skij D.O. Korpusniy podkhod k issledovaniyu frazeologii: novie rezultati po dannim parallelnikh korpusov // Vestnik SPbGU. Yazik i literatura. – 2020. – T. 17. – Vip. 3. – Pp. 398–411.
10. Dobrovol'skij D.O. Kriterii semanticheskoy chlenimosti idiom // RUDN Journal of Language Studies, Semiotics and Semantics. – 2025. – Vol. 16. – No. 3. – Pp. 638–655.
11. Hakimov Sh. TurkicNLP: An NLP Toolkit for Turkic Languages // arXiv preprint arXiv:2602.19174. – 2026.
12. Ide Y., Tanner J., Nohejl A., Hoffman J., Vasselli J., Kamigaito H., Watanabe T. COAM: Corpus of All-Type Multiword Expressions // arXiv preprint arXiv:2412.18151. – 2024.
13. Jurafsky D., Martin J.H. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. – Stanford University, 2025. – 640 p.
14. Nedumpozhimana V., Klubička F., Kelleher J.D. Shapley Idioms: Analysing BERT Sentence Embeddings for General Idiom Token Identification // Frontiers in Artificial Intelligence. – 2022. – Vol. 5. – Pp. 813967.
15. Rahmatullayev Sh. Nutqimiz ko'rki. – Toshkent: Fan, 1970. – 60 b.
16. Rahmatullayev Sh. O'zbek tilining izohli frazeologik lug'ati. – Toshkent: O'qituvchi, 1978. – 408 b.
17. Tanner J., Hoffman J. MWE as WSD: Solving Multiword Expression Identification with Word Sense Disambiguation // arXiv preprint arXiv:2303.06623. – 2023.
18. Yayavaram A., Yayavaram S., Upadhyay P., Das A. BERT-based Idiom Identification using Language Translation and Word Cohesion // Proceedings of MWE-2024. – 2024. – Pp. 110–118.
19. Yodgorov U.S. Formation of a database of phraseological units based on the corpus of the Uzbek language // American Journal of Philological Sciences. – 2025. – Vol. 5. – No. 3. – Pp. 148–151.
20. Zaitova I., Hirak V., Abdullah B.M., Klakow D., Möbius B., Avgustinova T. Attention on Multiword Expressions: A Multilingual Study of BERT-based Models with Regard to Idiomaticity and Microsyntax // arXiv preprint arXiv:2505.06062. – 2025.

СОЦИАЛЬНО ОРИЕНТИРОВАННАЯ ЦИФРОВАЯ ТРАНСФОРМАЦИЯ ЭКОНОМИКИ ЦЕНТРАЛЬНОАЗИАТСКИХ СТРАН

С. С. Гулямов[1], С. С. Гулямов[2]

[1] Заслуженный деятель науки Республики Узбекистан, академик Академии наук Республики Узбекистан, доктор экономических наук, профессор, Ташкентский международный университет «Кимё», Ташкент, Республика Узбекистан

[2] Доктор юридических наук, профессор кафедры киберправа, Ташкентский государственный юридический университет, Ташкент, Республика Узбекистан.

https://doi.org/10.5281/zenodo.23136319

Аннотация. *В статье проверяется распространённое допущение о том, что развитие цифровой инфраструктуры автоматически повышает уровень жизни населения. Данные по пяти странам Центральной Азии за 2018–2025 годы это допущение опровергают. Казахстан — региональный лидер по цифровым рейтингам (индекс развития электронного правительства 0,9009, 24-е место в мире) — в 2025 году зафиксировал снижение реальных доходов населения на 2,7 %. Узбекистан при более скромной цифровой оснащённости (индекс 0,7999) сократил бедность с 14,0 % в 2022 году до 5,8 % в 2025 году, а неравенство доходов (коэффициент Джини) снизил с 0,35 до 0,33. Цель работы — объяснить это расхождение и предложить практический инструмент управления. Авторы рассматривают экономику как систему из шести взаимосвязанных узлов с «социальным контуром обратной связи»: сигнал о положении домохозяйств должен фиксироваться цифровыми системами и превращаться в конкретную помощь. Введён индекс скорости социальной обратной связи (ИССК) и показатель «структурного разрыва» между технической оснащённостью и реальной социальной отдачей. Расчёт по девяти экономикам показал, что главная проблема региона — не нехватка инфраструктуры и не отставание законов, а разомкнутость контура: цифровые системы фиксируют бедность, но не превращают этот сигнал в автоматическое управленческое решение. Сценарный прогноз до 2030 года показывает, что замыкание этого контура позволит вывести из бедности дополнительно 2,73 млн жителей региона без пропорционального роста бюджетных расходов — за счёт лучшей организации управления, а не увеличения затрат.*

Ключевые слова: *цифровая трансформация, обратная связь, Центральная Азия, бедность, реальные доходы, электронное правительство, искусственный интеллект, социальная политика.*

Abstract. *The article tests the common assumption that developing digital infrastructure automatically raises living standards. Data on the five Central Asian countries for 2018–2025 refute it. Kazakhstan, the regional leader in digital rankings (e-government development index 0.9009, 24th worldwide), recorded a 2.7 % decline in real household income in 2025, whereas Uzbekistan, with a more modest index of 0.7999, cut poverty from 14.0 % in 2022 to 5.8 % in 2025 and reduced income inequality (the Gini coefficient) from 0.35 to 0.33. The study explains this gap and proposes a practical management tool. The authors treat the economy as a system of six interacting nodes with a "social feedback loop": the signal about households' condition*

must be captured by digital systems and turned into concrete support. A Social Cybernetic Feedback Velocity Index (SCFVI) and a "structural gap" indicator are introduced. Calculations across nine economies show that the region's key constraint is neither infrastructure deficit nor legislative lag, but an open loop: digital systems record poverty yet fail to convert the signal into an automatic control action. Scenario modelling to 2030 shows that closing the loop would lift an additional 2.73 million residents out of poverty without a proportional rise in budget spending.

Keywords: *digital transformation, feedback, Central Asia, poverty, real income, e-government, artificial intelligence, social policy.*

Annotatsiya. *Maqolada raqamli infratuzilmaning rivojlanishi aholi turmush darajasini avtomatik ravishda oshiradi, degan keng tarqalgan taxmin tekshiriladi. Markaziy Osiyoning besh mamlakati bo'yicha 2018–2025-yillardagi ma'lumotlar bu taxminni rad etadi. Raqamli reytinglar bo'yicha mintaqaviy yetakchi bo'lgan Qozog'iston (elektron hukumatni rivojlantirish indeksi 0,9009, dunyoda 24-o'rin) 2025-yilda aholining real daromadlari 2,7 % ga kamayganini qayd etdi. O'zbekiston esa nisbatan kamtarona raqamli jihozlanganlik sharoitida (indeks 0,7999) kambag'allikni 2022-yildagi 14,0 % dan 2025-yilda 5,8 % gacha qisqartirdi, daromadlar tengsizligini (Jini koeffitsiyenti) esa 0,35 dan 0,33 gacha pasaytirdi. Ishning maqsadi — ushbu tafovutni tushuntirish va amaliy boshqaruv vositasini taklif etishdir. Mualliflar iqtisodiyotni «ijtimoiy teskari aloqa konturi»ga ega o'zaro bog'liq oltita tugundan iborat tizim sifatida ko'rib chiqadilar: xonadonlarning ahvoli haqidagi signal raqamli tizimlar tomonidan qayd etilishi va aniq yordamga aylanishi lozim. Ijtimoiy teskari aloqa tezligi indeksi (ITATI) hamda texnik jihozlanganlik bilan real ijtimoiy samara o'rtasidagi «strukturaviy tafovut» ko'rsatkichi kiritilgan. To'qqizta iqtisodiyot bo'yicha hisob-kitob shuni ko'rsatdiki, mintaqaning asosiy muammosi infratuzilma yetishmasligi yoki qonunlarning orqada qolishi emas, balki konturning uzilganligidadir: raqamli tizimlar kambag'allikni qayd etadi, biroq bu signalni avtomatik boshqaruv qaroriga aylantirmaydi. 2030-yilgacha bo'lgan stsenariy prognozi shuni ko'rsatadiki, ushbu konturni yopish byudjet xarajatlarining mutanosib o'sishisiz mintaqaning qo'shimcha 2,73 mln aholisini kambag'allikdan chiqarish imkonini beradi — bu xarajatlarni ko'paytirish hisobiga emas, balki boshqaruvni yaxshiroq tashkil etish hisobiga amalga oshadi.*

Kalit so'zlar: *raqamli transformatsiya, teskari aloqa, Markaziy Osiyo, kambag'allik, real daromadlar, elektron hukumat, sun'iy intellekt, ijtimoiy siyosat*

1. ВВЕДЕНИЕ

За 2018–2025 годы пять государств Центральной Азии совершили заметный рывок в цифровой инфраструктуре. Узбекистан поднялся в глобальном индексе развития электронного правительства с 81-го места в 2018 году до 63-го в 2024 году, перейдя из группы стран с «высоким» в группу с «очень высоким» уровнем этого показателя. Казахстан закрепился на 24-й строке мирового рейтинга и стал первым среди стран, не имеющих выхода к морю. Кыргызстан утвердил Концепцию цифровой трансформации на 2024–2028 годы, а Таджикистан объявил 2025–2030 годы периодом развития цифровой экономики.

Однако при сопоставлении цифровых и социальных показателей возникает картина, которая не укладывается в привычную логику «инфраструктура порождает благосостояние». В Узбекистане, занимающем в регионе лишь вторую позицию по цифровой готовности, бедность снизилась с 14,0 % в 2022 году до 5,8 % в 2025 году. В

Казахстане, лидирующем по всем цифровым метрикам, реальные доходы населения в 2025 году, наоборот, уменьшились на 2,7 %, а доля расходов домохозяйств на продовольствие превысила 52 %.

Это расхождение — не статистическая случайность и не следствие разных методик подсчёта. Оно указывает на самостоятельную научную проблему: между цифровой оснащённостью государства и социальной отдачей цифровизации нет автоматической связи. Инвестиции в инфраструктуру и государственные информационные системы могут достигать высокой интенсивности, но не превращаться в улучшение жизни домохозяйств. Причина, как показано в работе, лежит не в отдельных элементах системы, а в устройстве связей между ними — в том, замкнут ли контур, по которому сигнал о положении общества возвращается к тем, кто принимает управленческие решения.

1.1. Актуальность и научная новизна

Существующие исследования цифровой трансформации региона сосредоточены в трёх плоскостях: инфраструктурно-рейтинговой (динамика мест в индексах, охват широкополосным доступом), нормативно-правовой (законодательство о данных, электронной подписи, искусственном интеллекте) и секторальной (цифровизация финансов, здравоохранения, сельского хозяйства). Ни один из этих подходов не объясняет, почему высокие инфраструктурные показатели не превращаются в социальный результат.

Кибернетическая традиция в экономике — от работ Н. Винера и У. Р. Эшби до модели жизнеспособной системы С. Бира и опыта проектирования общегосударственных систем управления В. М. Глушкова — предлагает язык, на котором эта проблема формулируется корректно. Экономика описывается как совокупность узлов, соединённых каналами передачи информации, а благосостояние трактуется как выходной параметр контура. Настоящая статья развивает этот подход в социальном измерении, ранее остававшемся периферийным.

Научная новизна работы состоит в следующем. **Во-первых,** предложено выделить социальное благосостояние в самостоятельный узел модели, ранее растворённый в обобщённом узле «общество». **Во-вторых,** введён индекс скорости социальной обратной связи (ИССК), измеряющий не оснащённость канала, а скорость превращения цифрового сигнала в изменение уровня жизни. **В-третьих,** количественно измерен феномен «структурного разрыва контура» — расхождения между технической готовностью канала и его фактической социальной отдачей. **В-четвёртых,** выполнен сравнительный анализ нормативных актов региона 2020–2026 годов и построен трёхсценарный прогноз до 2030 года, в котором различие между сценариями определяется не объёмом инвестиций, а качеством управления.

Цель исследования — построить социально ориентированную модель цифровой трансформации центральноазиатских государств, проверить её на статистике 2018–2025 годов и разработать сценарный прогноз до 2030 года с акцентом на динамику доходов и уровня бедности.

2. ТЕОРЕТИЧЕСКАЯ ОСНОВА

2.1. Кибернетический взгляд на экономику

Представление экономики как системы контуров обратной связи восходит к основателям кибернетики. Закон необходимого разнообразия У. Р. Эшби гласит: управляющая система способна удерживать объект в нужном состоянии только тогда,

когда её собственное разнообразие не уступает разнообразию воздействующих на объект возмущений. Для государственного управления это означает: если сложность социальной среды растёт быстрее, чем способность органов власти её анализировать и реагировать, управляемость теряется независимо от объёма расходуемых ресурсов.

С. Бир перевёл этот принцип на язык организационного проектирования в модели жизнеспособной системы. Ф. фон Хайек описал ценовую систему как распределённый процесс, агрегирующий рассеянное знание, — по сути классическую отрицательную обратную связь. Дж. М. Кейнс и Х. Мински, напротив, описали механизмы положительной обратной связи — самоусиливающиеся процессы, ведущие систему к нестабильности при отсутствии внешнего сдерживания. Кибернетический язык позволяет объединить обе логики, поскольку оперирует не допущениями о поведении агентов, а формальными характеристиками связей.

2.2. Почему социальное благосостояние нужно выделять в отдельный узел

Ранее разработанная авторами пятиузловая модель объединяла в узле «общество» две разные сущности: систему ценностей, норм и ожиданий — и материальное благосостояние домохозяйств. Для анализа платформенных монополий или доверия такое объединение допустимо, но для анализа социальных результатов цифровизации оно становится источником систематической ошибки по трём причинам.

- **Разная скорость изменений.** Ценности общества меняются на горизонте поколений, а уровень доходов может измениться на десятки процентов за один финансовый год. Объединение таких разных по инерции переменных искажает оценку.
- **Разные входные каналы.** Ценности формируются сигналами из технологического и политического узлов, а благосостояние — из рыночного (зарплата, занятость) и институционального (трансферты, налоги, пособия).
- **Разная измеримость.** Благосостояние измеряется напрямую (доход, бедность, неравенство), а ценности — только косвенно, через опросы с большой погрешностью.

Отсюда исследовательский пробел, который закрывает данная работа: прежняя модель не позволяла диагностировать ситуацию, когда технологический и институциональный узлы работают исправно, а благосостояние населения не улучшается или ухудшается. Именно эта ситуация наблюдается в отдельных экономиках Центральной Азии.

2.3. Мировой опыт: три модели управления контурами

Либертарианская модель (США) максимизирует скорость технологического узла при минимальном регулировании. Результат — самая капитализированная цифровая экономика мира, но и самоусиливающиеся перекосы: платформенные монополии, концентрация капитала, географическая поляризация. Антимонопольные разбирательства 2020-х годов — запоздалая попытка восстановить обратную связь.

Регуляторная модель Европейского союза — единственная в мире попытка заранее проектировать сдерживающие механизмы. Однако наложение множества регламентов породило вторичную проблему — регуляторную перегрузку. К 2025 году около 70 % базовых моделей искусственного интеллекта разрабатывались в США, а объём европейского венчурного капитала составлял лишь 5 % глобального против 52 % в США и 40 % в Китае. Инициатива Digital Omnibus (ноябрь 2025 года) впервые предусматривает

сознательное снижение регуляторной нагрузки не менее чем на 25 % для всех предприятий к 2029 году.

Модель государственного захвата (Китай) управляет обоими типами контуров: государственные венчурные фонды за десятилетие направили 912 млрд долларов США, а цифровой юань к концу 2025 года обработал свыше 3,4 млрд транзакций. Одновременно государство подавляет разнообразие общественных сигналов, что нарушает закон необходимого разнообразия.

Для государств Центральной Азии ни одна из трёх моделей не воспроизводима в чистом виде: либертарианская требует ёмкого венчурного рынка, регуляторная — наднациональной административной мощности, а модель захвата — масштаба внутреннего рынка. Отсюда необходимость собственного проектного решения.

3. МЕТОДОЛОГИЯ

3.1. Устройство модели и социальный контур обратной связи

Предлагаемая модель развивает пятиузловую конструкцию, разделяя обобщённый узел «общество» на два: узел общественных ценностей (S_1) и узел социального благосостояния (S_2). Итоговая архитектура включает шесть узлов:

- **T** — технологический узел: создание и распространение технических решений, включая искусственный интеллект;
- **I** — институциональный узел: установление и применение правил, включая механизмы адресной поддержки;
- **M** — рыночный узел: обмен, формирование зарплаты и занятости;
- **S_1** — узел общественных ценностей: доверие, ожидания, готовность осваивать новое;
- **S_2** — узел социального благосостояния: доходы, бедность, неравенство;
- **P** — политический узел: коллективный выбор, распределение бюджета, целеполагание.

Ключевое нововведение — выделение социального контура обратной связи, замкнутой цепи $S_2 \to T \to I \to P \to S_2$. В ней изменение положения домохозяйств фиксируется цифровыми системами (передача $S_2 \to T$), превращается в изменение правил поддержки ($T \to I$), утверждается как бюджетное решение ($I \to P$) и возвращается к людям в форме трансферта, услуги или нового рабочего места ($P \to S_2$).

Главное теоретическое утверждение модели: социальная отдача цифровой трансформации определяется **произведением**, а не суммой коэффициентов контура. Обнуление любого из четырёх звеньев обнуляет весь результат независимо от совершенства остальных. Государство может иметь первоклассную систему сбора данных, развитое законодательство о поддержке и достаточный бюджет, но если нет механизма автоматического превращения решения в выплату или рабочее место, социальная отдача цифровизации будет близка к нулю. Это прямой аналог закона необходимого разнообразия: результат ограничен самым слабым звеном цепи.

Четыре типа сбоев социального контура

Тип α — сенсорная слепота. Инфраструктура развёрнута, но положение самых уязвимых домохозяйств в неё не поступает: сельские территории вне покрытия, неформальная занятость не отражена в реестрах, доходы трудовых мигрантов не учитываются.

Тип β — трансляционный разрыв. Данные собираются, но нет правовых оснований их использовать или запрещён межведомственный обмен. Гражданин вынужден повторно подтверждать то, что государство уже знает.

Тип γ — бюджетная непроводимость. Решение выработано, но годовой цикл планирования не позволяет реагировать на изменение положения домохозяйств внутри финансового года.

Тип δ — исполнительная утечка. Решение принято и профинансировано, но до адресата доходит лишь часть ресурса из-за посредничества, дублирования получателей или отсутствия проверки фактического получения.

Отдельно выделяется тип ε — избыточное демпфирование, когда все звенья работают, но чрезмерная регуляторная нагрузка подавляет инновации и снижает долгосрочный потенциал роста доходов.

3.2. Индекс скорости социальной обратной связи (ИССК)

Существующие индексы цифрового развития (индекс электронного правительства, индекс сетевой готовности) измеряют оснащённость системы — наличие инфраструктуры и услуг на момент наблюдения. Ни один из них не измеряет, с какой скоростью цифровой сигнал превращается в изменение уровня жизни. Именно эту величину измеряет предлагаемый индекс ИССК, построенный на трёх компонентах:

- **D — плотность цифрового сигнала:** способность технологического узла воспринимать информацию о домохозяйствах (проникновение интернета, широкополосный доступ, доля госуслуг в электронной форме);
- **T — пропускная способность канала:** готовность превратить сигнал в адресное воздействие (охват единым социальным реестром, доля выплат с цифровой идентификацией, субиндекс онлайн-услуг ООН);
- **W — социальная отдача:** фактическая скорость изменения благосостояния (темп сокращения бедности, прирост реальных доходов, снижение неравенства).

Агрегация: ИССК = 0,25 · D + 0,35 · T + 0,40 · W (нормализация методом «минимум-максимум» по выборке). Распределение весов отражает главную идею работы: выходной сигнал контура (W) важнее его технической оснащённости (D), а пропускная способность канала (T) важнее плотности сигнала на входе. Этим ИССК принципиально отличается от индексов, где доминирует инфраструктура.

Методологическая оговорка. Компонент W измеряет скорость, а не уровень благосостояния. Поэтому зрелые экономики, приблизившиеся к пределу сокращения бедности, получают низкие значения W не из-за сбоя, а из-за насыщения. Это требует обязательного контроля по уровневому показателю (уровню бедности), что и выполнено ниже.

3.3. Данные и сценарии прогноза

Эмпирическую базу составили данные ООН (индекс электронного правительства, издания 2018–2024 годов), Всемирного банка, Международного валютного фонда, Международного союза электросвязи и национальных статистических ведомств пяти республик за 2018–2025 годы, а также нормативные акты региона за 2020–2026 годы. Часть показателей для Кыргызстана, Таджикистана и особенно Туркменистана представляет экспертную оценку и помечена в таблицах.

Прогноз до 2030 года построен в трёх сценариях, различающихся не объёмом инвестиций, а качеством управления контуром. Инерционный сценарий сохраняет

сложившуюся архитектуру связей. Базовый предполагает выполнение официально утверждённых стратегий. Кибернетический дополнительно реализует предложения раздела 6 — автоматическое назначение поддержки, региональный обмен социальными данными, внутригодовое бюджетное реагирование.

4. РЕЗУЛЬТАТЫ

4.1. Цифровой профиль региона

Индекс развития электронного правительства показывает устойчивое продвижение региона при значительной внутренней дифференциации (таблица 1).

Таблица 1. Индекс развития электронного правительства государств Центральной Азии, 2018–2024 годы

Государство	2018	2020	2022	2024	Индекс 2024
Казахстан	39	29	28	24	0,9009
Узбекистан	81	87	69	63	0,7999
Кыргызстан	91	83	81	78	0,7213*
Таджикистан	131	133	129	122	0,5460*
Туркменистан	147	145	142	138	0,4980*

*Источники: база данных ООН по индексу развития электронного правительства, издания 2018–2024 годов; расчёты авторов. Знаком * помечены значения, восстановленные авторами по позиции в рейтинге.*

Наиболее показателен Узбекистан: страна не только преодолела откат 2020 года, но и досрочно превысила целевой ориентир 0,75, установленный стратегией «Цифровой Узбекистан — 2030» на 2025 год. При этом телекоммуникационная составляющая индекса достигла 0,87, тогда как человеческий капитал оценён в 0,75, а онлайн-услуги — в 0,76. Это означает опережающее развитие технологического узла при отставании узлов, обеспечивающих переработку сигнала.

Инфраструктурные показатели региона за 2023–2025 годы (таблица 2) подтверждают значительную дифференциацию: если Казахстан и Узбекистан приблизились к уровню развитых экономик по проникновению интернета и доле цифровых госуслуг, то Таджикистан и Туркменистан отстают в несколько раз.

Таблица 2. Инфраструктурные и сервисные индикаторы цифровизации, 2023–2025 годы

Показатель	Узбеки-стан	Казах-стан	Кыргыз-стан	Таджи-кистан	Туркме-нистан
Проникновение интернета, % населения	89,0	93,0	78,0*	40,0*	35,0*
Фиксированный ШПД, подключений на 100 жителей	30,3	37,2	6,0*	6,5*	0,6*
Покрытие мобильным ШПД, % населённых пунктов	99,0	98,0*	85,0*	70,0*	55,0*
Доля госуслуг в электронной форме, %	82,0	92,0	60,0*	30,0*	15,0*

*Источники: Международный союз электросвязи; Всемирный банк; национальные ведомства; оценки авторов (помечены знаком *).*

К концу 2025 года мобильный широкополосный доступ в Узбекистане охватил 99 % населённых пунктов, сети пятого поколения развёрнуты во всех областных центрах, международная пропускная способность достигла 4200 Гбит/с, доля госуслуг в цифровом виде превысила 80 %. Вместе с тем сохраняются структурные разрывы: подключённость сельских территорий не достигает 50 %, а расходы на исследования и разработки удерживаются на уровне 0,13 % ВВП — более чем в тридцать семь раз ниже показателя Республики Корея (4,91 %).

Экспорт информационно-коммуникационных услуг Узбекистана вырос с 0,6 млн долларов США в 2017 году до 619,7 млн в 2024 году и около 1 млрд по оценке 2025 года — более чем в тысячу раз за восемь лет при старте с крайне низкой базы. К 2025 году сеть IT Park объединяла свыше 2600 компаний с совокупной выручкой резидентов порядка 1,6 млрд долларов США. Это почти образцовый пример самоусиливающегося положительного контура. Однако при численности занятых в отрасли около ста тысяч человек и рабочей силе свыше четырнадцати миллионов прямой вклад сектора в доходы населения пока ограничен.

4.2. Центральный парадокс: цифровое лидерство и социальная отдача

Сопоставление макроэкономических результатов пяти государств за 2025 год приведено в таблице 3.

Таблица 3. Макроэкономические индикаторы государств Центральной Азии, 2025 год

Показатель	Узбекистан	Казахстан	Кыргызстан	Таджикистан	Туркменистан
Прирост ВВП, %	7,7	6,3*	11,1	8,4	6,0*
ВВП на душу, долл. США	3600	14 770	2750	1430	13 340
Уровень бедности, %	5,8	5,7*	25,0*	20,5*	н/д
Изменение реальных доходов, %	рост	−2,7	рост	рост	н/д

*Источники: Международный валютный фонд; национальные статистические ведомства; расчёты и оценки авторов (помечены знаком *).*

Случай Узбекистана. Бедность сократилась с 14,0 % в 2022 году до 11,0 % в 2023-м, 8,9 % в 2024-м и 5,8 % в 2025-м. Только за 2025 год показатель снизился на 3,1 процентного пункта, из бедности выведено около 302 тысяч семей. Сокращение зафиксировано во всех регионах, а сельская бедность (6,4 %) лишь незначительно превышает городскую (5,2 %). Важно, что снижение бедности сопровождалось сокращением неравенства: доходы наименее обеспеченных групп выросли на 26 %, наиболее обеспеченных — на 13 %, вследствие чего коэффициент Джини снизился с 0,35 до 0,33. Это в кибернетических терминах прямое свидетельство работающего отрицательного контура: система реагирует на положение самых уязвимых домохозяйств управляющим воздействием.

Случай Казахстана. Республика удерживает первое место на постсоветском пространстве по душевому ВВП (14,77 тыс. долларов США) и лидирует в регионе по всем

цифровым индикаторам. Тем не менее реальные денежные доходы населения в 2025 году сократились на 2,7 %, реальная зарплата — на 3,2 %, а доля расходов на продовольствие превысила 52 %. Среднедушевые номинальные доходы выросли на 9,8 %, но инфляция полностью поглотила этот прирост.

Таким образом, регион демонстрирует обратную зависимость между уровнем цифровой готовности и темпом улучшения социальных показателей. Государство с наивысшим индексом (0,9009) показало отрицательную динамику реальных доходов; государство со значением 0,7999 — рекордное сокращение бедности. Это не доказывает вреда цифровизации. Это означает, что цифровая готовность — условие необходимое, но принципиально недостаточное, и вопрос о том, что именно превращает её в социальный результат, требует отдельного аппарата.

4.3. Расчёт индекса ИССК и структурный разрыв

Расчёт выполнен по девяти экономикам: пять государств Центральной Азии и четыре эталона — Эстония (модель тотальной цифровой идентичности), Республика Корея (высокая абсорбционная способность), Индия (прямые цифровые трансферты при умеренной инфраструктуре), Грузия (малая экономика с быстрым регуляторным циклом). Итоговые значения приведены в таблице 4.

Таблица 4. Компонентные оценки и итоговое значение индекса ИССК

Ранг	Экономика	D	T	W	ИССК
1	Узбекистан	0,776	0,844	1,000	**0,889**
2	Республика Корея	0,984	0,974	0,224	**0,676**
3	Грузия	0,697	0,752	0,598	**0,676**
4	Эстония	0,934	1,000	0,137	**0,639**
5	Индия	0,354	0,892	0,534	**0,615**
6	Казахстан	0,888	0,926	0,039	**0,562**
7	Кыргызстан	0,449	0,521	0,529	**0,506**
8	Таджикистан	0,130	0,207	0,421	**0,274**
9	Туркменистан	0,000	0,000	0,209	**0,083**

Источник: расчёты авторов. Нормализация методом «минимум-максимум» по выборке; веса 0,25 / 0,35 / 0,40.

Ранжирование требует осторожного прочтения. Первое место Узбекистана объясняется исключительной величиной компонента W: наивысший темп сокращения бедности, опережающий рост реальных доходов, наибольшее снижение неравенства. Это не значит, что у страны лучший контур в абсолютном выражении, — это значит, что контур работает с наибольшей скоростью относительно достигнутого уровня. Низкие значения W у Эстонии (0,137) и Кореи (0,224) отражают не сбой, а насыщение: обе экономики близки к пределу дальнейшего сокращения бедности.

Чтобы отделить насыщение от патологии, введён показатель структурного разрыва: G = (D + T) / 2 − W. Он измеряет расхождение между технической готовностью канала и его фактической отдачей (таблица 5). Контрольным показателем служит уровень бедности.

Таблица 5. Структурный разрыв контура и его интерпретация

Экономика	Оснащён-ность (D+T)/2	Отдача W	Разрыв G	Бедность 2025, %	Интерпретация
Казахстан	0,907	0,039	**+0,868**	5,7*	Патологический разрыв
Эстония	0,967	0,137	+0,830	2,0*	Насыщение
Республика Корея	0,979	0,224	+0,755	3,3*	Насыщение
Индия	0,623	0,534	+0,089	11,3*	Сбалансированный контур
Узбекистан	0,810	1,000	**−0,190**	5,8	Отдача существенно опережает оснащённость

*Источник: расчёты авторов. Уровни бедности приведены по национальным методологиям и не строго сопоставимы (помечены знаком *). Показаны ключевые случаи выборки.*

Таблица 5 позволяет сделать три вывода. Первый: положительный разрыв сам по себе не является диагнозом — Эстония и Корея показывают значения G, близкие к казахстанскому, но при бедности 2–3 % это нормальное состояние зрелой системы (насыщение). Второй: казахстанский разрыв патологичен — при сопоставимой бедности и более высокой оснащённости выходной сигнал контура отрицателен. Это соответствует комбинации сбоев типа γ (бюджетная непроводимость) и типа ε (избыточное демпфирование): цифровая система фиксирует положение домохозяйств, но инфляционное возмущение социальный контур не компенсирует. Третий: отрицательный разрыв Узбекистана (−0,190) означает, что социальная отдача превышает ожидаемую по технической оснащённости.

4.4. Механизм узбекистанского опережения: гибридный контур

Сокращение бедности в Узбекистане обеспечивалось не столько цифровыми технологиями как таковыми, сколько гибридной архитектурой социального контура, соединяющей цифровой реестр с низовым институтом территориального самоуправления — махаллёй. Реестр «Темир дафтар» («железная тетрадь»), формировавшийся с 2019 года на основе подворного обхода, обеспечил сенсорную функцию контура в условиях, когда чисто цифровые инструменты не охватывали сельские территории и неформально занятое население. Единый реестр социальной защиты устранил дублирование получателей. Вовлечение свыше трёх миллионов человек в доходную трудовую деятельность только за первое полугодие 2025 года обеспечило исполнительное звено — преимущественно через занятость, а не через трансферт.

Кибернетическое истолкование прямое. Чисто цифровой канал, охватывающий 89 % населения, обладает недостаточным разнообразием для управления обществом, где значительная часть уязвимых домохозяйств находится вне цифрового периметра. Гибридный канал, дополняющий цифровой реестр подворным обследованием, восполняет

дефицит разнообразия организационными средствами. Именно этот механизм — компенсация недостатка цифрового разнообразия организационным — объясняет, почему страна с более скромной инфраструктурой опередила регионального лидера по социальному результату.

Отсюда гипотеза гибридного контура: на этапе, когда цифровое покрытие ещё не полно, замыкание социального контура достигается быстрее комбинацией цифрового реестра с низовым организационным звеном, чем ожиданием полного цифрового охвата. Одновременно нужно учитывать риски: гибридный контур сохраняет человеческое посредничество (риск искажения сигнала при слабой проверке), ограниченно масштабируется и зависит от административной мобилизации.

4.5. Нормативная динамика региона

В Узбекистане нормативную основу образует последовательность актов, каждый из которых выполняет свою функцию в контуре. Указ № УП-6079 (2020) утвердил стратегию «Цифровой Узбекистан — 2030» и задал целевой ориентир по индексу электронного правительства. Указ № УП-158 (2023) о стратегии «Узбекистан — 2030» соединил технологическое целеполагание с социальным, включив цели по благосостоянию населения. Постановление № ПП-358 (2024) о развитии искусственного интеллекта предусмотрело редкий для региона механизм — разработку порядка оценки социальной и экономической эффективности проектов внедрения ИИ, то есть явное проектирование обратной связи. Постановление № ПП-257 (2026) о системе «Цифровое правительство» предусматривает создание Единой интегрированной цифровой платформы, что устраняет трансляционный разрыв типа β. Объявление 2026 года Годом развития махалли и общества прямо укрепляет то организационное звено, которое обеспечило опережающую социальную отдачу.

В Казахстане нормативная динамика 2025–2026 годов наиболее интенсивна: Закон «Об искусственном интеллекте» (17 ноября 2025 года) стал первым специализированным актом такого рода в Центральной Азии, а Общенациональная стратегия «Digital Qazaqstan» до 2029 года предусматривает инвестиции в диапазоне 10,5–16,5 млрд долларов США. Однако ни один из проанализированных актов не устанавливает измеримого целевого показателя социальной отдачи — прироста реальных доходов, сокращения бедности. Это принципиальное отличие от узбекистанской практики, где стратегия «Узбекистан — 2030» задаёт социальные индикаторы наравне с технологическими.

Обобщающее наблюдение: ни в одном нормативном акте региона, за исключением стратегии «Узбекистан — 2030» и постановления № ПП-358, социальная отдача цифровизации не выступает измеряемым и подотчётным параметром. Цифровизация рассматривается как самостоятельная ценность, а не как инструмент, эффективность которого оценивается по изменению уровня жизни. В этом — главный нормативный дефект региональной модели.

4.6. Сценарный прогноз до 2030 года

Результаты сценарного расчёта по ключевым показателям приведены в таблицах 6–9. Различие между сценариями определяется не объёмом инвестиций, а качеством управления контуром.

Таблица 6. Прогноз ВВП на душу населения до 2030 года, долл. США

Государство	2025	Инерцион-ный	Базовый	Кибернети-ческий	Разница киб./инерц., %
Узбекистан	3600	5290	5798	6344	+19,9
Казахстан	14 770	17 122	18 406	19 766	+15,4
Кыргызстан	2750	3857	4231	4634	+20,1
Таджикистан	1430	2006	2150	2303	+14,8
Туркменистан	13 340	14 728	15 465	16 230	+10,2

Источник: расчёты авторов. Наибольший относительный выигрыш от замыкания контура получают экономики с низкой стартовой базой.

Базовый сценарий для Узбекистана даёт 5798 долларов США на душу к 2030 году, что практически совпадает с обновлённым ориентиром национальной стратегии (5,8 тыс. долларов). Это означает, что заявленная цель достижима лишь при полной реализации утверждённых стратегических документов: инерционная траектория выводит на 5290 долларов, не достигая цели примерно на 9 %.

Таблица 7. Прогноз уровня бедности до 2030 года, % населения

Государство	2025	Инерцион-ный	Базовый	Кибернети-ческий
Узбекистан	5,8	4,0	3,1	2,2
Казахстан	5,7	6,5	5,0	3,9
Кыргызстан	25,0	18,8	15,2	12,1
Таджикистан	20,5	16,3	13,5	10,8
Туркменистан	12,0	11,1	9,8	8,3

Источник: расчёты авторов. Уровни бедности приведены по национальным методологиям и не строго сопоставимы между странами.

Таблица 8. Агрегированный социальный эффект сценариев по региону

Показатель	2025	Инерцион-ный 2030	Базовый 2030	Кибернети-ческий 2030
Население за чертой бедности, млн чел.	8,20	7,31	5,88	4,58
Выведено из бедности к 2030 году, млн чел.	—	0,89	2,32	3,62
Дополнительный эффект к инерционному, млн чел.	—	—	+1,43	+2,73

Источник: расчёты авторов на основе среднего варианта демографического прогноза.

Ключевой результат: разница между инерционной и кибернетической траекториями составляет 2,73 млн человек, выведенных из бедности к 2030 году. Эта разница обеспечивается не разным объёмом инвестиций в инфраструктуру, а разной организацией управления. Именно поэтому предложения раздела 6 не требуют

пропорционального роста бюджетных расходов: они адресуют архитектуру связей, а не мощность узлов.

Прогноз доли цифровой экономики в ВВП (таблица 9) показывает, что даже в кибернетическом сценарии регион остаётся дифференцированным. Для экспорта информационно-коммуникационных услуг Узбекистана расчёт даёт 2,59 млрд долларов США в инерционном сценарии, 4,01 млрд в базовом и 5,19 млрд в кибернетическом. Официально заявленная цель на 2030 год — 5 млрд долларов США — достижима, следовательно, лишь в кибернетическом сценарии, то есть при устранении узких мест в подготовке кадров и институциональном обеспечении экспортных операций.

Таблица 9. Прогноз доли цифровой экономики в ВВП до 2030 года, %

Государство	2024–2025	Инерцион-ный	Базовый	Кибернети-ческий
Узбекистан	5,5	8,0	10,5	13,0
Казахстан	5,0	7,5	10,0	12,0
Кыргызстан	4,0	6,0	8,0	10,0
Таджикистан	3,0	5,0	7,0	9,0
Туркменистан	2,0	3,0	4,0	5,0

Источник: расчёты и оценки авторов.

5. УРОКИ ЗАРУБЕЖНОГО ОПЫТА

Индия даёт наиболее полное воплощение замкнутого социального контура. Система прямых выплат бенефициарам, введённая в 2013 году, опирается на связку из банковских счетов с нулевым остатком, биометрической идентификации (охват свыше 1,4 млрд человек) и мобильной связи. За 2009–2024 годы совокупная экономия от устранения утечек составила около 42 млрд долларов США, доля субсидий в госрасходах сократилась с 16 % до 9 % при шестнадцатикратном расширении охвата. Примечательно, что страна с проникновением интернета порядка 55 % обеспечивает более эффективное замыкание контура (Т = 0,892), чем Казахстан с показателем 93 % (Т = 0,926 при почти нулевой отдаче). Причина — архитектура строилась вокруг единственной функции: доведения средств до идентифицированного получателя без посредников. Урок для региона: приоритетом должно быть не расширение перечня цифровых услуг, а сквозная связность одного критического маршрута «идентификация → определение права → перечисление → верификация получения». При этом нужны альтернативные каналы подтверждения права для тех, чья биометрия не считывается надёжно.

Эстония представляет противоположный полюс: замыкание достигается через универсальную цифровую идентичность и принцип однократного предоставления данных, при котором государство не вправе повторно запрашивать сведения, которыми уже располагает. Принципиально воспроизводимым является не технологический стек, а именно этот нормативный принцип: его закрепление в законодательстве прямо устраняет трансляционный разрыв и требует юридического, а не технологического решения, потому доступно экономикам с ограниченными ресурсами.

Республика Корея демонстрирует решающее значение абсорбционной способности: расходы на исследования и разработки 4,91 % ВВП (второй показатель в мире), корпоративный сектор обеспечивает 74 % этих расходов. Способность экономики

перерабатывать технологический сигнал определяется не объёмом импортируемых технологий, а численностью и квалификацией специалистов. Для Узбекистана разрыв с корейским уровнем в тридцать семь раз — количественная мера основного долгосрочного ограничения: сигнал, поступающий в узел, неспособный его переработать, рассеивается без результата.

Европейский союз показывает как работоспособность заранее спроектированных сдерживающих механизмов (в 2025 году наложены штрафы 500 млн евро на Apple и 200 млн евро на Meta по Закону о цифровых рынках), так и издержки их избыточности: разрыв в производительности труда с США вырос почти до 15 долларов в час к 2024 году, причём почти исключительно за счёт сектора ИКТ. Урок для региона: перенос регуляторного стека, спроектированного для экономики с гипертрофированным технологическим узлом, в экономику с развивающимся узлом способен подавить инновации до того, как они достигнут уровня, требующего сдерживания. Закон Казахстана об ИИ, воспроизводящий европейскую риск-ориентированную классификацию, требует мониторинга на предмет соразмерности.

Обобщение зарубежного опыта приведено в таблице 10. Узбекистанская гибридная конфигурация не является ухудшенным вариантом какой-либо из зарубежных моделей — это самостоятельный кибернетический тип, адекватный стадии, на которой цифровое покрытие ещё не полно, а потребность в замыкании социального контура уже актуальна.

Таблица 10. Сравнительная матрица моделей замыкания социального контура

Модель	Ведущий механизм	Сильная сторона	Ограничение
Индия	Прямые цифровые трансферты по биометрии	Высокая пропускная способность при умеренной инфраструктуре	Риск исключения лиц с нечитаемой биометрией
Эстония	Универсальная цифровая идентичность, однократное предоставление данных	Устранение трансляционного разрыва правовыми средствами	Малый масштаб системы
Республика Корея	Абсорбционная способность через человеческий капитал и R&D	Устойчивое преобразование технологий в производительность	Длительный горизонт формирования
Европейский союз	Заблаговременное проектирование сдерживающих контуров	Предотвращение патологий до их возникновения	Регуляторная перегрузка
Узбекистан (гибридная)	Цифровой реестр в сочетании с низовым звеном (махалля)	Компенсация дефицита цифрового разнообразия	Зависимость от административной мобилизации

Источник: систематизация авторов.

6. ОБСУЖДЕНИЕ И ОГРАНИЧЕНИЯ

Анализ выявил три критических наблюдения по региональной модели. Первое: цифровизация измеряется входами, а не выходами. Большинство целевых индикаторов стратегических документов относится к входным параметрам — числу оцифрованных услуг, протяжённости сетей, объёму вычислительных мощностей. Выходные параметры (изменение реальных доходов, сокращение бедности) либо отсутствуют, либо не связаны с цифровым контуром. Система, измеряющая входы, но не выходы, по определению разомкнута. Второе: институциональный демпфер строится по заимствованному образцу без учёта состояния технологического узла. Третье: региональная статистика не гармонизирована, что лишает регион общего измерительного контура.

Сопоставление официальных целей с расчётными траекториями обнаруживает расхождения. Цель Таджикистана по доле цифровой экономики 10 % ВВП к 2027 году не подтверждается расчётом ни в одном сценарии (базовый даёт 7,0 % к 2030 году). Цель Узбекистана по экспорту ИКТ-услуг в 5 млрд долларов США к 2030 году достижима только в кибернетическом сценарии (5,19 млрд); инерционная траектория даёт 2,59 млрд, то есть менее 52 % цели — и разрыв определяется не налоговым режимом, а ограничениями абсорбционной способности.

Ограничения исследования. Часть субиндикаторов для Кыргызстана, Таджикистана и особенно Туркменистана представляет экспертную оценку; выводы по Туркменистану предварительны. Индекс ИССК является потоковым и по построению занижает оценку зрелых систем — он не предназначен для ранжирования стран по уровню благополучия. Распределение весов 0,25 / 0,35 / 0,40 — теоретически обоснованная гипотеза; проверка по трём альтернативным наборам весов сохранила ранжирование для восьми из девяти экономик. Установленные соотношения носят характер структурных соответствий, а не доказанных причинно-следственных связей; гипотеза гибридного контура требует проверки методами каузального вывода на панельных данных.

7. ВЫВОДЫ

Проведённое исследование позволяет сформулировать семь основных выводов, каждый из которых сопровождается практическим предложением.

Вывод 1. Цифровая готовность не гарантирует социальной отдачи.

Данные региона за 2022–2025 годы опровергают допущение об автоматической связи между цифровой оснащённостью и уровнем жизни: Казахстан при индексе 0,9009 зафиксировал сокращение реальных доходов на 2,7 %, а Узбекистан при индексе 0,7999 сократил бедность с 14,0 % до 5,8 %. Структурный разрыв контура составил соответственно +0,868 и −0,190 пункта.

Предложение. *ввести в стратегическое планирование каждой страны обязательный индикатор социальной отдачи цифровизации — отношение прироста реальных доходов нижнего квинтиля к бюджетным расходам на цифровую трансформацию, с ежегодной публикацией наравне с показателями цифровой готовности.*

Вывод 2. Социальный результат определяется произведением, а не суммой коэффициентов контура.

Обнуление любого из четырёх звеньев — сенсорного, трансляционного, бюджетного или исполнительного — обнуляет весь результат независимо от

совершенства остальных. Инвестиции в самое развитое звено при наличии слабейшего не дают приращения отдачи.

Предложение. *перед утверждением любой программы цифровизации социальной сферы проводить диагностику слабейшего звена контура с раздельной оценкой четырёх коэффициентов и приоритетным финансированием звена с минимальным значением.*

Вывод 3. Гибридная архитектура контура эффективнее чисто цифровой при неполном цифровом охвате.

Опережающая отдача узбекистанской модели объясняется соединением цифрового реестра с низовым организационным звеном подворного обследования, компенсирующим дефицит цифрового разнообразия и охватывающим уязвимые домохозяйства вне цифрового периметра — прежде всего сельские и неформально занятые.

Предложение. *признать гибридную архитектуру самостоятельной проектной моделью, а не переходным состоянием, и закрепить её в методических документах; переход функций к цифровому звену осуществлять поэтапно, сохраняя избыточный канал подтверждения права не менее трёх лет.*

Вывод 4. Отсутствие социальных целевых индикаторов в актах о цифровизации — системный дефект региональной модели.

За исключением стратегии «Узбекистан — 2030» и постановления № ПП-358 социальная отдача цифровизации не выступает подотчётным параметром ни в одном акте региона. Закон Казахстана об ИИ, стратегия «Digital Qazaqstan», Концепция Кыргызстана устанавливают технологические и институциональные, но не социальные цели.

Предложение. *при пересмотре стратегических документов включить раздел социальных целевых индикаторов, содержащий как минимум целевое сокращение бедности за счёт цифровых механизмов, долю выплат, назначаемых без личного обращения гражданина, и средний срок от изменения положения домохозяйства до назначения поддержки.*

Вывод 5. Дефицит абсорбционной способности — основное долгосрочное ограничение региона.

Расходы на исследования и разработки в Узбекистане составляют 0,13 % ВВП против 4,91 % в Корее и 2,7 % в среднем по ОЭСР. При сохранении этого разрыва расширение инфраструктуры не приведёт к пропорциональному росту производительности: сигнал, поступающий в узел без способности к переработке, рассеивается.

Предложение. *установить траекторию наращивания расходов на исследования и разработки — не менее 0,5 % ВВП к 2028 году и 1,0 % к 2032 году, с приоритетом прикладных исследований предприятий и университетов, и ввести налоговый вычет по корпоративным расходам на R&D.*

Вывод 6. Замыкание социального контура даёт эффект без пропорционального роста бюджетных расходов.

Сценарное моделирование показало, что переход от инерционной к кибернетической траектории выводит из бедности к 2030 году 3,62 млн жителей региона против 0,89 млн — дополнительный эффект в 2,73 млн человек. Различие обеспечивается изменением коэффициентов связей, то есть организацией управления, а не наращиванием расходов.

Предложение. *ввести в бюджетный процесс механизм внутригодового реагирования — резервный фонд адресной поддержки не менее 0,3 % ВВП, расходуемый по автоматическим правилам на основании данных единого социального реестра без внесения изменений в закон о бюджете. Эта мера прямо устраняет бюджетную непроводимость (тип γ).*

Вывод 7. Управление требует перехода от отчётности о входах к измерению выходов контура.

Региональная модель построена как разомкнутая система: измеряются и вознаграждаются входные параметры (число услуг, протяжённость сетей, мощности), тогда как выходной параметр — изменение уровня жизни — не возвращается в контур управления. Именно эта разомкнутость, а не дефицит инфраструктуры или отставание законов, составляет главное ограничение социальной результативности цифровизации.

Предложение. *закрепить приоритет показателей социальной отдачи над показателями оснащённости в оценке ответственных органов, ежегодно публиковать доклад о социальных результатах цифровой трансформации и включать в задания по проектам цифровизации обязательный раздел об измеримом социальном эффекте. Методическая основа заложена постановлением № ПП-358; предлагается распространить этот подход на все направления и все государства региона.*

Дополнительно, учитывая несопоставимость национальных методологий измерения бедности и доходов, целесообразно учредить Региональную обсерваторию цифровой и социальной трансформации Центральной Азии для согласования методологий и ежегодной публикации сопоставимых показателей по пяти государствам на базе существующих механизмов межгосударственных консультаций.

ЗАКЛЮЧЕНИЕ

Исследование исходило из эмпирического наблюдения, не объяснимого в рамках преобладающих подходов: в Центральной Азии государство с наивысшим уровнем цифровой готовности продемонстрировало сокращение реальных доходов, тогда как государство с более скромной инфраструктурой обеспечило рекордное для региона сокращение бедности при одновременном снижении неравенства.

Предложенная социально ориентированная модель, выделяющая благосостояние в самостоятельный узел и вводящая понятие социального контура обратной связи, позволяет объяснить это наблюдение. Социальная отдача цифровой трансформации определяется произведением четырёх коэффициентов, вследствие чего слабейшее звено определяет результат независимо от совершенства остальных. Индекс скорости социальной обратной связи и показатель структурного разрыва позволяют различить две внешне схожие ситуации: насыщение зрелой системы и патологическую разомкнутость системы развивающейся.

Практическая значимость определяется результатом моделирования: замыкание социального контура выводит из бедности дополнительно 2,73 млн жителей региона к 2030 году относительно инерционной траектории, причём за счёт изменения архитектуры управления, а не наращивания расходов. Дальнейшие направления исследования включают эмпирическую калибровку весов индекса на расширенной выборке, проверку гипотезы гибридного контура методами каузального вывода и изучение возможностей искусственного интеллекта для автоматического выявления нуждающихся домохозяйств до момента их обращения за поддержкой.

СПИСОК ЛИТЕРАТУРЫ

Законы и акты Президента Республики Узбекистан

1. Указ Президента Республики Узбекистан от 5 октября 2020 года № УП-6079 «Об утверждении Стратегии "Цифровой Узбекистан — 2030" и мерах по её эффективной реализации».
2. Указ Президента Республики Узбекистан от 11 сентября 2023 года № УП-158 «О Стратегии "Узбекистан — 2030"».
3. Указ Президента Республики Узбекистан от 14 октября 2024 года № УП-157 «О дополнительных мерах по поддержке предприятий, занимающихся экспортной деятельностью в сфере цифровизации».
4. Постановление Президента Республики Узбекистан от 14 октября 2024 года № ПП-358 «Об утверждении Стратегии развития технологий искусственного интеллекта до 2030 года».
5. Постановление Президента Республики Узбекистан от 16 февраля 2026 года № ПП-64 «О дополнительных мерах по развитию креативной экономики».
6. Постановление Президента Республики Узбекистан от 8 июля 2026 года № ПП-257 «О мерах по дальнейшему развитию системы "Цифровое правительство"».

Иные нормативные акты государств региона и Европейского союза

7. Закон Республики Казахстан от 17 ноября 2025 года № 230-VIII «Об искусственном интеллекте».
8. Указ Президента Республики Казахстан от 13 октября 2025 года № 1042 «Об утверждении Общенационального плана мероприятий по реализации Послания Главы государства народу Казахстана от 8 сентября 2025 года».
9. Указ Президента Республики Казахстан от 9 июня 2026 года «Об утверждении Общенациональной стратегии масштабной цифровизации и тотального внедрения технологий искусственного интеллекта "Digital Qazaqstan" до 2029 года».
10. Концепция цифровой трансформации Кыргызской Республики на 2024–2028 годы: утв. решением Кабинета Министров Кыргызской Республики.
11. Указ Президента Республики Таджикистан от 3 апреля 2021 года № 147 «О создании Агентства инноваций и цифровых технологий при Президенте Республики Таджикистан».
12. Постановление Правительства Республики Таджикистан от 28 марта 2022 года № 127 «О перечне государственных услуг, оказываемых в электронной форме».
13. Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) // Official Journal of the European Union. — 2024.
14. Regulation (EU) 2022/1925 on contestable and fair markets in the digital sector (Digital Markets Act) // Official Journal of the European Union. — 2022.

Научная литература и источники данных

15. Гулямов С. С., Гулямов С. С. Кибернетико-экономическая теория цифровой трансформации: монография. — Ташкент — Москва, 2026. — 638 с.
16. Эшби У. Р. Введение в кибернетику / пер. с англ. — М.: Издательство иностранной литературы, 1959. — 432 с.
17. Винер Н. Кибернетика, или Управление и связь в животном и машине / пер. с англ. — М.: Наука, 1983. — 344 с.

18. Бир С. Мозг фирмы / пер. с англ. — М.: Радио и связь, 1993. — 416 с.
19. Глушков В. М. Основы безбумажной информатики. — М.: Наука, 1987. — 552 с.
20. Хайек Ф. А. Использование знания в обществе // Индивидуализм и экономический порядок. — М.: Изограф, 2000. — С. 89–105.
21. Кейнс Дж. М. Общая теория занятости, процента и денег / пер. с англ. — М.: Гелиос АРВ, 2012. — 352 с.
22. Ashby W. R. Requisite variety and its implications for the control of complex systems // Cybernetica. — 1958. — Vol. 1, No. 2. — P. 83–99.
23. Beer S. The Heart of Enterprise. — Chichester: John Wiley & Sons, 1979. — 596 p.
24. Comin D., Mestieri M. If technology has arrived everywhere, why has income diverged? // American Economic Journal: Macroeconomics. — 2018. — Vol. 10, No. 3. — P. 137–178.
25. Acemoglu D. The simple macroeconomics of AI. NBER Working Paper No. 32487. — Cambridge, MA: NBER, 2024. — 42 p.
26. Zuboff S. The Age of Surveillance Capitalism. — New York: PublicAffairs, 2019. — 704 p.
27. Nardo M., Saisana M., Saltelli A. et al. Handbook on Constructing Composite Indicators: Methodology and User Guide. — Paris: OECD Publishing, 2008. — 162 p.
28. United Nations Department of Economic and Social Affairs. UN E-Government Survey 2024: Accelerating Digital Transformation for Sustainable Development. — New York: United Nations, 2024. — 328 p.
29. World Bank. World Development Indicators Database. — Washington, D.C.: World Bank, 2025.
30. International Monetary Fund. World Economic Outlook Database, October 2025. — Washington, D.C.: IMF, 2025.
31. International Telecommunication Union. Measuring Digital Development: Facts and Figures 2025. — Geneva: ITU, 2025. — 42 p.
32. Draghi M. The Future of European Competitiveness: Report to the European Commission. — Brussels: European Commission, 2024. — 393 p.
33. Центр экономических исследований и реформ при Администрации Президента Республики Узбекистан. Анализ уровня бедности и доходов населения Республики Узбекистан по итогам 2025 года. — Ташкент, 2025.
34. Бюро национальной статистики Агентства по стратегическому планированию и реформам Республики Казахстан. Среднедушевые номинальные и реальные денежные доходы населения, IV квартал 2025 года. — Астана, 2026.
35. BlueKraft Digital Foundation. A Quantitative Assessment of India's Direct Benefit Transfer System: A Policy Analysis. — New Delhi, 2025. — 68 p.

CORPUS-BASED LINGUOSTATISTICAL ANALYSIS OF UZBEK SCIENTIFIC DISCOURSE: LEXICAL UNITS AND DISCOURSE MARKERS

Ibodullayeva Kunduz Rustam qizi
Lecturer of the Department of Computational Linguistics and Applied Linguistics
ibodullayeva_k@nuu.uz
National University of Uzbekistan
https://doi.org/10.5281/zenodo.23136366

Annotatsiya. *Maqolada o'zbek ilmiy diskursining leksik tarkibi va diskursiv xususiyatlari korpusga asoslangan lingvostatistik yondashuv asosida tadqiq etildi. Tadqiqot materiali O'zbekiston Milliy universiteti ilmiy nashrlaridan tanlangan 150 ta maqoladan shakllantirilgan 233475 tokenli korpusdan iborat. Chastota va nisbiy chastota tahlili asosida yuqori chastotali leksik birliklar hamda ilmiy argumentatsiyada faol qo'llanuvchi diskurs markerlari aniqlanib, funksional guruhlar bo'yicha tasniflandi. Tadqiqot natijalari korpusga asoslangan lingvostatistik yondashuv o'zbek ilmiy diskursining leksik va diskursiv xususiyatlarini empirik jihatdan tavsiflashda samarali vosita ekanligini ko'rsatadi.*

Abstract. *The article examines the lexical composition and discursive features of Uzbek scientific discourse based on a corpus-based linguostatistical approach. The research material consists of a 233,475-token corpus formed from 150 articles selected from scientific publications of the National University of Uzbekistan. Based on frequency and relative frequency analysis, high-frequency lexical units and discourse markers actively used in scientific argumentation were identified and classified into functional groups. The research results show that the corpus-based linguostatistical approach is an effective tool for empirically describing the lexical and discursive features of Uzbek scientific discourse.*

Kalit so'zlar: *o'zbek ilmiy diskursi, korpus lingvistikasi, diskurs markerlari, leksik birliklar, ilmiy argumentatsiya, korpus tahlili, lingvostatistik tahlil, chastota tahlili, kollokatsiya, leksik boylik, avtomatik tahlil, korpus asosida yondashuv.*

Keywords: *Uzbek scientific discourse, corpus linguistics, discourse markers, lexical units, scientific argumentation, corpus analysis, linguostatistical analysis, frequency analysis, collocation, lexical richness, automatic analysis, corpus-based approach.*

In modern linguistics, scientific discourse is regarded as a distinct communicative system that ensures the process of creating, systematizing, and communicating scientific knowledge to the scientific community. Its linguistic features are manifested through the frequency of use of language units, functional-semantic possibilities, terminological composition, syntactic structure, and methods of logical and argumentative expression of thought. In this regard, the study of scientific discourse is one of the important directions not only for theoretical linguistics, but also for applied and computer linguistics [Abdurahmanova M., Rayimjonova M., 2021]

Terminological units hold a distinct place within the lexical-semantic structure of scientific discourse. By precisely naming, delimiting, classifying, and contextually defining scientific concepts, terms ensure the semantic accuracy of scientific information. Concurrently, discourse markers also have a significant functional importance in scientific texts. Specifically,

units such as *for instance (masalan), including (jumladan), in particular (xususan), as well as (shuningdek), because (chunki), but (ammo), however (biroq), thus (shunday qilib), therefore (demak),* and *as a result (natijada)* serve to express the semantic-pragmatic relationships between ideas within the text. They establish discursive relations such as exemplification and clarification, addition, cause-and-effect, contrast, and conclusion, thereby ensuring the logical coherence and argumentative integrity of the scientific exposition. Determining the frequency of use and contextual distribution of these units on a corpus basis makes it possible to empirically identify the primary linguistic means for organizing thought in scientific discourse.

In world linguistics, the works of researchers such as R. Bailey, R. Allan, R. Oakman, V. Cherubini, R. Merritt, V. Kuraskievich, and G. Altmann are focused on the statistical and mathematical modeling of linguistic phenomena. In Russian linguistics, the research of V. Petrov, Aleksandrov, V. A. Bogoroditsky, A. M. Peshkovsky, and A. Bulakhovsky used statistical methods to study the sound composition and linguistic features of texts of various genres. [Urinbayeva D., 2022: 25]

In Turkology, the creation of frequency, alphabetic-frequency, and inverse dictionaries based on materials from Kazakh, Kyrgyz, Karakalpak, Azerbaijani, and Turkish languages demonstrates the practical significance of statistical methods in the study of Turkic languages.

In Uzbek linguistics, a distinct scientific tradition has been established in the field of linguostatistics. The research of V. V. Reshetov, A. K. Borovkov, V. M. Popov, I. A. Kissen, R. Kungurov, A. Tikhonov, S. Karimov, F. Abdullayev, M. Imomnazarov, N. Yakubova, E. Khojaniyozov, S. Muhammadov, S. Musakhodjaev, S. Rizaev, S. Mukhamedova, and A. P. Ibragimov has illuminated various aspects of statistical analysis. M. Yuldashev, S. Mukhamedova, and M. Saparniyazova associate linguostatistical analysis with observing the usage patterns of language units and statistically processing the results of these observations. [Yo'ldoshev M., 2021:11] Furthermore, the research of D. Urinbaeva demonstrates the potential of linguostatistical methods in modeling lexical-morphological structure and the language system. [Urinbayeva D., 2022: 25]

A significant portion of existing research is based on literary and journalistic texts, while the corpus-based linguostatistical study of scientific discourse has not been systematically explored in sufficient detail. Specifically, the comprehensive analysis of the frequency, distribution, and functional characteristics of lexical units and discourse markers in Uzbek scientific discourse, based on authentic corpus materials, emerges as a pressing scientific problem.

The aim of the research is to provide a quantitative characterization of the lexical composition and discursive organization of Uzbek scientific discourse on a corpus basis, and to determine the functional activity of high-frequency lexical units and discourse markers.

The research material consisted of a small corpus in TXT format, which was compiled from 150 scholarly articles selected from the scientific publications of the National University of Uzbekistan [Abdurakhmonova, Alisher, & Toirova, 2022]

ID	Muallif	Davr (yoki yil)	Janr	Uslub	Matn turi	Mavzu (Tematika)	Manba (Nashr)	Yozuv (Alifbo)	Til	Hajm (Token)
1	Inomjon Azimov, Nilufar Abduraxmonova, Ochilbek Yulbarsov	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	Yangi so‘z va atamalarning o‘zbekcha muqobilini sun’iy intellekt asosida tahlil qilishning	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1748
2	Shoira Amonturdiyeva	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	Og‘zaki va yozma diniy matnlarning fonetik hamda	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1118
3	Mohigul Asronova	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	O‘zbek tilida rasmiy-iqtisodiy diskursning lingvistik modeli va	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1166
4	Axadjon Ataboyev	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	Ijtimoiy ongning tuzilishi va rivojlanishi xususida	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1457
5	Maxmud Babayarov	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	O‘zbek tilshunosligida sport leksikasining o‘rganilishi	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1459
6	Dildora Begmatova	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	“Devonu lug‘atit-turk” maqollarida hayvon nomlarining ishtiroki	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1568
7	Aidaxon Bumatova	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	Mumtoz g‘azallarda yor obrazining talqini	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1612
8	Dilorom Karimova	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	Odamning lisoniy manzarasida lingvomadaniyat va antroposentrik birliklarning mohiyati	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1723
9	Rayhona Kenjayeva	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	Hikmat istilohining turfa ma’nolari	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1878
10	Nilufar Maxsudova	Mustaqillik davri	ilmiy maqola	ilmiy	yozma	Xalq maqollarida ingliz milliy qiyofasi	“O‘ZMU xabarlari” (2024)	lotin	o‘zbek	1801
11	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Ingliz tilida aspektual ma’no	“O‘ZMU xabarlari”	[illegible]	[illegible]	1452

Table 1. Metadata of the corpus compiled for the study of Uzbek scientific discourse

As the table indicates, the research corpus is composed of scientific articles, with each text described by metadata such as author, publication year, genre, style, text type, topic, source, script type, language, and token volume. This metadata provides the basis for systematizing the corpus, parametrically grouping the texts, and conducting subsequent linguostatistical analyses. The analysis revealed that the corpus, comprising 150 scientific articles, has a total size of 233,475 tokens. The 40,801 unique token forms within the corpus resulted in a total Type-Token Ratio (TTR) of 17.48%. On average, each article contains 1,556.50 tokens and 95.72 sentences. The average sentence length of 16.26 tokens indicates that ideas in scientific articles are expressed through relatively broad syntactic units. However, interpreting this indicator as a universal syntactic parameter of scientific discourse requires methodological caution, as sentence length can be influenced by factors such as the article's topic, genre, authorial style, and scientific field.

Indicator	Value
Number of articles	150
Jami token	233 475
Total tokens	40 801
Total TTR	40 801
Total sentences	17,48 %
Average tokens per article	1556,50
Average sentence per article	95,72
Average sentence length	16,26 token

Table 2. General statistical overview of the corpus

This result indicates that in scientific articles, ideas are expressed through relatively broad syntactic units. However, interpreting this indicator as a universal syntactic norm of scientific discourse requires a comparative study with other genres and corpora.

High-frequency functional and semantic units in the corpus were identified, and their usage features in scientific discourse were studied [Abdurakhmonova, N., Nurmanova, S., 2026]. Specifically, the frequency of units such as *va* (and), *bu* (this), *bilan* (with), *ham* (also), *o‘zbek*

(Uzbek), *tahlil* (analysis), and *til* (language) was determined. Discourse markers that shape scientific argumentation, such as *masalan* (for example), *shuningdek* (also), *chunki* (because), *ammo* (but), *jumladan* (including), *xususan* (in particular), *biroq* (however), *demak* (therefore), *shuning uchun* (for this reason), and *natijada* (as a result), were analyzed as a separate group.

№	Unity (Birlik)	Frequency (Chastota)	Share (%) (Ulushi) (%)
1	va	7413	3,17
2	bu	1945	0,83
3	bilan	1851	0,79
4	ham	1748	0,75
5	bir	1 304	0,56
6	o‘zbek	1 216	0,52
7	uchun	1 089	0, 47
8	yoki	943	0,40
9	tahlil	837	0,36
10	ushbu	810	0,35
11	mumkin	799	0,34
12	til	788	0,34
13	uning	787	0,34
14	orqali	779	0,33
15	sifatida	768	0,33

Table 3. High-frequency lexical units

One of the key features of scientific discourse is the clear expression of logical relationships between ideas. In the corpus, units that perform this function were observed to occur with significant frequency.

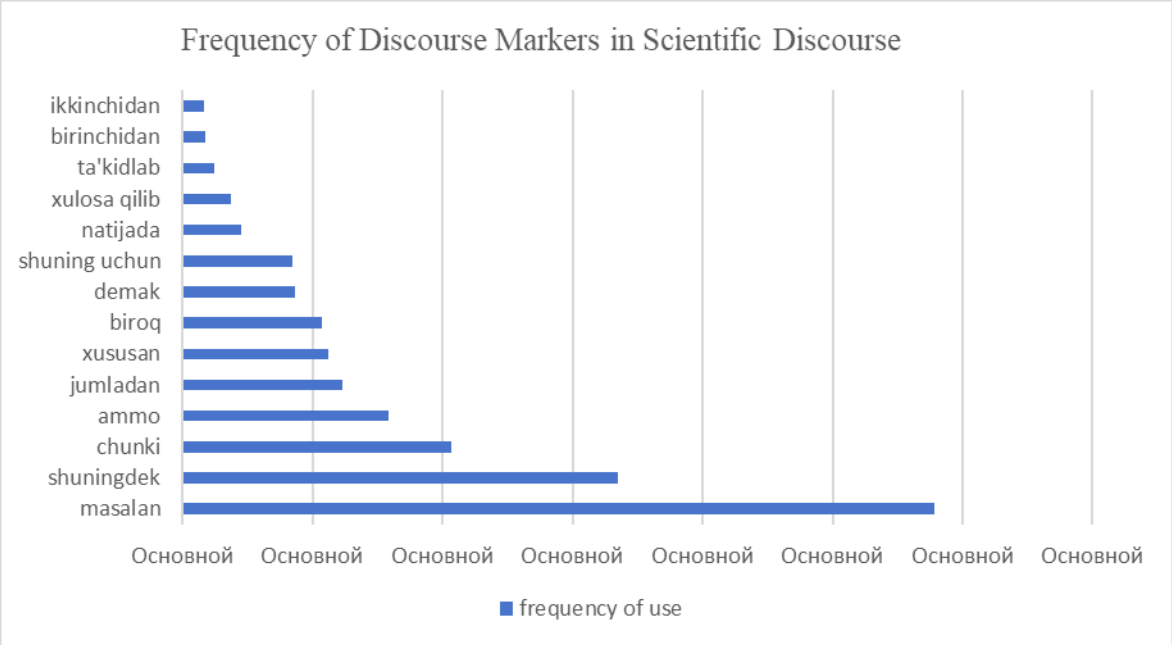


Table 4. Discourse Markers in Scientific Discourse

The results of the corpus analysis showed that the frequency of use of different functional units in scientific discourse is not uniform. Among the units examined, those performing the function of exemplification and clarification recorded the highest frequency, with 813 occurrences (42.1%). This group includes the Uzbek discourse markers masalan (for example), jumladan (including), and xususan (in particular). Their high frequency of use indicates a strong need in scientific texts to specify ideas and to clarify general statements or theoretical reasoning through factual examples.

By functionally grouping the given markers, some mechanisms of scientific argumentation can be seen.

	Functional group (Funksional guruh)	Discourse units (Diskursiv birliklar)	Sum frequency (Yig'indi chastotasi)
1	Exemplification and clarification (Misollash va aniqlashtirish)	Masalan, jumladan	813
2	Additional comment (Qo'shimcha fikr)	shuningdek	335
3	Cause and effect (Sabab-natija)	Chunki, shuning uchun, natijada	336
4	Contradiction (Qarama-qarshilik)	Ammo, biroq	265
5	Conclusion-drawing (Xulosalash)	Demak, xulosa qilib	123
6	Organization of arguments (Argumentlarni tartiblash)	birinchidan, ikkinchidan	33
7	Emphasis (Ta'kidlash)	ta'kidlab	24

Figure 5. Frequency of scientific discourse markers by functional groups

This distribution shows that the mechanisms of illustration and clarification are particularly active in the argumentative structure of scientific discourse. In this case, the author first outlines a general scientific idea and then concretizes it with an example, evidence, or a specific situation.

The results of the corpus analysis showed that there are certain patterns in the lexical composition and discursive organization of Uzbek scientific discourse. The predominance of functional units such as *va*, *bu*, *bilan*, *ham*, *uchun* in the structure of high-frequency units indicates the high need to syntactically connect semantic components and ensure the consistency of thoughts in the scientific text. At the same time, the active use of semantic units such as *o'zbek*, *tahlil*, *til* is explained by the fact that the corpus is formed from scientific texts of a linguistic direction. Analysis of discourse markers showed that scientific argumentation is

formed not on the basis of identical means, but through a system of units representing different functional relations.

The corpus-based linguostatistical analysis made it possible to characterize the lexical composition and discursive organization of Uzbek scientific discourse on the basis of empirical and quantitative indicators. The analysis of a corpus comprising 150 scientific articles, with a total size of 233,475 tokens, demonstrated not only the high frequency of functional units but also the active use of field-specific content vocabulary. The predominance of exemplification and clarification devices in the functional distribution of discourse markers indicates the active use of strategies for concretizing ideas, clarifying general statements through factual examples, and substantiating scientific claims in Uzbek scientific discourse.

The research results allow for the evaluation of the linguostatistical approach as an effective practical method for determining the lexical and functional features of Uzbek scientific discourse. At the same time, the fact that the corpus is limited to specific scientific publications limits the possibility of generalizing the results. In future research, a comparative analysis of the corpus of various fields of science and scientific genres will allow for a more precise definition of the general and specific linguistic patterns of scientific discourse.

REFERENCES

1. Mengliev, D., Barakhnin, V., & Abdurakhmonova, N. (2021). Development of intellectual web system for morph analyzing of uzbek words. Applied Sciences, 11(19), 9117.
2. Муҳамедов С. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дис.… канд. филол. наук. -Ташкент, 1980. – С.25.20
3. Urinbayeva D.B "Xalq og'zaki ijodi asarlari matnining janriy-lisoniy va lingvostatistik tadqiqi" "Qamar media" nashriyoti Toshkent – 2022
4. Yo'ldoshev M, Muhammedova. S, Saparniyozova. M "Matn lingvistikasi" Toshkent "Ishonchli hamkor" 2021
5. Кириллов М.А. Лингвостатический анализ художественного текста (На материале коротких Ф.С. Фицжеральда) 2002
6. Йўлдошев Б. Матнни ўрганишнинг лингвостатистик методлари. Самарқанд, 2008. –29.
7. Abdurahmanova, M., & Rayimjonova, M. (2021). KORPUS LINGVISTIKASIDA TARJIMASHUNOSLIK MASALASI. *COMPUTER LINGUISTICS: PROBLEMS, SOLUTIONS, PROSPECTS*, *1*(1).
8. Abdurakhmonova, N., & Rajabova, K. (2025, September). Stages of Creating Uzbek Language Multimedia Corpus. In *2025 10th International Conference on Computer Science and Engineering (UBMK)* (pp. 1628-1633). IEEE.
9. Abduraxmonova, N. (2025). LEKSIKOGRAFIYADA TERMINOLOGIK LUG'ATLARNING O'RNI VA AHAMIYATI. *YANGI DAVR ILM-FANI: INSON UCHUN INNOVATSION G 'OYA VA YECHIMLAR*, *1*(9), 47-49.
10. Abdurakhmonova, N., Alisher, I., & Toirova, G. (2022, September). Applying web crawler technologies for compiling parallel corpora as one stage of natural language processing. In *2022 7th International Conference on Computer Science and Engineering (UBMK)* (pp. 73-75). IEEE.
11. Abdurakhmonova, N., Nurmanova, S., Kozimjonova, M., Otamurodova, S., Toshpulatova, M., & Abidjanov, A. (2026, June). Classification of Uzbek Banking Texts by Document Types. In *2026 IEEE 27th International Conference of Young Professionals in Electron Devices and Materials (EDM)* (pp. 1-5). IEEE.

ELEKTRON LUG'AT BAZASI UCHUN SO'Z BIRIKMASI – FRAZEOLOGIK BIRLIK JUFTLIGINING IDENTIFIKATSIYA NAMUNALARI

Iskandarova Aybibi Rakhimbayevna
Associate Professor, Ph.D. (Philology)
Department of Computational Linguistics and applied Linguistics
National University of Uzbekistan
https://doi.org/10.5281/zenodo.23136383

Annotatsiya. *Ushbu maqolada badiiy asarlar lingvistik bazasi uchun tarjima birliklarini kodlash, asliyatda erkin so'z birikmasi, tarjimada frazema shaklida xoslanish holatlari, ularning funksional ekvivalentligi haqida mulohaza yuritiladi. Shuningdek, tarjima birliklarining elektron lug'at bazasi uchun identifikatsiyalashning ahamiyati haqida fikr bildiriladi.*

Kalit so'zlar: *elektron lug'at baza, tarjima birliklari, mashina tarjimasi, frazema, identifikatsiya namunalari, kodlash.*

Аннотация. *В данной статье рассматривается кодирование переводческих единиц для лингвистической базы художественных произведений, случаи свободного словосочетания в оригинале, обособления в форме фраземы в переводе, их функциональная эквивалентность. Также высказывается мнение о важности процесса идентификации для электронной словарной базы переводных единиц.*

Ключевые слова: *электронная словарная база, единицы перевода, машинный перевод, фразема, образцы идентификации.*

Abstract. *This article discusses the coding of translation units for the linguistic base of literary works, cases of free phrase formation in the original, phraseme formation in translation, and their functional equivalence. It also expresses an opinion on the importance of the identification process for the electronic dictionary database of translation units.*

Keywords: *electronic dictionary database, translation units, machine translation, phrase, identification samples.*

KIRISH

Frazeologizm, frazeologik birlik, frazema – ikki yoki undan ortiq so'zdan tashkil topgan, ma'noviy jihatdan o'zaro bog'liq, shaklan so'z birikmasi yoki gapga teng keladigan, yaxlitligicha ko'chma ma'noda qo'llanadigan barqaror bog'lanmalarning nomidir. Tilshunoslikda frazeologizmlar ibora, frazeologik birikma, turg'un birikma, barqaror birikma atamalari bilan yuritiladi.

ADABIYOTLAR TAHLILI VA METODOLOGIYA

Parallel korpus yaratishda biz N.Abduraxmonova tomonidan olib brilgan ishlarni tahlilga tortdik hamda uning ilmiy-nazariy metodologik nazariyasiga asoslandik [1, 2, 3, 4, 5, 6, 7].

Shuningdek, Sh.Rahmatullayev frazeologik birliklarning tarkibida kamida ikkita leksema qatnashishi; o'zaro sintaktik bog'langan bu leksemalar mohiyatiga ko'ra birikmaga yoki gapga teng bo'lishi; frazeologik ma'no, uni sintaktik birlik (nutq birligi) deb emas, balki semantik birlik (til birligi) deb qarashga olib kelishini asoslaydi: "Iboralarning asosiy qismini fe'l frazeologik birliklar tashkil etadi. Frazeologik birliklarning leksik tarkibida, fe'l so'z - komponentidan

tashqari, boshqa turkum so‘zi ham qatnashadi. Bunday komponentlar ko‘pincha ot, sifat, ravish turkumi bilan ham ifodalanadi [8]

V.A.Maslova frazeologizmlarning aksariyatida aniqlanishi lozim bo‘lgan milliy madaniyatning "izi" mavjudligi; frazeologik birliklarning ichki shaklida obyektiv olamning u yoki bu hodisalarini obrazli tasavvuri bo‘lgan madaniy informatsiya saqlanishi va u frazeologizmga madaniy-milliy bo‘yoq berishi; ularning tahlili jarayonida milliy-madaniy konnotatsiyani ochish muhimligini ta'kidlaydi [9]

Rus va o‘zbek badiiy matnlaridagi frazemalar orasida asliyat va tarjima variantida muqobil va muqobil bo‘lmagan shakllari uchraydi. Parallel korpus doirasidagi tadqiqotlarda asosan muqobil frazemalar, ularning lingvistik bazasiga doir nazariy va amaliy ishlar ko‘p bajarilgan. Tarjimada interpretatsiya holatlari ko‘p uchraydi, maqsad tiliga o‘girish jarayonida erkin birikmalarning frazemalarga aylanish holatlari sodir bo‘ladi:**"**Qiyomat" romanida ***задуманное дело*** birligi– ***ko‘nglimga tukkan ish*** deb tarjima qilingan, rus varianti sifatdosh+otdan hosil bo‘lgan erkin so‘z birikmasi. O‘zbek tilidagi variantida u frazemaga aylangan. "Ko‘nglimga tukkan ish" birligining frazeologik yadrosi–ibora: "ko‘ngliga tugmoq".

TAHLIL VA NATIJALAR

Maqolamizda lingvistik baza uchun maqsad tilida frazemaga aylangan birliklar uchun identifikatsiya namunalarini tahlil qildik. Rus tilida erkin so‘z birikmasi, tarjima varianti frazema namunasi bo‘lgan birliklarni lingvistik baza uchun quyidagicha teglaymiz:

1. *Звуки реального мира* - olamning butun qiylu qoli.

И тут вновь ворвались в ее сознание ***звуки реального мира***

*Shunda yana uning qulog‘i ochildi****, olamning butun qiylu qolini*** *eshita boshladi*

Звуки реального мира–rus tilida erkin birikmasi, frazema emas. Grammatik modeli: Noun+Adjective+Noun. Tarjimada frazeologik birlikka aylangan: olamning butun qiylu qoli. Noun+Adjective+Phraseological Noun. Lingvistik baza uchun quyidagicha identifikatsiya bilan tizimlashtiramiz:

<unit lang= "ru"
<unit type= "erkin _ birikma"
pos= "noun">
Звуки реального мира
<unit type= "phraseologizm"
pos= "noun">
Olamning butun qiylu qoli
</unit>
type= "free"
type= "phraseological"

2.Ничего– quruq gap

Ничего inkor olmoshi bo‘lib, "hech narsa", "hechqisi yo‘q" kabi ma'nolarni anglatadi. Grammatik modeli: Pronoun. ***Quruq gap***–o‘zbek tilida frazeologik birlik. Ot frazema. Grammatik modeli: Adjective+Noun. Behuda va'da yoki mazmunsiz mulohaza ma'nosini anglatadi.

Lingvistik baza uchun quyidagicha identifikatsiya bilan tizimlashtiramiz:

<unit lang= "ru"
type= "word"
pos= "pronoun"

model= "Pronoun>

Ничего

</ unit>

<unit lang= "uz"

type= "phraseologism"

pos= "noun"

model= "Adjective+Noun">

Quruq gap

</ unit>

type= "word"

type= "phraseologism"

3.То все равно–ikki dunyo bir qadam

Я не могу их соединить, а это мне одно нужно. А если этого нет, то ***все равно****.*

Agar qovushtirolmasam – u vaqtda menga ***ikki dunyo bir qadam.***

То все равнонинг lug'aviy ma'nosi "baribir", "farqi yo'q", tarjimada "ikki dunyo bir qadam" iborasi shaklida o'girilgan. Rus variantining grammatik modeli: Paticle+Pronoun+Predicative. Bu ikki birlikning biri predikativ shaklga ega, ikkinchisi ibora. Ikki dunyo bir qadam–ot frazema. Grammatik modeli: Numeral+Noun+Numeral+Noun

<unit

lang= "ru"

pos= "predicative"

model= "Particle+ Pronoun+ Predicative"

То всё равно

</ unit>

</ source>

< target>

<unit

lang= "uz"

type = "phraseologism"

pos= "noun"

model= "Numeral+Noun+Numeral+Noun

Ikki dunyo bir qadam

</unit>

</target>

<equivalence type = "functional"/>

<translation_metod>phraseological_substitution</translation_metod> Ya'ni tarjimon ruscha erkin birikmani o'zbek tilida frazeologizm bilan almashtirgan.

4.Он бонвиван большой– rosa oladigan xilidan-da

Он бонвиван большой–erkin predikativ birikma, ibora emas. Grammatik modeli: Pronoun+Noun+Adjective. Kauf-safoni yaxshi ko'radigan odamlarga nisbatan ishlatiladi. "Rosa oladigan xilidan-da" o'zbek tilida turg'un birikma. Ot frazema. Grammatik modeli: Adverb +Verb (Participle)+ Noun+Particle

Lingvistik baza uchun quyidagicha identifikatsiya bilan tizimlashtiramiz

lang= "ru"

< unit type= "free_phrase"

pos= "predicative"
model= "Pronoun+ Noun+Adjective">
Он бонвиван большой
</ unit>
</ source>
< target>
<unit
lang= "uz"
type = "phraseologism"
pos= "noun"
model= "Adverb+ Participle+Noun+ Particle
Rosa oladigan xilidan-da
</unit>
</ target>
<equivalence type= "funksional"/>
Bu ikki birlikning ekvivalentlik turi–kontekstual ekvivalentlikdir.
5.Еще слово–bir og'iz so'z

- ***Еще слово****: во всяком случае, советую решить вопрос скорее. Нынче не советую говорить, – сказал Степан Аркадьич. – Поезжай завтра утром, классически, делать предложение, и да благословит тебя Бог – Yana* ***bir og'iz so'z****: har nima bo'lganda ham, masalani tezroq hal qilishing maslahat, bugun og'iz ochma, maslahat bermayman, – dedi Stepan Arkadich. – Ertaga ertalab butun dabdaba bilan bor-u, og'iz solaver, xudo yor bo'lsin...*

Еще слово–rus tilida frazeologik birlik sanalmaydi. Grammatik modeli: Adverb+Noun. Otli erkin birikma. Uning asardagi ekvivalenti– bir og'iz so'z–frazeologik birlik bo'lib, qisqa fikr, qisqa gap ma'nosini bildiradi. Grammatik modeli: Numeral +Noun+Noun.

< source>
< unit
lang= "ru"
type= "free_ phrase"
pos= "noun"
model= "Adverb+ Noun">
Еще слово
</unit>
</source>
<target>
<unit
lang: "uz"
type= "phraseologism"
pos= "noun"
model: "Numeral+Noun+Noun">
Bir og'iz so'z
</unit>
</ target>
<equivalence type= "functional"/>
Ekvivalentlik turi–kontekstual ekvivalentlik.

6. Желание иметь детей –bola tirnog'iga zorlik

*Его **желание иметь детей** она объясняла себе тем, что он не дорожил ее красотой.* - *Vronskiyning **bola tirnog'iga zorligini**: «U husnimning qadriga yetmaydi» deb tushunardi.* Желание иметь детей– erkin so'z birikmasi. Grammatik modeli: Noun+ Verb (infinitive) +Noun. Bola tirnog'iga zorlik– o'zbek tilida ot frazema. Semantik yaxlitlik, turg'unlik, obrazlilik kabi frazeologik belgilarga ega. Grammatik modeli: Noun+Noun+Noun.

< source>
< unit
lang= "ru"
type= "free_ phrase"
pos= "noun"
model= "Noun+Verb (Infinitive)+ Noun">
Желание иметь детей
</unit>
</source>
<target>
<unit
Lang: "uz"
type= "phraseologism"
pos= "Noun"
model: "Noun+Noun+Noun">
Bola tirnog'iga zorlik
</unit>
</ target>
<equivalence type= "functional"/>

7. Не имеет смысла–kurakda turmaydigan narsalar

*... но ведь то, что ты говоришь, **или не имеет смысла**, или имеет очень дурной смысл.*

*...lekin sen aytgan narsalar yo bema'ni narsalar yoki **kurakda turmaydigan narsalar.***

Не имеет смысла–inkor ma'nosini bildiruvchi predikativ birikma, muayyan harakat, fikr yoki vaziyatning mantiqiy, amaliy asosga ega emasligini anglatadi.Grammatik modeli:Particle+ Verb+Noun. Kurakda turmaydigan narsalar– ot frazema. Mantiqqa to'g'ri kelmaydigan ishlarga nisbatan qo'llaniladi. Grammatik modeli: Noun+Participle +Noun

< source>
< unit
lang= "ru"
type= "free_ phrase"
pos= "predicative"
model= "Particle+Verb+Noun">
Не имеет смысла
</unit>
</source>
<target>
<unit

lang: "uz"
type= "phraseologism"
pos= "Noun"
model: "Noun+Participle+Noun">
Kurakda turmaydigan narsalar
</unit>
</ target>
<equivalence type= "functional"/>
<translation method> phraseologization </translation_ method>

Elektron lug'atda mazkur juftlikni free_phrase hamda equivalence= erkin birikmaning iboraga aylanishi, kontekstual ekvivalent sifatida qayd qilish ilmiy jihatdan asoslidir.

XULOSA

Elektron lingvistik bazada frazemalarni identifikatsiya qilish, ularni boshqa til birliklaridan ajratish, strukturaviy va semantik xususiyatlarini tizimlashtirish, til birliklarini elektron muhitda qayta ishlash imkonini yaratuvchi muhim bosqichdir. Identifikatsiya jarayonida frazemaning shakliy tarkibi, grammatik modeli, lug'aviy ma'nosi, semantik yaxlitligi, variantlari va funksional xususiyatlarini maxsus belgilar orqali kodlash, mashina tarjimasida aniq va izchil, qayta ishlanadigan birlik sifatida shakllanishini ta'minlaydi. Ayniqsa, |rus-o'zbek lingvistik bazasida frazemalarni identifikatsiyalash tarjima ekvivalentlarini aniqlash, ularning korpusdagi qo'llanish chastotasini kuzatish, kontekstual holatlarini qiyoslash, avtomatik qidiruv imkoniyatlarini kengaytirishga xizmat qiladi.

Xulosa qilib aytganda, frazemalarni elektron lingvistik bazada identifikatsiyalash bazaning strukturaviy yaxlitligi, ma'lumotlarning izchilligi va lingvistik qiymatini ta'minlovchi asosiy mezonlardan biridir.

FOYDALANILGAN ADABIYOTLAR

1. Abdurakhmonova, N., Alisher, I., & Sayfulleyeva, R. (2022, September). MorphUz: Morphological analyzer for the Uzbek language. In 2022 7th International Conference on Computer Science and Engineering (UBMK) (pp. 61-66). IEEE.
2. Abduraxmonova, N. (2023). Korpus lingvistikasi. Toshkent: Globe.
3. Абдураҳмонова, Н. (2020). Замонавий корпусларнинг компьютер моделлари. Ўзбекистонда хорижий тиллар.-2020, 1(30), 50-58.
4. Abdurakhmonova, N. Z. (2019). A two-level morphological analysis of the uzbek corpus. In Иностранная филология. Социальная и национальная вариативность языка и литературы (pp. 425-430).
5. Abduraxmonova, N. (2021). O 'zbek tili elektron korpusining kompyuter modellari. Filologiya fanlari doktori (DSc) dissertatsiyasi avtoreferati.
6. Abdurakhmonova, N. (2021). Formal-Functional Models of The Uzbek Electron Corpus. ANGLISTICUM. Journal of the Association-Institute for English Language and American Studies, 10(8), 59-66.
7. Abduraxmonova, N. Z. Q., & Urazaliyeva, M. Y. (2022). O'ZBEK TILI ELEKTRON KORPUSIDA (http://uzbekcorpus. uz/) OG'ZAKI MATNLAR KORPUSINI YARATISHNING NAZARIY VA AMALIY MASALALARI. Academic research in educational sciences, 3(3), 644-650.
8. Ш. Раҳматуллаев. Ўзбек тилининг фразеологик луғати. Тошкент. 1992. 378 б.

9. Маслова В.А. Лингвокультурология: Учеб.пособия для студ. высщ. учеб. заведений. - М. Издательский центр "Академия", 2001. – С.82
10. Л. Толстой. Анна Каренина. Т. 2019. 1-4 қисм: 636 б; 5-8 қисм: 556 б.
11. Л. Толстой. Анна Каренина. Т. 2020. С.800.
12. Айтматов Ч. Плаха. Алма-Ата. 1987. –302 б.
13. Айтматов Ч. Қиёмат. Т. 2024.–334 б.

O‘ZBEK TILI NUTQ SINTEZATORI

Iskandarova Iroda Bakir qizi
O‘zbekiston milliy universiteti, Jurnalistika va filologiya fakulteti, Kompyuter lingvistikasi mutaxassisligi I bosqich magistranti
https://doi.org/10.5281/zenodo.23136405

Annotatsiya. *Mazkur maqolada o‘zbek tili nutq sintezatorining mohiyati, rivojlanish bosqichlari va kompyuter lingvistikasidagi ahamiyati yoritilgan. Nutq sintezatorining yozma matnni audio shaklga aylantirish imkoniyatlari, uning lingvistik va fonetik asoslari hamda sintezlangan nutqning tabiiyligini ta’minlash masalalari ko‘rib chiqilgan. Shuningdek, o‘zbek tili uchun nutq sintezatorini yaratishda tilning o‘ziga xos xususiyatlarini hisobga olish zaruriyati tahlil qilingan.*

Kalit so‘zlar: *Nutq sintezatori, o‘zbek tili, kompyuter lingvistikasi, tabiiy tilni qayta ishlash, nutq texnologiyalari, fonetik tahlil.*

Annotation. *This article examines the concept, development, and significance of the Uzbek speech synthesizer in computational linguistics. It discusses the ability of speech synthesizers to convert written text into audio, their linguistic and phonetic foundations, and the factors that contribute to the naturalness of synthesized speech. The article also highlights the importance of considering the specific linguistic features of Uzbek in developing speech synthesis technologies.*

Key words: *Speech synthesizer, Uzbek language, computational linguistics, natural language processing, speech technologies, phonetic analysis.*

Аннотация. *В статье рассматриваются сущность, развитие и значение синтезатора речи узбекского языка в компьютерной лингвистике. Раскрываются возможности преобразования письменного текста в аудио формат, лингвистические и фонетические основы синтеза речи, а также факторы, обеспечивающие естественность синтезированной речи. Особое внимание уделяется необходимости учета специфических особенностей узбекского языка при разработке технологий синтеза речи.*

Ключевые слова: *Синтезатор речи, узбекский язык, компьютерная лингвистика, обработка естественного языка, речевые технологии, фонетический анализ.*

Kirish

Matn bilan ishlashga mo‘ljallangan lingvistik dasturlar va mobil ilovalarga talab hamda ehtiyoj tobora kuchaymoqda. Ana shunday dasturlar sirasida nutq sintezatori alohida ahamiyatga egadir [3]. Bugungi kunda axborot-kommunikatsiya texnologiyalarining jadal rivojlanishi natijasida inson va kompyuter o‘rtasidagi tabiiy muloqotni ta’minlashga xizmat qiluvchi nutq texnologiyalariga bo‘lgan ehtiyoj ortib borayotgan bo‘lib, nutq sintezatori yozma matnni tabiiy va tushunarli inson nutqiga aylantirish imkonini berishi bilan ta’lim, axborot xizmatlari, elektron lug‘atlar, mobil ilovalar hamda boshqa raqamli platformalarda keng qo‘llanish imkoniyatiga ega. O‘zbek tili uchun nutq sintezatorini yaratish jarayonida esa tilning fonetik, fonologik, morfologik va leksik xususiyatlarini hisobga olish muhim bo‘lib, ayniqsa so‘zlarning to‘g‘ri talaffuzi, bo‘g‘in tuzilishi, urg‘u, intonatsiya va tovushlarning o‘zaro ta’siri kabi lingvistik hodisalarni modellashtirish sintez qilingan nutqning tabiiyligi va aniqligini ta’minlashda muhim omil

hisoblanadi [7]. Shu sababli o‘zbek tilining o‘ziga xos lingvistik xususiyatlarini chuqur o‘rganish, nutqiy ma’lumotlarni tizimlashtirish va zamonaviy raqamli texnologiyalar asosida qayta ishlash orqali o‘zbek tili nutq sintezatorining lingvistik hamda texnologik asoslarini ishlab chiqish bugungi kunning dolzarb ilmiy-amaliy masalalaridan biri sifatida namoyon bo‘lmoqda.

Asosiy qism

Nutq sintezatori-bu muayyan bir tildagi matnlarni audio shaklda o‘qib beruvchi dastur [6]. Mazkur texnologiya yozma ko‘rinishdagi axborotni inson nutqiga yaqin bo‘lgan tovushli shaklga aylantirish orqali foydalanuvchiga ma’lumotni eshitish imkonini berib, ta’lim, elektron lug‘atlar, mobil ilovalar, avtomatlashtirilgan axborot tizimlari va boshqa raqamli xizmatlarda qo‘llanilishi bilan zamonaviy til texnologiyalarining muhim yo‘nalishlaridan birini tashkil etadi. O‘zbek tili nutq sintezatorini yaratishda esa umumiy texnologik jarayon bilan bir qatorda, o‘zbek tilining fonetik va fonologik tizimi, tovushlarning talaffuz xususiyatlari, bo‘g‘in tuzilishi, so‘z urg‘usi, intonatsiya hamda yozma shaklning og‘zaki nutqqa aylanish jarayonidagi o‘ziga xos jihatlarini hisobga olish talab etiladi. Shu jihatdan qaralganda, o‘zbek tili nutq sintezatorining samarali ishlashi uchun nafaqat yetarli hajmdagi nutqiy ma’lumotlar bazasini shakllantirish, balki ushbu ma’lumotlarni lingvistik jihatdan qayta ishlash, tartibga solish va kompyuter tomonidan to‘g‘ri talqin qilinadigan shaklga keltirish ham muhim ahamiyat kasb etadi.

Tabiiy tillarni qayta ishlash nafaqat lisoniy, balki nutqiy faolyat bilan bog‘liq muammolarni hal etishda ham o‘ziga xos o‘rin tutadi. Bunda, ayniqsa, nutq sintezatorlari muhim vosita sifatida xizmat qiladi [4]. Nutqni tushunishda ham, nutqni ishlab chiqarishda ham hozirgi nutq texnologiyasidagi tendensiya bilimlarni aniq shakllantirishdan qochish va tizimni rivojlantirishga yordam beradigan avtomatik usullardan foydalanish muhim ahamiyatga ega [1], shu sababli o‘zbek tili nutq sintezatorini ishlab chiqishda lingvistik bilimlarni zamonaviy avtomatik va kompyuter texnologiyalari bilan uyg‘unlashtirish, katta hajmdagi nutqiy ma’lumotlardan foydalanish hamda ularni qayta ishlash orqali tabiiy va tushunarli sun’iy nutqni shakllantirish muhim vazifalardan biri hisoblanadi.

1980-1990 yillar davomida sintezlangan nutqning tabiiyligi va tushunarliligini yaxshilashda sezilarli yutuqlarga erishildi. Bu lingvistik modellarni takomillashtirish, prosodiyani (ritm, intonatsiya va stress) o‘z va odamga o‘xshash ovozlarni yaratish uchun mashinani o‗rganish usullaridan foydalanishni o‘z ichiga olgan [2], natijada nutq sintezatorlari tomonidan hosil qilinadigan sun’iy nutqning inson nutqiga yaqinligi tobora ortib, ushbu texnologiyalarning amaliy qo‘llanish doirasi ham kengayib bordi. Keyingi bosqichlarda kompyuter texnologiyalari va mashinani o‘rganish algoritmlarining rivojlanishi nutq sintezida tovushlarning ketma-ketligini, so‘zlarning talaffuzini, urg‘u va intonatsiyani aniqroq modellashtirish imkonini berib, tabiiy eshitiladigan nutq yaratish jarayonini sezilarli darajada takomillashtirdi. Shu bilan birga, har bir tilning fonetik, fonologik va grammatik xususiyatlari turlicha bo‘lgani sababli, boshqa tillar uchun ishlab chiqilgan sintezator modellarini o‘zbek tiliga bevosita tatbiq etish kutilgan natijani bermasligi mumkin, bu esa o‘zbek tilining o‘ziga xos xususiyatlarini hisobga olgan holda maxsus nutq sintezatori yaratishni taqozo etadi [8].

Bugungi kunga qadar kompyuter lingvistikasi sohasida bir qator ilmiy va amaliy natijalarga erishildi: tabiiy tilda avtomatik tarjima tizimi yaratildi, matndagi ma’lumotlarning avtomatik qidiruv tizimi ishlab chiqildi, og‘zaki nutqning avtomatik analizi va sintezi tizimi yaratildi, bir qator lingvistik muammolarni hal etuvchi kompyuter dasturlari ishlab chiqildi, inson va mashina (kompyuter) muloqoti optimallashtirildi, tabiiy tilni qayta ishlash (Natural Language Processing) tizimi shakllantirildi [5]. Mazkur natijalar tilshunoslik va axborot

texnologiyalarining o‘zaro integratsiyasi natijasida tabiiy tilni kompyuter vositasida qayta ishlash, tahlil qilish va hosil qilish imkoniyatlari kengayib borayotganini ko‘rsatib, nutq sintezatorlarining ham zamonaviy kompyuter lingvistikasi doirasidagi muhim yo‘nalishlardan biriga aylanganini ko‘rsatadi. O‘zbek tilida bunday texnologiyalarni rivojlantirish yo‘lida yaratilgan "UzNutq Sintezator" dasturi yozma matnni avtomatik ravishda nutqqa aylantirish imkoniyatini ta'minlab, o‘zbek tilining raqamli muhitdagi qo‘llanish doirasini kengaytirishga xizmat qiluvchi amaliy ishlanmalardan biri sifatida ahamiyat kasb etadi [9]. Shu nuqtayi nazardan, o‘zbek tili nutq sintezatorini yanada takomillashtirish uchun mavjud lingvistik va texnologik yechimlarni rivojlantirish, o‘zbek tilining fonetik, leksik va grammatik xususiyatlarini hisobga olish hamda tabiiy va tushunarli sun'iy nutq hosil qilishga xizmat qiluvchi ma'lumotlar bazasini shakllantirish muhim vazifalardan biri hisoblanadi.

Xulosa

Xulosa qilib aytganda, o‘zbek tili nutq sintezatori zamonaviy kompyuter lingvistikasi va tabiiy tilni qayta ishlash texnologiyalarining muhim yo‘nalishlaridan biri bo‘lib, yozma matnni inson nutqiga yaqin bo‘lgan tabiiy va tushunarli audio shaklga aylantirish imkonini beradi. Nutq sintezatorlarining rivojlanishi davomida lingvistik modellar, fonetik tahlil, prosodik xususiyatlar, mashinani o‘rganish va zamonaviy sun'iy intellekt texnologiyalarining takomillashuvi natijasida sintezlangan nutqning tabiiyligi va aniqligi sezilarli darajada oshdi. O‘zbek tilida nutq sintezatori yaratish jarayonida esa tilning fonetik, fonologik, leksik va grammatik xususiyatlarini, xususan, so‘zlarning talaffuzi, bo‘g‘in tuzilishi, urg‘u va intonatsiya kabi jihatlarni hisobga olish muhim bo‘lib, bunday yondashuv sintez qilingan nutqning tabiiyligini ta'minlashga xizmat qiladi. O‘zbek tilida yaratilgan "UzNutq Sintezator" kabi dasturlar ushbu yo‘nalishda muhim amaliy qadam bo‘lib, kelgusida nutqiy ma'lumotlar bazasini kengaytirish, lingvistik resurslarni boyitish va zamonaviy sun'iy intellekt usullaridan samarali foydalanish orqali o‘zbek tili nutq sintezatorlarining imkoniyatlarini yanada rivojlantirish mumkin. Shu sababli o‘zbek tili uchun sifatli, tabiiy va foydalanuvchi ehtiyojlariga mos nutq sintezatorini yaratish nafaqat kompyuter lingvistikasi rivoji, balki o‘zbek tilining raqamli muhitdagi mavqeini mustahkamlash va undan zamonaviy axborot-kommunikatsiya texnologiyalarida samarali foydalanish uchun ham muhim ilmiy-amaliy ahamiyatga ega.

FOYDALANILGAN ADABIYOTLAR

1. Carlson, R. (n.d.). *[Nutqni tushunish va ishlab chiqarishda nutq texnologiyalari]*. Department of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm, Sweden.
2. John Wiley & Sons Ltd. (n.d.). *[Nutq sintezi bo‘yicha material]*. Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, England. https://www.wiley.com
3. Kurbanova, M. A., & Abjalova, M. A. (2020). Raqamli iqtisodiyot davrida nutq sintezatorining zarurati. *"O‘zbekistonda ilm-fan va ta'lim" 3-son konferensiyasi*.
4. Kurbanova, M. A. (2022). O‘zbek nutq sintezatori va uning funksional imkoniyatlari. In *O‘zbek tilshunosligining zamonaviy yo‘nalishlari: Muammo va yechimlar* [Konferensiya materiallari]. Toshkent, O‘zbekiston.
5. Qurbonova, M., Maqsudov, B., & Isoyev, E. (2021). *"UzNutq Sintezator" kompyuter dasturi*. Toshkent.
6. O‘rinboyeva, R. T., & Allaberdiyeva, D. G. (n.d.). O‘zbek dialektal nutq sintezatorini yaratishda xorijiy tajribalar tahlili.

7. Abdurakhmonova, N., & Rajabova, K. (2025, September). Stages of Creating Uzbek Language Multimedia Corpus. In *2025 10th International Conference on Computer Science and Engineering (UBMK)* (pp. 1628-1633). IEEE.
8. Abdurakhmonova, N., Astanova, G., Akhmedov, A., Mengliev, D., Ibragimov, B., & Abdullayev, A. (2025). Development of a lexical dataset and a rule-based algorithm for the analysis of Khorezm dialects of the Uzbek language. *Data in Brief*, 112216.
9. Mengliev, D., Abdurakhmonova, N., Kholmurodova, I., Ibragimov, B., Latipova, G., & Kadirova, Z. (2025, October). Bridging dialectal variations in Uzbek texts: A comparative evaluation of modern approaches. In *AIP Conference Proceedings* (Vol. 3377, No. 1, p. 070001). AIP Publishing LLC.

TURIZM LEKSIKASINING N-GRAM MODELI VA GIPONIM-GIPERONIMIK MUNOSABATLARI

Adizova Nilufar Istamovna
Buxoro davlat universiteti O'zbek tilshunosligi
va jurnalistika kafedrasi dotsenti, f.f.f.d.(PhD)
nilufaradizova991@gmail.com
https://doi.org/10.5281/zenodo.23136419

Annotatsiya. *Kompyuter lingvistikasi doirasida leksik birliklar o'rtasidagi semantik munosabatlarni avtomatik aniqlash masalasi nazariy va amaliy jihatdan alohida ahamiyatga ega bo'lib, ular orasida giponim–giperonimik munosabatlar tilning semantik iyerarxiyasini ochib beruvchi muhim mexanizm sanaladi. Mazkur maqolada turizmga oid milliy-madaniy birliklarning n-gram modeli, giponim-giperonimik munosabatlari bilan bog'liq lingvistik xususiyatlari, semantik aloqalari hamda nutqda qo'llanish chastotasi, o'zaro birikuvchanligi va kontekstual xususiyatlarini aniqlashda n-gram modelining imkoniyatlari tadqiq etiladi.*

Kalit so'zlar: *turizm leksikasi, milliy-madaniy birliklar, n-gram modeli, korpus lingvistikasi, giponim, giperonim, leksik-semantik munosabat, terminologiya, semantik tahlil...*

Аннотация. *В рамках компьютерной лингвистики задача автоматического выявления семантических отношений между лексическими единицами имеет особое теоретическое и практическое значение. Среди них гипонимико-гиперонимические отношения рассматриваются как важный механизм, раскрывающий семантическую иерархию языка. В данной статье исследуются возможности n-граммной модели в выявлении лингвистических особенностей, семантических связей, частотности употребления в речи, сочетаемости и контекстуальных характеристик национально-культурных единиц, относящихся к туристической лексике, а также в изучении их гипонимико-гиперонимических отношений.*

Ключевые слова: *туристическая лексика, национально-культурные единицы, n-граммная модель, корпусная лингвистика, гипоним, гипероним, лексико-семантические отношения, терминология, семантический анализ…*

Abstract. *Within the field of computational linguistics, the task of automatically identifying semantic relations between lexical units is of particular theoretical and practical significance. Among these relations, hyponymic-hyperonymic relations are regarded as an important mechanism for revealing the semantic hierarchy of a language. This article examines the potential of an n-gram model for identifying the linguistic features, semantic relations, frequency of use in speech, collocational patterns, and contextual characteristics of national-cultural units belonging to tourism-related vocabulary, as well as for studying their hyponymic-hyperonymic relations.*

Keywords: *tourism vocabulary, national-cultural units, n-gram model, corpus linguistics, hyponym, hyperonym, lexico-semantic relations, terminology, semantic analysis…*

Leksikografik birliklarni n-gram modeli va giponim-giperonimik munosabatlarini o'rganish N.Abduraxmonova ishlarida o'z ifodasini topgan [Abduraxmonova N. 2021]

Giponimiya semantik munosabatning bir turi bo‘lib, giperonim, giponim so‘zlari jins-tur munosabatini ifodalovchi atama sifatida 1968-yili ingliz tilshunosi J.Layonz tomonidan fanga olib kirilgan[Lyonz J.,1996;125]. Lingvistik adabiyotlarda jins tushunchasini ifodalovchi so‘zlar giperonimlar (yunoncha “hyper” – “ustida”, “onyma” – “nom”), tur tushunchasini ifodalovchi so‘zlar esa giponimlar deb ataladi. Giperonim jins belgisini bildiruvchi predmet, shaxs yoki hodisaning nomini ifodalovchi ko‘pgina ma’noni semantik jihatdan o‘zida birlashtirib turuvchi umumiy bir tizimning markaziy leksemasi, dominantasi sifatida namoyon bo‘ladi. Giponim esa ma’lum jins turining nomini hamda o‘zining semantik tarkibida implitsit tarzda jins ma’nosini ham ifodalovchi semantik jihatdan giperonimga nisbatan boy bo‘lgan leksik birlikdir [Mengliyev B., 2018;173].

Mazkur hodisa keyinchalik fransuz tilshunosi V.G.Gak, rus tilshunoslari A.A.Ufimseva, L.A.Novikov, M.V.Nikitin, S.B.Sakovtseva, A.I.Shein, L.D. Dyachenko, Y.Y.Kotsova, I.N.Gridina, M.V.Lisyakovalar tomonidan keng tadqiq etildi. O‘zbek tilshunosligida giponimiyaning mazmun-mohiyati R.Safarova, N.Sabirova, O‘.Xoliyorov, N.Jumabayeva, E.Yaxshiyeva kabi tadqiqotchilar ishlarida yoritib berilgan. Jumladan, *Y.Kotsova*ning fikricha, “Giperonim tabiiy tilda keng tarqalgan bo‘lib, nutq jarayonida tur-jins munosabatli so‘zlarni yanada faol ravishda almashtirish imkonini beradi. Xususan, giperonim turli so‘z turkumlaridagi jins-tur iyerarxiyasi darajalarida ham kuzatiladi.” [Kotsova Y.,2010;18].

I.N.Gridina tadqiqotlari bolalar mental leksikonida gipo-giperonimik munosabatlarning shakllanish xususiyatlarini aniqlashga bag‘ishlangan bo‘lib, u giponimiya hodisasiga “tur-jins farqlarini ifodalovchi leksik paradigmaning bir bo‘lagi – dominanta (giperonim) va unga bo‘ysunuvchi bir nechta komponentlardan (giponimlardan) iborat uya deb ta’rif beradi.”[Gridina I.,2010;10].

O‘zbek tilshunosligida *gipo-giperonim* munosabatlariga doir nazariy qarashlar *R.Safarova* tadqiqotlarida batafsil yoritib berilgan.[Safarova R.,1996;8.]. Tadqiqotchi izlanishlarida partonimiya, funksionimiya, graduonimiya, iyerarxionimiya, sinonimiya, antonimiya kabi hodisalarning o‘zbek tili semantik munosabatlardagi o‘ziga xos turi bo‘lgan giponimiyadan farqli jihatlarini hamda giponimik munosabatga ega so‘z guruhlarini o‘rganishning asosiy yo‘nalishlarini asoslab bergan. Shuningdek, olima jins-tur munosabatini ifodalashda tilning ikki nominativ vositasi – so‘z hamda so‘z birikmasidan foydalanish mumkinligini ta’kidlaydi:

1)tur tushunchasi jins va tur belgilarini ifodalovchi so‘zlarning bevosita birikuvi orqali ifodalanadi: *cho‘rtan baliq, xitoy olmasi, turk choyi;*

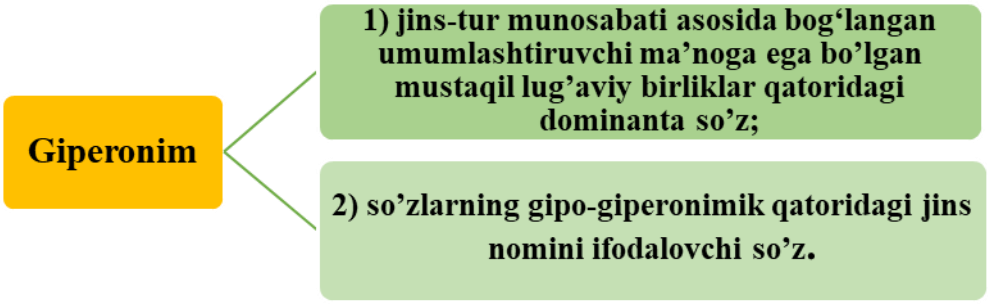
2)tur tushunchasi jins belgisi bevosita ifodalanmagan holda, alohida lug‘aviy birliklar orqali ifodalanadi: *ot, qulun, do‘nan.*

Umuman olganda, *giperonim* – jins belgisini bildirgan predmetning nomini ifodalovchi ko‘pgina ma’nolarni semantik jihatdan umumlashtiruvchi mikrotizimning markaziy so‘zi, dominantasi sifatida namoyon bo‘luvchi lug‘aviy birlik bo‘lsa, g*iponim* – ma’lum jins turlarining nomlarini ifodalovchi hamda o‘zinnng semantik tarkibida implitsid tarzda jins ma’nosini ifodalovchi, tarkibda so‘zni biriktirib kelgan, semantik jihatdan giperonimga nisbatan boy bo‘lgan lug‘aviy birlikdir.[Safarova R., 1996; 12].

Giponimiya hodisasida giperonimning ma’nosi giponimga nisbatan kengroq va xiraroq (noaniqroq), giponimning ma’nosi esa giperonimnikiga nisbatan *torroq va* aniqroq ifodalanadi. Shu boisdan giperonimning ma’noviy mohiyati barcha giponimning ma’noviy mohiyati yig‘indisiga tengligini qayd etish mumkin va mazkur hodisa pragmatik sathda yaqqolroq ko‘zga

tashlanadi. Masalan: *daraxt – terak – qayin – archa.* Bu o‘rinda *daraxt* so‘ziga nisbatan uning barcha boshqa turlarning nomlarini ifodalovchi so‘zlar teng qiymatga ega. Tur nomini ifodalovchi har bir so‘z shu turning o‘zi mansub bo‘lgan jins nomini ifodalovchi so‘z bilan semantik jihatdan birikib, gipo-giperonimik paradigmani hosil qila oladi.

Logik-semantik munosabat doirasida barcha turdosh aniq otlar o’zaro jins-tur munosabatini ifodalash asosida bog’lanadi. Lekin giperonimga ta’rif berganda muayyan tilning o‘ziga xos xususiyatini inobatga olgan holda lug‘aviy birliklar va ularning ma’lum guruhdagi o’zaro munosabatidan kelib chiqib yondashish kerak. R.Safarova giperonimni quyidagicha tasniflaydi (1-chizmaga qarang):



1-chizma. R. Safarovaning giperonimga doir tasnifi

Giponimik qatorning barcha elementlari yagona, umumiy ma’no asosida birikadi. Bu ma’no giponimik paradigmaning dominantasi sifatida giperonim zamirida aks etadi. Shunga ko‘ra, tilda giperonim ifodalanmas ekan, giponim ham mavjud bo‘lmaydi.[Safarova R.,1996; 15].

So‘nggi yillarda statistik va ehtimollik modellarga asoslangan yondashuvlar, xususan, **n-gram modeli** tabiiy tilni qayta ishlash (NLP) jarayonlarida keng qo‘llanilmoqda. N-gram modeli yordamida matnda ketma-ket kelgan birliklarning takrorlanish chastotasi va kontekstual bog‘liqligi aniqlanadi. Bu esa leksik birliklar o‘rtasidagi semantik munosabatlarni, jumladan, giponim–giperonimik aloqalarni avtomatik aniqlash imkonini beradi. Turizm diskursida qo‘llaniladigan leksik birliklar, ko‘pincha, umumiy tushunchadan xususiy tushunchaga yo‘naltirilgan semantik iyerarxiyani tashkil etadi. Mazkur holat turizm leksikasini *giponim–giperonimik munosabatlar* asosida avtomatik ravishda tizimli o‘rganish zaruratini yuzaga keltiradi.

Turizm diskursida giponim–giperonimik munosabatlar, ayniqsa, faol bo‘lib, ular axborotni tizimli va mantiqiy tarzda taqdim etishga xizmat qiladi. Turizm matnlarida umumiy tushunchalarni ifodalovchi giperonimik birliklar bilan bir qatorda, aniq va xususiy tushunchalarni anglatuvchi giponimik birliklar ham keng qo‘llanadi. Bunday munosabatlar nafaqat lug‘aviy darajada, balki murakkab so‘z birikmalari (bigram va trigramlar) darajasida ham namoyon bo‘ladi.

Turizmga oid leksik birliklarni mazmuniy va funksional jihatdan bir necha guruhlarga ajratish mumkin. Ushbu tasnif turizm diskursining murakkab semantik tuzilishini aniqlashga xizmat qiladi:

- **umumiy turizm terminlari.** Bular *“turizm”, “sayyoh”, “sayohat”, “turistik xizmat”, “turistik marshrut”* kabi keng ma’noli, giperonimik xarakterga ega birlik sifatida boshqa leksik birliklarni o‘z ichiga oladi.
- **madaniy va milliy komponentli leksik birliklar**. Bular *“milliy bayram”, “urf-odat”, “an’ana”, “marosim”, “xalq sayli”, “festival”* yoki *“milliy taomlar”, “milliy kiyimlar”, “hunarmandchilik buyumlari”* kabi birliklar bo‘lib, turizmning madaniy yo‘nalishini ifodalaydi.

Ushbu birliklar turizm diskursida, ayniqsa, muhim o‘rin tutadi va ular asosida giponimik qatorlar shakllanadi.

Masalan, *milliy bayram → Navro‘z, Mustaqillik kuni…*

- **obyektga oid leksik birliklar**. Bular *“tarixiy obida”,“ziyoratgoh”, “muzey”, “me’moriy yodgorlik”* kabi so‘zlardir. Bu birliklar ham giperonimik munosabat asosida tizimlanadi.
- **xizmat va infratuzilma bilan bog‘liq leksik birliklar.** Ular *“mehmonxona”, “gid”, “transport xizmati”, “turpaket”* kabi birliklarni o‘z ichiga oladi.

O‘zbek tilidagi turizm leksikasining muhim xususiyatlaridan biri unda milliy-madaniy komponentning kuchli namoyon bo‘lishidir. Masalan, *“Navro‘z”, “sumalak”, “Beshik to‘yi”, “Sunnat to‘yi”, “Kelin salom”* kabi birliklar nafaqat lingvistik, balki etnomadaniy ahamiyatga ham ega. Ular tilda gipo-giperonim munosabati asosida tizimlanadi. Turizmga oid milliy-madaniy birliklarni aniqlash uchun olib borilgan tadqiqot uchun quyidagi manbalar asosida **turizm korpusi** shakllantirildi:

- turistik broshyuralar;
- turistik taqdimnomalar va kataloglar;
- sayyohlik veb-saytlari matnlari;
- madaniy turizmga oid ilmiy risolalar;
- milliy bayram va marosimlarni tasvirlovchi matnlar;
- milliy taomlar, milliy kiyimlar haqidagi matnlar;
- mashhur va tarixiy shaxslar haqidagi matnlar.

Korpus asosida quyidagi *n-gramlar* ajratildi:

Bigramlar: *milliy bayramlar, milliy taomlar, milliy urf-odatlar, tarixiy obida, muqaddas qadamjolar, Minorayi Kalon, Ulug‘bek madrasasi, Ichan qal’a, Dishan qal’a, Tillakori madrasasi, Registon majmuasi…*

Trigramlar: *madaniy meros obyektlari, an’anaviy xalq bayramlari, Labi hovuz majmuasi, Shohi Zinda majmuasi, Go‘ri Amir maqbarasi, Sitorai Mohi Xossa, Chor bakr majmuasi…*

Ushbu n-gramlar turizm leksikasida **giperonimik birliklar** sifatida faol ishlatilishi aniqlandi. Bu birliklar, ko‘pincha, o‘z atrofida semantik maydon hosil qiladi va iyerarxik tuzilishga ega bo‘ladi. N-gram modeli yordamida aniqlangan birliklar asosida quyidagi iyerarxiya tuzildi. Misollarni kuzatamiz (1-jadvalga qarang):

1-jadval:

Turizmga oid milliy-madaniy birliklarning gipo-giperonim munosabati

№	Giperonim (jins)	Giponim (tur)
1	milliy marosimlar	*beshik to‘yi, sunnat to‘yi, kelin salom, challandi*
2	milliy bayramlar	*Navro‘z, Ramazon hayiti, Qurbon hayiti, Mustaqillik kuni, Xotira va qadrlash kuni*
3	milliy taomlar	*palov, shirguruch, moshxo‘rda, sho‘la (shovla)*
4	milliy kiyimlar	*to‘n, yaktak, nimcha, chopon, guppy, kamzul*
5	muqaddas qadamjolar	*ziyoratgoh, muqaddas maskan, qadamjo, maqbara*
6	hukmdor	*podshoh, amir, xon, toj-u taxt sohibi, sohibqiron*
7	taqinchoqlar	*bilakuzuk, bargak, zebigardon, tillaqosh, isirg‘a*

8	me'moriy obidalar	*Ulug'bek madrasasi, Nodir devonbegi xonaqosi*
9	majmua	*"Labi hovuz" majmuasi, Xoja Zayniddin majmuasi, Hazrati Imom majmuasi, Bolo Hovuz majmuasi.*

N-gram modeli – matnda ketma-ket kelgan *n* ta so'z birliklarining statistik ehtimolini aniqlashga asoslangan kompyuter modeli. N-gram modeli yordamida turizmga oid matnlarda barqaror so'z birikmalari**,** terminologik konstruksiyalar va semantik bog'lanishlar aniqlanadi. Shunga ko'ra, n-gram modeli quyidagi shakllarda aks etadi:

- **unigram** — 1 ta so'z (asos);
- **bigram** — 2 ta so'z (asos);
- **trigram** — 3 ta so'z (asos).

N-gram modeli lokal, ya'ni qo'shni so'zlar o'rtasidagi ehtimoliy bog'lanishni yaxshi ko'rsatadi, lekin butun jumlaning sintaktik tuzilishi va ma'nosini to'liq ifodalay olmaydi. Masalan, *"O'g'lon darvozasi yonidagi masjid, madrasa, mozor buzib tashlanadi."* jumlasini n-gramm modeli asosida tahlil qilishda natija quyidagicha aks etadi:

1. Jumla tokenlarga ajratib olinadi (jumlani so'zlarga ajratish):

O'g'lon, darvozasi, yonidagi, masjid, madrasa, mozor, buzib, tashlanadi.

Demak, 8 ta token mavjud.

2. Unigram birliklar aniqlanadi. Unigram modelida har bir so'z alohida ko'riladi. Unigram model so'zning oldingi so'zlarga bog'liqligini hisobga olmaydi.

3. Bigram modelida ketma-ket ikkita so'z olinadi. Ushbu jumlada 7 ta bigram mavjud: *O'g'lon darvozasi, darvozasi yonidagi, yonidagi masjid, masjid madrasa, madrasa mozor, mozor buzib, buzib tashlanadi.*

- *O'g'lon darvozasi=O'g'lon(ADJ)+darvozasi(NOUN);*
- darvozasi(NOUN)+yonidagi(ADV).

Turizmga oid matnlarni tahlil qilishda *unigram*, *bigram*, ayniqsa juda muhim, chunki u turizmga oid birliklardan qaysi biri eng ko'p uchrashi va individual qo'llanish chastotasini aniqlashga yordam beradi.

Tadqiqot natijalari turizm diskursining semantik iyerarxiyaga ega ekanligini va bu iyerarxiyani statistik modellash orqali aniqlash mumkinligini ko'rsatdi. Unga ko'ra giperonimik birliklar, asosan, **bigram** va **trigram** shaklida, giponimlar esa **unigram** yoki **antroponimik birliklar** sifatida ko'p uchrashi aniqlandi. Bu holat turizm matnlarida umumiy tushunchalar, ko'pincha, murakkab leksik birliklar orqali ifodalanishini ko'rsatdi.

FOYDALANILGAN ADABIYOTLAR

1. Agostini, A., Usmanov, T., Khamdamov, U., Abdurakhmonova, N., & Mamasaidov, M. (2021, January). Uzwordnet: A lexical-semantic database for the uzbek language. In *Proceedings of the 11th Global Wordnet conference* (pp. 8-19).
2. Ismailov, A. S., Shamsiyeva, G., Abdurakhmonova, N., & Navoi, A. (2021). Statistical machine translation proposal for Uzbek to English. *Science and Education*, *2*(12), 212-219.
3. Abdurakhmonova, N., Begimov, A., Hamroyeva, M., Aliyeva, E., Kurbanboyeva, S., & Ibragimov, B. (2025, November). Sentiment mining of patient experience data using T5 architectures. In *2025 IEEE XVII International Scientific and Technical Conference on Actual Problems of Electronic Instrument Engineering (APEIE)* (pp. 1-5). IEEE.

4. Abduraxmonova, N. (2021). O'zbek tili elektron korpusining kompyuter modellari. *Filologiya fanlari doktori (DSc) dissertatsiyasi avtoreferati*.
5. Inglizcha matnlarni o'zbek, A. N. (2018). tiliga tarjima qilish dasturining lingvistik taminoti (Sodda gaplar misolida). Filol. fan. bo 'yicha falsafa doktori (PhD)… dis. avtoref.
6. Abduraxmonova, N. (2023). Korpus lingvistikasi. *Toshkent: Globe*.
7. Абдураҳмонова, Н. (2020). Замонавий корпусларнинг компьютер моделлари. *Ўзбекистонда хорижий тиллар.-2020, 1*(30), 50-58.
8. Abdurakhmonova, N. Z. (2019). A two-level morphological analysis of the uzbek corpus. In *Иностранная филология. Социальная и национальная вариативность языка и литературы* (pp. 425-430).
9. Гридина И.Н. Гипо-гиперонимические отношения в ментальном лексиконе детей и взрослых. Автореф. дисс. …канд. филол. наук. – Санкт Петербург: 2010;
10. Котцова Е.Е. Гипонимия в лексической системе русского языка (на материале глагола): Дисс. д-ра ... филол. наук. – Архангельск: 2010. – 18 с.
11. Lyons J. Linguistic Semantics: An Introduction. – Cambridge: Cambridge University Press, 1996. – P. 125;
12. Mengliyev B. Hozirgi o'zbek tili. –Toshkent: Tafakkur bo'stoni. 2018. – Б. 173;
13. Сафарова Р. Ўзбек тилида гипонимия // Ўзбек тили ва адабиёти, 1987. № 1 – Б. 60 – 62;
14. Сафарова Р. Гипонимик қаторда полисемия ва омонимия. Ўзбек тилини синхроник ўрганиш масалалари. – Тошкент, 1987;
15. Safarova R. Leksik-semantik munosabatning turlari. Monografiya. Toshkent: O'qituvchi – 1996.B.8.
16. Сафарова Р. Гипонимия в узбекском языке (на материале общеупотребительных зоонимов): Автореф. дисс. канд. филол. наук. – Ташкент: 1990.

BREND NOMLARINING LINGVISTIK MOHIYATI

Jamolova Sarvinoz Shuhratjon qizi
O‘zMU Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi
1-bosqich magistranti
https://doi.org/10.5281/zenodo.23136450

Annotatsiya*. Mazkur maqolada brend nomlarining lingvistik mohiyati zamonaviy tilshunoslik nuqtayi nazaridan tahlil qilinadi. Brend nomi tilning nominativ imkoniyatlari mahsuli sifatida talqin etilib, uning semantik, fonetik va lingvokulturologik xususiyatlari yoritiladi. Tadqiqot davomida brend nomlarining iste’molchi ongida assotsiativ tasavvurlarni shakllantirishdagi o‘rni, madaniy identitet bilan bog‘liqligi hamda kommunikativ vazifalari tahlil qilinadi. Shu bilan birga, ularning til va jamiyat o‘rtasidagi o‘zaro munosabatni aks ettiruvchi muhim lingvistik hodisa ekanligi asoslab beriladi. Tadqiqot natijalari brend yaratish jarayonida til omilining ahamiyatini chuqurroq anglashga xizmat qiladi.*

Kalit so‘zlar: *nominatsiya, semantika, fonetika, assotsiatsiya, lingvokulturologiya, madaniy identitet, kommunikativ vazifa, ergonim.*

Annotation. *This article analyzes the linguistic nature of brand names from the perspective of modern linguistics. Brand names are interpreted as products of the nominative potential of language, and their semantic, phonetic, and linguocultural features are examined. The study explores the role of brand names in shaping associative perceptions in the consumer’s mind, their connection with cultural identity, and their communicative functions. It is substantiated that brand names constitute an important linguistic phenomenon reflecting the interaction between language and society. The findings of the research contribute to a deeper understanding of the significance of linguistic factors in the process of brand creation.*

Keywords: *nomination, semantics, phonetics, association, linguoculturology, cultural identity, communicative function, ergonym.*

Bugungi globallashuv va bozor munosabatlarining jadal rivojlanishi sharoitida brend nomlari oddiy nomlash vositasi doirasidan chiqib, tilning ijtimoiy, madaniy va psixologik imkoniyatlarini o‘zida mujassam etgan murakkab lingvistik hodisaga aylandi. Zamonaviy jamiyatda brend nomi mahsulot yoki xizmatni farqlash bilangina cheklanmay, balki iste’molchi ongida muayyan tasavvur va munosabatni shakllantirishda muhim omil sifatida namoyon bo‘lmoqda. Shu jihatdan brend nomlari til va tafakkur, til va jamiyat o‘rtasidagi uzviy bog‘liqlikni ifodalovchi muhim birliklar sirasiga kiradi. Har bir muvaffaqiyatli brend nomi ortida til birliklarining ongli va maqsadga yo‘naltirilgan tanlovi yotadi. Nomning semantik yuklamasi, fonetik shakli hamda madaniy konnotatsiyasi uning auditoriyaga ta’sir ko‘rsatish mexanizmini belgilaydi. Ayniqsa, qisqa va esda qoluvchan nomlar iste’molchi xotirasida tez mustahkamlanib, mahsulot yoki xizmatga nisbatan ijobiy munosabat shakllanishiga xizmat qiladi. Bu esa brend nomlarini lingvistik jihatdan chuqur tadqiq etish zaruratini yanada dolzarblashtiradi. Zamonaviy tilshunoslikda brend nomlari nominatsiya jarayonining alohida ko‘rinishi sifatida talqin qilinib, ergonimlar qatlamiga mansub sun’iy yaratilgan nominativ birliklar sifatida o‘rganilmoqda. Ularning vazifasi faqat nomlash bilan chegaralanmay, balki kommunikativ va pragmatik ta’sir ko‘rsatishdan ham iboratdir.[Abduraxmonova, N., 2025] Shu bilan birga, brend nomlari lingvokulturologik jihatdan ham muhim ahamiyat kasb etib, milliy madaniyat, tarixiy xotira va

ijtimoiy qadriyatlar bilan uzviy bog‘liq holda shakllanadi. Ayniqsa, mahalliy bozorga mo‘ljallangan brendlarda milliy til birliklarining qo‘llanilishi iste’molchida yaqinlik va ishonch hissini kuchaytiradi. Shunga ko‘ra, mazkur maqolada brend nomlarining lingvistik mohiyati nominatsiya jarayoni, semantik tuzilishi, fonetik shakli hamda lingvokulturologik omillar asosida tahlil qilinadi. Tadqiqot brend nomlarining iste’molchi ongida assotsiativ tasavvurlarni shakllantirishdagi o‘rnini yoritishga hamda ularning til va jamiyat o‘rtasidagi o‘zaro munosabatni aks ettiruvchi muhim lingvistik hodisa ekanligini asoslashga qaratilgan.

Tilning eng asosiy va qadimiy vazifalaridan biri nominatsiya, ya’ni borliqdagi predmet, hodisa va tushunchalarni nomlashdan iboratdir. Inson atrof-muhitni idrok etish jarayonida ushbu birliklarni nomlash orqali ularni ongida tizimlashtiradi. Shu ma’noda nominatsiya jarayoni tafakkur, til va real voqelik o‘rtasidagi uzviy munosabat mahsuli sifatida namoyon bo‘ladi. Tilshunoslikda bu jarayon tilning kognitiv va kommunikativ funksiyalari bilan bog‘liq holda talqin qilinadi [Nurmonov, 2000]. Zamonaviy jamiyatda nominatsiya jarayoni yangi ijtimoiy va iqtisodiy sharoitlar ta’sirida o‘ziga xos shakllarda rivojlanib, brend nomlari ana shunday maxsus nominativ birliklar sifatida baholanmoqda. Brend nomlari iqtisodiy, ijtimoiy va madaniy ehtiyojlar asosida yaratilgan bo‘lib, aniq pragmatik maqsadga yo‘naltirilganligi bilan tavsiflanadi. Ular nafaqat obyektni nomlaydi, balki ushbu obyekt haqidagi tasavvur va munosabatni shakllantirishga ham xizmat qiladi. Shu jihatdan brend nomlari an’anaviy leksik birliklardan o‘zining vazifaviy yo‘nalganligi bilan farqlanadi. Oddiy so‘zlardan farqli o‘laroq, brend nomlari ongli ravishda yaratiladi va muayyan auditoriyaga mo‘ljallanadi. Ularning shakllanishida til birliklarining semantik, fonetik va estetik imkoniyatlari hisobga olinadi. Natijada brend nomi kommunikativ va psixologik ta’sir vositasi sifatida iste’molchi ongida mahsulot yoki xizmatga nisbatan muayyan baho va munosabatni shakllantiradi. Bu holat brend nomlarini faqat nominativ birlik sifatida emas, balki tilning ta’sirchanlik funksiyasini amalga oshiruvchi vosita sifatida baholash mumkin. Tilshunoslik tadqiqotlarida sun’iy yaratilgan nomlar alohida nominativ guruh sifatida ko‘rib chiqiladi. Brend nomlari ham mazkur guruh tarkibiga kiradi. O‘zbek tilshunosligida nominatsiya jarayoniga oid izlanishlarda bunday birliklarning tilning amaliy ehtiyojlari bilan chambarchas bog‘liq ekani qayd etiladi.[Zaynobiddin qizi, N. A., & Asliddin qizi, S. G., 2024] Shu bilan birga, brend nomlari jamiyat taraqqiyoti bilan uzviy aloqada shakllanib, ijtimoiy ongdagi yangi tushunchalar, bozor munosabatlari hamda madaniy o‘zgarishlarni til vositasida aks ettiradi. Natijada brend nomlari til, jamiyat va iqtisod o‘rtasidagi o‘zaro ta’sirni ifodalovchi muhim lingvistik birlik sifatida namoyon bo‘ladi [Crystal, 2003]. Ularni o‘rganish zamonaviy tilshunoslikning dolzarb yo‘nalishlaridan biri bo‘lib kelmoqda.

Dastlab, brend nomlarining tahlilini semantik jihatdan ko‘rib chiqish o‘rinlidir. Ma’lumki, brend nomlarida salbiy bo‘yoqdorlikka ega birliklar deyarli uchramaydi. Chunki bu holat tilning emotsional va pragmatik imkoniyatlari bilan chambarchas bog‘liqdir. Inson tafakkurida ijobiy ma’no bilan bog‘langan so‘zlar tezroq qabul qilinadi va ishonch uyg‘otadi. Shu sababli, brend nomlarini yaratishda ijobiy semantik komponentlarning ustuvorligi muhim tamoyil sifatida qaraladi. Ijobiy semantika, avvalo, mahsulot yoki xizmatga nisbatan ishonch hissini kuchaytiradi. Nom orqali berilgan ijobiy ma’no iste’molchida ishonchlilik, sifat va barqarorlik haqidagi tasavvurlarni shakllantiradi. Masalan, til birliklaridagi “*yangi”, “toza”, “tez”, “mustahkam*” kabi semalar ong ostida ijobiy baho bilan qabul qilinadi. Bu holat tilshunoslikda so‘zning emotsional-ekspressiv bo‘yoqdorligi bilan izohlanadi [Yo‘ldoshev, 2012]. Shuningdek, ijobiy semantik yuklama tanlov jarayonini tezlashtiradi. Bozor sharoitida iste’molchi ko‘pincha qisqa vaqt ichida qaror qabul qiladi va bunda brend nomi muhim rol

o‘ynaydi. Ijobiy assotsiatsiya uyg‘otuvchi nomlar tanlov jarayonida ustunlikka ega bo‘lib, raqobatchi nomlar orasida ajralib turadi. Bu holat brend nomining faqat nominativ emas, balki baholovchi vazifani ham bajarishini ko‘rsatadi. Masalan, *Shifo Medline, Toza-choy, Tez.uz, Yashil Makon, Yangi Bank, Nurli Xonadon, Chitir-chitir somsa, Natural Beaty, Clear, Fresh Juice, Bright Smile, Soft touch, Happy Baby, Smart Technology, Golden Grain, Barakali dasturxon, Pokiza food.*

Bundan tashqari, ijobiy semantik komponent nomning esda qolish darajasini oshiradi. Inson xotirasi ko‘pincha emotsional jihatdan ahamiyatli axborotni tezroq va uzoqroq saqlaydi. Ijobiy bo‘yoqqa ega brend nomlari iste’molchi xotirasida mustahkamlanib, qayta-qayta esga tushish imkoniyatiga ega bo‘ladi. Shu jihatdan semantika brend nomining kognitiv qabul qilinishida muhim omil hisoblanadi [Aaker, 1991].

Nomlashdagi metaforik va metonimik ko‘chishlar ham ko‘plab brend nomlari bevosita mahsulotni emas, balki u bilan bog‘liq bo‘lgan hissiy holat, sifat yoki natijani ifodalaydi. Bu holat iste’molchi ongida ijobiy holat va tasavvurlarni uyg‘otuvchi konseptlar sifatida xizmat qiladi. Ya’ni bitta nom orqali keng ma’nolar doirasi ifodalanadi. Shuningdek, brend nomlari nominatsiyasida vaqt omili ham alohida ahamiyat kasb etadi. Zamonaviy bozor sharoitida nom tez qabul qilinadigan, tez tarqaladigan va tez taniladigan bo‘lishi talab etiladi. Shu sababli, brend nomlari ko‘pincha tilning faol lug‘at tarkibiga yaqin birliklar asosida yoki yangi, lekin oson o‘zlashtiriladigan shakllarda yaratiladi. Bu esa brend nomlarining til taraqqiyoti bilan birga o‘zgarib borishini ko‘rsatadi. Nominatsiya jarayonida auditoriya omili ham muhim mezon hisoblanadi. Brend nomi kimga mo‘ljallanganiga qarab uning shakli va mazmuni belgilanadi. Masalan, yoshlar auditoriyasiga yo‘naltirilgan brend nomlari ko‘pincha zamonaviy va xorijiy unsurlarga boy bo‘lsa, keng omma yoki oilaviy iste’molchiga mo‘ljallangan nomlarda an’anaviylik va tushunarlilik ustuvor bo‘ladi. Bu holat brend nomlarini ijtimoiy qatlam bilan bog‘liq nominativ birliklar sifatida ham talqin qilish imkonini beradi.

Brend nomining yana bir muvaffaqiyati ko‘pincha uning eshtilish uslubiga va fonetik xususiyatlariga bog‘liq bo‘ladi. Tadqiqotlar shuni ko‘rsatadiki, iste’molchi ongida qisqa, talaffuzi oson, tovush uyg‘unligiga ega nomlar tezroq yodda qoladi. Shu jihatdan brend nomlari nafaqat semantik, balki fonetik jihatdan ham kommunikativ vosita sifatida ahamiyatlidir[Abduraxmonova, N., & Shamsiyeva, G., 2026]. Fonetik shaklning roli quyidagi asosiy qonuniyatlarda namoyon bo‘ladi:

- Ochiq unlilar: Masalan, nomlaridagi “o” va “a” unlilari nomni ohangli qiladi, iste’molchi uchun eshitishda yoqimli his yaratadi: Artel, *Amazon, Adidas, Zara, Samsung, Avon, Sephora, Nivea, Prada, Versace.*
- Jarangli undoshlar: Masalan, *Google, Dolce & Gabbana, L’Oreal, Maybelline, Alibaba, Barilla* kabi nomlarda “g”, “m”, “l”, “r”, “d” “b” kabi tovushlari nomga dinamik va energiyali ohang beradi.
- Takroriy tovushlar ritmik effekt hosil qiladi va nomni esda qolishini yaxshilaydi: *Coca-Cola, Chupa Chups, Tic Tac, KitKat, Miu Miu, Lululemon, ZiZi, PayPal, Dunkin’ Donuts.*
- Qisqa va bir-ikki bo‘g‘inli nomlar talaffuzi oson va tez esda qoladi. Shu sababdan ko‘pchilik muvaffaqiyatli brend nomlari 1-3 bo‘g‘inli bo‘lib, murakkab talaffuzdan saqlanishadi: *Mars, Ford, Jeep, Dell, Vans, Shell, Slack.*

Fonetik qonuniyatlar shuni ko‘rsatadiki, nomning esda qoluvchanligi nafaqat semantik, balki tovush tizimi bilan chambarchas bog‘liqdir. Bu tilning fonetik mexanizmlarini real muloqot jarayonida qanday ishlatish mumkinligini ko‘rsatadi.

Brend nomlari nafaqat iqtisodiy vosita, balki madaniy makon bilan chambarchas bog‘liq lingvistik fenomen ham hisoblanadi. Ayniqsa, milliy bozorga yo‘naltirilgan brendlar mahalliy til birliklarini va madaniy kodlarni o‘zida aks ettiradi, bu esa iste’molchida yaqinlik, ishonch va identifikatsiya hissini uyg‘otadi. Bu jarayonda an’anaviy madaniy belgilar zamonaviy iste’mol muhitiga moslashib, yangi semantik qatlam bilan boyiydi. Natijada tarixiy yoki milliy birliklar faqat ramziy ahamiyat kasb etib qolmay, balki amaliy kommunikativ vazifani ham bajaradi.

O‘zbek tilida yaratilgan brend nomlarida uchraydigan asosiy lingvokulturologik xususiyatlar quyidagilar:

- Tarixiy leksika. Brend nomlari ko‘pincha o‘zbek tarixiy va adabiy manbalaridan yoki qadimiy leksik birikmalardan foydalanadi. Masalan, *Navbahor* nomi “bahor” faslini va yangilik, yangilanish ramzini o‘zida aks ettiradi. Undan tashqari, *Barakali dasturxon, Amirsoy Resort, Besh Qozon, Caravan, STONE CITY [tosh shahar], Café Navro‘z, Do‘ppi restaurant, Sogdiana Café, Adras Republic, 7SABER (qilich nomi), Choyxona Uyi* kabilar nafaqat nomni belgilaydi, balki milliy qadriyat va tarixiy makonni ham eslatadi.
    - Milliy qadriyatlar. Brend nomi orqali o‘zbek madaniyatiga xos qadriyatlar, masalan, mehmondo‘stlik, halollik, hurmat va sog‘liqni asrash kabi tushunchalar ifodalanadi. Masalan, *Sarpo, Boboy Restaurant, Arrows & Sparrows [ov], Diydor restaurant, Do‘stlik café, Darvozai Safed Kafesi, Registan, Mehmon, Valide Xanim, Orom hotel* kabi nomlar mazmunan milliylik, samimiylik, mehmondo‘stlik, tarixiy-madaniy qadriyat tushunchalarini ifodalamoqda.

Brend nomlarining lingvokulturologik xususiyatlari milliy til va madaniyatga oid qadriyatlarning saqlanishi hamda avloddan avlodga yetkazilishida muhim ahamiyat kasb etmoqda. Ayniqsa, globallashuv sharoitida xorijiy nomlash tendensiyalari kuchayib borayotgan bir paytda, milliy til birliklari asosida yaratilgan brendlar madaniy qarshilik va identitetni saqlash vositasi sifatida namoyon bo‘ladi [Crystal, 2003] Bu holatni brend nomlarini madaniy xotirani mustahkamlovchi lingvistik belgilar sifatida talqin qilinish mumkin. Shu jihatdan, brend nomlarining lingvokulturologik tahlili ularning madaniy mazmunini aniqlash bilangina cheklanmay, balki jamiyatning o‘zini anglash jarayonida tilning qanday rol o‘ynashini ham yoritishga xizmat qiladi. Brend nomlari orqali milliy madaniyat kundalik iste’mol makoniga kirib boradi va zamonaviy ijtimoiy muloqotning ajralmas qismiga aylanadi.

Brend nomi tilning nominativ imkoniyatlarini amaliyotga tatbiq etib, maqsadli va strategik axborot yetkazuvchi birlik sifatida shakllanadi. Bu jarayon semantik, fonetik va lingvokulturologik elementlarning uyg‘unligini talab qiladi, chunki nomning samaradorligi uning lingvistik xususiyatlari bilan chambarchas bog‘liq. Shuningdek, brend nomi kommunikativ vosita sifatida murakkab funksiyalarga ega. U axborot yetkazish, mahsulot yoki xizmatni identifikatsiyalash, auditoriya bilan psixologik va madaniy aloqa o‘rnatish kabi jarayonlarni bir vaqtning o‘zida amalga oshiradi. Bu jihatdan brend nomi tilning strategik resursidir. Uning yaratish jarayonida lingvistik bilim va madaniy kontekstni hisobga olish zarur. Brend nomlarini chuqur lingvistik tahlil qilish nafaqat marketing uchun, balki tilshunoslik nazariyasini boyitish uchun ham muhimdir.

FOYDALANILGAN ADABIYOTLAR

1. Abduraxmonova N. Kompyuter lingvistikasi (darslik), Toshkent: Nodirabegim, 2021. - B. 400

2. Aaker D. A. Managing Brand Equity, Free Press. 1991. - B. 299
3. Crystal D. Language and Culture, Cambridge University Press. 2003. - B. 212
4. Nurmonov A., Yo'ldoshev B. Tilshunoslik va tabiiy fanlar, Sharq. - B. 180
5. Yo'ldoshev B. Ekspressivlik va baholash kategoriyasi, Fan, 2012. - B. 120
6. Abdurakhmonova, N., & Rajabova, K. (2025, September). Stages of Creating Uzbek Language Multimedia Corpus. In *2025 10th International Conference on Computer Science and Engineering (UBMK)* (pp. 1628-1633). IEEE.
7. Mengliev, D., Abdurakhmonova, N., Shirinova, R., Ibragimov, B., Khudayberganova, D., & Suyunova, M. (2025, October). A comparative study of annotation schemes for named entity recognition in Uzbek legal texts. In *AIP Conference Proceedings* (Vol. 3377, No. 1, p. 070005). AIP Publishing LLC.
8. Abdurahmanova, M., & Rayimjonova, M. (2021). KORPUS LINGVISTIKASIDA TARJIMASHUNOSLIK MASALASI. *COMPUTER LINGUISTICS: PROBLEMS, SOLUTIONS, PROSPECTS*, *1*(1).
9. Abduraxmonova, N., & Shamsiyeva, G. (2026). KO 'P TILLI KORPUSLAR ASOSIDA MASHINA TARJIMASI TIZIMLARINI TAKOMILLASHTIRISH BOSQICHLARI. *COMPUTER LINGUISTICS: PROBLEMS, SOLUTIONS, PROSPECTS*, *1*(1).
10. Abduraxmonova, N. (2025). LEKSIKOGRAFIYADA TERMINOLOGIK LUG'ATLARNING O'RNI VA AHAMIYATI. *YANGI DAVR ILM-FANI: INSON UCHUN INNOVATSION G 'OYA VA YECHIMLAR*, *1*(9), 47-49.
11. Zaynobiddin qizi, N. A., & Asliddin qizi, S. G. (2024, November). Theoretical Foundations of Corpus-based Uzbek-English Machine Translation. In *2024 IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE)* (pp. 1650-1653). IEEE.

AVTOMATIK NUTQNI TANISH TEXNOLOGIYASINING ZAMONAVIY YONDASHUVLARI

Jo'raeva Zilola Obidjon qizi
O'zbekiston Milliy universiteti, Jurnalistika va o'zbek filologiyasi fakulteti,
Kompyuter lingvistikasi mutaxassisligi I bosqich magistranti
https://doi.org/10.5281/zenodo.23136509

Annotatsiya. *Ushbu maqolada avtomatik nutqni tanish (ASR) tizimlarining shakllanish tarixi, tasniflanishi va zamonaviy rivojlanish tendensiyalari tahlil qilingan. Tadqiqotda an'anaviy nutqni aniqlash usullaridan zamonaviy o'z-o'zini boshqaradigan (Self-Supervised Learning – SSL) modellariga o'tish bosqichlari hamda ularning o'zbek tili nutq texnologiyalaridagi o'rni yoritilgan. Shuningdek, maqolada o'zbek tili nutq korpuslarining yaratilishi, xususan, professor N. Abduraxmonova rahbarligidagi loyihalar va Mozilla Common Voice ma'lumotlar bazasining SSL modellarini (Wav2Vec 2.0, Whisper) o'qitishdagi ahamiyati asoslab berilgan. Tadqiqot natijalari SSL yondashuvi resursi kam tillar uchun matnga bog'lanmagan audio ma'lumotlardan samarali foydalanish imkonini berishini ko'rsatadi.*

Kalit so'zlar: *ASR, avtomatik nutqni tanish, Self-Supervised Learning (SSL), Wav2Vec 2.0, Whisper, o'zbek tili milliy korpusi, Mozilla Common Voice, WER (Word Error Rate), kompyuter lingvistikasi.*

ABSTRACT. *This article analyzes the historical evolution, classification, and contemporary development trends of Automatic Speech Recognition (ASR) systems. The study explores the transition from traditional speech recognition methods to modern Self-Supervised Learning (SSL) models and their role in Uzbek speech technologies. Furthermore, the paper highlights the development of Uzbek speech corpora, specifically projects led by Professor N. Abduraxmonova, and the significance of the Mozilla Common Voice database in training SSL models such as Wav2Vec 2.0 and Whisper. The research findings indicate that the SSL approach provides an efficient solution for low-resource languages by leveraging unlabelled audio data effectively.*

Keywords: *ASR, Automatic Speech Recognition, Self-Supervised Learning (SSL), Wav2Vec 2.0, Whisper, Uzbek National Corpus, Mozilla Common Voice, WER (Word Error Rate), Computational Linguistics.*

Аннотация. *В данной статье анализируются история формирования, классификация и современные тенденции развития систем автоматического распознавания речи (ASR). В исследовании освещаются этапы перехода от традиционных методов распознавания речи к современным моделям самообучения (Self-Supervised Learning – SSL), а также их роль в развитии речевых технологий для узбекского языка. Кроме того, в статье обосновывается значимость создания речевых корпусов узбекского языка, в частности, проектов под руководством профессора Н. Абдурахмоновой и базы данных Mozilla Common Voice в обучении таких SSL-моделей, как Wav2Vec 2.0 и Whisper. Результаты исследования показывают, что подход SSL является наиболее эффективным решением для языков с ограниченными ресурсами (low-resource languages), позволяя продуктивно использовать неразмеченные аудиоданные.*

Ключевые слова: *ASR, автоматическое распознавание речи, Self-Supervised Learning (SSL), Wav2Vec 2.0, Whisper, национальный корпус узбекского языка, Mozilla Common Voice, WER (Word Error Rate), компьютерная лингвистика.*

Kirish.

Bugungi tezkor axborotlar almashinuvi har bir sohani qamrab olgan zamonda og‘zaki nutqni tushuna olish, to‘g‘ridan to‘g‘ri javob berish tabiiy tilni qayta ishlash (NLP) sohasining eng asosiy maqsadlaridan biri hisoblanadi.

Og‘zaki nutqni qayta ishlash kompyuterlar yaratilishidan o‘nlab yillar oldin paydo bo‘lganini ko‘rishimiz mumkin. O‘tgan asrning 20-yillarida nutqni taniy oladigan birinchi qurilma 500 Gts chastotali akustik energiya ta’sirida prujinalar orqali harakatlanadigan “Rex” laqabli kuchukcha o‘yinchoq shaklida yaratilgan [1].

Zamonaviy dunyoda insonlar texnologik tizimlardan yanada ko‘proq qulaylik kutmoqda. Foydalanuvchilarning aksariyati avtomatik jihozlarga jismoniy harakatsiz buyruq berishni, ehtiyojlarini tez va oson shaklda qondirishni istaydilar. Ushbu vazifalarni hal qilishda ASR (Automatic Speech Recognition)- Avtomatik nutqni tanish texnologiyasi markaziy o‘rin tutadi. ASR texnologiyasi nutq signalini avtomatik traskripsiya qilish, ya’ni matn ko‘rinishiga o‘tkazish tamoyiliga asoslanadi. Garchi bu soha to‘laligicha o‘z yechimini topgan deb bo‘lmasa-da, u ko‘plab amaliy vazifalarni bajarish uchun mos keladigan darajada rivojlangan [4]. Bugungi kunda ommalashib borayotgan Gemini, Chatgpt, Copilot kabi sun’iy intellekt ilovalari ham nutqni aniqlashning yuqori darajasiga ko‘tarilgan. Shunga qaramay, nutqni taniy olish va kontekstni tushunish va mantiqiy javob berishda muammolar talaygina. ASR texnologiyasi kuchaytiruvchi aloqaning bir qismi sifatida kompyuter va nogironligi bo‘lgan odamlar o‘rtasidagi muloqot uchun muhimdir. Insoniyat har doim nutqni matnga ko‘chirish ehtiyojini sezib kelgan. Bugungi kunda bu talab yanada ortib bormoqda.

ASR tizimlari nutqning tabiiyligi, leksik murakkabligi va akustik o‘zgaruvchanligiga ko‘ra quyidagi uchta asosiy turlarga bo‘linadi:

1. **Raqamlarni tanib olish (Digit recognition)** – bu kompyuter yoki sun’iy intellekt tizimining aytilgan yoki yozilgan raqamlarni (0 dan 9gacha) avtomatik tarzda tanish va ularni raqamli formatga o‘tkazish qobiliyatidir. Raqamli aniqlashning ikki xil turi mavjud: tasvir orqali (visual) va ovoz orqali (akustik). Ushbu tur ASRning eng sodda va aniqlik darajasi juda yuqori bo‘lgan turi hisoblanadi. Uning lug‘at boyligi cheklangan – atigi 10-11 ta so‘z mavjud.
2. **O‘qilgan nutq (Read speech)** – ASR sohasida nutqning eng tartibli va tushunish oson bo‘lgan turi hisoblanadi. Bu insonning oldindan tayyorlangan matnni (masalan, kitob, matn, maqola, ssenariy va boshqalar) ovoz chiqarib o‘qishidir. Bunda tayyor matn o‘qilganligi sababli nutq “toza” bo‘ladi. Bu kabi nutq aniq talaffuz, to‘xtalishlarning yo‘qligi, grammatik tartibning mavjudligi, nutqni buzuvchi omillar (parazit so‘zlar, vulgarizmlar) deyarli uchramasligi bilan STT (Speech-to-text) jarayonini osonlashtiradi.
3. **Suhbat nutqi (Conversational speech)** – bu ikki yoki undan ortiq insonning o‘zaro erkin muloqoti bo‘lib, ASR tizimlari uchun eng qiyin va murakkab vazifa hisoblanadi. Suhbat nutqida to‘liqsiz jumlalar, to‘ldiruvchi tovushlar, ustma-ust gapirish (Overlap), noaniq talaffuz, imo-ishoralarning nutq mazmuniga ta’sir qilishi kabi omillar mavjudligi bilan ASR jarayonini texnik jihatdan murakkablashtiradi.

ASR tizimlarining murakkablik darajalariga ko‘ra tasnifi

Xususiyat	**Raqamlarni tanish**	**O‘qilgan nutq**	**Suhbat nutqi**
lug‘at boyligi	cheklangan (10-11 so‘z)	o‘rtacha/katta	juda katta
nutq sifati	yuqori	toza va tartibli	shovqinli, tartibsiz
aniqlik darajasi	99% gacha	95% - 98%	80% - 90%

ASR tizimining dastlabki bosqichi audiomatnlarni transkripsiya qilish murakkab va ko‘p mehnat talab qiluvchi jarayon sanaladi [2]. Nutqni aniqlash tizimlari uchun **“o‘qilgan nutq” (Read speech)** nazorat ostidagi sinov muhiti bo‘lsa, **“erkin suhbat nutqi” (Conversational speech)** barcha lingvistik va akustik qiyinchiliklar namoyon bo‘ladigan real amaliy muhit- “jang maydoni”dir. Suhbat nutqiga xos bo‘lgan murakkabliklar uzoq vaqt davomida ASR tizimlari uchun asosiy to‘siq bo‘lib keldi. Biroq so‘nggi yillarda **o‘z-o‘zini boshqarish (Self-Supervised Learning- SSL)** texnologiyasining paydo bo‘lishi bu muammolarni hal qilishda yangi sahifa ochdi. Dastlabki SSL modeli Meta AI tomonidan 2019-yilda Wav2vec modeli taqdim etilgan bo‘lib, haqiqiy inqilob 2020-yilda Wav2Vec2.0 ning yaratilishi bilan yuz berdi [3]. Aynan shu nuqtada kompyuterlar inson kabi – tayyor matnlarsiz, shunchaki nutq sadosini tinglash orqali tilning murakkab akustik qonuniyatlarini mustaqil o‘rganish imkoniyatiga ega bo‘ldi (Baevski et al., 2020). SSL tizimlari an’anaviy tizimlardan farqli o‘laroq, inson tomonidan tayyorlangan transkripsiyalarsiz ham tilning akustik va lingvistik xususiyatlarini mustaqil o‘rganish qobiliyatiga ega. SSL modellari – bu sun’iy intellektning “eshitish orqali o‘rganish” qobiliyatidir. Ular matn tayyorlash uchun ketadigan yillab vaqt va millionlab mablag‘ tejaydi. D.Jurafsky va J.Martin nutqning akustik o‘zgaruvchanligini ASR tizimlari uchun asosiy qiyinchilik sifatida ko‘rsatib, SSL modellari bu muammoni ulkan hajmdagi xom audio ma’lumotlarni dastlabki o‘qitish (pre-training) orqali hal qilish mumkinligini ta’kidlaydi [1]. Bu jarayonda SSL modeli tilning fonetik va prosodik xususiyatlarini mustaqil ravishda o‘rganib, akustik noaniqliklarni kamaytirishga xizmat qiladi. SSL yondashuvi nutq signallarini “yashirin tasvirlar” ko‘rinishida o‘rganishi tufayli, u hattoki murakkab suhbat nutqi (conversational speech) sharoitida ham mantiqiy izchillikni saqlab qolish imkonini beradi. SSLni shunchaki yangi texnologiya emas, balki lingvistik resurslar yetishmovchiligi muammosiga texnik yechimdir. Quyida Self-supervised learning modellarini tasniflari bilan tanishamiz:

Self-supervised learning (o‘z-o‘zini boshqaradigan nutq) modellarining tasnifi

Model turi	**Asosiy vakillari**	**Ishlash prinsipi (Mexanizmi)**	**Afzalligi**
kontrastiv (Contrastive)	**wav2vec, wav2vec 2.0**	Audio signalning “niqoblangan” qismini boshqa o‘xshash bo‘lmagan namunalar ichidan ajratib olishni o‘rganadi.	Nutqning umumiy strukturasini juda aniq o‘rganadi.
bashorat qiluvchi	**HuBERT, WavLM**	Nutqni oldindan guruhlangan	Semantik (ma’no) va akustik ma’lumotlarni

(Predictive)		"klasterlarga" bo‘ladi va yashirilgan qismlarda qaysi klaster (tovush birligi) borligini bashorat qiladi.	yaxshi farqlaydi.
generativ (Generative)	**AudioLM, SpeechResNet**	O‘tkazib yuborilgan yoki keyingi keladigan audio signallarini qayta tiklash (generatsiya qilish) orqali o‘rganadi.	Nutq sintezi (tovush chiqarish) va tozalashda juda samarali.
ko‘p vazifali (Multitask)	**SeamlessM4T, Whisper***	Bir vaqtning o‘zida ham SSL, ham transkripsiya ma’lumotlaridan foydalanadi (Gibrid yondashuv).	Bir vaqtning o‘zida tarjima, matnga aylantirish va tushunish imkoniyati.

Zamonaviy Self-supervised learning (SSL) modellari o ‘zbek tili kabi raqamli resurslari kam tillar uchun nutqni aniqlash (ASR) tizimlarini yaratishda eng samarali yechim hisoblanadi. O‘zbekistonda ushbu yo‘nalishdagi ishlar uchta asosiy bosqichda amalga oshrilmoqda:

1. Lingvistik poydevor va ma’lumotlar bazasini yaratish. Har qanday SSL modeli sifatli ishlashi uchun katta hajmdagi til resurslariga tayanadi. O‘zbekistonda bu boradagi fundamental ishlarni professor N. Abduraxmonova amalga oshirmoqda. Olima rahbarligida yaratilgan “O‘zbek tili milliy korpusi” va multimedia resurslari SSL modellarini o ‘zbek tilining fonetik xususiyatlariga moslashtirish (fine-tuning) uchun asosiy manba bo‘lib xizmat qiladi [9, 10, 11, 12, 13].
2. Ochiq korpuslar bilan ishlash. SSL modellarini amaliyotga tatbiq etishda Mozilla common voice loyihasining o‘rni katta. Ushbu ochiq bazadagi audio yozuvlar yordamida Whisper va Wav2Vec2.0 kabi modellar o‘zbekcha nutqni, ayniqsa, turli sheva va uslubni aniqlashga o‘rgatilmoqda. Tadqiqotchi M.Abduvaxidov va boshqa izlanuvchilarning tajribalari modellarning xatolik darajasini (WER-Word Error Rate) kamaytirishga yordam bermoqda [14].
3. Akademik tadqiqotlar va natijalar. Bugungi kunda Toshkent axborot texnologiyalari universiteti olimlari, xususan, Z.Saparov jamoasi SSL yondashuvi asosida nutqni aniqlash modellarini ishlab chiqmoqda. Tadqiqot natijalari shuni ko‘rsatadiki, SSL modellari an’anaviy usullarga qaraganda kamroq matn talab qilgan holda, murakkab va shovqinli muhiotlarda ancha aniqroq natija bermoqda [15].

Xulosa

Ushbu maqolada avtomatik nutqni tanish (ASR) tizimlarining tarixiy shakllanish bosqichlaridan tortib, zamonaviy o‘z-o‘zini boshqaradigan (SSL) modellarigacha bo‘lgan tadrijiy yo‘li tahlil qilindi. Tadqiqotlar shuni ko‘rsatadiki, nutqni qayta ishlash texnologiyalari oddiy mexanik o‘yinchoqlardan boshlab, bugungi kunda inson nutqining murakkab akustik va lingvistik xususiyatlarini tushuna oladigan murakkab sun’iy intellekt tizimlarigacha rivojlandi.

Olib borilgan tahlillar natijasida quyidagi muhim xulosalarga kelindi:

Nutqning murakkabligiga ko‘ra "o‘qilgan nutq" va "erkin suhbat nutqi" o‘rtasidagi farqlar ASR tizimlari uchun hamon asosiy texnik to‘siq bo‘lib qolmoqda. Biroq, zamonaviy SSL algoritmlari ushbu "jang maydoni"da (conversational speech) akustik noaniqliklarni sezilarli darajada kamaytirish imkonini beradi.

SSL texnologiyasining inqilobiy roli: Wav2Vec 2.0 va Whisper kabi modellar nutqni transkripsiyalarsiz, faqat eshitish orqali o‘rganishi tufayli, til resurslarini tayyorlashda ketadigan millionlab mablag‘ va vaqtni tejash imkoniyatini berdi. Bu, ayniqsa, o‘zbek tili kabi "low-resource" (resursi kam) tillar uchun texnologik sakrash (leapfrogging) imkoniyatini yaratadi.

Mahalliy olimlar, xususan N. Abduraxmonova rahbarligida yaratilgan milliy korpuslar va Mozilla Common Voice loyihasi doirasida to‘plangan audio bazalar o‘zbek tili nutq texnologiyalarining lingvistik poydevorini shakllantirdi. Bu resurslar zamonaviy SSL modellarini fine-tuning qilish orqali o‘zbekcha nutqni tanish aniqligini (WER) xalqaro darajaga olib chiqish imkonini bermoqda.

Xulosa o‘rnida ta’kidlash lozimki, o‘zbek tili nutq texnologiyalarining kelajagi bevosita fundamental korpus lingvistikasi va zamonaviy neyron tarmoqlarining integratsiyasiga bog‘liq. Kelajakdagi tadqiqotlar o‘zbek tilining turli shevalari va akustik shovqinli muhitlarda nutqni aniqlash aniqligini yanada oshirishga, shuningdek, nutqni nafaqat matnga aylantirish, balki uning semantik va hissiy (emotion recognition) qatlamlarini tushunishga qaratilishi maqsadga muvofiqdir. Bu esa o‘zbek tilining global raqamli ekotizimda to‘laqonli va raqobatbardosh til sifatida yashashini ta’minlaydi. Bizning fikrimizcha, o‘zbek olimlari tomonidan yaratilgan boy korpuslar va zamonaviy SSL modellarining birikmasi o‘zbek tili nutq texnologiyalarini rivojlantirishning eng to‘g‘ri yo‘lidir. Bu hamkorlik o‘zbek tilining global raqamli makonda munosib o‘rin egallashini ta’minlaydi.

FOYDALANILGAN ADABIYOTLAR

1. Jurafsky, D., Martin, J. H. (2023). Speech and Language Processing (3rd ed. draft). Stanford University. Chapter 15: Automatic Speech Recognition, pp. 334-345.
2. Abduraxmonova N. (2023). Korpus lingvistikasi. Oliy o‘quv yurtlari talabalari uchun darslik. Toshkent: "Akademnashr" nashriyoti, 297-bet.
3. Baevski, A., Zhou, Y., Mohamed, A., Auli, M. (2020). wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations. Advances in Neural Information Processing Systems (NeurIPS), 33, pp. 12449-12460.
4. Rabiner L., Juang, B. H. (1993). Fundamentals of Speech Recognition. Prentice Hall, pp. 1-15.
5. Dixon, M. (2010). The History of Dictation and its Role in Literature. Journal of Historical Biography, Vol. 7, pp. 112-118.
6. Mozilla Common Voice Dataset. (2023). Uzbek Language Corpus Statistics and Community Contributions. [Online resource]. URL: https://commonvoice.mozilla.org/uz/datasets
7. Abduvaxidov M., Olimov, S. (2023). Fine-tuning Whisper and Wav2Vec2.0 models for Uzbek Speech Recognition using Common Voice. International Conference on Artificial Intelligence and Natural Language Processing (AINL), pp. 45-52.
8. Saparov Z., Qodirov N. (2023). Implementation of Self-Supervised Learning Models for Uzbek Speech Recognition. Muhammad al-Xorazmiy avlodlari (TATU ilmiy jurnali), №3 (25), 88-94-betlar.
9. Abdurakhmonova, N., Alisher, I., & Sayfulleyeva, R. (2022, September). MorphUz: Morphological analyzer for the Uzbek language. In 2022 7th International Conference on Computer Science and Engineering (UBMK) (pp. 61-66). IEEE.
10. Abduraxmonova, N. (2023). Korpus lingvistikasi. Toshkent: Globe.

11. Абдураҳмонова, Н. (2020). Замонавий корпусларнинг компьютер моделлари. Ўзбекистонда хорижий тиллар.-2020, 1(30), 50-58.
12. Abdurakhmonova, N. Z. (2019). A two-level morphological analysis of the uzbek corpus. In Иностранная филология. Социальная и национальная вариативность языка и литературы (pp. 425-430).
13. Abduraxmonova, N. (2021). O ‘zbek tili elektron korpusining kompyuter modellari. Filologiya fanlari doktori (DSc) dissertatsiyasi avtoreferati.
14. Abduvaxidov, M., & Olimov, S. (2023). Fine-tuning Whisper and Wav2Vec2.0 models for Uzbek Speech Recognition using Common Voice. International Conference on Artificial Intelligence and Natural Language Processing (AINL), pp. 45-52.
15. Saparov, Z., & Qodirov, N. (2023). Implementation of Self-Supervised Learning Models for Uzbek Speech Recognition. Muhammad al-Xorazmiy avlodlari (TATU ilmiy jurnali), №3 (25), 88-94-betlar.

KORPUS MATNLARINI SINTAKTIK TEGLASHNING AHAMIYATI

Madina Jo'raqulova
O'zMU Kompyuter lingvistikasi va amaliy
tilshunoslik kafedrasi o'qituvchisi
madinasubxonovna0993@gmail.com
https://doi.org/10.5281/zenodo.23136528

Annotatsiya. *Ushbu maqolada korpus matnlarini sintaktik teglashning nazariy va amaliy ahamiyati, morfologik teglashning sintaktik tahlil uchun zarurati, hamda korpus matnlaridagi sintaktik birliklar va ular o'rtasidagi munosabatlarni avtomatik aniqlash imkonini beruvchi grammatikalar, xususan, tobelanish grammatikasi tahlil qilinadi. Shuningdek, sintaktik teglangan korpuslarning tabiiy tilni qayta ishlash, avtomatik tarjima, axborot qidiruv tizimi, lingvistik tadqiqotlar va elektron lug'atlar yaratishdagi ahamiyati ham yoritilgan. Bundan tashqari, korpus matnlarini sintaktik teglash uchun tobelanish grammatikasi modellarini takomillashtirish zarurati masalasi ham tahlilga tortilgan.*

Kalit so'zlar: *bevosita tarkib toptiruvchi grammatika, tobelanish grammatikasi, sintaktik teglash, lingvistik teglash, anaforik teglash, diskursiv teglash*

Abstract. *This article discusses the theoretical and practical significance of the syntactic tagging of corpus texts, the necessity of morphological tagging for syntactic analysis, and the grammars—specifically dependency grammar—that enable the automatic identification of syntactic units and the relationships between them within corpus texts. It also highlights the importance of syntactically tagged corpora in natural language processing, machine translation, information retrieval systems, linguistic research, and the creation of electronic dictionaries. Furthermore, the article analyzes the need to refine dependency grammar models for the syntactic tagging of corpus texts.*

Key words: *constituency parsing, dependency parsing, linguistic tagging, syntactic tagging, anaphoric tagging, discourse tagging.*

Аннотация. *В данной статье анализируется теоретическое и практическое значение синтаксической разметки корпусных текстов, необходимость морфологической разметки для синтаксического анализа, а также грамматики, позволяющие автоматически идентифицировать синтаксические единицы и связи между ними в корпусных текстах, в частности, грамматика подчинения. Подчеркивается также важность корпусов с синтаксической разметкой в обработке естественного языка, автоматическом переводе, системах информационного поиска, лингвистических исследованиях и создании электронных словарей. Кроме того, анализируется необходимость совершенствования моделей грамматики подчинения для синтаксической разметки корпусных текстов.*

Ключевые слова: *синтаксический анализ на основе составляющих, синтаксический анализ на основе зависимостей, лингвистическая разметка, синтаксическая разметка, анафорическая разметка, дискурсивная разметка.*

Lingvistik teglash (annotatsiyalash) – matn tarkibidagi lingvistik birliklarga ularning grammatik, sintaktik, semantik, diskursiv va boshqa xususiyatlarini ifodalovchi maxsus ma'lumotlarni biriktirish jarayonidir. Bunda matn birliklariga beriladigan izohlar so'z turkumi haqidagi izohdan tortib, murakkab sintaktik tuzilma, semantik munosabat va diskursiv aloqalargacha ifodalashi mumkin. Shu jihatdan lingvistik teglash tabiiy tilni kompyuter yordamida qayta ishlash, korpuslarni shakllantirish va lingvistik tadqiqotlar natijalarini tekshirishning muhim vositalaridan biri hisoblanadi.

Lingvistik teglashning ilmiy va amaliy ahamiyati, avvalo, matnni avtomatik tahlil qilish imkoniyati bilan belgilanadi. Teglangan matn kompyuter dasturlariga uning strukturaviy va mazmuniy xususiyatlarini aniqlash uchun zarur bo'lgan qo'shimcha ma'lumotni taqdim etadi. Bu esa tabiiy tilni qayta ishlash (Natural Language Processing – NLP) doirasidagi turli vazifalarni amalga oshirish imkonini yaratadi. Xususan, lingvistik jihatdan izohlangan ma'lumotlar asosida matnning grammatik tuzilishi, so'zlar o'rtasidagi munosabatlar va ayrim semantik xususiyatlarni avtomatik aniqlash imkoniyati vujudga keladi.

Shuningdek, teglangan korpuslar kompyuter modellarini o'qitish uchun asosiy ma'lumotlar bazasi bo'lib xizmat qiladi. Bu kabi izohlangan korpuslardan mashina tarjimasi, sentiment tahlil, nomlangan obyektlarni aniqlash (Named Entity Recognition – NER), sintaktik tahlil va boshqa NLPga oid amallarni bajarishga mo'ljallangan modellarni ishlab chiqishda foydalaniladi

Bundan tashqari, teglangan korpuslar til birliklarining qo'llanish chastotasi, ularning grammatik va semantik xususiyatlari, o'zaro munosabatlari hamda kontekstual qo'llanish qonuniyatlarini katta hajmdagi material asosida o'rganish imkonini beradi. Natijada tilshunos turli lingvistik hodisalarni tizimli ravishda tadqiq etishi, muayyan qonuniyatlarni aniqlashi va ilgari surilgan ilmiy farazlarni korpus ma'lumotlari asosida tekshirishi mumkin.

Lingvistik teglash bir qancha asosiy turlarga ajratiladi. Jumladan, *morfologik teglash* so'zlarning so'z turkumi va ularga xos grammatik kategoriyalarni izohlab beradi. Bunda so'z turkumlari, egalik, kelishik, shaxs/son, zamon, mayl, nisbat va boshqa morfologik belgilar avtomatik aniqlanadi. *Sintaktik teglash* esa gap tarkibidagi birliklarning o'zaro sintaktik munosabatlarini ifodalaydi. Bunda ega, kesim, to'ldiruvchi, aniqlovchi, hol kabi gap bo'laklari hamda so'zlar o'rtasidagi grammatik bog'lanishlar aniqlanadi.

Semantik teglash esa matn birliklarining ma'noviy xususiyatlarini va ular o'rtasidagi semantik munosabatlarni ifodalashga xizmat qiladi. Bunda sinonim, antonim, giperonim va boshqa semantik aloqalar aniqlanadi. *Diskursiv teglangan korpusda* esa matnning yuqori darajadagi tuzilishi, ya'ni mavzu, argument, gaplar o'rtasidagi mantiqiy va mazmuniy aloqalar avtomatik aniqlanadi. *Anaforik teglash* matndagi koreferent munosabatlarni belgilash jarayoni bo'lib, olmoshlarning matndagi qaysi birlikka ishora qilayotganini aniqlash imkonini beradi. *Prosodik teglash* esa og'zaki nutq birliklarining urg'u, intonatsiya va boshqa prosodik xususiyatlarini izohlaydi [Yuldashev A., 2025:341].

Shunday qilib, lingvistik teglashni faqat so'z turkumlarini belgilash bilan chegaralangan jarayon sifatida talqin qilish mutlaqo noto'g'ri. U matn birliklarining *morfologik, sintaktik, semantik, diskursiv, anaforik va prosodik* xususiyatlarini tavsiflash imkonini beruvchi ko'p darajali annotatsiya tizimini tashkil etadi. Korpus matnlarida bunday lingvistik ma'lumotlarning mavjudligi, ayniqsa, korpusning qidiruv imkoniyatlarini kengaytirish, murakkab so'rovlarni amalga oshirish va til birliklarini turli mezonlar asosida izlash imkonini beruvchi qidiruv tizimlarini yaratishda muhim ahamiyat kasb etadi.

Lingvistik teglash jarayonining samaradorligi va ilmiy qiymati uning qanday tamoyillar asosida teglanganligiga ham bog‘liq. V. Zaxarov ta’kidlaganidek, lingvistik teglashni tashkil etishda bir qator metodologik tamoyillarga amal qilish zarur. Jumladan, teglash sxemasi va uning nazariy asoslari aniq tavsiflanishi, unda lingvistik tushunchalarning umumqabul qilingan tizimidan foydalanilishi, foydalanuvchiga tushunarli va amaliyotda tanish bo‘lgan tahlil sxemasi tanlanishi, kiritilayotgan parametrlarning ilmiy jihatdan asoslangan bo‘lishi, teglash tizimi nazariy jihatdan imkon qadar neytral va an’anaviy yondashuvga asoslanishi hamda xalqaro standartlarga muvofiqligi ta’minlanishi lozim [Захаров В.П. , Богданова С.Ю., 2015:2].

Morfologik teglash korpus lingvistikasida muhim o‘rin egallaydi. Birinchidan, mavjud yirik korpuslarning katta qismi morfologik jihatdan teglangan bo‘lib, bu ulardan grammatik mezonlar asosida foydalanish imkonini beradi. Ikkinchidan, morfologik tahlil keyingi bosqichlarda amalga oshiriladigan sintaktik va semantik tahlil uchun boshlang‘ich lingvistik baza vazifasini bajaradi. Uchinchidan, kompyuter morfologiyasi sohasidagi yutuqlar katta hajmdagi matnlarni yuqori darajada avtomatik teglash imkonini bermoqda.

Lingvistik teglash tarixida Braun korpusi (Brown Corpus) alohida ahamiyat kasb etadi. 1980-yilda uning lingvistik jihatdan teglangan versiyasi yaratilib, unda matn tarkibidagi so‘z shakllarini lemmatizatsiya qilish imkoni yaratildi hamda sintaktik xususiyatlari belgilangan. Shu tariqa mazkur korpus tajribasi katta hajmdagi tabiiy til materiallarini standartlashtirilgan lingvistik teglar yordamida qayta ishlash va ulardan avtomatlashtirilgan lingvistik tahlil vositasi sifatida foydalanishning muhim amaliy asoslaridan birini shakllantirgan.

Morfologik teglash esa lingvistik teglashning asosiy bosqichi bo‘lib, keyingi sintaktik, semantik va boshqa darajadagi tahlillar uchun zarur lingvistik asosni yaratadi. Shu sababli korpus yaratishda teglash tizimining nazariy jihatdan asoslanganligi, terminologik izchilligi, xalqaro standartlarga mosligi va avtomatik qayta ishlash uchun qulay shaklda tashkil etilishi korpusning ilmiy va amaliy qiymatini belgilovchi muhim omillardan hisoblanadi.

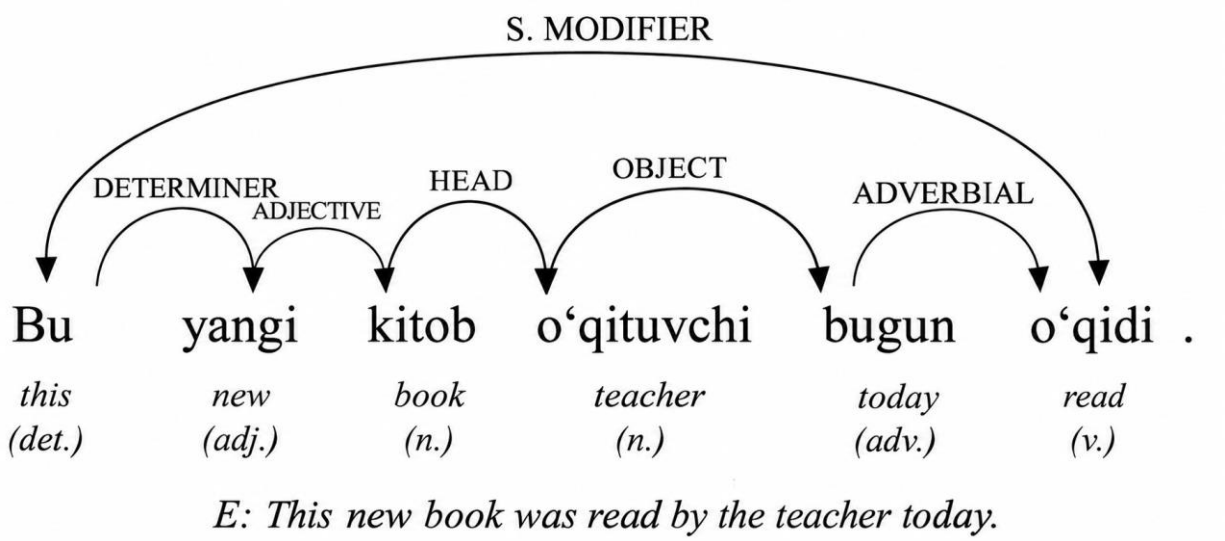
Sintaktik teglash morfologik tahlil jarayonida olingan ma’lumotlarga tayangan holda amalga oshiriladigan lingvistik parsing bosqichi bo‘lib, unda matn tarkibidagi leksik birliklarning o‘zaro sintaktik munosabatlari hamda ularning turli sintaktik tuzilmalarning xususiyatlari aniqlanadi va formal ravishda ifodalanadi..

Morfologik teglashdan farqli ravishda, sintaktik teglashda gap bo‘laklari va ular o‘rtasidagi munosabatlarni ifodalashning yagona universal modeli mavjud emas. Turli lingvistik nazariyalar va grammatik yondashuvlarga muvofiq sintaktik tuzilmani modellashtirishning bir nechta usullari shakllangan. Jumladan, so‘zlar o‘rtasidagi sintaktik munosabatlarni ifodalovchi tobelanish (dependency) grammatikasi, tarkib toptiruvchi (constituency) grammatika, qurilish sxemalari grammatikasi, gap bo‘laklariga asoslangan an’anaviy sintaktik nazariya hamda funksional grammatika kabi bir qator sintaktik tahlil usullari mavjud. Mazkur yondashuvlarning har biri sintaktik birliklarning tuzilishi, ularning o‘zaro aloqasi va gap tarkibidagi funksional vazifasini turlicha talqin qilishga asoslanadi.

Sintaktik teglash – tabiiy tilning sintaktik tuzilishini, ayniqsa, sintaktik munosabatlarni va gaplarning tarkibiy qismlari joylashuvini avtomatik belgilash. Yuqorida ta’kidlanganidek, turli grammatik nazariyalar gaplarning sintaktik tuzilishini tavsiflash uchun turli xil qarashlarni taklif qilgan. Kompyuter lingvistikasida eng keng tarqalgan tahlil usullari tobelanish grammatikasi va tarkib toptiruvchi grammatika bo‘lib, har ikki teglash turi uchun turli xil algoritmlar va muammolarga yechim sifatida turli yondashuvlar talab qilinadi.

Tobelanish grammatikasi asosidagi sintaktik tahlil turida har bir so‘zning boshqa so‘zga bo‘lgan tobe munosabati asosida tahlil amalga oshiriladi. Bu tahlil usuli tilni gapdagi bog‘lanishlar orqali modellashga qaratilgan. Misol uchun *Men kitob o‘qiyman* jumlasida *o‘qiyman* – fe’l bo‘lib, *Men* subyekt sifatida, *kitob* esa obyekt sifatida qaraladi. Bu yerda *Men va kitob* so‘zi *o‘qiyman*ga bog‘lanadi, shu tarzda, bu tahlil orqali jumladagi so‘zlarning o‘zaro bog‘lanishlari va ularning roli aniq belgilanadi.

Tobelanish asosidagi sintaktik tahlilda har bir so‘z boshqa biror so‘zga bog‘lanadi va bu bog‘lanish tizimi kesim atrofida birlashadi. O‘zbek tilida gapdagi markaziy bo‘lak kesim bo‘lib, boshqa barcha so‘zlar shu markazga bevozita yoki bilvosita tobelanadi. Masalan, *Men kitob o‘qidim* degan jumlada *o‘qidim* markaziy *”root”* so‘z bo‘lib, *men* va *kitob* so‘zlari unga bog‘langan. Natijada bog‘lanishlar yo‘naltirilgan chiziqlar shaklida ifodalanadi. Har bir bog‘lanish bir so‘zni boshqasiga bog‘laydi va bu bog‘lanish yo‘naltirilgan bo‘ladi. Ya’ni bir so‘z ikkinchisiga bog‘lanib, uni boshqaradi. Masalan, *Men o‘qidim* jumlasida *o‘qidim* so‘zi *men* so‘zini boshqaradi. Har bir bog‘lanish sintaktik munosabatni, masalan, subyekt, obyekt, predikat va boshqa sintaktik rollarni ifodalaydi. Diagrammalarni ko‘rish orqali tilshunoslar va kompyuter tizimlari grammatik tuzilmalarni osongina tahlil qilishi mumkin.



1-rasm. Tobelanish grammatikasi asosidagi tahlil natijasi

Tobelanish grammatikasi asosidagi tahlilni amalga oshirishda ilg‘or yondashuvlar va algoritmlar mavjud. Sun’iy intellekt va chuqur o‘rganish texnologiyalarining rivojlanishi bilan, tobelanish grammatikasi yanada aniqroq va samaraliroq bo‘ldi. Ayniqsa, BERT, GPT kabi transformer arxitekturalari tilning sintaktik va semantik strukturalarini tushunishda katta yordam beradi. Jahon tilshunosligida tobelanish grammatikasi uchun turli xil vositalar va tahlil usullari ishlab chiqilgan. Bu tahlil usullari turli tillarda so‘zlar o‘rtasidagi grammatik bog‘lanishlarni aniqlashga yordam beradi va turli maqsadlarda ishlatiladi.

Barcha tillarda bir xil sintaktik struktura yaratishga intilishlar orasida Universal tobelanish (Universal Dependencies) tashkiliy tizimi katta ahamiyatga ega bo‘ldi [Universal Dependencies [Elektron resurs]. – URL: https://universaldependencies.org/]. UD tizimi, turli tillarda tobelanish grammatikasi asosidagi sintaktik tahlilni umumiy sintaktik modelga asoslangan holda birlashtirishga yordam beradi. Bu, tildan qat’iy nazar, bir xil sintaktik bog‘lanishlarni aniqlashni maqsad qilgan. UD tilshunoslar va dasturchilar uchun ko‘p tilli sintaktik tahlil modellarini ishlab chiqish va ko‘plab tillardagi tobelik munosabatlarini tahlil

qilishni yanada osonlashtirdi. Xususan, Universal Dependencies bu ochiq va bepul korpus hisoblanib, bir necha tillarga mansub jumlalarni sintaktik va morfologik jihatdan tahlil qiluvchi dasturdir. Ushbu dasturda turk, qozoq, tatar, uyg‘ur tillariga doir tahlillarni ham amalga oshirish mumkin.

Hozirgi kompyuter lingvistikasi va korpus lingvistikasi borasida qilingan ishlar natijasida ko‘plab tabiiy tillarning sintaktik tuzilishini avtomatik aniqlash va lingvistik jihatdan teglashga mo‘ljallangan maxsus dasturiy vositalar ishlab chiqilgan. Ushbu vositalar yordamida matndagi sintaktik birliklar, ularning o‘zaro bog‘lanish turlari hamda grammatik vazifalari avtomatik yoki yarimavtomatik tarzda aniqlanadi.

Shu bilan birga, turli tillar uchun ishlab chiqilgan sintaktik teglash tizimlarida umumiy va universallashtirilishi mumkin bo‘lgan lingvistik xususiyatlar ham mavjud bo‘lib, ular ko‘p tilli korpuslar va universal grammatik modellarni yaratishda muhim metodologik asos vazifasini bajaradi. Demak, sintaktik teglash – bu tabiiy tillarni qayta ishlash va lingvistik tadqiqotlar uchun muhim vositadir. Korpus interfeysida lingvistik teglashni aks ettirish tadqiqotchining ish qulayligi va samaradorligini belgilovchi muhim jihatdir.

FOYDALANILGAN ADABIYOTLAR

1. Abdurakhmonova N. Dependency parsing based on Uzbek Corpus. InProceedings of the International Conference on Language Technologies for All (LT4All) 2019.
2. Abdurakhmonova, N. Z., Ismailov, A. S., & Mengliev, D. (2022, November). Developing NLP tool for linguistic analysis of turkic languages. In 2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON) (pp. 1790-1793). IEEE.
3. Abdurakhmonova, N., Astanova, G., Akhmedov, A., Mengliev, D., Ibragimov, B., & Abdullayev, A. (2025). Development of a Lexical Dataset and a Rule-Based Algorithm for the Analysis of Khorezm Dialects of the Uzbek Language. Data in Brief, 112216
4. Abdurakhmonova, N., Adalı, E., & Shamieva, N. (2025, September). Semantic Relationships Between Synset Adjectives for English-Uzbek Bilingual Thesaurus. In 2025 10th International Conference on Computer Science and Engineering (UBMK) (pp. 305-310). IEEE.
5. Mengliev, D., Barakhnin, V., Abdurakhmonova, N., & Eshkulov, M. (2024). Developing named entity recognition algorithms for Uzbek: Dataset insights and implementation. Data in Brief, 54, 110413.
6. Mengliev, D., Abdurakhmonova, N., Barkhnin, V., Ibragimov, B., Jurakulova, M., Urazaliyeva, M., & Islombekov, B. (2025, October). Integrating morphological stemming and syntactic parsing for low-resource Uzbek texts. In AIP Conference Proceedings(Vol. 3377, No. 1, p. 040003). AIP Publishing LLC.
7. Yuldashev A.U. O‘zbek tili korpusida lingvistik va ekstralingvistik qidiruv parametrlarining lingvistik va dasturiy asoslari: 10.00.11 – Til nazariyasi. Amaliy va kompyuter lingvistikasi: filologiya fanlari bo‘yicha falsafa doktori (PhD) ilmiy darajasini olish uchun yozilgan dissertatsiya. – Toshkent: Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti, 2025. – 172 b.
8. https://studfile.net/preview/1848098/
9. Universal Dependencies [Elektron resurs]. – URL: https://universaldependencies.org/]

NAZARIY VA AMALIY DARS MASHG'ULOTLARINING UYG'UNLIGI HAMDA TA'LIM YO'NALISHLARIDA JORIY ETILISHI

Kamolova Ozodaxon Muzaffar qizi
a doctoral researcher (tayanch doktorant) at the
Department of Pedagogy and Psychology at the
Denov Institute of Entrepreneurship and Pedagogy.
https://doi.org/10.5281/zenodo.23136615

Annotatsiya. *Ushbu tadqiqotda oliy ta'lim muassasalarida biologiya fanlarini o'qitish jarayonida nazariy va amaliy dars mashg'ulotlarining o'zaro uyg'unligini ta'minlash jihatlari yoritilgan. Shuningdek, "4+2" tartibidagi dual ta'lim tamoyillari asosida talabalarning kasbiy ko'nikmalarini rivojlantirish, ishlab chiqarish hamda amaliyot hududlari bilan integratsiyani kuchaytirish yo'llari ilmiy-metodik jihatdan tahlil etilgan. Shu nuqtayi nazardan, "4+2" tartibini birgina haftaalik darslarni taqsimlash shakli sifatida emas, balki bo'lajak pedagogning kasbiy tayyorgarligini amaliyotga yo'naltiruvchi yaxlit pedagogik tizim sifatida baholanishi haqida ma'lumot berib o'tilgan. Mazkur tizimning sifatini oliy ta'lim muassasalarining professor o'qituvchilari va amaliyot zonalaridagi rahbarlari bilan o'zaro hamkorlikda olib borishi va talabaning amaliy faoliyatini muntazam monitoring qilish hamda amaliyot natijalarini shaffof baholash kabi omillarga keng yoritilgan.*

Kalit so'zlar*: ta'lim, mashg'ulot, dual ta'lim, tamoyil, ko'nikma, metod.*

Аннотация. *В данном исследовании освещены аспекты обеспечения взаимосвязи теоретических и практических занятий в процессе преподавания биологических дисциплин в высших учебных заведениях. Также научно-методически проанализированы пути развития профессиональных навыков студентов на основе принципов дуального образования в формате «4+2», усиления интеграции с производством и сферой практики. Исходя из этого, отмечается, что формат «4+2» следует оценивать не просто как форму распределения недельных занятий, а как целостную педагогическую систему, направленную на практическую подготовку будущих педагогов. В данном контексте широко освещены такие факторы, как осуществление этой системы в сотрудничестве профессорско-преподавательского состава высших учебных заведений с руководителями зон практики, регулярный мониторинг практической деятельности студентов и прозрачная оценка результатов практики.*

Ключевые слова*: образование, занятие, дуальное образование, принцип, навык, метод.*

Abstract. *This study highlights the aspects of ensuring the interconnection between theoretical and practical classes in the process of teaching biological disciplines in higher education institutions. Furthermore, it scientifically and methodologically analyzes ways to develop students' professional skills based on the dual education principles of the "4+2" format, and to strengthen integration with production and practical fields. From this perspective, it is noted that the "4+2" format should be evaluated not merely as a form of weekly class distribution, but as an integral pedagogical system oriented toward practical training of future educators. In this context, extensive coverage is given to factors such as the*

implementation of this system in cooperation between the professors and teachers of higher education institutions and the heads of practical zones, regular monitoring of students' practical activities, and transparent assessment of practical results.

***Keywords**: education, class, dual education, principle, skill, method.*

Introduction

Today, fundamentally improving the quality of personnel training in the higher education system and closely connecting the educational process with practice is one of the most urgent tasks. Ensuring the harmony of theoretical and practical training classes based on qualification requirements, curricula, and programs serves as a legal and methodological foundation guaranteeing indicators of educational quality. In particular, applying acquired theoretical knowledge in practice during the teaching of biological sciences forms professional competencies in students. According to researchers' scientific studies, the harmony of theoretical and practical classes is formed on the basis of "4+2" dual education principles across all educational areas, where students spend certain days of the week acquiring theoretical knowledge at the university and the remaining days directly in production enterprises, factories, educational institutions, or in general, in practical training zones [Umurov Z.L., 2025:186]. In teaching biological sciences, this harmony is achieved by reinforcing the information received in theoretical classes through laboratory sessions, practical assignments, and projects, as well as by effectively utilizing interactive pedagogical technologies to develop students' independent thinking [Ahmadjonova U. T., 2026: 80]. The dual education process was developed by German scientists and relies on the "Duale Ausbildung" (two-way education) concept [Masharipova B.E., et al., 2025: 462]. Ensuring the harmony of theoretical and practical classes in the field of biology education sharply increases students' preparedness for professional activity [Salybekova, N., et al., 2021: 1106]. Furthermore, the implementation of the "4+2" principle serves as an integral bridge between theory and production, ensuring the emergence of competitive professionals in the future [Mirzakulova N.I., 2025: 211]. According to pedagogical analyses, it has been fully proven scientifically that a person retains only 10-20% of information in memory through listening, 30% through seeing, and up to 90% through practice [Djalalov, B., et. al., 2022: 26]. Therefore, it is advisable to harmonize theoretical and practical training classes based on qualification requirements and curricula [Numanova S.E., 2021: 585].

According to the Decree of the President of the Republic of Uzbekistan No. PF-134 dated May 11, 2022, "On approval of the national program for the development of public education in 2022–2026" in addition to paragraph 7, paragraph 2 of Resolution No. PQ-289 dated June 21, 2022, "On approval of the national program to improve the quality of pedagogical education and further develop higher education institutions training pedagogical personnel" introduced the implementation of weekly educational classes in a "4+2" format for full-time students studying in the 2nd-4th stages of higher education institutions training pedagogical personnel. Below, the process of organizing theoretical and practical training classes in biology education based on the "4+2" format is described in detail (Figure 1).

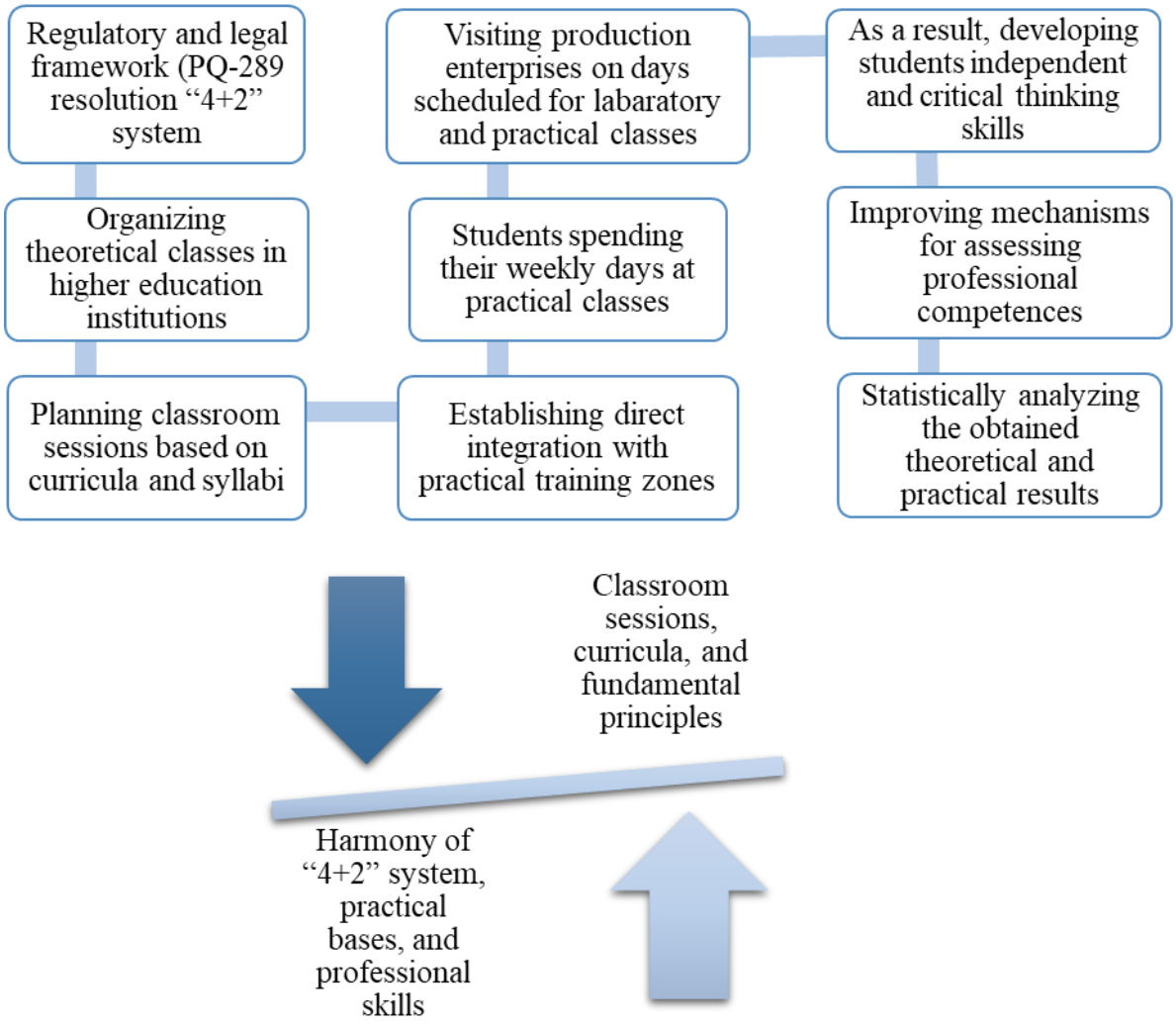


Figure 1. The process of integrating theoretical and practical training classes based on the "4+2" system.

Based on the model-based schema of this systematic structure (Figure 1), several stages are outlined to improve the quality indicators of teaching biological sciences in higher education institutions. Specifically, relying on the requirements of Resolution No. PQ-289, the "4+2" principle, syllabi, and working curricula in organizing the educational process is of great importance for mastering theoretical knowledge in classrooms, through which mechanisms for assessing professional competencies and training personnel are improved. The following figure illustrates the advantages of the "4+2" principle of the dual education system (Figure 2). In this principle, cooperation among higher education institutions, the practical training base, and the student plays a crucial role. Theoretical knowledge provided in higher education institutions is further reinforced during practical training classes. The repetition of theoretical knowledge during the practical process enables students to continuously improve their activities, enhance independent learning and critical thinking skills, identify their shortcomings, and eliminate them. The "4+2" principle serves precisely to harmonize theory and practice within higher education institutions.

Furthermore, the "4+2" principle makes it possible to ensure the continuity of forming professional competencies. In addition, competence manifests not merely as a collection of theoretical knowledge, but serves for the regular observation, evaluation, and analysis of knowledge, skills, abilities, and professional experience in practical activity. Practical classes

develop a student's competencies in lesson planning, making appropriate decisions for pedagogical processes, establishing quality communication with students, utilizing educational tools, and evaluating their own pedagogical activity.

The opportunities of the "4+2" principle in teaching biology are wide; specifically, biology education requires practical activities such as observation, experiments, laboratory work, working with natural objects, and analyzing ecological and biological processes alongside theoretical knowledge.

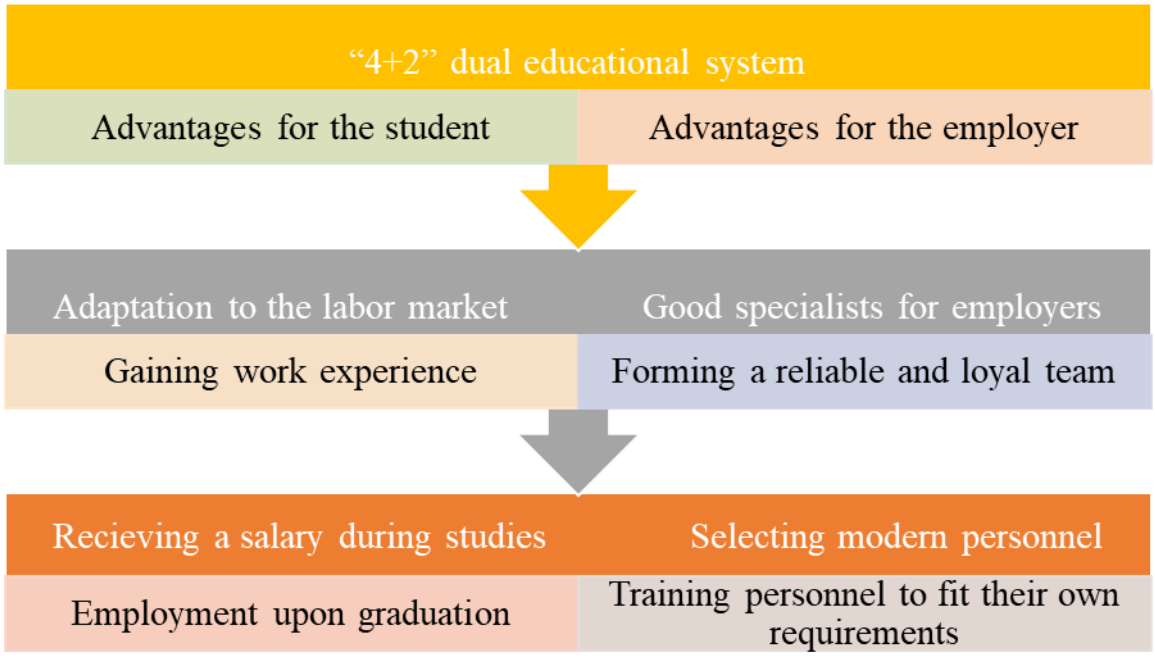


Figure 2. Main advantages of the system

In this regard, this system has a number of advantages, which include:

- the graduate having not only theoretical knowledge but also at least one to two years of actual work experience upon receiving their diploma;
- salaries being paid regularly to students working in enterprises based on dual education in the manner established by legislation;
- three-way contracts being concluded between higher education institutions and practical training zones, fully adapting educational programs directly to market demands.

There are also advantages of this system for employer zones, which constitute the following:

- training personnel to fit their own requirements;

involving students in enterprise projects and benefiting from their initiative;

- incentives provided by the state, including the allocation of subsidies by the state to the enterprise for fully employing the student.
- deduction of expenses related to the education of students from the taxable base;
- there are several advantages, such as the formation of a reliable and loyal team.

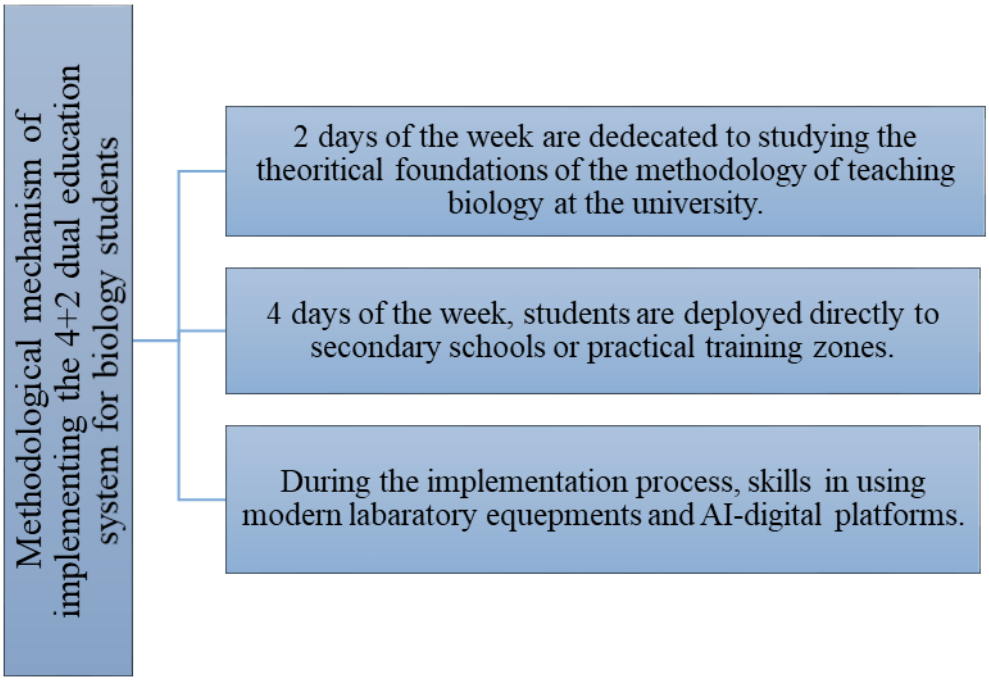


Figure 3. Methodological mechanism of implementing the "4+2" dual education system for biology students.

By applying these stages to the educational process, firstly, students develop practical skills by independently applying biological competencies in a school environment; secondly, they gain the opportunity to independently use modern technologies when teaching biology at school; thirdly, while teaching lessons to students, their psychological and pedagogical competencies in the classroom develop; fourthly, students' communication skills are further shaped through the process of working with school colleagues; fifthly, the skill of working with a number of school regulatory documents (Kundalik.com system, quarterly plans, subject programs, curricula, calendar work plans) is formed at an early stage.

Conclusion

Today, the principle of the harmony of theory and practice is of great importance in teaching biological sciences, and this principle is implemented across the "Science + Education + Production" cross-section. As a result of the introduction of "4+2" dual education, the student becomes an independent researcher in the school environment. Formulating the harmony of theoretical and practical training on the basis of the "4+2" dual education system in the biology education direction radically develops the level of professional preparedness of future educators. The application of this principle to the educational process equips students not only with theoretical knowledge, but also with competencies in independently conducting biological experiments in practical training zones, applying interactive methods, and managing a class community. This accomplishes a number of tasks, such as ensuring continuity between higher education and secondary schools and supplying competitive personnel to the labor market. The implementation of the "4+2" principle in higher education institutions serves the harmony of theoretical and practical training classes and their introduction.

REFERENCES

1. Umurov, Z.L. The Harmony of Dual Education and Production is a Modern Education Conception // Inter education & global study // Vol. 3, no. 3, 2025, – P. 185-192.

2. Ahmadjonova, U.T. 4+2 Educational system harmony of theory and practice // Экономика и социум // No. 4-2 (143), 2026, – P. 80-83.

2. Masharipova, B.E., et al. Opportunities of Dual Education in the Continuous Education System: On the Example of Uzbekistan and Foreign Countries // Partner Conferences // 1(13), 2025, – P. 462–466.

1. https://academicsbook.com/index.php/konferensiya/article/view/1660

3. Salybekova, N., et al. Pupils' research skills development through project-based learning in biology. Cypriot Journal of Educational Science, 16(3), 2021. – P. 1106-1121. https://doi.org/10.18844/cjes.v16i3.5829

4. Mirzakulova, N.I. Competences Developed in Students During the 4+2 Practical Internship // Theoretical Aspects of the Formation of Pedagogical Sciences, 4(4), 2025, – P. 211-218. https://doi.org/10.71337/inlibrary.uz.tafps.70178

5. Djalalov, B., et al. Psychological and Pedagogical Features of Intellectual Development of Students // Oriental renaissance: Innovative, educational, natural and social sciences, vol. 2, no. 12, 2022, – P. 26-37. https://cyberleninka.ru/article/n/o-quvchilarni-intellektual-rivojlantirishning-psixologik-pedagogik-xususiyatlari

7. Numanova, S.E. Methodology for Developing Students' Practical Knowledge in the Educational Process // Oriental renaissance: Innovative, educational, natural and social sciences, vol. 1, no. 9, 2021, pp. 585-589. https://cyberleninka.ru/article/n/ta-lim-jarayonida-talabalarning-amaliy-bilimlarini-rivojlantirish-metodikasi

8. Ivan Beloev, et al. Formation of Professional Skills of Agricultural Engineers during Laboratory Practice when Studying Fundamental Science // 2024, – P. 144-156. DOI:10.53656/str2023-3-2-stu

NOMLANGAN BIRLIKLARNI ANIQLASH TIZIMLARINING NAZARIY ASOSLARI VA EVOLYUTSIYASI

Kanayeva Ozoda Esirgapovna
O‘zbekiston Milliy universiteti,
Jurnalistika va o‘zbek filologiyasi fakulteti,
Kompyuter lingvistikasi mutaxassisligi I bosqich magistranti
https://doi.org/10.5281/zenodo.23136637

Annotatsiya. *Mazkur maqolada nomlangan birliklarni aniqlash tizimlarining nazariy asoslari, rivojlanish bosqichlari va zamonaviy yondashuvlari tahlil qilingan. Tadqiqotda ushbu sohaning qoida asosidagi dastlabki modellaridan tortib statistik va chuqur o‘rganishga tayanuvchi tizimlargacha bo‘lgan evolyutsiyasi yoritilgan. Shuningdek, nomlangan birliklarni aniqlash tizimlarining samaradorligi nafaqat matematik model, balki ma’lumotlar to‘plamining sifati, annotatsiya aniqligi va tilning lingvistik xususiyatlariga ham bevosita bog‘liq ekani ko‘rsatib berilgan.*

Kalit so‘zlar: *nomlangan obyektlar, kompyuter lingvistikasi, tabiiy tilni qayta ishlash, korpus, annotatsiya, o‘zbek tili, gibrid model.*

Annotation. *This article analyzes the theoretical foundations, stages of development, and modern approaches of Named Entity Recognition (NER) systems. The study highlights the evolution of this field, from early rule-based models to statistical and deep learning–based systems. It is also demonstrated that the effectiveness of NER systems depends not only on the mathematical model, but also directly on the quality of the dataset, the accuracy of annotation, and the linguistic characteristics of the language.*

Keywords: *named entities, computational linguistics, natural language processing, corpus, annotation, Uzbek language, hybrid model.*

Аннотация. *В данной статье проанализированы теоретические основы, этапы развития и современные подходы систем распознавания именованных сущностей (NER). В исследовании освещена эволюция данной области — от ранних правил-ориентированных моделей до статистических и основанных на глубоком обучении систем. Также показано, что эффективность систем распознавания именованных сущностей зависит не только от математической модели, но и непосредственно от качества набора данных, точности аннотирования и лингвистических особенностей языка.*

Ключевые слова: *именованные сущности, компьютерная лингвистика, обработка естественного языка, корпус, аннотация, узбекский язык, гибридная модель.*

Kirish. Tabiiy tilni qayta ishlash va kompyuter lingvistikasining muhim vazifalaridan biri matn tarkibidagi shaxs, tashkilot, joy, sana, geosiyosiy birlik va boshqa maxsus obyektlarni avtomatik aniqlashdan iborat. Ushbu vazifa ilmiy adabiyotlarda nomlangan mohiyatlarni aniqlash deb yuritiladi va u axborot qazib olish, savol-javob tizimlari, mashinali tarjima, matnlarni indekslash, bilim bazalari yaratish hamda ma’lumotlarni anonimlashtirish kabi ko‘plab amaliy yo‘nalishlarda qo‘llaniladi.

Jurafsky va Martinning ta'kidlashicha, tabiiy tilni qayta ishlash jarayonida nomlangan birliklarni aniqlash matndagi muhim birliklarni strukturaviy va semantik jihatdan ajratib ko'rsatishga xizmat qiluvchi tayanch vazifalardan biri hisoblanadi. Ularning fikricha, bunday tizimlar axborotni avtomatik qazib olish, matnni tushunish va bilim bazalarini shakllantirish kabi yuqori bosqichli vazifalarning asosiy tarkibiy qismidir [3].

Nomlangan birliklarni aniqlash yo'nalishi dastlab asosan ingliz tili kabi resurslari boy tillar doirasida shakllangan. Keyinchalik xalqaro baholash dasturlari, ayniqsa CoNLL-2003 doirasida olib borilgan tadqiqotlar ushbu sohani tizimlashtirishga xizmat qildi va turli tillar bo'yicha modellarni qiyosiy baholash imkonini yaratdi [1]. Mazkur bosqichdan boshlab nomlangan mohiyatlarni aniqlash nafaqat amaliy, balki mustaqil nazariy masala sifatida ham qarala boshlandi.

Biroq tilga bog'liq bo'lmagan yondashuvlar barcha tillar uchun bir xil darajada samarali emas. Ayniqsa o'zbek tili kabi agglyutinativ tillarda qo'shimchalarning serqatlamligi, so'z shakllarining ko'pligi, erkin sintaktik qurilish va morfologik murakkabliklar tayyor modellarni bevosita qo'llashni cheklaydi. Shu sababli o'zbek tili uchun maxsus belgilangan korpuslar, annotatsiyalangan ma'lumotlar to'plamlari va tilning ichki xususiyatlariga mos algoritmlarni ishlab chiqish masalasi dolzarb bo'lib bormoqda [2].

Metodologiya. Nomlangan birliklarni aniqlash tizimlarining rivojlanish tarixini tahlil qilish shuni ko'rsatadiki, bu yo'nalish uch asosiy metodologik bosqich orqali shakllangan.

Birinchi bosqich: qoida va lug'atga asoslangan tizimlar. Dastlabki davrda nomlangan mohiyatlarni aniqlash asosan lingvistik qoidalar, imlo belgilari, bosh harf, tayyor nomlar ro'yxati va sintaktik andozalarga tayangan holda amalga oshirilgan [6]. Bunday tizimlarning afzalligi ularning tushunarli va nazorat qilinadigan bo'lishidadir. Agar qoidalar aniq tuzilsa, ular muayyan sohada yuqori aniqlik bera oladi. Biroq bu yondashuv katta hajmdagi qo'lda tayyorlangan resurslarni talab qiladi, yangi birliklarni aniqlashda sust ishlaydi hamda boshqa mavzular yoki janrlarga ko'chirilganda samaradorligini yo'qotadi.

Bikel, Schwartz va Weischedel nomlangan birliklarni aniqlashning dastlabki bosqichlarida lingvistik qoidalar, imlo xususiyatlari va maxsus nomlar bazasi katta ahamiyatga ega ekanini ko'rsatadi. Olimlarning fikricha, qoida va lug'atga asoslangan yondashuvlar ayniqsa ma'lum soha doirasida yuqori aniqlik bera oladi, biroq ularning moslashuvchanligi cheklangan bo'lib, yangi matnlar va noma'lum birliklar bilan ishlashda samaradorlik pasayadi [6].

Ikkinchi bosqich: statistik modellar. Keyingi bosqichda nomlangan birliklarni aniqlash matematik ehtimollik va statistik o'rganishga tayanuvchi modellar yordamida amalga oshirila boshlandi. Bu jarayonda yashirin Markov modellari, shartli tasodifiy maydonlar va boshqa ketma-ketlikni belgilash usullari keng qo'llanildi [5]. Aynan shu yondashuvlar orqali so'zning o'z shakli bilan birga uning atrofidagi konteksti, oldingi va keyingi birliklar bilan bog'lanishi ham hisobga olinadigan bo'ldi. CoNLL-2003 doirasida ishlab chiqilgan baholash mezonlari ana shu modellarni turli tillarda solishtirishga imkon bergan [1].

Nadeau va Sekine nomlangan birliklarni aniqlash tizimlari evolyutsiyasini tahlil qilar ekan, statistik modellar ushbu sohani yangi bosqichga olib chiqqanini ta'kidlaydi. Ularning fikricha, statistik yondashuvlarning ustunligi shundaki, ular faqat tayyor qoidalarga emas, balki korpusda kuzatilgan takroriy qonuniyatlarga tayangan holda birliklarni aniqlaydi. Shu bilan birga, bunday tizimlar natijasining barqarorligi to'g'ridan-to'g'ri o'quv ma'lumotlarining sifati va hajmiga bog'liq [5].

Uchinchi bosqich: chuqur o‘rganish va kontekstual modellar. So‘nggi bosqich chuqur o‘rganishga asoslangan modellar bilan bog‘liq. Bunda ikki yo‘nalishda ishlovchi xotirali neyron tarmoqlar, kontekstual vektor ifodalari va keyinchalik Transformerlarga asoslangan yondashuvlar keng qo‘llanila boshladi [4], [7]. Ushbu tizimlar so‘zning faqat lug‘aviy ma’nosini emas, balki uning kontekst ichidagi vazifasini ham o‘rganadi. Natijada ko‘p ma’noli birliklar, murakkab nomlar va uzoq masofali bog‘lanishlar ancha aniqroq tahlil qilinadi.

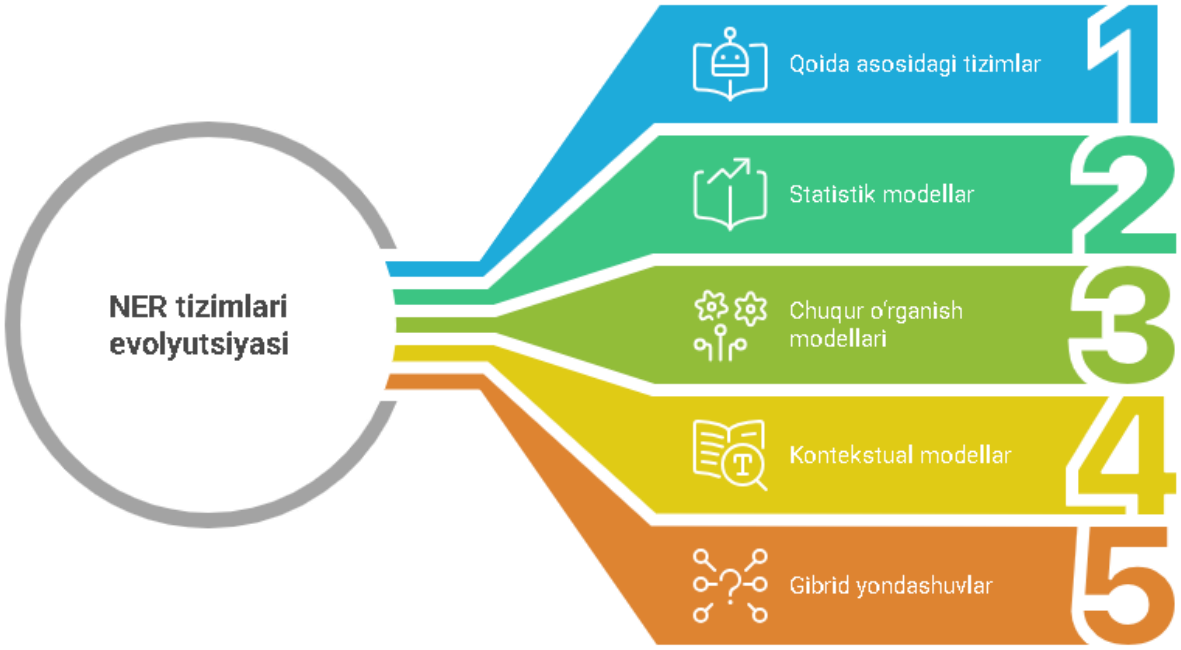
Lample va hammualliflar nomlangan birliklarni aniqlashda chuqur o‘rganishga asoslangan arxitekturalar, ayniqsa ikki yo‘nalishli neyron tarmoqlar bilan ketma-ketlikni nazorat qiluvchi qatlamlarning uyg‘unligi yuqori natija berishini asoslab beradi. Ularning fikricha, bunday modellar so‘zning shakliy xususiyatlari bilan birga uning kontekstdagi funksiyasini ham hisobga olgani sababli murakkab nomlarni aniqlashda an’anaviy yondashuvlarga nisbatan ustunlik qiladi [4].

Shu bilan birga, zamonaviy model qanchalik kuchli bo‘lmasin, uning natijasi o‘qitilayotgan ma’lumotlar sifatiga bog‘liq bo‘lib qoladi. Annotatsiya sifati past bo‘lgan yoki morfologik jihatdan noto‘g‘ri tuzilgan korpuslar eng ilg‘or tizimlarning ham samaradorligini cheklaydi [2]. Ayniqsa o‘zbek tili kabi shakl o‘zgarishi kuchli tillarda morfologik tahlil va to‘g‘ri belgilash jarayonlari asosiy tayanch vazifasini bajaradi.

1-jadval. Nomlangan birliklarni aniqlash tizimlarining rivojlanish bosqichlari

Bosqich	Asosiy yondashuv	Xususiyati	Afzalligi	Kamchiligi
Dastlabki bosqich	Qoida va lug‘atga asoslangan tizimlar	Tilshunoslar tuzgan qoidalar, tayyor nomlar ro‘yxati va andozalarga tayanadi	Muayyan sohada yuqori aniqlik beradi	Yangi birliklarni topishda sust, ko‘p qo‘lda mehnat talab qiladi
O‘rta bosqich	Statistik modellar	Ehtimollik va ketma-ketlikka asoslangan usullar orqali ishlaydi	Kontekstni qisman hisobga oladi	Korpus sifati past bo‘lsa samaradorligi tushadi
Zamonaviy bosqich	Chuqur o‘rganish va kontekstual modellar	So‘zning keng kontekstdagi ma’nosini o‘rganadi	Yuqori aniqlik va moslashuvchanlik beradi	Hisoblash resursi ko‘p talab etadi
Istiqbolli bosqich	Gibrid yondashuv	Lingvistik qoidalar va sun’iy intellekt usullari birlashtiriladi	Kam resursli tillar uchun samaraliroq	Tuzilishi murakkabroq

Nazariy tahlil va evolyutsiya. Nomlangan birliklarni aniqlash tizimlarining evolyutsiyasi “qoida asosidagi yondashuv”dan “ma’lumotga tayanuvchi yondashuv”ga o‘tish jarayonini yaqqol namoyon etadi. Ilk bosqichda tizimlar tilshunoslar tomonidan ishlab chiqilgan qoidalar asosida faoliyat yuritgan bo‘lsa, keyingi bosqichda korpus va statistik ma’lumotlar markaziy o‘ringa chiqdi. Hozirgi bosqichda esa modelga oldindan tayyor qoidalar berishdan ko‘ra, unga katta hajmdagi ma’lumotlar orqali qonuniyatlarni mustaqil o‘rgatish yondashuvi ustunlik qilmoqda [4].



1-rasm. NER tizimlari evolyutsiyasi

Shu bilan birga, bu evolyutsiya "lingvistikaning o‘rni kamaydi" degani emas. Aksincha, chuqur o‘rganish modellarining muvaffaqiyati ham lingvistik jihatdan puxta tayyorlangan ma'lumotlar bazasiga tayanadi. Masalan, korpusda entity chegaralari, toifalar o‘rtasidagi farq, ko‘p ma'nolilik, qisqartmalar, imloviy variantlar va morfologik shakllar aniq ko‘rsatilmagan bo‘lsa, model noto‘g‘ri umumlashtirishlar qiladi. Shu sababli zamonaviy nomlangan mohiyatlarni aniqlash tizimlari aslida matematik modellar bilan lingvistik tayyorgarlikning uyg‘unlashgan shaklidir.

NER tizimlari nazariyasida yana bir muhim jihat – ularni baholash mezonlari masalasidir. Bu yerda odatda aniqlik, qamrov va ularning garmonik o‘rtachasi bo‘lgan F1 ko‘rsatkichi qo‘llanadi [1]. Ammo baholashda alohida token emas, balki to‘liq nomlangan birlik hisobga olinishi kerak. Chunki tizim bitta murakkab tashkilot nomining ikki qismini to‘g‘ri topib, bir qismini xato qilgan bo‘lsa, bu butun birlikni noto‘g‘ri aniqlagan deb qaraladi. Mazkur yondashuv tizim samaradorligini yanada realroq baholash imkonini beradi.

O‘zbek tilida nomlangan mohiyatlarni aniqlash tizimlarini yaratishda asosiy qiyinchilik tilning agglyutinativ tuzilishi bilan bog‘liq. Biror shaxs yoki joy nomiga kelishik, egalik, ko‘plik va boshqa qo‘shimchalar qo‘shilganda uning yuzaki shakli o‘zgaradi. Natijada oddiy token darajasidagi model bunday birlikni yangi so‘z sifatida qabul qilishi mumkin. Bu esa morfologik segmentatsiya va lemmatizatsiya masalasini alohida dolzarb qiladi.

Yana bir muhim omil – o‘zbek tilidagi resurslarning cheklanganligi. Ingliz tilida millionlab belgilangan misollar mavjud bo‘lsa, o‘zbek tilida bunday korpuslar endigina shakllanmoqda. Shu nuqtai nazardan, Mengliev va hammualliflar tomonidan ishlab chiqilgan o‘zbek tili uchun dataset ushbu yo‘nalishda muhim qadam hisoblanadi [2]. Mazkur ish shuni ko‘rsatadiki, agar ma'lumotlar to‘plami sifatli va izchil annotatsiya asosida tuzilsa, o‘zbek tilida ham xalqaro standartlarga yaqin natijalarga erishish mumkin.

O‘zbek tilidagi nomlarni aniqlashda yana bir muammo – kontekstga kuchli bog‘liqlikdir. Ayrim so‘zlar shaxs nomi, joy nomi yoki tashkilot nomi sifatida turli vaziyatlarda qo‘llanishi mumkin. Bunday holatda faqat lug‘atga tayangan tizimlar yetarli emas, balki keng kontekstni

ham hisobga oluvchi modellar zarur bo‘ladi. Transformerlarga asoslangan zamonaviy tizimlar aynan shu jihatdan ustunlikka ega [7].

Natijalar. Mavjud ilmiy manbalar tahlili nomlangan mohiyatlarni aniqlash tizimlarida nazariy yondashuvlarning almashinuvi amaliy natijalarda ham yaqqol aks etganini ko‘rsatadi. CoNLL-2003 doirasidagi eng yaxshi statistik modellar ingliz tili uchun 88 foiz atrofida F1 ko‘rsatkichini qayd etgan [1]. Keyinchalik chuqur o‘rganish va Transformerlarga asoslangan modellar kirib kelishi bilan bu ko‘rsatkich 92–93 foizgacha ko‘tarilgan [7].

O‘zbek tili bo‘yicha olib borilgan tajribalar ham optimistik natijalarni ko‘rsatmoqda. Mengliev va boshqalar tomonidan ishlab chiqilgan ma’lumotlar to‘plami va modellar asosida o‘tkazilgan sinovlar o‘zbek tilidagi nomlangan mohiyatlarni avtomatik aniqlashning amaliy jihatdan samarali bo‘lishini tasdiqlaydi [2, 9, 10]. Biroq bu yerda bir narsani alohida ta’kidlash zarur: natijaga faqat model emas, balki korpus sifati, annotatsiya izchilligi va tilga xos morfologik xususiyatlarni qamrab olgan tayyorlov bosqichlari ham bevosita ta’sir ko‘rsatadi.

Shu nuqtai nazardan, o‘zbek tilidagi NER tizimlarida yuqori natijaga erishish uchun quyidagi omillar ustuvor hisoblanadi:

a) sifatli va standartlashtirilgan dataset yaratish;
b) morfologik tahlil vositalarini integratsiya qilish;
c) lingvistik qoidalarni neyron model bilan birlashtirish;
d) turli janrdagi matnlarni qamrab oluvchi korpuslarni kengaytirish.

Muhokama. Nazariy va amaliy tahlillar shuni ko‘rsatadiki, NER tizimlarini rivojlantirish faqat yangi model topish bilan hal bo‘ladigan masala emas. Bu jarayon lingvistik tayanch, sifatli ma’lumotlar bazasi va zamonaviy hisoblash usullarining o‘zaro uyg‘unligiga tayanadi. Agar bitta omil yetishmasa, umumiy tizim barqaror ishlamaydi.

Ayniqsa o‘zbek tili uchun gibrid yondashuvlar istiqbolli ko‘rinadi. Bunda an’anaviy lingvistik qoidalar, morfologik tahlil va chuqur o‘rganish modellarini birlashtirish mumkin. Bunday yondashuv bir tomondan tilning ichki qonuniyatlarini saqlab qolsa, ikkinchi tomondan modelga katta kontekstni o‘rganish imkonini beradi. Shu sababli o‘zbek tili uchun tayyor xorijiy modelni ko‘chirib ishlatishdan ko‘ra, mahalliy resurslar asosida moslashtirilgan gibrid tizimlar samaraliroq bo‘lishi kutiladi.

Xulosa. Xulosa qilib aytganda, nomlangan mohiyatlarni aniqlash tizimlarining evolyutsiyasi qoida asosidagi modellar, statistik yondashuvlar va chuqur o‘rganishga tayanuvchi zamonaviy tizimlar o‘rtasidagi uzviy rivojlanish jarayonini aks ettiradi. Ushbu rivojlanish natijasida tizimlarning aniqligi sezilarli darajada oshgan bo‘lsa-da, ularning haqiqiy samaradorligi baribir ma’lumotlar sifati va lingvistik moslashuvga bog‘liq bo‘lib qolmoqda.

O‘zbek tili misolida bu holat yanada aniqroq namoyon bo‘ladi. Agglyutinativ tuzilish, morfologik murakkablik va resurslarning kamligi tayyor universal modellarni cheklaydi. Shu sababli o‘zbek tilida nomlangan mohiyatlarni aniqlash bo‘yicha kelgusidagi eng muhim vazifa – sifatli korpuslar va annotatsiyalangan datasetlarni kengaytirish, shuningdek, lingvistik va sun’iy intellekt yondashuvlarini birlashtirgan gibrid modellarni ishlab chiqishdan iborat.

FOYDALANILGAN ADABIYOTLAR

1. Tjong Kim Sang E. F., De Meulder F. Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition // Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003. – 2003. – 142-147 - B.

2. Mengliev D., Barakhnin V., Abdurakhmonova N., Eshkulov M. Developing named entity recognition algorithms for Uzbek: Dataset insights and implementation // Data in Brief. – Vol. 54. – 2024.
3. Jurafsky D., Martin J. H. Speech and Language Processing. – 3rd ed. draft. – Pearson, 2023. – 600 - B.
4. Lample G., Ballesteros M., Subramanian S., Kawakami K., Dyer C. Neural Architectures for Named Entity Recognition // Proceedings of NAACL-HLT. – 2016. – 260-270 - B.
5. Nadeau D., Sekine S. A survey of named entity recognition and classification // Lingvisticae Investigationes. – 2007. – Vol. 30, No. 1. – 3-26 - B.
6. Bikel D. M., Schwartz R., Weischedel R. M. An Algorithm that Learns What's in a Name // Machine Learning. – 1999. – 211-231 - B.
7. Devlin J., Chang M. W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proceedings of NAACL-HLT. – 2019. – 4171-4186 - B.
8. Xolmanova Z. Kompyuter lingvistikasi. O‘quv qo‘llanma. – Toshkent: O‘zbekiston, 2020. – 150-155 - B.
9. Abdurakhmonova, N., Mengliev, D., Mardonova, S., Asadov, T., Xalilov, F., & Maxsudova, F. (2025, September). A Hybrid Named-Entity Recognition Algorithm for Ecological Documentation in Uzbekistan. In 2025 10th International Conference on Computer Science and Engineering (UBMK) (pp. 1778-1781). IEEE.
10. Abdurakhmonova, N., Adizova, N., Jamolitdinova, D., Begimov, A., Sultanova, A., & Juraeva, G. (2025, November). Development of Hybrid Algorithm for Named Entity Recognition in Uzbek Public Administration Texts. In 2025 IEEE XVII International Scientific and Technical Conference on Actual Problems of Electronic Instrument Engineering (APEIE) (pp. 1-6). IEEE.

A CONTEXT-AWARE NOISY-CHANNEL MODEL FOR AUTOMATIC SPELLING CORRECTION OF UZBEK USER-GENERATED TEXTS

Sevinch Kenjayeva
Master's student
National University of Uzbekistan named after Mirzo Ulugbek
Scientific supervisor: N. Z. Abdurakhmonova, Doctor of Philological Sciences, Professor
https://doi.org/10.5281/zenodo.23136658

Abstract. *User-generated Uzbek texts — social media comments, customer reviews and citizens' appeals to e-government portals — contain a high proportion of spelling errors that degrade the quality of downstream social-media monitoring tasks. The paper formulates spelling correction for Uzbek as a noisy-channel decoding problem in which the source model is an interpolation of a word n-gram model and a BERT-type contextual neural language model, and the channel model is a set of error probabilities estimated from a corpus of aligned erroneous–correct word pairs. Candidate generation is constrained by a finite-state affix analyser, which keeps the search space small for an agglutinative language. The paper presents the formal model, a beam-search decoding algorithm, an evaluation methodology that measures not only word-level accuracy but also the effect of correction on a downstream sentiment classifier, and a worked example illustrating how the channel and source models interact.*

Keywords: *noisy-channel model, spelling correction, Uzbek language, user-generated content, contextual neural language model, social media monitoring, error model.*

Annotatsiya: *Foydalanuvchilar tomonidan yaratilgan o'zbekcha matnlar — ijtimoiy tarmoq izohlari, mijozlar sharhlari va elektron hukumat portallariga yo'llangan murojaatlar — ijtimoiy media monitoringi vazifalarining sifatini pasaytiruvchi ko'p miqdordagi imloviy xatolarga ega. Maqolada o'zbek tili uchun imlo xatolarini tuzatish noisy channel modeli asosida decoding masalasi sifatida qo'yiladi: source modeli so'z n-gram modeli va BERT turidagi masked language modelning interpolatsiyasidan, channel modeli esa xato–to'g'ri so'z juftliklari korpusidan baholangan xato ehtimolliklaridan iborat. Candidate generation finite-state affiks analizatori bilan cheklanadi, bu agglutinativ til uchun qidiruv fazosini ixcham saqlaydi. Maqolada rasmiy model, beam search asosidagi decoding algoritmi, tuzatishning nafaqat so'z darajasidagi aniqligini, balki downstream vazifadagi sentiment tasniflagichiga ta'sirini ham o'lchovchi baholash metodikasi hamda channel va source modellarining o'zaro ta'sirini ko'rsatuvchi ishlangan misol keltirilgan.*

Kalit so'zlar: *noisy channel model, imlo xatolarini tuzatish, o'zbek tili, foydalanuvchi matnlari, masked language model, ijtimoiy media monitoringi, xato modeli.*

Аннотация. *Пользовательские тексты на узбекском языке — комментарии в социальных сетях, отзывы и обращения граждан на порталы электронного правительства — содержат высокую долю орфографических ошибок, снижающих качество задач мониторинга социальных медиа. В статье исправление орфографии для узбекского языка формулируется как задача декодирования в модели зашумлённого канала: модель источника представляет собой интерполяцию n-граммной и контекстной нейронной языковой модели типа BERT, а модель канала — вероятности ошибок,*

оценённые по корпусу выровненных пар «ошибочное–правильное слово». Генерация кандидатов ограничивается конечно-автоматным анализатором аффиксов. Представлены формальная модель, алгоритм декодирования на основе лучевого поиска, методика оценки, учитывающая влияние исправления на классификатор тональности, а также разобранный пример взаимодействия модели канала и модели источника.

Ключевые слова: *модель зашумлённого канала, исправление орфографии, узбекский язык, пользовательский контент, контекстная нейронная языковая модель, мониторинг социальных медиа, модель ошибок.*

Kirish

Raqamli transformatsiya jarayonida ijtimoiy tizimlarni o‘rganishning asosiy manbalaridan biri foydalanuvchi matnlari — ijtimoiy tarmoq izohlari, sharhlar, davlat portallariga yo‘llangan murojaatlar — bo‘lib qoldi. O‘zbek tilida bunday matnlar ustida mavzularni ajratish [Sayfullaeva R. va b., 2025b: 1450], bemorlar sharhlarining sentiment tahlili [Sayfullaeva R. va b., 2025a: 2240] va soliq to‘lovchilar murojaatlarini tasniflash [Mengliev D. va b., 2025: 2090] kabi amaliy vazifalar yechilmoqda. Biroq bu tadqiqotlarda umumiy to‘siq qayd etiladi: foydalanuvchi matnlari tahrirdan o‘tmagan bo‘lib, bir xil so‘z o‘nlab yozuv variantlarida uchraydi. Bu lug‘atga tayanadigan tasniflagichlar uchun leksik siyraklikni oshiradi va neyron modellarning umumlashtirish qobiliyatini pasaytiradi.

Shu bois imlo xatolarini avtomatik tuzatish ijtimoiy media monitoringi quvurining zaruriy dastlabki bosqichi hisoblanadi. O‘zbek tili uchun bu masala ikki jihatdan murakkab. Birinchidan, til agglutinativ bo‘lgani uchun so‘zshakllar to‘plami amalda cheksiz va xato so‘z ko‘pincha lug‘atda yo‘q, lekin morfologik jihatdan yig‘iladigan shaklga o‘xshaydi [Abdurakhmonova N. va b., 2022: 61; Sharipov M., Sobirov O., 2022: 154]. Ikkinchidan, foydalanuvchi matnlaridagi xatolar tasodifiy emas: diakritikaning tushirilishi, tutuq belgisining yozilmasligi, h/x va q/k chalkashuvi kabi tizimli sabablarga ega, ya’ni ularning taqsimoti ingliz tili uchun ishlab chiqilgan xato modellaridan sezilarli farq qiladi.

Mazkur ishning maqsadi — o‘zbek tilidagi foydalanuvchi matnlari uchun imlo xatolarini tuzatishni noisy channel modeli doirasida rasmiylashtirish, channel modelini o‘zbek tiliga xos xatolar korpusidan o‘rganish, source modeliga kontekstni kirituvchi til modelini integratsiyalash va tizimni baholashning downstream ijtimoiy tahlil vazifasiga ta’sirni ham hisobga oluvchi metodikasini taklif etishdan iborat. Ishning yangiligi shundaki, xato modeli qo‘lda beriladigan vaznlar o‘rniga haqiqiy foydalanuvchi xatolaridan baholanishi ko‘zda tutiladi va tuzatish sifati nafaqat so‘z darajasidagi aniqlik, balki sentiment tasniflash aniqligining o‘zgarishi orqali ham o‘lchanadi.

Mavzuga oid ishlar

Noisy channel modeli imlo xatolarini tuzatishga Kernighan, Church va Gale tomonidan tatbiq etilgan bo‘lib, unda xato so‘z w to‘g‘ri so‘z c ning «shovqinli» kuzatuvi sifatida qaraladi va eng ehtimolli c Bayes formulasi orqali topiladi [Kernighan M. va b., 1990: 205]. Brill va Moore channel modelini harflar ketma-ketligi darajasidagi almashtirishlar bilan boyitdilar [Brill E., Moore R., 2000: 286]. Kukich sharhida xatolar lug‘atda uchramaydigan va real-word turlarga ajratilib, ikkinchisi faqat kontekst orqali aniqlanishi ko‘rsatilgan [Kukich K., 1992: 377].

Morfologik boy tillarda channel modeli odatda n-gram til modeli bilan birlashtiriladi: amazig‘ tilida Damerau–Levenshteyn metrikasi va n-gramlar [Chaabi Y., Ataa Allah F., 2022: 6116], urdu tilida real-word xatolari uchun kontekstli tasniflagichlar [Aziz R. va b., 2023:

100948] qo‘llanilgan. So‘nggi yillarda source modeli sifatida BERT turidagi masked language modellar ishlatilmoqda, chunki ular so‘zning chap va o‘ng kontekstini bir vaqtda hisobga oladi [Devlin J. va b., 2019: 4171].

O‘zbek tili uchun imlo tekshiruviga oid bevosita ishlar kam. Yaqinda mBERT modelini token darajasida sozlash orqali xatolarni aniqlashda 94 % aniqlikka erishilgan [Axmedova X., Mamayusufova T., 2026: 102], biroq bu ish tuzatish bosqichini qamrab olmaydi. Morfologik ta’minot sohasida MorphUz analizatori [Abdurakhmonova N. va b., 2022: 61], finite state machine asosidagi lemmatizatsiya [Sharipov M., Sobirov O., 2022: 154] va morfologik tahlilning veb-tizimi [Mengliev D. va b., 2021: 9117] ishlab chiqilgan; sheva shakllarini adabiy shaklga keltirish algoritmi [Mengliev D. va b., 2023: 1440] ham foydalanuvchi matnlarini normallashtirishga yaqin masala hisoblanadi. UzBERT [Mansurov B., Mansurov A., 2021] hamda RoBERTa va T5 modellarining o‘zbek matnlarida tinish belgilarini tiklashda qo‘llanilishi [Adilova F. va b., 2024] mavjud. Biroq channel modeli haqiqiy xatolar korpusidan o‘rganilgan va downstream vazifasiga ta’siri baholanadigan yaxlit yondashuv o‘zbek tili uchun hozircha taklif etilmagan.

Masalaning noisy channel model doirasida qo‘yilishi

Kuzatilgan (xato bo‘lishi mumkin bo‘lgan) so‘z w va uning konteksti h berilgan bo‘lsin. Tuzatish masalasi nomzodlar to‘plami C(w) ichidan quyidagi ifodani maksimallashtiruvchi c so‘zni topishdan iborat:

$$\hat{c} = \arg\max P(c \mid h) \cdot P(w \mid c), \; c \in C(w) \quad (1)$$

Bu yerda P(c | h) — source model (til modeli), nomzodning kontekstga mosligini baholaydi; P(w | c) — channel model (xato modeli), to‘g‘ri so‘z c ning w ko‘rinishida buzilish ehtimolini beradi. Agar w o‘zi ham C(w) ga kiritilsa va P(w | w) uchun «xatosiz o‘tish» ehtimoli berilsa, (1) ifoda real-word xatolarini ham yagona mexanizm bilan qamrab oladi.

Channel model. Ingliz tili uchun ishlab chiqilgan yondashuvlardan farqli o‘laroq, channel modeli o‘zbek foydalanuvchi matnlaridan baholanadi. Buning uchun ijtimoiy tarmoq izohlari va sharhlaridan xato so‘z va uning to‘g‘ri shaklidan iborat (w, c) juftliklar korpusi tuziladi; tuzatishni ikki mutaxassis mustaqil bajarib, kelishuv darajasi Koen κ koeffitsiyenti bilan nazorat qilinadi. Har bir juftlik uchun w va c orasidagi minimal edit path topilib, undagi elementar amallar (almashtirish, qo‘shish, o‘chirish, o‘rin almashtirish) sanaladi. Kanal ehtimoli, Kernighan va boshqalar taklifiga muvofiq, amalning kuzatilgan chastotasi shu kontekstning umumiy chastotasiga nisbati sifatida, Laplace smoothing bilan hisoblanadi:

$$P(w \mid c) = \prod_k \left[(count(op_k) + 1) / (count(ctx_k) + |\Sigma|) \right] \quad (2)$$

bu yerda op_k — tahrir yo‘lidagi k-amal, ctx_k — shu amal qo‘llanishi mumkin bo‘lgan kontekst (masalan, c dagi tegishli harf), Σ — alifbo. Muhim jihat: o‘ va g‘ harflari hamda tutuq belgisi alohida alifbo birliklari sifatida qaraladi; aks holda to‘g‘ri → togri o‘tishi ikkita mustaqil o‘chirish sifatida hisoblanib, ehtimoli sun’iy pasayadi. 1-jadvalda o‘zbek foydalanuvchi matnlarida channel modeli alohida hisobga olishi lozim bo‘lgan asosiy amallar va ularning lisoniy sabablari keltirilgan.

1-jadval. O‘zbek foydalanuvchi matnlariga xos channel amallari

Channel amali (c → w)	**Misol (to‘g‘ri → xato)**	**Lisoniy sababi**
o‘ → o, g‘ → g (diakritika tushishi)	to‘g‘ri → togri	Mobil klaviaturada diakritikali harfni terish qiyinligi

’ → ∅ (tutuq tushishi)	ma’no → mano	Tutuq belgisining klaviaturada alohida belgi ekani
h ↔ x	hisob → xisob	Ikki tovushning talaffuzda yaqinligi, kirill an’anasi ta’siri
q ↔ k	quyosh → kuyosh	Sheva talaffuzi ta’siri
k → g, q → g‘ (o‘zak oxirida)	tilagim → tilakim	Qo‘shimcha oldidagi tovush almashinuvining bilinmasligi
Qo‘shni harflar o‘rin almashuvi	maktab → makatb	Tez terish (tipografik)
Harf ikkilanishi / tushishi	hammasi → hamasi	Tez terish (tipografik)
Qo‘shni klavishlar almashuvi	kitob → kutob	Klaviatura joylashuvi (i–u, o–p, n–m)

Jadvaldagi dastlabki uch amal tasodifiy klaviatura xatosi emas, balki yozuv tizimidan kelib chiqadigan tizimli hodisa: foydalanuvchilar telefon klaviaturasida diakritikali harflarni terishga vaqt sarflamaydi, shu bois ularning channel ehtimoli oddiy almashtirishlardan yuqori bo‘lishi kutiladi. Tekis (barcha amallar teng vaznli) channel modeli aynan shu asimmetriyani hisobga olmaydi, natijada togri so‘zi uchun to‘g‘ri va tog‘ri kabi nomzodlar teng baholanadi.

Source model. Kontekstli ehtimol ikki modelning chiziqli interpolatsiyasi sifatida hisoblanadi:

$$P(c \mid h) = \lambda \cdot P_ng(c \mid c_{-2}, c_{-1}) + (1 - \lambda) \cdot P_bert(c \mid h) \quad (3)$$

P_ng — tahrirdan o‘tgan matnlar korpusida [Abdurakhmonova N. va b., 2021: 1] o‘qitilgan, Kneser–Ney smoothing qo‘llanilgan trigram modeli; P_bert — masked language modelning w o‘rnida c so‘zi turishi ehtimoli, u UzBERT [Mansurov B., Mansurov A., 2021] asosida foydalanuvchi matnlarida qo‘shimcha o‘qitilgan. Trigram modeli tez ishlaydi va chastotali birikmalarni ishonchli baholaydi, masked language model esa uzoq kontekstni hisobga oladi va noyob birikmalarda ustunlik qiladi. λ koeffitsiyenti validatsiya to‘plamida tanlanadi.

Nomzodlarni morfologik cheklash. Agglutinativ tilda edit distance 2 gacha bo‘lgan nomzodlar soni bir necha yuzga yetishi mumkin, bu esa masked language modelni har bir nomzod uchun chaqirishni qimmatlashtiradi. Shu sababli generatsiya ikki bosqichda bajariladi: dastlab deletion index orqali sirtqi jihatdan yaqin satrlar olinadi, so‘ngra ular finite-state affiks analizatoridan [Abdurakhmonova N. va b., 2022: 61; Sharipov M., Sobirov O., 2022: 154] o‘tkazilib, faqat o‘zagi lug‘atda [Madvaliyev A., Begmatov E., 2013] mavjud va qo‘shimchalar zanjiri morfotaktik jihatdan to‘g‘ri [Abduraxmonova N., 2021] bo‘lgan shakllar qoldiriladi. Bu filtr nomzodlar sonini qisqartiradi va lug‘atda yo‘q, lekin grammatik jihatdan to‘g‘ri so‘zshakllarni (kitoblarimizdagilar) saqlab qoladi.

Dekodlash algoritmi. Jumla darajasida (1) ifoda har bir so‘z uchun mustaqil yechilmaydi, chunki qo‘shni so‘zlarning tuzatilishi bir-biriga bog‘liq. Biz kengligi B bo‘lgan beam search algoritmidan foydalanamiz:

```
KIRISH: S = s₁…sₙ jumla; channel modeli Ch; source modeli LM;
affiks avtomati A; B
CHIQISH: eng ehtimolli tuzatilgan jumla Ŝ
1 Beam ← {(∅, 0)}
```

```
2 i = 1 … n uchun:
3   C ← DeletionIndex(sᵢ, 2) ∪ {sᵢ}
4   C ← {c ∈ C : MorphValid(c, A)}
5   NewBeam ← ∅
6   har bir (prefix, score) ∈ Beam, har bir c ∈ C uchun:
7     s ← score + log Ch(sᵢ | c) + log LM(c | prefix,
s_{i+1..i+2})
8     NewBeam ← NewBeam ∪ {(prefix + c, s)}
9   Beam ← TopB(NewBeam)
10 Ŝ ← argmax Beam; qaytarish Ŝ va {(sᵢ, ĉᵢ) : sᵢ ≠ ĉᵢ}
```

7-qatorda masked language model uchun o‘ng kontekst sifatida hali tuzatilmagan keyingi ikki so‘z ishlatiladi; xatoli so‘z ham odatda umumiy ko‘rinishini saqlagani uchun bu soddalashtirish maqbul.

Baholash metodikasi

Taklif etilayotgan modelni baholash uchun quyidagi metodika ko‘zda tutiladi. Test to‘plami ijtimoiy tarmoq izohlari, xizmatlar haqidagi sharhlar va ochiq murojaatlardan iborat bo‘lib, channel modelini o‘qitish uchun ishlatilgan juftliklar bilan kesishmasligi shart. Har bir jumla ikki mutaxassis tomonidan mustaqil tuzatiladi; real-word xatolari alohida belgilanadi, chunki ular kontekstsiz modellar uchun ko‘r nuqta hisoblanadi. So‘z darajasidagi metrikalar: birinchi taklifning to‘g‘riligi (Acc@1), to‘g‘ri variantning dastlabki beshlikda bo‘lishi (Acc@5) hamda real-word xatolari uchun alohida aniqlik, to‘liqlik va F1.

Metodikaning asosiy jihati — downstream ta’sirni o‘lchash: [Sayfullaeva R. va b., 2025a: 2240] ishidagi kabi sentiment tasniflagichi bir xil test matnlarida xom holda va tuzatilgandan keyin baholanadi. Ikki ko‘rsatkich orasidagi farq imlo tuzatish modulining ijtimoiy media monitoringi quvuriga qo‘shgan real hissasini ko‘rsatadi. Modelning qaysi komponenti qanday hissa qo‘shishini ajratish uchun 2-jadvaldagi ablation sxemasi qo‘llaniladi: har bir keyingi konfiguratsiya oldingisiga bitta komponent qo‘shadi.

2-jadval. Ablation sxemasi: konfiguratsiyalar va ular tekshiradigan gipotezalar

Konfiguratsiya	**Tekshiriladigan gipoteza**
K0. Tuzatishsiz (xom matn)	Downstream ko‘rsatkichi uchun boshlang‘ich nuqta
K1. Tekis channel + unigram	Klassik lug‘aviy yondashuvning bazaviy darajasi
K2. O‘rganilgan channel + unigram	O‘zbek xatolarining tizimli tabiatini hisobga olish foyda beradimi
K3. K2 + trigram	Chap kontekst real-word xatolarini qay darajada qamrab oladi
K4. K2 + masked language model	Ikki tomonlama kontekst trigramdan ustunmi
K5. To‘liq (K2 + interpolatsiya + morfologik filtr)	Komponentlar bir-birini to‘ldiradimi; morfologik filtr aniqlikni saqlagan holda tezlikni oshiradimi

Ishlangan misol

Modelning ishlashini «Xizmat yaxshi, hamasi togri ishladi» jumlasi misolida ko‘rib chiqamiz. hamasi so‘zi uchun morfologik filtrdan o‘tgan nomzodlar: hammasi (hamma + -si),

hamisha, hamrohi (hamroh + -i). Channel modeli bo‘yicha hammasi eng yuqori ehtimolga ega (harf ikkilanishining tushishi — chastotali amal), source modeli ham «yaxshi, hammasi» birikmasini yuqori baholaydi; ikki omil bir yo‘nalishda ishlagani uchun tanlov ishonchli.

togri so‘zi uchun holat murakkabroq: morfologik filtrdan to‘g‘ri va tog‘ri variantlari o‘tadi. Tekis channel modeli ikkalasini teng baholaydi, chunki har ikkisi bitta amal bilan hosil bo‘ladi. O‘rganilgan channel modeli esa o‘ → o o‘tishiga (to‘g‘ri → togri) diakritika tushishining yuqori ehtimolini beradi va to‘g‘ri variantini oldinga chiqaradi; source modeli «togri ishladi» kontekstida to‘g‘ri so‘zini yana kuchaytiradi. Shunday qilib, channel modeli yozuvga xos tizimli xatolarni, source modeli esa kontekstni hisobga oladi.

Real-word xatosiga misol: «Xona butunlay bosh edi» jumlasida bosh so‘zi lug‘atda mavjud, shu bois kontekstsiz tizimlar uni tekshirmaydi. Taklif etilayotgan modelda w ning o‘zi nomzodlar to‘plamiga kiritilgani uchun bosh va bo‘sh variantlari (1) bo‘yicha qiyoslanadi; source modeli bu kontekstda bo‘sh variantiga ustunlik beradi va tuzatish taklif etiladi, «bosh og‘rig‘i» kabi kontekstda esa, aksincha, asl so‘z saqlanadi.

Xulosa

Maqolada o‘zbek tilidagi foydalanuvchi matnlari uchun imlo xatolarini tuzatish noisy channel modeli doirasida rasmiylashtirildi. Channel modelini haqiqiy foydalanuvchi xatolari korpusidan baholash usuli taklif etildi va o‘zbek yozuviga xos tizimli xatolar (diakritika, tutuq belgisi, h/x, q/k) uchun alifboni kengaytirish zarurligi asoslandi. Source modeli sifatida trigram va masked language modelning interpolatsiyasi, nomzodlarni cheklash uchun morfologik filtr, jumla darajasida decoding uchun beam search taklif qilindi. Baholash metodikasi so‘z darajasidagi metrikalar bilan birga tuzatishning sentiment tasniflagichiga ta’sirini o‘lchashni nazarda tutadi. Keyingi bosqichda xato juftliklari korpusini yig‘ish, taklif etilgan ablation sxemasi bo‘yicha tajriba o‘tkazish va sheva normalizatsiyasi bilan integratsiya rejalashtirilgan.

FOYDALANILGAN ADABIYOTLAR

1. Abdurakhmonova N., Alisher I., Sayfulleyeva R. MorphUz: Morphological analyzer for the Uzbek language // UBMK 2022. — IEEE, 2022. — P. 61–66.
2. Abdurakhmonova N., Tuliyev U., Gatiatullin A. Linguistic functionality of Uzbek Electron Corpus: uzbekcorpus.uz // ICISCT 2021. — IEEE, 2021. — P. 1–4.
3. Abduraxmonova N. Kompyuter lingvistikasi (darslik). — Toshkent: Nodirabegim, 2021. — 400 b.
4. Adilova F., Davronov R., Safarov R., Abdurakhmonova N., Kushmurotov S. Using the RoBERTa and T5 models for analyzing texts in Uzbek: NER tasks and punctuation prediction // MSBC 2024. CCIS, vol. 2211. — Cham: Springer, 2024. — https://doi.org/10.1007/978-3-031-72260-8_7
5. Axmedova X. I., Mamayusufova T. O. Detection of spelling errors in the Uzbek language using the mBERT model // Science and Innovation. — 2026. — Vol. 5, No. 5. — P. 102–110. — https://doi.org/10.5281/zenodo.20523532
6. Aziz R., Anwar M. W., Jamal M. H., Bajwa U. I. et al. Real word spelling error detection and correction for Urdu language // IEEE Access. — 2023. — Vol. 11. — P. 100948–100962.
7. Brill E., Moore R. C. An improved error model for noisy channel spelling correction // Proceedings of the 38th Annual Meeting of the ACL. — 2000. — P. 286–293.

8. Chaabi Y., Ataa Allah F. Amazigh spell checker using Damerau-Levenshtein algorithm and N-gram // Journal of King Saud University — Computer and Information Sciences. — 2022. — Vol. 34, No. 8. — P. 6116–6124.
9. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding // Proceedings of NAACL-HLT 2019. — 2019. — P. 4171–4186.
10. Kernighan M. D., Church K. W., Gale W. A. A spelling correction program based on a noisy channel model // Proceedings of COLING 1990. — 1990. — Vol. 2. — P. 205–210.
11. Kukich K. Techniques for automatically correcting words in text // ACM Computing Surveys. — 1992. — Vol. 24, No. 4. — P. 377–439.
12. Madvaliyev A., Begmatov E. O‘zbek tilining imlo lug‘ati. — Toshkent: Akademnashr, 2013.
13. Mansurov B., Mansurov A. UzBERT: pretraining a BERT model for Uzbek // arXiv preprint arXiv:2108.09814. — 2021.
14. Mengliev D. B., Abdurakhmonova N., Hayitbayeva D., Barakhnin V. B. Automating the transition from dialectal to literary forms in Uzbek language texts // APEIE 2023. — IEEE, 2023. — P. 1440–1443.
15. Mengliev D., Abdurakhmonova N., Fattaxova D. et al. Development and comparative analysis of classification algorithms of Uzbek taxpayer complaints and questions // EDM 2025. — IEEE, 2025. — P. 2090–2095.
16. Mengliev D., Barakhnin V., Abdurakhmonova N. Development of intellectual web system for morph analyzing of Uzbek words // Applied Sciences. — 2021. — Vol. 11, No. 19. — P. 9117.
17. Sayfullaeva R., Abdurakhmonova N., Ganiyeva S. Development of sentiment analysis algorithms of Uzbek patient reviews // EDM 2025. — IEEE, 2025a. — P. 2240–2244.
18. Sayfullaeva R., Shirinova R., Abdurakhmonova N. Unveiling themes in social media data: a two-stage hashtag-driven clustering and classification method in Uzbek language // EDM 2025. — IEEE, 2025b. — P. 1450–1454.
19. Sharipov M., Sobirov O. Development of a rule-based lemmatization algorithm through Finite State Machine for Uzbek language // CEUR Workshop Proceedings. — 2022. — Vol. 3315. — P. 154–159.

O‘ZBEK TILI PARSING TIZIMI UCHUN SODDA GAPLARNING SINTAKTIK MODELLARI: TADQIQOTLAR SHARHI

Sanobarkhon Khabibullaeva
Master's student National University of Uzbekistan
https://doi.org/10.5281/zenodo.23136671

Annotatsiya. *Ushbu maqolada o‘zbek tili parsing tizimi uchun sodda gaplarning sintaktik modellariga oid ilmiy tadqiqotlar sharhlanadi. Tadqiqotlarda qo‘llangan sintaktik modellar, parsing yondashuvlari va gap bo‘laklari o‘rtasidagi munosabatlar tahlil qilinadi. Shuningdek, mavjud yondashuvlarning parsing tizimlari uchun ahamiyati va muammoli jihatlari umumlashtiriladi.*

Kalit so‘zlar: *parsing, sodda gap, sintaktik model, dependency parsing, sintaktik teglash.*

Abstract. *This article reviews three studies on syntactic models of simple sentences for Uzbek language parsing systems. The syntactic models, parsing approaches, and relationships between sentence constituents used in these studies are comparatively analyzed. The relevance and limitations of existing approaches for Uzbek parsing systems are also summarized.*

Keywords: *parsing, simple sentence, syntactic model, dependency parsing, syntactic tagging.*

Аннотация. *В статье рассматриваются три научных исследования, посвящённых синтаксическим моделям простых предложений для системы парсинга узбекского языка. Проводится сравнительный анализ синтаксических моделей, подходов к парсингу и отношений между членами предложения, представленных в данных исследованиях. Также обобщаются значение и проблемные аспекты существующих подходов для систем парсинга узбекского языка.*

Ключевые слова: *парсинг, простое предложение, синтаксическая модель, dependency parsing, синтаксическая разметка.*

Kirish

Bugungi kunda tabiiy tilni qayta ishlash (NLP) texnologiyalarining rivojlanishi o‘zbek tilini avtomatik tahlil qilish masalasini dolzarb yo‘nalishlardan biriga aylantirmoqda. Sintaktik parsing muayyan sintaktik birlik va strukturalarni tahlil qiladi. Tahlilning bu bosqichi tabiiy tilni qayta ishlashda juda muhim bosqich sanaladi. U kompyuter lingvistikasi va NLPning deyarli barcha yo‘nalishlari (mashina tarjimasi, savol-javob, axborot qidiruv tizimi, sentiment analiz)da foydalaniladi. Chunki gapdagi har bir komponent muayyan vazifani bajarib, u yaxlit holatda mazmunni aks ettiradi. [Abdurahmonova N., 2019:138].

Adabiyotlar sharhi

O‘zbek tili parsing tizimi uchun sodda gaplarning sintaktik modellarini ishlab chiqish kompyuter lingvistikasi va tabiiy tilni qayta ishlash yo‘nalishlari bilan bog‘liq. Mavjud tadqiqotlarda sodda gaplarning sintaktik tuzilishini formal ifodalash, gap bo‘laklari o‘rtasidagi munosabatlarni aniqlash va ularni avtomatik parsingda qo‘llash masalalari o‘rganilgan.

Shunday izlanishlardan biri N. Abdurahmonovaning **“**Dependency parsing based on uzbek corpus**”** tadqiqotida o‘zbek tili korpusi asosida dependency parsing masalasi tadqiq etilib, gapdagi so‘zlar o‘rtasidagi sintaktik munosabatlarni aniqlashda morfologik belgilar muhimligi

ta'kidlangan. Tadqiqotda Universal Dependencies va CoNLLU formatlaridan foydalanish hamda o'zbek tili uchun sintaktik treebank yaratish imkoniyatlari ko'rsatilgan.

Abdullayeva, Nizomova va Qodirovning "O'zbek tilida gap bo'laklari tartibini avtomatik aniqlash (sodda gaplar misolida)" maqolasida esa sodda gaplarning sintaktik tuzilishi, gap bo'laklari va ularning tartibini avtomatik aniqlash masalasi o'rganilgan. O'zbek tilida gap bo'laklarining joylashish tartibi, asosan, erkin bo'lsa-da, ular gapda ma'lum grammatik qoidalar asosida o'rinlashadi. [Qodirov, 2014: 101-102]. Shu sababli sintaktik tahlilda faqat so'zlarning joylashuviga emas, balki ularning grammatik va morfologik xususiyatlariga ham e'tibor berish zarurligi izohlab berilgan.

"O'zbek tilida sintaktik parsing masalasi: muammo va takliflar" maqolasida o'zbek tili uchun sintaktik teglar, dependency parsing va treebank yaratish masalalari yoritilgan. O'zbek tili morfoanalizatorida 35 000 ta turli uzunlikdagi gaplar POS teglangan bo'lib, keyingi bosqich o'zbek tilida sintaktik teglash tizimini ishlab chiqish teglarni taklif etish hisoblanadi. [Oqila Abdullayeva., 2025]. Tadqiqot natijalari o'zbek tilining agglyutinativ tuzilishi va erkin so'z tartibi parsing jarayonida morfologik hamda sintaktik ma'lumotlarni birgalikda hisobga olishni talab qilishini ko'rsatadi.

Umuman olganda**,** mazkur tadqiqotlar o'zbek tili sintaktik parsingining muhim jihatlarini ochib berilgan. Biroq sodda gaplarda gap bo'laklari va ularning sintaktik munosabatlarini avtomatik aniqlash masalasi yanada takomillashtirishni talab qiladi. Shu sababli ushbu tadqiqotda sodda gaplarning sintaktik tuzilishini aniqlashga e'tibor qaratilgan.

Tadqiqot metodologiyasi

Mazkur tadqiqotlar o'zbek tilidagi sodda gaplarning sintaktik tuzilishini avtomatik aniqlashga qaratilgan.

Tadqiqotda o'zbek badiiy, akademik va rasmiy matnlardan tashkil topgan korpus asosida UD va CoNLL-U yondashuvlari qo'llanilgan. Morfologik tahlilda o'zbek tilining lotin va kirill alifbolari, grammatik qoidalari va leksikonidan foydalanilgan. Asosiy vositalar sifatida HFST, LexC, TWOLC, Python va UDPipe ishlatilgan. Jarayon tokenizatsiya, morfologik tahlil, noaniq shakllarni qo'lda tuzatish, CoNLL-U formatiga o'tkazish va dependency parsing bosqichlaridan iborat. Natijada so'zlarning `nsubj`, `obj`, `obl`, `amod`, `acl`, `root` kabi sintaktik munosabatlari aniqlangan. Mazkur metodologiya o'zbek tilidagi sodda gaplarning sintaktik tuzilishini avtomatik aniqlash uchun asos bo'lib xizmat qiladi.

O'zbek tilida gap bo'laklari tartibini avtomatik aniqlash (sodda gaplar misolida) tadqiqotida o'zbek tilidagi sodda gaplarning odatiy va noodatiy so'z tartibini aniqlash uchun grammatik qoidalar va sintaktik modellashtirish usulidan foydalanilgan. Gaplar tarkibidagi ega, kesim, aniqlovchi, to'ldiruvchi va hol kabi bo'laklar teglar orqali belgilangan. Shuningdek, VB, ADJ, NUM, ADV, PRON, AUX kabi POS teglaridan foydalanilib, sodda gaplarning N+VB, N+ni+VB, N+ga+VB, PRON+N kabi modellari shakllantirilgan. Ushbu modellar asosida noto'g'ri tartibdagi gaplarni aniqlash va me'yoriy tartibga keltirish algoritmi ishlab chiqilgan. Natijalarni amaliy qo'llash uchun web-sayt yaratish va undan o'zbek tili korpuslari, ta'lim hamda mashina tarjimasida foydalanish imkoniyatlari ko'rib chiqilgan.

O'zbek tilida sintaktik parsing masalasi: muammo va taklifilari maqolasida o'zbek tilining sintaktik xususiyatlari va so'z tartibining erkinligi hisobga olinib, POS teglash, morfologik tahlil va dependency parsing bosqichlari asos qilib olingan. O'zbek tili ta'limiy korpusidagi 14 ta so'z turkumi (POS) tegi hamda turkiy tillar uchun yaratilgan treebank va UD teglar tizimi tahlil qilindi. Shuningdek, **Meaning-Text Theory (MTT)**ning ko'p qatlamli

modeli asosida o‘zbek tilida sintaktik munosabatlarni aniqlash imkoniyatlari ko‘rib chiqilgan. Izlanish davomida o‘zbek tilidagi gap bo‘laklarining grammatik shakli, kelishik qo‘shimchalari va so‘z tartibi sintaktik vazifani aniqlashdagi ahamiyati baholangan. Natijada o‘zbek tilidagi sintaktik parsing uchun 18 ta sintaktik teg taklif etish va ularni UD formatiga moslashtirish metodologik asos sifatida belgilandi.

Tahlil va muhokama

Maqolalar tahlili shuni ko‘rsatadiki, o‘zbek tilida sintaktik parsingni amalga oshirishda morfologik tahlil va POS teglash muhim asos hisoblanadi. Tadqiqotida 35 000 ta gap POS teglar bilan belgilangan bo‘lib, morfologik xususiyatlarning sintaktik munosabatlarni aniqlashdagi ahamiyati ko‘rsatilgan. Universal Dependencies (UD), CoNLL-U, HFST, LexC va TWOLC kabi vositalardan foydalanish imkoniyatlari ko‘rib chiqilgan. O‘zbek tilida so‘z tartibining nisbatan erkinligi sababli faqat SOV modeliga tayanish yetarli emasligi, SOV va OSV kabi tartiblarni ham hisobga olish zarurligi aniqlangan.

Shuningdek, o‘zbek tili uchun taklif qilingan va ularni UD standartiga moslashtirish maqsadga muvofiq deb ko‘rsatilgan. Parser yaratishda kamida 20 000 ta sintaktik belgilangan gap zarurligi qayd etilgan. Umuman, uchala tadqiqot o‘zbek tili uchun samarali parser yaratishda morfologik belgilar, sintaktik teglar va katta hajmdagi belgilangan korpusni birgalikda qo‘llash zarurligini ko‘rsatadi.

XULOSA

Xulosa qilib aytganda, o‘zbek tilidagi sodda gaplarni avtomatik sintaktik tahlil qilishda morfologik va sintaktik ma’lumotlarni birgalikda qo‘llash zarurligini ko‘rsatildi. O‘zbek tilining agglyutinativligi va so‘z tartibining nisbatan erkinligi sababli faqat so‘zlarning joylashuviga asoslangan tahlil yetarli emas. Maqolalar tahlili asosida POS teglash, gap bo‘laklarini aniqlash va dependency munosabatlarini belgilash parser yaratishning asosiy bosqichlari sifatida belgilangan. Kelgusida ushbu yondashuvlar asosida sodda gaplarning sintaktik tuzilishini avtomatik aniqlovchi samarali parser ishlab chiqish mumkin.

FOYDALANILGAN ADABIYOTLAR

1. Abduraxmonova N. *Kompyuter lingvistikasi* (darsliToshkent, 2019. — B. 138
2. Qodirov M. va boshqalar. *Ona tili*: o‘rta maktablarning 8-sinfi uchun darslik. Toshkent. — B. 201.
3. Abduraxmonova N. *Dependency parsing based on uzbek corpus*. Proceedings of the Language Technologies for All (LT4All), Paris, UNESCO Headquarters, 5–6 December, 2019. — P. 164–168.
4. Abdullayeva O., Nizomova F., Qodirov K. *O‘zbek tilida gap bo‘laklari tartibini avtomatik aniqlash (sodda gaplar misolida). Kompyuter lingvistikasi: muammolar, yechimlar, istiqbollar*, № 1, Toshkent, 2024. — B. 262–268.
5. Abdullayeva O. *O‘zbek tilida sintaktik parsing masalasi: muammo va takliflar*. *O‘zbekiston Milliy universiteti xabarlari. Filologiya*, 2025, № 11/31. — B. 216–219.

OLIY TA'LIM MUASSASALARIDA TADBIRKORLIK EKOTIZIMINI SHAKLLANTIRISH VA UNING MINTAQA IQTISODIY O‘SISHIGA TA'SIRINI TIZIMLI MODELLASHTIRISH

Khajimuratov Abdukakhkhor Abdumutalovich
Iqtisodiyot fanlari doktori,
Farg‘ona jamoat salomatligi tibbiyot instituti
https://doi.org/10.5281/zenodo.23136695

Annotatsiya. *Ushbu maqolada oliy ta'lim muassasalarida (OTM) "University 3.0" va "University 4.0" konsepsiyalari doirasida, tadbirkorlik ekotizimini shakllantirish va uning mintaqaviy iqtisodiy o‘sishga ta'sirini baholashda tizimli modellashtirish hamda makon-zamon tahlili yondashuvlari, tadqiq etilgan. Tadqiqot natijasida intellektual kapitalning hududiy sanoat klasterlari barqarorligiga ko‘rsatadigan ijobiy multiplikativ samarasi asoslangan va O‘zbekiston hududlari sharoitida qo‘llash bo‘yicha amaliy tavsiyalar ishlab chiqilgan.*

Kalit so‘zlar: *tadbirkorlik ekotizimi, tizimli modellashtirish, makon-zamon tahlili, mintaqaviy iqtisodiyot, startap akseleratorlari, University 3.0, innovatsion infratuzilma.*

Abstract. *This article investigates system modeling and spatio-temporal analysis approaches in shaping the entrepreneurial ecosystem within higher education institutions (HEIs) under the framework of "University 3.0" and "University 4.0" concepts, as well as assessing its impact on regional economic growth. As a result of the research, the positive multiplicative effect of intellectual capital on the sustainability of regional industrial clusters is substantiated, and practical recommendations are developed for implementation in the regions of Uzbekistan.*

Keywords: *entrepreneurial ecosystem, system modeling, spatio-temporal analysis, regional economy, startup accelerators, University 3.0, innovation infrastructure.*

Аннотация. *В данной статье исследуются подходы системного моделирования и пространственно-временного анализа в формировании предпринимательской экосистемы в высших учебных заведениях (ВУЗах) в рамках концепций "University 3.0" и "University 4.0", а также в оценке ее влияния на региональный экономический рост. В результате исследования обоснован положительный мультипликативный эффект интеллектуального капитала на устойчивость региональных промышленных кластеров, а также разработаны практические рекомендации по их применению в условиях регионов Узбекистана.*

Ключевые слова: *предпринимательская экосистема, системное моделирование, пространственно-временной анализ, региональная экономика, стартап-акселераторы, University 3.0, инновационная инфраструктура.*

Kirish. Keyingi yillarda, O‘zbekistonda raqamli iqtisodiyot va global texnologik o‘zgarishlar sharoitida oliy ta'lim muassasalarining iqtisodiy tizimdagi roli tubdan o‘zgarib bormoqda. Bugungi kunda OTMlar faqatgina malakali kadrlar yetkazib beruvchi muassasa emas, balki innovatsiyalar, intellektual mulk va startap loyihalarni generatsiya qiluvchi asosiy drayverga — tadbirkorlik ekotizimining markaziga aylanmoqda. O‘zbekistonda oliy ta'lim tizimini transformatsiya qilish, oliy ta'lim muassasalarini tijoratlashtirish darajasini oshirish

hamda "University 3.0" konseptini joriy etish davlat siyosatining ustuvor yo‘nalishlaridan biri etib belgilangan. Biroq, OTMlarda shakllanayotgan tadbirkorlik muhitining mintaqaviy iqtisodiy o‘sishga, bandlikka va hududiy sanoat klasterlari shakllanishiga ta'sirini baholash murakkab chiziqli bo‘lmagan aloqalar (*non-linear feedback loops*) va vaqtinchalik kechikishlar (*time delays*) sababli an'anaviy ekonometrik usullar bilan to‘liq ochib berolmaydi. Shu bois, ushbu jarayonni tizimli modellashtirish (*System Dynamics*) va makon-zamon tahlili (*Spatial-Temporal Analysis*) vositalari yordamida o‘rganish alohida dolzarblik kasb etadi.

Adabiyotlar tahlili. Akademik tadbirkorlik va universitetlar bazasidagi ekotizimlar kontseptsiyasi dastlab Etzkowitz va Leydesdorff tomonidan taklif etilgan "Uchlama spiralsimon model" (Triple Helix Model) doirasida keng o‘rganilgan [1]. Ushbu model davlat, biznes va oliy ta'lim tizimining o‘zaro sinergik aloqalarini tushuntirib beradi. Keyinchalik, Isenberg o‘zining tadbirkorlik ekotizimi modelida universitetlarni inson kapitali va texnologik resurslar yetkazib beruvchi asosiy subyekt sifatida belgilagan [2]. Tizimli dinamika (*System Dynamics*) asoschisi Jay Forrester yondashuvlaridan foydalangan holda, hududiy iqtisodiy tizimlarda innovatsiyalar diffuziyasini modellashtirish bo‘yicha Sterman bir qancha chiziqli bo‘lmagan modellarni taklif etgan [3]. Mintaqaviy makon-zamon tahlili borasida esa Krugman va yangi iqtisodiy geografiya namoyandalari oliy ta'lim muassasalari atrofida innovatsion klasterlarning aglomeratsiya samarasi hisobiga shakllanishini isbotlaganlar [4].

O‘zbekistonlik olimlardan: S. S. G‘ulomov ta’limoti O‘zbekiston mustaqilligining dastlabki pallasida tadbirkorlik ta'limining shakllanishi va mintaqaviy iqtisodiy o‘sish modellarini baholashda asosiy poydevor hisoblangan klassik asarlardan biridir[5]; Xodiyev, B. Yu., Qosimova, M. S., & Samadov, A. N. lar tomonidan dastlabki o‘quv qo‘llanmasida kichik biznes va innovatsion tadbirkorlik subyektlarining infratuzilmaviy shakllanishi hamda ularni boshqarish mexanizmlari tizimli tarzda yoritilgan[6]; Shu mualliflar keying arslikda oliy ta'lim tizimi bitiruvchilarining tadbirkorlik kompetensiyalarini shakllantirish, biznes g‘oyalarni tanlash va ularni startap darajasida texnik-iqtisodiy jihatdan modellashtirish usullari bayon etilgan[7]; Bexmurodov, Q. A., & Muftaydinov, Q. Mualliflar tomonidan raqamli transformatsiya davrida korporativ va akademik tadbirkorlik integratsiyasi hamda innovatsion infratuzilmalarni hududiy klasterlar tarkibida rivojlantirish masalalari tahlil qilingan[8]; Shuningdek, Bexmurodov, Q. A. Oliy ta'lim muassasalarida 'University 3.0' modelini joriy etish talabalar loyihalarini inkubatsiya markazlari orqali real ishlab chiqarish va venchur fondlari bilan bog‘lashni taqozo etadi degan[9].

Umumiy hulosa qilish mumkinki, respublikamizda oliy ta'lim muassasalarining innovatsion faoliyatini va raqamli tadbirkorlikni rivojlantirish modellarini tadqiq etishgan, biroq tizimli dinamika va hududiy makon-zamon tahlilini yaxlit tizim sifatida modellashtirish masalalari hali mukammal darajada birlashtirilmagan ko‘rinadi.

Metodologiya. Ushbu tadqiqotda OTM tadbirkorlik ekotizimining mintaqa iqtisodiyotiga ta'sirini baholash uchun Tizimli dinamika (System Dynamics) va Saboqlar-oqibatlar diagrammasi (Causal Loop Diagram - CLD) metodologiyasidan foydalanildi. Modellashtirish uchta asosiy blokni qamrab oladi: *1. Universitet intellektual bloki:* Talabalar soni, startap g‘oyalar, inkubatsiya markazlarining quvvati, ilmiy ishlanmalar hajmi. *2. Mintaqaviy tadbirkorlik bloki:* OTM bitiruvchilari yaratgan startaplar, jalb qilingan venchur investitsiyalar, yangi ish o‘rinlari. *3.Mintaqaviy makroiqtisodiy blok:* Yalpi hududiy mahsulot (YHM) o‘sishi, soliq tushumlari, innovatsion mahsulotlar ulushi. Modelda vaqtinchalik kechikish (*Time Delay*) parametri 3 yil qilib belgilandi, ya'ni talabaning universitet tadbirkorlik ekotizimiga kirishidan

uning tijoriy loyihasi mintaqa YHMga real hissa qo‘shishigacha bo‘lgan o‘rtacha davr inobatga olindi.

Asosiy qism. "Iqtisodiyotda tizimli modellashtirish va makon-zamon tahlili" yo‘nalishi doirasida O‘zbekiston oliy ta'lim muassasalarida (OTM) tadbirkorlikni o‘qitishning "Raqamli ekotizim va hududiy-dinamik modellashtirish" (University 3.0 va 4.0 konsepti asosida) modeli shakllanmoqda. Bu model an'anaviy nazariy ta'limdan farq qilib, talabalarga biznes jarayonlarini nafaqat iqtisodiy, balki fazoviy (makon) va vaqtinchalik (zamon/dinamika) o‘lchamlarda prognozlashni o‘rgatishga qaratilgan. Ushbu shakllanayotgan yangi ta'lim modelining asosiy ustunlari va o‘ziga xos xususiyatlari quyidagilardan iborat:

O‘qitish modelining asosiy tarkibiy qismlari:

*Makon-zamon tahlili (Spatial-Temporal Analysis): Talabalar tadbirkorlik subyektlarini joylashtirish (geografik va mintaqaviy optimal nuqtalar), infratuzilma logistikasi va bozor konyunkturasining vaqt bo‘yicha o‘zgarishini (dinamikasini) GIS (Geografik axborot tizimlari) va ekonometrik modellar yordamida tahlil qilishni o‘rganishmoqda.

*Tizimli modellashtirish va Big Data: Tadbirkorlik risklarini va bozor muvozanatini matematik, imitatsion hamda prognozlashtirish modellar orqali raqamli simulyatsiya qilish ta'lim mazmuniga kiritilmoqda.

*Amaliy "Startap-Ekotizim" yondashuvi: O‘zbekiston OTMlarida joriy etilayotgan "Tadbirkorlik asoslari" fanlari doirasida quruq ma'ruzalardan voz kechilib, hududiy muammolarga yechim bo‘luvchi amaliy startap loyihalar (inkubatsiya va akseleratsiya markazlari orqali) ishlab chiqilmoqda [10].

Kichik va o‘rta biznes (KOB) klasterizatsiyasi: Davlar rahbarlarining tadbirkorlikni rivojlantirish va KOB rolini oshirish bo‘yicha strategik qarorlaridan kelib chiqib, ta'lim mintaqaviy sanoat klasterlari bilan integratsiya qilinmoqda. "University 3.0" — bu oliy ta’lim muassasalarining klassik vazifalari bo‘lgan ta’lim berish va ilmiy tadqiqotlar olib borish missiyasiga uchinchi asosiy missiya — tijoratlashtirish va tadbirkorlikni qo‘shish kontseptsiyasidir.

Universitet evolyutsiyasini tushunish uchun uning uchta bosqichini taqqoslash kifoya: *University 1.0:* Faqat Ta'lim (Bilim ulashish va kadrlar tayyorlash); *University 2.0*: Ta'lim + Ilm-fan (Bilim yaratish va fundamental tadqiqotlar o‘tkazish); *University 3.0*: Ta'lim + Ilm-fan + Tijoratlashtirish (Bilimni iqtisodiy qiymatga/biznesga aylantirish).

"University 3.0" modelining 4 ta asosiy ustuni:

***Biznes-inkubatorlar va akseleratorlar*: Talabalar va professor-o‘qituvchilarning startap g‘oyalari universitet ichida shakllantiriladi va bozorga tayyorlanadi.

***Intellektual mulk va patentlar:* Universitetda yaratilgan yangi ixtirolar va ilmiy ishlanmalar shunchaki tokchada qolib ketmaydi, balki patentlanib, biznesga sotiladi (*technology transfer*).

***Venchur moliyalashtirish*:** Universitetlar qoshida istiqbolli startap loyihalarni to‘g‘ridan-to‘g‘ri moliyalashtiruvchi maxsus investitsiya fondlari va biznes-farishtalar tarmog‘i ishlaydi.

***Talaba-tadbirkor maqomi*:** Talaba o‘qish davomida nafaqat diplom olishni, balki o‘zining real korxonasini yoki startapini yaratishni maqsad qiladi. Ko‘pincha bitiruv malakaviy ishi o‘rniga real startap loyiha himoya qilinadi.

Bularga muvofiq aniqriq qilib aytganda, "University 3.0" — bu ilmiy kashfiyotlarni to‘g‘ridan-to‘g‘ri bozor mahsulotiga aylantiradigan, o‘zini-o‘zi moliyalashtira oladigan va mintaqa iqtisodiyotining drayveriga aylanadigan ***tadbirkorlik universiteti*** modelidir.

Ma’lumki, O‘zbekistonda "Oliy ta'lim tizimini 2030-yilgacha rivojlantirish konsepsiyasi" doirasida barcha davlat OTMlarini bosqichma-bosqich "University 3.0" modeliga o‘tkazish strategik maqsad qilib belgilangan.

Hozirgi vaqtda O‘zbekiston oliy ta’lim tizimida an’anaviy ta’lim modeli - University 2.0 hamda davlat dasturlari darajasida jadal tatbiq etilayotgan tadbirkorlik universiteti - University 3.0 bosqichi o‘zaro kesishgan holda ustunlik qilmoqda. Bugungi kunda respublikamizda ushbu model elementlarini (inkubatorlar, startaplar va tijoratlashtirish) amaliyotga joriy etishni boshlagan va aniq natijalarga erishayotgan asosiy yetakchi OTMlar mavjud. Bular: Toshkent davlat iqtisodiyot universiteti (TDIU), Toshkent davlat tehnika universiteti (TDTU) - "Yilning tadbirkor universiteti" mukofot sovrundori, Samarqand davlat universiteti (SamDU), TEAM University (Xususiy sektor drayveri) - O‘zbekistondagi *ilk xususiy tadbirkorlik universiteti* hisoblanadi.

Yuqoridagilarga asoslanib Oliy ta'lim muassasasida tadbirkorlik ekotizimini tizimli modellashtirish doirasida quyidagi asosiy bog‘liqlik zanjirlari aniqlandi:

OTM Infratuzilmasi → Startaplar soni → Mintaqaviy bandlik → YHM (Yalpi hudud mahsuloti) → OTMga Investitsiyalar.

Ushbu teskari aloqa zanjiri (*Reinforcing Loop*) iqtisodiy o‘sishning multiplikativ samarasini namoyon etadi. O‘zbekiston iqtisodiyoti sharoitida (masalan, Toshkent, Samarqand va Farg‘ona vodiysi mintaqalari misolida) makon-zamon tahlili ko‘rsatganidek, oliy ta'lim muassasalari geografik jihatdan sanoat klasterlariga qanchalik yaqin joylashgan bo‘lsa, bilimlar diffuziyasi (*knowledge spillover*) shunchalik tez amalga oshadi.

Tizimli modellashtirish simulyatsiyasi natijalari shuni ko‘rsatadiki, agar OTMda inkubatsiya markazlarini moliyalashtirish 20% ga oshirilsa, keyingi 5 yillik tsiklda mintaqadagi innovatsion kichik biznes subyektlarining umumiy ulushi 8.4% ga o‘sadi, hududiy ishsizlik darajasi esa 1.2% ga kamayadi.

Muhokama. Olingan natijalar shuni ko‘rsatadiki, O‘zbekistonda oliy ta'lim muassasalarida tadbirkorlikni o‘qitishning hozirgi modeli ko‘proq nazariy darslar berish bilan cheklanib qolmoqda. Tizimli tahlil shuni tasdiqlaydiki, haqiqiy tadbirkorlik ekotizimi shakllanishi uchun universitetlarda *"Talaba-tadbirkor" (Student Entrepreneur Status)* maqomi qonuniy mustahkamlanishi hamda *"Startap — bitiruv malakaviy ishi"* mexanizmi keng joriy etilishi lozim.

Shuningdek, makon tahlili natijalari (*Spatial Analysis*) O‘zbekiston viloyatlari o‘rtasida raqamli infratuzilma va venchur kapital yetishmovchiligi sababli jiddiy nomutanosibliklar borligini ko‘rsatadi. Toshkent shahridan tashqaridagi hududiy OTMlarda startap akseleratorlarining real ishlab chiqarish sektori bilan aloqalari zaif, bu esa intellektual resurslarning hududdan "poytaxtga oqib ketishi" (*brain drain*) effektini keltirib chiqarmoqda.

Xulosa va tavsiyalar. Oliy ta'lim muassasalarida tadbirkorlik ekotizimini shakllantirish shunchaki akademizm emas, balki mintaqalar iqtisodiy raqobatbardoshligini oshirishning eng asosiy drayveridir. “OTM Infratuzilmasi → Startaplar soni → Mintaqaviy bandlik → YHM (Yalpi hudud mahsuloti) → OTMga Investitsiyalar” Tizimli modellashtirish shuni isbotladiki, OTM, biznes va davlat o‘rtasidagi uchlama uzviylik hamda raqamli texnologiyalar integratsiyasi hududiy iqtisodiyot dinamikasini tubdan yaxshilaydi.

Mintaqaviy barqaror o‘sishni ta’minlash uchun quyidagilarni tavsiya etamiz:

1)Hududiy OTMlar qoshida sanoat va agrar yo‘nalishlarga moslashtirilgan "Makon-zamon tahlil laboratoriyalari" va biznes-inkubatorlarni kengaytirish;

2) Talabalar startaplarini hududiy hokimliklarning innovatsion jamg‘armalari orqali to‘g‘ridan-to‘g‘ri grandlar va imtiyozli subsidiyalar bilan ta'minlash tizimini yo‘lga qo‘yish.

FOYDALANILGAN ADABIYOTLAR

1. Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations. *Research Policy*, 29(2), 109-123.
2. Isenberg, D. J. (2010). How to start an entrepreneurial revolution. *Harvard Business Review*, 88(6), 40-50.
3. Sterman, J. D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill.
4. Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3), 483-499.
5. G‘ulomov, S. S**.** (2002). *Tadbirkorlik va kichik biznes*. Toshkent: "Sharq" nashriyoti.
6. Xodiyev, B. Yu., Qosimova, M. S., & Samadov, A. N. (2010). *Kichik biznes va xususiy tadbirkorlik*. O‘quv qo‘llanma. Toshkent: Toshkent davlat iqtisodiyot universiteti (TDIU).
7. Xodiyev, B. Yu., va boshqalar. (2021). *Kichik biznes va tadbirkorlik*. Darslik. Toshkent: Zenodo platformasi raqamli nashri.
8. Bexmurodov, Q. A., & Muftaydinov, Q. (2019). *Raqamli iqtisodiyot sharoitida tadbirkorlik faoliyatini innovatsion rivojlantirish modellari*. Monografiya. Toshkent: "Iqtisod-Moliya".
9. Bexmurodov, Q. A. (2022). OTMlarda biznes va tadbirkorlik ta'limini raqamlashtirish strategiyalari. *Biznes va tadbirkorlik oliy maktabi ilmiy axborotnomasi*, 2(4), 12-18.
10. O‘zbekiston OTMlarida joriy etilayotgan "Tadbirkorlik asoslari" fanlari
11. TADBIRKORLIKKA KIRISH. KA Abdumutalovich - Scientific Journal of Actuarial Finance and Accounting, 2024/Vol. 4 Issue 10 | pp. 396-406 | ISSN: 2181-1865/https://finance.tsue.uz/index.php/afa
12. A.A. Hadzhimuratov. Evolution of the entrepreneurship in Uzbekistan | E3S Web of Conferences.19 july 2023

DIGITAL TRANSFORMATION AND GOVERNANCE OF HIGHER EDUCATION INSTITUTIONS IN UZBEKISTAN: INSTITUTIONAL CAPACITY, DATA INFRASTRUCTURE AND INTERNATIONAL COMPETITIVENESS

Mukhlisa Khamidova
Westminster International University in Tashkent
https://doi.org/10.5281/zenodo.23136709

Abstract. *This paper examines the concept of digital transformation and governance in the higher education sector of Uzbekistan, analyzing how institutional capacity, the institutional data ecosystem, funding models, and institutional autonomy impact universities' ability to meet the challenges of international competition. Research materials are gathered as qualitative evidence through semi-structured interviews with leaders of 11 Uzbek Higher Education Institutions (HEIs) and an official responsible for the national National Quality Assurance Agency for Educators (NQAAE) ranking function and the data from the official QS World University Rankings 2026, while the higher-education policy documents serve as reference materials. The analysis builds the analytical concept of a capability gradient to understand how institutional responses to the ranking requirements are influenced by the different financial, administrative, human, internationalisation and data-related capabilities. The results showed that digital and data infrastructure is not just a technological tool but is also an integral element of the governance capacity that is used to gather, validate, analyse and use information for strategic decision-making by the universities. The article suggests that digital transformation could be a facilitator towards more sustainable institutional development in the presence of capability sensitive funding, meaningful institutional autonomy, shared data infrastructure and inter-university benchmarking. The draft document connects digital governance with institutional capacity and international competitiveness and offers policy suggestions for the transformation of the higher education system in Uzbekistan in the future.*

Keywords*: digital transformation; higher education; institutional capacity; data infrastructure; international rankings; institutional autonomy; Uzbekistan;*

Аннотация. *В данной статье рассматривается концепция цифровой трансформации и управления в секторе высшего образования Узбекистана, анализируется, как институциональный потенциал, экосистема институциональных данных, модели финансирования и институциональная автономия влияют на способность университетов справляться с вызовами международной конкуренции. В качестве качественных доказательств были собраны учебные материалы посредством полуструктурированных интервью с руководителями 11 узбекских вузов и должностным лицом, ответственным за национальную функцию оценки качества высшего образования/рейтинга, а также данные из официального рейтинга QS World University Rankings 2026, в то время как документы по политике в области высшего образования служили справочными материалами. Анализ формирует аналитическую концепцию градиента возможностей, позволяющую понять, как институциональные реакции на требования рейтинга зависят от различных финансовых, административных, кадровых, интернационализационных и связанных с данными возможностей. Результаты показали,*

что цифровая и информационная инфраструктура — это не просто технологический инструмент, но и неотъемлемый элемент управленческого потенциала, используемый университетами для сбора, проверки, анализа и использования информации в целях принятия стратегических решений. В статье предполагается, что цифровая трансформация может способствовать более устойчивому институциональному развитию при наличии финансирования, учитывающего потенциал, значимой институциональной автономии, общей инфраструктуры данных и межуниверситетского сравнительного анализа. В проекте документа рассматривается взаимосвязь цифрового управления с институциональным потенциалом и международной конкурентоспособностью, а также предлагаются политические рекомендации по трансформации системы высшего образования в Узбекистане в будущем.

Ключевые слова: *цифровая трансформация; высшее образование; институциональный потенциал; инфраструктура данных; международные рейтинги; институциональная автономия; Узбекистан;*

Annotatsiya. *Ushbu maqolada O'zbekiston oliy ta'lim sohasida raqamli transformatsiya va boshqaruv konsepsiyasi ko'rib chiqiladi, institutsional salohiyat, institutsional ma'lumotlar ekotizimi, moliyalashtirish modellari va institutsional avtonomiya universitetlarning xalqaro raqobat muammolarini hal qilish qobiliyatiga qanday ta'sir qilishini tahlil qiladi. O'quv materiallari sifatli dalil sifatida 11 ta O'zbekiston OTMlari rahbarlari va Ta`lim Sifatini Ta`minlash Milliy Agentiligi tomonidan mas'ul mansabdor shaxs bilan yarim tuzilgan intervyular hamda QS World University Rankings 2026 rasmiy ma'lumotlari orqali to'planadi, oliy ta'lim siyosati hujjatlari esa ma'lumotnoma materiallari sifatida xizmat qiladi. Tahlil reyting talablariga institutsional javoblarning turli moliyaviy, ma'muriy, insoniy, xalqarolashtirish va ma'lumotlar bilan bog'liq imkoniyatlarga qanday ta'sir qilishini tushunish uchun qobiliyat gradientining analitik kontseptsiyasini yaratadi. Natijalar shuni ko'rsatdiki, raqamli va ma'lumotlar infratuzilmasi nafaqat texnologik vosita, balki universitetlar tomonidan strategik qarorlar qabul qilish uchun ma'lumotlarni to'plash, tasdiqlash, tahlil qilish va ulardan foydalanish uchun foydalaniladigan boshqaruv salohiyatining ajralmas elementi hamdir. Maqolada raqamli transformatsiya imkoniyatlarga sezgir moliyalashtirish, mazmunli institutsional avtonomiya, umumiy ma'lumotlar infratuzilmasi va universitetlararo benchmarking mavjud bo'lganda yanada barqaror institutsional rivojlanishga yordam berishi mumkinligi ta'kidlangan. Hujjat loyihasi raqamli boshqaruvni institutsional salohiyat va xalqaro raqobatbardoshlik bilan bog'laydi va kelajakda O'zbekistonda oliy ta'lim tizimini o'zgartirish bo'yicha siyosiy takliflarni taqdim etadi.*

Kalit so'zlar *: raqamli transformatsiya; oliy ta'lim; institutsional salohiyat; ma'lumotlar infratuzilmasi; xalqaro reytinglar; institutsional avtonomiya; O'zbekiston;*

1. Introduction

The rankings of universities have become a more prominent part of higher education policy, especially at the international level. Rankings are used to communicate about international competitiveness by the government, as a tool to compare institutional performance by universities, and as one source of comparative information by prospective students and partners. However, rankings are not just created in isolation. However, there are significant differences in the financial resources, research infrastructure, staffing, international connections, administrative capacity and experience in the ranking methodologies of universities. These

differences are important because many ranking criteria are based on outputs that are accrued over time, as opposed to activities that can be modified on the spot. It's an especially interesting topic to explore in the context of Uzbekistan. The national higher education strategy set out a clear target for the internationalisation and competitiveness of the country's higher education institutions, the goal of which was to have at least ten Uzbek universities among the first 1,000 universities in international rankings by 2030, and ambitions to make this a reality for flagship universities, the National University of Uzbekistan and Samarkand State University. The policy framework is also designed to include a structure of support for institutions aiming at these targets. The official data of the government proves the growth of participation in the international ranking. In 2025, seven Uzbek universities participated in the QS World University Rankings, in 2026, the results of the QS subject rankings appeared. The 2026 QS World University Rankings, meanwhile, show that institutional capacity is vital to 2026. TIIAME-NRU was ranked as #469 in the 2026 QS World University Rankings, while the National University of Uzbekistan was in the #721–730 range. The effectiveness of the implementation of these results indicates the improvement but also indicates that the Uzbek system is not starting from a uniform institutional baseline. This is not to say that rankings are a bad method for policymaking. Instead, it is necessary to determine if a common performance incentive can yield an equitable and sustainable result in cases where universities have differing capacities. The study poses three questions: (1) What is university leaders' perception of the correlation between the financial and administrative resources they have had in the past and their competitiveness in the ranking? (2) how much institutional autonomy is substantive and (3) how can the difference in capabilities be diminished without sacrificing incentives for high performance? The article contributes in two ways. It sets up a concept of a capability gradient and explains how it can be used to account for the differences in the institutional capacity accumulated by universities, and thus their responsiveness to ranking pressures. Second, it links this empirical observation with the design of performance-based funding, stating that a policy can inadvertently perpetuate institutional inequalities as it pays for measured outcomes but does not also provide funding for the competencies that can produce the outcomes.

2. Literature Review and Theoretical Framework

The paper used four theoretical approaches to analyze. The first is resource dependence theory. The interdependence of organisation behaviour and autonomy is dependent on the usage of outside resources, according to Pfeffer and Salancik (1978). Institutions of higher education may have formally gained autonomy in decision-making, but this does not always lead to any substantive autonomy if they do not have the discretionary financial power, specialist staff or access to external networks. This is relevant especially to ranking competition, which needs data management, research development and international partnerships and specialised managerial skills. The second view is that of path dependence. The initial institutional setting can create increasing returns, as North (1990) and Pierson (2000) demonstrate by focusing on the early advantages that can make later development easier, or on the institutional disadvantage that can lead to higher costs of catching up. A university that has a history of research groups, grant managers, with international co-authorship networks and a proven track record of successful conversions of new investments will be faster to convert new investments in higher education than a university that has yet to develop these skills and networks. Third, the literature on university autonomy includes a distinction between formal (procedural) autonomy and substantive autonomy. Berdahl (1990) and later in the field of higher education governance

research, the capacity for making decisions has meaning only if there exists institutional resources and organisational authority to put them into practice. The university has the authority to set priorities, enter partnerships and/or redesign their own processes, but is not free to do so if processes are not allowed for budget reasons, staffing regulations or central approval. Fourth, research in rankings shows the accumulated and reactive nature of ranking systems. Hazelkorn (2015) and Marginson (2006) demonstrate the effect of rankings on institutional behaviour, policy priorities and resource allocation. Ranking thus turns into not just a measuring tool, but a governance tool as well. For instance, the current QS methodology for World University Rankings is heavily based on academic reputation, citations per faculty, employer reputation, faculty-student ratio, international research networks and other measures that rely on the cumulative achievements of the institution. Academic reputation is now cited as being 30% of the criteria, and citations per faculty as 20% – the methodology is reviewed annually. Combined, these viewpoints imply a theoretical mechanism: the unequal resource endowment leads to unequal organisational capabilities, which in turn affect the speed and quality with which universities can react to the requirements of the rankings; the observed performance is then rewarded, and if this reward does not reflect the initial resource endowment, then the resulting allocation can actually exacerbate the initial differences. This mechanism is referred to as the capability-gradient effect in this article.

3. Research Methodology

This study uses a qualitative, comparative case-oriented design. The empirical material is made up of semi-structured in-depth interviews with leaders from 11 Uzbek HEIs, and an interview with an official in charge of the national quality-assurance/ranking function. The institutional sample was structured to represent the variation in institutional mission, structure, comprehensive, special, pedagogical, regional, agrarian institutions and new institutions. Sampling was analytical and not statistical. The interviews focused on four themes: (1) financial and administrative capacities of institutions; (2) practical autonomy in resource allocation and strategic decision-making; (3) institutional reactions to the adoption of international rankings; and (4) perceptions of ranking related policy and potential reforms. Official information on the QS World University Rankings 2026 from the QS World University Rankings and public documents on higher education policy in Uzbekistan were also used in the analysis. Multiple evidence types were employed to differentiate participants' perceptions from policy and ranking information that was observable. Thematic analysis and cross-case synthesis were used to analyze the interview data. The concepts found were repeated and contrasted between institutions. The analysis was geared toward understanding the mechanisms, rather than numerically estimating them, and the goal was to understand how leaders account for the differences in institutional capacity and how those differences influence the feasibility of ranking strategies. Causality claims are avoided, and when the evidence is in the form of participant perceptions, the article clearly indicates that these are perceptions or interpretations, rather than statistically determined effects. The model is consistent with the capability-gradient argument developed in this article. Digital information systems have the potential to be more effective tools for benchmarking, planning and performance enhancement for universities that have existing digital data systems, specific staff trained in how to use them and research-management expertise. The advantages of digital transformation may not be realized until institutions have first put in place foundational data governance, staff knowledge and organizational processes. Digital transformation should therefore be considered as a process of the socio-organizational

nature and not only as technology implementation. Digital transformation is not only the adoption of digital tools, but the use of data, capabilities and governance processes within institutional decision-making as defined in this study. By analyzing the empirical material, a model is proposed that sees the institutional capacity, the digital infrastructure, the governance autonomy, and the strategic allocation of resources as mutually influencing each other in the way they affect the capacity of universities to turn information into institutional performance. In this model, the information base is digital infrastructure, the governance arrangements are the rules and regulations that govern the use of the information, and institutional capabilities are the abilities of the institutions to turn digital resources into sustainable research, internationalization and management outcomes.

3.1. Analytical model of digital transformation and institutional capacity

There is one important evidentiary limitation of the study. Does not include a longitudinal institutional panel or financial data at the institutional level for statistical testing of the causal effect of prior expenditure on the ranking improvement. The notion of a capability gradient should thus be considered as an analytical construct that is based on the empirical analysis and that may serve as an indicator for the direction of empirical research in the future, rather than as a causal constant that is derived from the present study.

4. Findings

4.1. A common funding formula can mask very different institutional costs

The apparent costs at institutions can be very different, even when the same funding formula is used. Interview responses suggest a worry of a 'log of funding‘ not equaling 'equal capacity to compete'. One university leader compared the consequences of a seemingly neutral formula for a university with a more radical student body and mission with the consequences of the same formula for a different size, mission, and research intensity university. The question is not just how much money will be made available; it's also the balance between funding and institutional mission and the costs of developing research and internationalization capacity. This result is further in line with resource dependence theory. Expenditures for international ranking competitiveness must be proportional to the number of students, but they also have fixed or semi-fixed costs, such as research infrastructure, doctoral supervision, data systems, international research travel, grant development, and professional research support and recruitment of internationally active academics. A formula which primarily reflects enrolment or standardized categories will, therefore, lead to underfunding of activities of capacity building, which are the most important for international competitiveness.

4.2. The capability gradient

The interviews also indicate that the universities of Uzbekistan differ not only in their existing resources but also in their organizational capacities, developed thanks to earlier investments. The one participant compared an institution with an established ranking team and expertise with universities where the ranking function, internationalization and data work are distributed, with one or several employees. This distinction is of analytical significance. A ranking strategy is not just a listing of the activities, it is an organizational ability that demands specialized knowledge, knowledge management, research management, international networking and the capacity to keep up these initiatives over several years. The capability gradient is therefore a measure of the unequal distribution of the institutional capabilities that are needed to turn strategic intentions into measurable outcomes. A university with some established research teams and international networks can apply a new funding for increased activity to existing

teams. For a less resourced university, this allocation may need to be used first to hire staff, establish procedures and develop partnerships before measurable outputs can be seen. The additional funding may then have varying marginal impacts between institutions, etc. This is also in line with the present framework of QS ranking indicators. There are processes that create the academic reputation, research impact, international research networks and other indicators associated with faculty that typically take time to build up, rather than change on a whim. The indicators explicitly stated by QS are related to research quality, citation impact, academic reputation and sustained international partnerships, including research networks based on repeated joint publication. reputation and sustained international partnerships including research networks based on repeated joint publication.

4.3. TIIAME as an illustrative case of accumulated capacity

TIIAME-NRU illustrates a case in keeping with the capability-gradient argument. The university has climbed a spot to become #469 in the QS World University Rankings 2026, from #547 in 2025. The university and government statistics also demonstrate good results on several QS-related aspects and a strong international presence. The present study does not assert that the ranking position of TIIAME was solely due to its historic resource base. Instead, the case illustrates the importance of building up organizational capacity. A university with existing research infrastructure, expertise, international connections and ranking history has a stronger chance to leverage further resources into ranking outcomes. The case is thus an example of path dependence, but not a proof of a causal-determined link. This distinction is important. Success in ranking is multi-faceted and cannot be boiled down to just funding. All these factors are interdependent, such as academic reputation, research quality, citations and international involvement, student to staff ratios, and other factors. The added value of the capability gradient approach is that it provides an explanation of why institutions with different starting points could react differently to the same policy incentives.

4.4. Institutional autonomy: formal permission versus substantive capacity

Interview evidence also suggests a difference between the formal autonomy and the ability to be autonomous. An institution can have strategic priorities but be bound by central rules regarding staffing, procurement, generating revenue and/or the use of public funds. In such cases, autonomy is not a substantive but a procedural one. This is important for ranking policy. To ensure international competitiveness, decisions must be made quickly concerning research recruitment, international partnerships, data systems, conferences, research incentives and specialized support. The formal autonomy may not result in strategic responsiveness if these decisions must be passed through various levels of approval and/or are not expected to be funded using discretionary resources. Empirical evidence thus indicates that autonomy should be assessed not just in terms of the number of decisions that are legally delegated to universities, but also in terms of their financial and managerial autonomy, or the extent to which they can implement these decisions. This is in line with the resource-dependence view: authority with ineffective resources leads to ineffective autonomy.

4.5. The risk of reinforcing differences through performance-based incentives

The combination of capability differences and performance-based incentives brings a central policy implication to the fore. It is expected that if funding is made chiefly based on observed ranking performance, institutions with high research and internationalization capacity are likely to be able to achieve the performance criteria. The results of the investment might take a longer time to be seen in institutions with less initial capabilities. It's not about performance-

based funding. Performance-based rewards can enhance the accountability and prevent inefficient spending. The issue is that the policy fails to connect the concept of performance with capability. In a path-dependent system, an outcome-based mechanism can be self-reinforcing: the more well they do the more resources that are provided; the more resources that are provided the better they do; the less capacity they have the less well they do. This is especially significant, given that Uzbekistan's policy goal is expressed on the sectoral level but must be implemented selectively. The official concept 2030 foresee at least ten institutions to be in the top 1,000 of internationally recognized rankings, as well as the modernization of state support systems. According to the current government report, the number of universities in Uzbekistan included in the QS World University Rankings in 2025 is 7, among which 3 are in the top 1000.According to the current government report, 7 universities in Uzbekistan are currently listed in the QS World University Rankings, among which 3 institutions are in the top 1000.

5. Discussion

The findings provide an argument for moving from a competition model to a capability-sensitive competition model. It is assumed that universities play on a relatively level playing field in the simple model and that performance incentives will foster efficiency. The capability sensitive model acknowledges that the universities start at different points and that the outputs that contribute to the production of rankings are cumulative. This separation has governance consequences that are directly felt. Policy makers wishing to use rankings to improve a system need two instruments, not one: a performance instrument for demonstrating results, and a capability-building instrument for funding the institutional underpinnings necessary to achieve results. The two instruments are not substitutes for each other but are complementary tools. The argument is also a contribution to the institutional isomorphism literature. There are significant pressures on universities to have similar structures, offices and strategies because of rankings. But the same formal set up doesn't lead to the same results. The organizational structure of an employee-based ranking office, based in a resource-challenged university, differs from that of a multidisciplinary ranking and research analytics office in a well-funded institution. These formal institutional structures can thus be converged at the institutional level while the capabilities may diverge. Finally, the results indicate that the ranking policy ought to be related to the institutional mission differentiation. It is not reasonable to expect ranking results to be the same at the same time at research intensive universities, regional teaching-oriented universities, specialized professional universities, and newly established universities. A differentiated system can keep national standards and be differentiated in terms of developmental pathways and resource packages.

6. Policy Recommendations

6.1. Introduce mission- and capability-sensitive funding

Funding formulas should include base funding, mission-specific and performance incentives. The performance element should be based on achievement, progress and results, while the capabilities building element should be linked with research infrastructure, internationalization, doctoral education, research management and research data capacity in research institutions where these capacities are missing. The aim should be to not duplicate the reward for the same capability, i.e. where it already exists and where the policy is relying on its outputs as the key driver of further funding.

6.2. Build shared ranking and data infrastructure

A central or shared service may help lower the costs of smaller and less-resourced universities. This mechanism could offer methodological training, bibliometric analysis, data validation, intelligence and tools for international partnerships and benchmarks. This would not take the burden of institutions away from them, it would only help to bring down the cost of acquiring specialized knowledge.

6.3. Establish formal inter-university benchmarking

Structured benchmarking reports of the number of publications, citations, international cooperation, faculty ratios and other factors should be sent out to universities comparing their research output with other universities having a comparable mission and comparable size. A league table only represents a form of benchmarking. It should be used for developmental purposes to find out what capabilities are lacking and what investments are necessary to develop them.

6.4. Link autonomy to accountability rather than treating them as opposites

There should be more flexibility for universities in decisions where a quick strategic action is required and reports made to the public should be transparent with independent assessment. The purpose of measuring autonomy should therefore be not only in terms of formal delegation of autonomy but also in terms of capability to make and implement strategic decisions.

6.5. Use multi-year ranking policy horizons

Investments should be safeguarded according to the rank via multi-year institutional plans. Since the reputation and research markers build up over time, having an annual budgeting cycle can lead to short-term projects that do not actually help measure the desired research. Cumulative nature of research work and building up international reputation would be better matched with three to five years' time horizon.

7. Limitations and Future Research

There are several limitations in the study. First, the sample is not statistically representative of the higher education sector in Uzbekistan as it includes 11 university leaders and one official at the national level. Second, the study is cross-sectional and cannot directly look at the impacts of the 2026–2027 policy environment on universities over time. Third, the financial information for institutions were not available in a form that would allow for a quantitative estimation of the link between expenditure and ranking outcomes. Third, not all the information in the policies comes from published documents, so is conveyed as institutional perspectives, not as incontrovertible legal facts. Future studies should develop a longitudinal data panel for Uzbek universities that integrates annual spending, staffing, research production, international cooperation and student-faculty ratios as well as ranking measures of universities. This type of data would enable researchers to determine whether institutional capacity is a determinant of the impact of extra funding. Cross-system comparisons of Central Asian systems might also explore the effects of differentiated funding mechanisms on the level of ranking inequality. Another area of research is to differentiate between the more specific concept of "ranking capacity" and overall institutional quality. It's possible that universities can secure a better ranking by investing in highly relevant ranking-based measures that are less relevant to other education outcomes. Future research should thus explore whether ranking-based investment strategies correlate with increased teaching quality, enhanced graduate results, better research quality, and increased public value.

8. Conclusion

Improving the competitiveness of Uzbek universities in international rankings is not just a matter of strategic objectives. The qualitative evidence analysed in this article shows that universities go into the competition with different financial, administrative, human and internationalisation capacities. These disparities result in a capability gradient which influences the rate at which institutions can realize measurable results in response to policy. The central governance issue is that it is inappropriate to develop performance-based incentives as if all institutions started from the same point. In addition to enhancing accountability, performance funding can help to exacerbate current institutional advantages if there is no corresponding system of capacity building. More sustainable approach would involve mission differentiated baseline funding, capability investments, joint ranking and data infrastructure, institutional autonomy and developmental benchmarking this meaningfully. The positive results of the national policy in the rankings highlight its effectiveness. A more important step in the reform agenda, however, should be aimed not only at rewarding institutions which are already good performers, but also at the creation of the capacity of a larger number of institutions to become internationally competitive. In this regard it is more apt to view the goal of ranking policy as not only a move up in a table but as reinforcing the institutional bases for continued international performance.

REFERENCES

1. Pfeffer, J., & Salancik, G. R. (1978). The external control of organizations: A resource dependence perspective. Harper & Row.
2. North, D. C. (1990). Institutions, institutional change and economic performance. Cambridge University Press.
3. Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. American Political Science Review, 94(2), 251–267.
4. Berdahl, R. (1990). Academic freedom, autonomy and accountability in British universities. Studies in Higher Education, 15(2), 169–180.
5. Hazelkorn, E. (2015). Rankings and the reshaping of higher education: The battle for world-class excellence (2nd ed.). Palgrave Macmillan.
6. Marginson, S. (2006). Dynamics of national and global competition in higher education. Higher Education, 52(1), 1–39.
7. Salmi, J. (2009). The challenge of establishing world-class universities. World Bank.
8. Altbach, P. G., & Salmi, J. (Eds.). (2011). The road to academic excellence: The making of world-class research universities. World Bank.
9. DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147–160.
10. Khamidova, M. (2025a). The role of internationalization in enhancing Uzbekistan's higher education sector. Iqtisodiy Taraqqiyot va Tahlil, 144–149.
11. Khamidova, M. (2025b). Governance and strategic responses to global rankings: Evidence from Uzbek universities. Ilg'or Iqtisodiyot va Pedagogical Technologiyalar, 499–506.
12. Quacquarelli Symonds. (2025). QS World University Rankings 2026: Methodology and results. QS.

13. President of the Republic of Uzbekistan. (2019). Decree No. PF-5847: On approval of the Concept for the Development of the Higher Education System of the Republic of Uzbekistan until 2030.
14. Ministry of Higher Education, Science and Innovation of the Republic of Uzbekistan. (2025–2026). Official information on international rankings and higher education reform.

POETRY TRANSLATION AND THE PROBLEM OF EQUIVALENCE

Begoyim Kholbekova
Associated Professor
UzSWLU
https://doi.org/10.5281/zenodo.23136735

Abstract. *Poetic translation remains one of the most theoretically complex and aesthetically demanding areas within translation studies due to the inseparable unity of meaning, rhythm, imagery, sound, and cultural symbolism. This article examines the problem of translational adequacy in poetic translation and argues that direct translation from the source language constitutes the most reliable criterion for achieving equivalence.*
Keywords: *poetic translation; direct translation; equivalence; translation as recreation; literary translation; cultural specificity*

Among the various forms of literary translation, poetic translation occupies a particularly complex and controversial position. Poetry is traditionally defined as emotionally rhythmic speech organized according to a specific melodic and structural order. As noted in A Dictionary of Literary Studies Terms by N. Khotamov and B. Sarimsoqov, poetry represents a synthesis of emotion, rhythm, imagery, and sound. This synthesis renders poetic translation uniquely challenging, as the translator must negotiate not only semantic content but also aesthetic form and cultural symbolism.

It must be acknowledged that no translation can fully reproduce the totality of artistic components present in the original poetic text. The attempt to preserve all formal and semantic elements simultaneously often leads to aesthetic compromise. If absolute equivalence were established as a strict requirement, poetic translation would become theoretically and practically impossible. Therefore, poetic translation must be approached as an exceptional creative act, which this article defines as recreation. Although scholarly discourse continues to use the term "translation," its essence in the poetic domain lies in re-creating the work within another linguistic and cultural system.

One of the most persistent challenges of direct translation lies in recreating metaphorical and symbolic components of the source text in the target language. Each culture possesses a distinct metaphorical system shaped by historical experience, geography, climate, and worldview. Consequently, figurative expressions that appear natural and aesthetically pleasing in one language may become inappropriate, obscure, or even negatively connoted in another. This problem is vividly illustrated in Robert Burns's poem "John Anderson," where the poet compares his friend's black hair to a raven. In Scottish cultural perception, the raven is a familiar and aesthetically neutral bird, and the comparison primarily evokes color rather than symbolic meaning. However, in Uzbek cultural tradition, the raven carries negative connotations associated with misfortune and darkness, rendering the metaphor unacceptable.

The Russian translation by Samuil Marshak replaces the raven metaphor with pitch (смоль), a substance traditionally valued in Russian culture and commonly associated with deep black color. The Kazakh translation by Quvandıq Shangitbayev further adapts this imagery by

comparing the hair to coal (көмір), a culturally significant symbol of warmth and vital energy. In Uzbek poetic tradition, black hair is conventionally compared to night (тун), a metaphor associated with beauty, softness, intimacy, and aesthetic depth.

Thus, across four languages – English, Russian, Kazakh, and Uzbek – the same poetic idea is realized through culturally distinct metaphors: raven, pitch, coal, and night. Despite these differences, the semantic core of the original image is preserved. This example demonstrates that poetic equivalence is achieved not through literal reproduction but through culturally appropriate recreation.To provide a clearer sense and facilitate comparison, we present below the original versions of those lines alongside their translations in three languages:

John Anderson my jo, John,
When we were first acquent,
Your locks were like the raven,
Your bonie brow was brent.
Robert Burns

Джон Андерсон, мой старый друг,
Подумайка, давноль,
Густой, крутой твой локон
Был черен, точно смоль.
Samiul Marshak

Джон Андерсон, Джон досым,
Мынауна жол босын,
Самай шашын тогілген
Қара еді ғой комірден.
Quandiq Shang'itbayev

Жон Андерсон, қадрдон дўстим,
Бизлар илк бор танишган онда,
Тундай қаро қуюқ сочларинг
Барқ урарди дўнг пешонангда.
Begoyim Kholbekova

A central argument of this study concerns the consequences of indirect translation. In several cases examined, Kazakh and Uzbek translations of Burns's poetry were produced through Russian intermediary texts rather than directly from English. Such mediation significantly increases the risk of semantic loss and aesthetic distortion. The omission of the line "Your bonie brow was brent" in both Russian and Kazakh translations illustrates how indirect translation can result in the loss of poetically significant details that contribute to character portrayal and aesthetic imagery. Similar omissions are observed in translations of Burns's poem "The Lass That Made the Bed to Me," where the metaphor "Her teeth were like the ivorie" is absent in Russian, Uzbek, and Kazakh versions.

In the source text:
Her hair was like the links o'gowd,
Her teeth were like the ivorie,

Her cheeks like lilies dipt in wine,
The lass that made the bed to me.
Robert Burns
In the Russian translation:
Был мягок шёлк её волос,
И завивался, точно хмель,
Она была душистей роз,
Та, что постлала мне постель.
(Samiul Marshak)

In the Uzbek translation:
Сочлари-чи, бирам жингалак,
Ипакдан ҳам майинлиги бор,
Таратишда хуш бўй ул малак,
Атиргулни айларди бекор.
(Muhammad Ali)
In the Kazakh translarion:
Есіліп шашы қулмақтай,
Төсегін салған қарағым,
Анқиды исі гул бақтай-
Елжіреп балқып барамын .
(Quandiq Shang'itbayev)

These omissions cannot be explained solely by cultural differences. Rather, they reflect the cumulative effect of translational distance, whereby each additional mediating layer increases the likelihood of semantic erosion. Direct translation, by contrast, enables the translator to engage directly with the original imagery and to seek culturally appropriate equivalents without relying on secondary interpretations.

Below, we attempt to reconstruct the omitted line – left out in the Russian, Kazakh, and Uzbek indirect translations – by translating it directly from the original:

Майин эди сочи ипакдан,
Тишларин дур қатори дейсиз.
Менга ўрин - жой солиб берган
Гул юзни гул ифори дейсиз.
(Begoyim Kholbekova)

National literary traditions play a decisive role in shaping translational strategies. In Uzbek classical and folk poetry, a woman's white teeth are traditionally compared to pearls rather than ivory. This metaphor emphasizes beauty, purity, and value in a manner that aligns with Uzbek aesthetic norms. Accordingly, direct translation from English into Uzbek allows for the recreation of the original image in a culturally resonant form, whereas indirect translation often suppresses such imagery altogether.

Accordingly, this article reaffirms direct translation as a fundamental criterion of adequacy in poetic translation and supports the view that each poetic translation must be regarded as a unique creative exception rather than a mechanical reproduction.

REFERENCES

1. Kholbekova B. She'riy tarjima mezoni. Jahon Adabiyoti. Tashkent. 2009
2. Marshak, S. Selected translations. Moscow: Khudozhestvennaya Literatura. 1965
3. Salomov, G. Tarjima nazariyasi va amaliyoti. Tashkent: Fan. 1983
4. Роберт Бёрнс. Қўшиқ ва балладалар. Т., 1971 34 - бет.
5. Роберт Бёрнс. "ЖАЗУШЫ " БАСПАСЫ О., 1979 138 - бет

A REVIEW OF SELECTED STUDIES ON THE DEVELOPMENT OF AN ARABIC-UZBEK SEMANTIC SEARCH SYSTEM

Zebo Gulboyevna Kodirova
Doctor of Philosophy (PhD) in Philology
National University of Uzbekistan
https://doi.org/10.5281/zenodo.23136768

Annotation. *This article examines existing approaches to semantic search within the Qur'anic text and their linguistic foundations. It discusses the significance of ontological modeling, morphological annotation, parallel corpus development, and the identification of translation equivalents for semantic search. Particular attention is given to the linguistic resources required to develop a semantic search model based on the Arabic Qur'anic text and its Uzbek translations, as well as the possibilities for integrating these resources.*

Key words: *semantic search, linguistic resources, ontology, parallel corpus, annotation, translation equivalence, arabic-uzbek model.*

Аннотация. *В статье анализируются существующие подходы к организации семантического поиска в тексте Корана и их лингвистические основы. Рассматривается значение онтологического моделирования, морфологической разметки, создания параллельного корпуса и определения переводных соответствий для семантического поиска. Особое внимание уделяется лингвистическим ресурсам, необходимым для разработки модели семантического поиска на основе арабского текста Корана и его узбекских переводов, а также возможностям интеграции этих ресурсов.*

Ключевые слова: *семантический поиск, лингвистическое обеспечение, онтология, параллельный корпус, аннотирование, переводное соответствие, арабско-узбекская модель.*

Annotatsiya. *Maqolada Qur'on matnida semantik qidiruvni tashkil etishga doir mavjud yondashuvlar va ularning lingvistik asoslari tahlil qilinadi. Ontologik modellashtirish, morfologik annotatsiya, parallel korpus yaratish va tarjima muqobillarini aniqlashning semantik qidiruvdagi ahamiyati yoritiladi. Arabcha Qur'on matni va uning o'zbekcha tarjimalari asosida semantik qidiruv modelini ishlab chiqish uchun zarur lingvistik resurslar hamda ularni o'zaro birlashtirish imkoniyatlariga alohida e'tibor qaratiladi.*

Kalit so'zlar: *semantik qidiruv, lingvistik ta'minot, ontologiya, parallel korpus, annotatsiya, tarjima muqobilligi, arabcha-o'zbekcha model.*

Introduction. Advances in corpus linguistics and natural language processing technologies have expanded opportunities for studying texts in electronic form and retrieving information from them. Characterized by semantic complexity, distinctive grammatical structures, and an extensive exegetical tradition, the Qur'anic text has become a significant object of research in computational linguistics. Simple electronic searches are insufficient for working with the Qur'anic text because the concept sought by a user may not be expressed through the exact word entered in the search query. A single concept may be expressed across different verses through alternative lexical units, grammatical forms, semantic relationships, or explanatory translations.

Most existing search systems rely on identifying direct textual matches for the units contained in a user's query. Although this approach is useful for locating a particular word form, it does not adequately account for synonymy, polysemy, contextual equivalence, or cross-linguistic relationships. Establishing semantic search between the Arabic Qur'anic text and its Uzbek translations requires an integrated analysis of the typological differences between the two languages, their writing systems, morphological structures, and translation strategies.

Numerous studies have examined the Qur'anic text from the perspectives of computational linguistics and natural language processing. Several studies relevant to the development of a semantic search system are discussed below. M. H. Bashir, A. M. Azmi, H. Nawaz, W. Zaghouani, M. Diab, A. Al-Fuqaha, and J. Qadir summarize research on the analysis of the Qur'anic text using natural language processing methods. Their work addresses areas such as automatic morphological analysis, semantic search, and the identification of errors in Qur'anic recitation through speech recognition technologies [Bashir M. H., et al., 2023, 1]. A. Al-Shidi, F. H. Nordin, I. I. Mohamed, M. A. Younis, I. E. Elmutasim, and E. E. Abusham examine the grammatical and semantic analysis of the Qur'anic text, automatic recitation recognition, the study of translations, ontology development, and the design of search systems [Al-Shidi A., et al., 2025, 2]. They also describe the potential application of topic modeling, paragraph vectors, and clustering methods to identify themes and semantic relationships among them.

The linguistic nature of semantic search. Semantic search is an information retrieval approach that seeks to identify the meaning underlying a user's query rather than relying solely on the formal correspondence of letters or words. For example, when a user submits the query *"imtihonga tayyorlanish"* the system should not limit its results to texts containing only the words *"imtihon"* and *"tayyorlanish*". The same meaning may also be conveyed through expressions such as *"nazorat ishiga hozirlik ko'rish", "sinov oldidan mavzularni takrorlash", "test yechish", "savollarga javob tayyorlash",* or *"bilimlarni mustahkamlash".*

These examples demonstrate that semantic search involves identifying the concepts, intentions, context, and semantic relationships embedded in a user's query rather than merely locating the external forms of words.

When applied to the Qur'anic text, semantic search must consider a word's form within a verse, its root, grammatical function, semantic field, relationships with other concepts, and representation in translation. For example, when a user submits the Uzbek query *"ota-onaga yaxshilik qilish"* the system should not restrict its results to translations containing the exact words *"ota-ona"* or *"yaxshilik"*. In relevant verses, the same concept may be expressed through units and phrases such as *"ehson", "yaxshi muomala", "hurmat", "marhamat", "rahm-shafqat", "ularga 'uff' demaslik",* or *"keksayganlarida g'amxo'rlik qilish".* The concept *"Yetimlarga g'amxo'rlik qilish"* may appear in different forms, including *"yetimni asrash", "yetim molini muhofaza qilish", "yetimning haqini yemaslik", "yetimga ehson qilish", "yetimni xorlamaslik", and "yetimlarga adolatli munosabatda bo'lish".* Therefore, the system must associate the query not merely with a single lexeme but with units belonging to a shared conceptual field related to the rights and protection of orphans.

In the Arabic text, the same concept may be conveyed through a different morphological or syntactic structure rather than through a direct formal equivalent of the Uzbek words. Accordingly, the central function of semantic search is to move beyond formal correspondence toward semantic equivalence.

M. H. Bashir, A. M. Azmi, H. Nawaz, W. Zaghouani, M. Diab, A. Al-Fuqaha, and J. Qadir emphasize the importance of tokenization, morphological segmentation, root identification, text alignment, ontologies, and conceptual search in Qur'anic natural language processing research. They also demonstrate that the Qur'anic text possesses distinctive orthographic characteristics, diacritical marks, and features associated with Classical Arabic [Bashir M. H., et al., 2023, 1].

The root-and-pattern morphology of Arabic is also important for semantic search. A single orthographic form may incorporate multiple morphological components. In Uzbek, by contrast, grammatical meanings are frequently expressed through sequences of affixes. Consequently, the number of words or grammatical segments in Arabic and Uzbek texts does not always correspond directly. A bilingual platform should therefore avoid relying exclusively on one-to-one lexical alignment.

Ontology-based search approaches. Ontological modeling constitutes one of the principal approaches to semantic search within the Qur'anic text. An ontology represents concepts within a particular domain, their properties, and their relationships in a form suitable for computational processing. In a Qur'anic ontology, verses, concepts, religious terms, individuals, events, and themes can be interconnected.

M. Alqahtani and E. Atwell propose an ontology-based search tool for the Qur'anic text. Their work demonstrates the need for a conceptual data layer in semantic search. The value of this approach lies in its ability to expand a user's query through related concepts rather than restricting the search to a single word form [Alqahtani M., Atwell E., 2016, 3]. R. Bentrcia, S. Zidat, and F. Marir investigate the extraction of semantic relationships from Qur'anic Arabic using Arabic conjunction patterns. Their research demonstrates that semantic relationships can be derived not only from predefined lexical resources but also from the syntactic structure of the text [Bentrcia R., et al., 2018, 4].

H. Afzal and T. Mukhtar associate Qur'anic concepts with an enhanced semantic search model based on the principles of English WordNet. This approach demonstrates the contribution of lexical-semantic networks, synonym sets, and conceptual relationships to search effectiveness. However, an English WordNet model cannot be transferred directly to Uzbek because the Uzbek representation of religious concepts, established patterns of usage, and semantic boundaries require separate linguistic description [Afzal H., Mukhtar T., 2019, 5]. F. Beirade, H. Azzoune, and D. E. Zegour propose a semantic search model that expands user queries through a Qur'anic ontology. Their system integrates morphological, syntactic, and semantic databases. Words are classified according to semantic fields, general meanings, related fields, and individual semantic units. The ontology was developed in OWL format using Protégé, while the search module was implemented through Apache Lucene [Beirade F., et al., 2021, 6].

I. K. Alshammari, E. Atwell, and M. A. Alsalka examine the integration of Qur'anic and Hadith topics using ontologies and word embeddings. Their study illustrates the possibility of linking Qur'anic concepts with other religious sources [Alshammari I. K., et al., 2024, 7]. In an Arabic-Uzbek model, the relationships among verses, translations, and exegetical interpretations may likewise constitute interconnected knowledge layers rather than isolated texts.

The quality of a semantic search platform depends substantially on its underlying corpus and level of annotation. A corpus is not merely a collection of electronic texts; it is a system that preserves structural relationships among textual units, metadata, and linguistic features. When developing a corpus for the Qur'anic text, the levels of surah, verse, word, and morphological

segment must be clearly defined. Each unit should be associated with a unique identifier, the original form of the Arabic text must be preserved, and subsequent annotation layers should be linked through the corresponding identifier.

In an Arabic-Uzbek model, annotation should be organized separately for each language while maintaining explicit relationships between the two layers. The Arabic layer should record word forms, lemmas, roots, morphological features, parts of speech, syntactic functions, and semantic concepts. The Uzbek layer should include translation units, lemmas, parts of speech, morphological forms, syntactic functions, religious terms, and conceptual tags. In addition, information concerning translation equivalence is required to connect the two layers.

Translation correspondence does not always imply that one Arabic word matches one Uzbek word. A single Arabic unit may be represented by several Uzbek words, multiple Arabic units may be translated through a single Uzbek expression, or certain meanings may be conveyed through explanatory expansion. Therefore, the parallel corpus should identify 1:1, 1:N, N:1, and N:N alignment patterns. Direct equivalents, contextual correspondences, explanatory expansions, and omitted units should also be explicitly marked.

Such annotation enables the retrieval of relevant verses and translations regardless of whether a user's query is formulated in Arabic or Uzbek. For example, if a meaning expressed in an Uzbek translation as *"yaxshilik qilish"* appears in the Arabic text through a different grammatical construction, the system can rely on alignment information connecting the corresponding units.

Conclusion. Semantic search within the Qur'anic text requires more than the identification of word forms; it must also account for meaning, grammatical properties, and relationships among concepts. Existing studies demonstrate that ontologies, morphological analysis, and semantic annotation expand information retrieval capabilities. Linking the Arabic Qur'anic text with Uzbek translations through a parallel corpus, translation equivalents, and religious concepts provides the linguistic foundation for a bilingual semantic search model.

REFERENCES

1. Bashir M. H., Azmi A. M., Nawaz H., Zaghouani W., Diab M., Al-Fuqaha A., Qadir J. Arabic natural language processing for Qur'anic research: a systematic review // Artificial Intelligence Review. – 2023. – Vol. 56. – P. 6801–6854. – DOI: 10.1007/s10462-022-10313-2.
2. Al-Shidi A., Nordin F. H., Mohamed I. I., Younis M. A., Elmutasim I. E., Abusham E. E. A literature review on natural language processing techniques for Qur'anic studies: challenges and insights // Frontiers in Signal Processing. – 2025. – Vol. 5. – Article 1535166. – DOI: 10.3389/frsip.2025.1535166.
3. Alqahtani M., Atwell E. Arabic Quranic Search Tool Based on Ontology // Natural Language Processing and Information Systems. Lecture Notes in Computer Science. – 2016. – Vol. 9612. – P. 478–485. – DOI: 10.1007/978-3-319-41754-7_52.
4. Bentrcia R., Zidat S., Marir F. Extracting semantic relations from the Quranic Arabic based on Arabic conjunctive patterns // Journal of King Saud University – Computer and Information Sciences. – 2018. – Vol. 30. – P. 382–390. – DOI: 10.1016/j.jksuci.2017.09.004.
5. Afzal H., Mukhtar T. Semantically Enhanced Concept Search of the Holy Quran: Qur'anic English WordNet // Arabian Journal for Science and Engineering. – 2019. – Vol. 44. – P. 3953–3966. – DOI: 10.1007/s13369-018-03709-2.

6. Beirade F., Azzoune H., Zegour D. E. Semantic query for Quranic ontology // Journal of King Saud University – Computer and Information Sciences. – 2021. – Vol. 33. – No. 6. – P. 753–760. – DOI: 10.1016/j.jksuci.2019.04.005.
7. Alshammari I. K., Atwell E., Alsalka M. A. Linking Quran and Hadith Topics in an Ontology using Word Embeddings and Cellfie Plugin // Proceedings of the 7th International Conference on Natural Language and Speech Processing. – 2024. – P. 449–455.

O‘QUV KORPUSINI SHAKLLANTIRISH METODIKASI VA CHASTOTALI LUG‘AT TUZISH TAMOYILLARI

Kuvondikova Gavhar Isomiddinovna
Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi mustaqil izlanuvchisi
https://doi.org/10.5281/zenodo.23253252

Annotatsiya. *Ushbu maqolada nofilologik yo‘nalish talabalarining yozma nutqi bazasida o‘quv korpusini shakllantirish metodikasi hamda uning negizida o‘quv-chastotali lug‘at tuzish tamoyillari tadqiq qilingan. Tadqiqotning empirik bazasini 10 ta fakultetdan yig‘ilgan 1300 ta tashkil etadi. Yig‘ilgan qo‘lyozma matnlar OCR texnologiyalari va AntConc dasturi vositasida qayta ishlanib, absolyut va nisbiy chastotalar hamda leksik xilma-xillik (TTR) ko‘rsatkichlari aniqlangan.*

Kalit so‘zlar: *korpus lingvistikasi, o‘quv korpusi, chastotali lug‘at, AntConc, OCR, leksik xilma-xillik, TTR.*

Аннотация. *В данной статье исследуются методика формирования учебного корпуса на основе письменной речи студентов нефилологических специальностей и принципы составления частотного словаря. Эмпирическую базу исследования составили 1300 эссе, собранных с 10 факультетов. Сканированные рукописные тексты были обработаны с помощью технологий OCR и программы AntConc, рассчитаны показатели абсолютной и относительной частоты, а также лексического разнообразия (TTR).*

Ключевые слова: *корпусная лингвистика, учебный корпус, частотный словарь, AntConc, OCR, лексическое разнообразие, TTR.*

Abstract. *This article explores the methodology of developing a learner corpus based on the written language of non-philology students and the principles of compiling a frequency dictionary. The empirical database consists of 1,300 essays collected across 10 faculties. Scanned handwritten texts were processed using OCR technologies and AntConc software to determine absolute and relative frequencies, as well as the Type-Token Ratio (TTR).*

Keywords: *corpus linguistics, learner corpus, frequency dictionary, AntConc, OCR, lexical diversity, TTR.*

KIRISH

Zamonaviy amaliy va nazariy tilshunoslikda lisoniy birliklarni avtomatlashtirilgan tahlil qilish hamda xolis empirik ma’lumotlarga tayanib xulosalar chiqarish yetakchi tamoyilga aylandi. Korpus lingvistikasi kompyuter texnologiyalari vositasida matnlar bazasini shakllantirish, ularga tarkibiy va leksik-grammatik ishlov berish hamda obyektiv statistik natijalarni olishning eng maqbul metodologiyasini taklif etadi (Sinclair, 1991; McEnery & Hardie, 2012). J. Sinclair ta’kidlaganidek, korpus vositasida tahlil qilish intuitiv yondashuvlardan voz kechib, matndagi haqiqiy lisoniy dalillar va so‘z qo‘llash qonuniyatlarini ochib beradi (Sinclair, 1991).

So‘nggi yillarda o‘zbek tilshunosligida ham kompyuter lingvistikasi va korpus tuzish borasida salmoqli tadqiqotlar amalga oshirilmoqda (Abdurahmonova, 2021). Biroq, muayyan maqsadli o‘quv korpuslari (learner corpora) va sohadan tashqari yo‘nalishlar yozma nutqini raqamli tahlil qilish masalalari hali ham yetarlicha tadqiq etilmagan metodologik yo‘nalishlardan

biri bo‘lib qolmoqda. O‘quv korpusi - til o‘rganuvchilar yoki muayyan guruh talabalari tomonidan yaratilgan og‘zaki va yozma matnlarning elektron bazasidir. S.Granger o‘quv korpuslari o‘zlashtirish jarayonidagi lisoniy o‘zgarishlarni (interlanguage) o‘rganishda va darslik hamda lug‘atlar uchun obyektiv material to‘plashda beqiyos manba ekanligini ko‘rsatib o‘tadi (Granger, 2008; Granger, 2015). An’anaviy leksikografiyada so‘zliklar ko‘pincha mutaxassislarning subyektiv tanlovi yoki umumiy matnlar asosida shakllantirilgan bo‘lsa, pedagogik leksikografiyada korpusga asoslangan (corpus-based) va korpus tomonidan yo‘naltirilgan (corpus-driven) yondashuvlar so‘zliklarni qat’iy chastotali va tarqalish ko‘lamiga ko‘ra saralash imkonini beradi (Hunston, 2002; Götz & Granger, 2024).

Oliy ta’lim muassasalarida nofilologik yo‘nalish (kimyo, fizika, biologiya va b.) talabalarining akademik va sohaviy lug‘at boyligini baholashda bir qator tizimli muammolar kuzatiladi:

1. Empirik bazaning yetishmasligi: Talabalarning haqiqiy akademik yozma nutqini aks ettiruvchi, raqamlashgan va tekshirilgan yozma korpuslarning mavjud emasligi.

2. Subyektivlik va noaniqlik: Akademik yozuvdagi leksik kamchiliklar hamda leksik xatoliklar (lexical errors) aniq miqdoriy ko‘rsatkichlarsiz, faqat umumiy kuzatishlar darajasida baholanishi.

3. Leksik minimum va lug‘atlar muammosi: Nofilologik yo‘nalishlar uchun tayyorlanadigan o‘quv darsliklari va sohaviy chastotali lug‘atlarda talabalarning o‘ziga xos nutqiy ehtiyojlari hamda faol ishlatiladigan leksik birliklar doirasi inobatga olinmasligi.

O‘quv korpusini yaratish va uning negizida chastotali lug‘at (frequency dictionary) shakllantirish ushbu muammolarni bartaraf etishning empirik yechimidir. Chastota ko‘rsatkichlari matnda qaysi so‘zlar ustuvor ekanligini, qaysi sohaviy va akademik leksika talabalar tomonidan samarali qo‘llanilayotganini aniq statistik raqamlarda ko‘rsatib beradi.

Ushbu tadqiqotning asosiy maqsadi - nofilologik yo‘nalish talabalari yozma matnlari bazasida o‘quv korpusini shakllantirish metodikasini ishlab chiqish hamda korpus tahlil dasturlari (masalan, AntConc) vositasida talabalar akademik nutqining chastotali lug‘atini loyihalashdan iborat. Ushbu maqsadga erishish uchun quyidagi vazifalar belgilanadi:

a) Nofilologik talabalar yozma matnlarni yig‘ish (sampling criteria) va ularni optik belgilarni aniqlash (OCR) texnologiyasi vositasida elektron plain-text (.txt) formatiga o‘tkazish;

b) Korpus dasturiy vositalaridan foydalanib chastotali so‘zliklar (Frequency lists) va kalit so‘zlar (Keywords) ro‘yxatini shakllantirish;

d) Shakllantirilgan chastotali so‘zliklar bazasida nofilologik yo‘nalishlar uchun mo‘ljallangan o‘quv chastotali lug‘ati modelini yaratish.

METODOLOGIYA

O‘quv korpusini yaratish va uning negizida chastotali lug‘at tuzish tadqiqoti uchta asosiy bosqichdan iborat bo‘lib, ular qat’iy mezonlar va raqamli ishlov berish standartlariga tayanadi.

Tadqiqotning empirik bazasini shakllantirishda nofilologik yo‘nalish (kimyo, fizika, biologiya, matematika, geografiya va b.) talabalarining akademik yozma ishlari (esselar) tanlab olindi. Matnlarni yig‘ishda namunaviylik (representativeness) va muvozanatlashganlik (balance) tamoyillariga rioya qilindi.

1-jadval

O‘quv korpusining empirik bazasi parametrlari

Ko‘rsatkich / Parametr	**Miqdoriy qiymati**	**Metodologik tavsifi**
Qamrab olingan fakultetlar	10 ta	Nofilologik yo‘nalishlar balansi
Matnlar soni (Insholar)	1 300 ta	Ikki bosqichda yig‘ilgan yozma ishlar
Umumiy so‘z qo‘llanishi (Tokens)	~185 000	Korpusning umumiy hajmi
Kam ishlatiladigan so‘z shakllari (Types)	~14 200	Leksik xilma-xillik ko‘rsatkichi
Hujjat formati	Hand-written PDF / Scanned	Qo‘lda yozilgan asl nusxalar
Matn tili va alifbosi	O‘zbek tili (Lotin alifbosi)	Standard UTF-8 kodirovkasi

Qo‘lda yozilgan va skaner qilingan matnlarni kompyuterda qayta ishlash formatiga o‘tkazish quyidagi texnologik zanjir asosida amalga oshirildi:

• **Optik belgilarni aniqlash (OCR):** Skanerlangan PDF va tasvir ko‘rinishidagi matnlar Google Drive OCR va ixtisoslashtirilgan matnni tanish dasturlari vositasida matnli formatga o‘tkazildi. Shakllantirilgan .txt fayllar bazasi korpus lingvistikasining standart dasturiy vositasi bo‘lgan AntConc (Version 4.2+) platformasiga yuklandi va tahlil parametrlari sozlandi.

• **Post-redakturlash (Manual Verification):** OCR jarayonidagi texnik xatoliklar (harf va tishlarning noto‘g‘ri aniqlanishi) qo‘lda tekshirilib, asl nusxaga moslashtirildi.

• **Formatlashtirish va UTF-8 standartiga o‘tkazish:** Barcha matnlar tinish belgilari va ortografik xatolardan tozalanib, korpus dasturlari o‘qiy oladigan .txt (Plain Text) formatida hamda UTF-8 kodirovkasida saqlandi.

Chastotali lug‘at tuzishda ikki xil chastota ko‘rsatkichidan foydalanildi:

1. Absolyut chastota (f): So‘zning korpusda necha marta takrorlanganligining aniq soni.

2. Nisbiy (normallashtirilgan) chastota (N_f): Har xil hajmdagi matnlarni o‘zaro taqqoslash uchun har 10 000 so‘zga to‘g‘ri keladigan chastota qiymati quyidagi formula bo‘yicha hisoblandi:

$$N_f = (f / N) \times 10\,000$$

(Bu yerda: f — so‘zning absolyut chastotasi, N — korpusdagi umumiy so‘zlar soni / tokens)

Tokenizatsiya jarayonida Word Characters sozlamalarida o‘zbek alifbosiga xos belgilardan tutuq belgisi (') qat’iy biriktirildi. Yordamchi so‘zlar filtri orqali va, ham, uchun, bilan, chunki kabi grammatik ko‘makchi va bog‘lovchilar alohida Stop-word list orqali ajratilib, chastotali so‘zlikning asosiy leksik va sohaviy qismidan guruhlandi.

Korpusning leksik boyligini aniqlash uchun Type-Token Ratio (TTR) ko‘rsatkichi quyidagi formula bo‘yicha hisoblandi:

$$TTR = (V / N) \times 100\ \%$$

(Bu yerda: V — Types / noyob so‘zlar soni, N — Tokens / umumiy so‘zlar soni)

NATIJALAR VA TAHLIL

Nofilologik yo‘nalish talabalarining 1300 ta akademik insholari asosida shakllantirilgan o‘quv korpusini AntConc dasturida qayta ishlash natijasida olingan miqdoriy va sifat ko‘rsatkichlari tahlil qilindi.Statistik tahlil natijasida korpusning umumiy leksik hajmi va xilma-xillik indeksi belgilandi.

2-jadval

Korpusning umumiy statistik ko‘rsatkichlari va TTR tahlili

Ko‘rsatkich nomi	Miqdoriy qiymat	Metodologik tahlil va izoh
Umumiy so‘z qo‘llanilishi (N / Tokens)	185,420	Korpusning umumiy leksik hajmi
Noyob so‘z shakllari (V / Types)	14,230	Lug‘at tarkibining takrorlanmas birliklari
Leksik xilma-xillik (TTR)	7.67%	TTR = (14,230 / 185,420) × 100
O‘rtacha insho hajmi	142.6 so‘z	Bir talabaning o‘rtacha yozuv hajmi
Yordamchi va funksional so‘zlar	22.4%	Bog‘lovchi, ko‘makchi va yuklamalar ulushi

Grammatik va yordamchi so‘zlar (stop-words) filtrlanganidan so‘ng, talabalar yozma nutqida eng yuqori chastota ko‘rsatgan mustaqil so‘z turkumlari (ot, sifat, fe’l) saralab olindi.

3-jadval

Eng yuqori chastotali 10 ta mustaqil so‘z birligi (Top-10)

Rank	So‘z shakli (Lemma)	So‘z turkumi	Absolyut chastota (f)	Nisbiy chastota (N_f / 10,000)
1	Ta'lim	Ot	1,420	76.58
2	Rivojlanish	Ot / Fe’l	1,185	63.90
3	Muhim	Sifat	982	52.96
4	Tizim	Ot	874	47.13
5	Jarayon	Ot	810	43.68
6	O‘rganish	Fe’l	765	41.25
7	Zamonaviy	Sifat	720	38.83
8	Natija	Ot	685	36.94
9	Jamiyat	Ot	640	34.51
10	Vosita	Ot	590	31.82

Tahlil: Nofilologik yo‘nalish talabalari nutqida ta’lim, rivojlanish, tizim, jarayon kabi umumakademik leksika ustuvorlik qiladi. Bu holat talabalarning mutaxassislik matnlaridan tashqari umumiy akademik va intellektual kontekstdagi so‘z birikmalarini faol qo‘llashini ko‘rsatadi.

Fakultetlar bo‘yicha o‘tkazilgan kalit so‘zlar (Keywords / Keyness) tahlili quyidagi leksik qatlamlarni ajratib ko‘rsatdi:

1. Umumakademik leksika (General Academic Vocabulary): Barcha 10 ta fakultet matnlarida bir xil chastotada va yuqori tarqalish ko‘lamida (dispersion) uchraydi (tahlil, tadqiqot, samaradorlik, obyekt, metod).

2. Tor sohaviy leksika (Field-specific Vocabulary): Aniq va tabiiy fanlar (kimyo, fizika, biologiya) talabalari nutqida sohaviy atamalar absolyut chastotaning 12–15% ini tashkil etadi (reaktsiya, molekula, energiya, organizm, tenglama).

3. Leksik takroriylik (Overuse phenomenon): Talabalar yozma nutqida ba’zi abstrakt sifat va ravishlar (juda, juda ham, muhim, katta, zo‘r) me’yordan ortiq qo‘llanilgani va bu holat akademik stilistik cheklanishni keltirib chiqarayotgani aniqlandi.

O‘quv korpusi ma’lumotlari negizida loyihalashtirilgan va chop etishga mo‘ljallangan chastotali lug‘at maqolasi quyidagi strukturada shakllantiriladi:

[LEMMA] (So‘z turkumi) | Absolyut chastota: f | Nisbiy chastota: Nf
├─ Grammatik ko‘rsatkichlar va birikuvchanlik (Collocations)
├─ Kontekstual misollar (Korpusdan olingan autentik gaplar)
└─ Leksik xatolik statistikasi (agar mavjud bo‘lsa)

4-jadval

Lug‘at maqolasi tarkibi (Namuna)

Komponent	**Lug‘at maqolasi tarkibi (Namuna)**
Bosh so‘z (Lemma)	JARAYON (ot)
Chastota ko‘rsatkichi	f = 810 | N_f = 43.6 (10,000 so‘zga)
Ustuvor kollokatsiyalar	O‘quv jarayoni (f=312), uzluksiz jarayon (f=115), kimyoviy jarayon (f=84)
Korpusdan misol (Concordance)	“Zamonaviy texnologiyalarni o‘quv jarayoniga tatbiq etish muhim natija beradi.”

XULOSA VA AMALIY TAVSIYALAR

Nofilologik yo‘nalish talabalarining 1 300 ta yozma ishi asosida shakllantirilgan o‘quv korpusi va uning negizida olingan chastotali so‘zliklar quyidagi ilmiy-amaliy xulosalarga kelish imkonini beradi:

1. Empirik bazaning yaratilganligi: O‘zbek tilshunosligida birinchi marta nofilologik yo‘nalish talabalarining ~185 000 so‘z qo‘llanishidan (tokens) iborat raqamlashgan va OCR vositasida tayyorlangan o‘quv korpusi shakllantirildi.

2. Leksik tarkib xususiyatlari: Tahlillar talabar akademik yozma nutqida umumakademik leksika (ta’lim, rivojlanish, tizim, jarayon) asosiy o‘rinni egallashini, tor sohaviy atamalar esa umumiy lug‘at tarkibining 12–15% ini tashkil etishini ko‘rsatdi.

3. Metodologik samaradorlik: AntConc dasturiy platformasi va nisbiy chastota (N_f) formulasini qo‘llash talabalar yozma nutqidagi faol va passiv leksikani xolis baholash hamda leksik takroriylik (overuse) holatlarini aniqlashda yuqori aniqlikni ta’minladi.

Tadqiqot natijalaridan kelib chiqqan holda, ta’lim va leksikografiya sohasi mutaxassislari uchun quyidagi amaliy tavsiyalar ilgari suriladi:

• Nofilologik yo‘nalishlar uchun “O‘zbek tili” va “Akademik yozuv” fan dasturlariga korpus tahlilida aniqlangan Top-500 umumakademik so‘zlikni majburiy leksik minimum sifatida kiritish.

• Nofilologik talabalar uchun o‘quv-chastotali va kollokatsion (so‘z birikuvchanligi) lug‘atlarini aynan real o‘quv korpusi va autentik misollar bazasida nashr etish.

• Korpusni kengaytirish, lingvistlar o‘quv korpusini boshqa oliy ta’lim muassasalari va fan yo‘nalishlari matnlari bilan to‘ldirib borish hamda avtomatlashtirilgan leksik-grammatik razmetka (POS-tagging) tizimini joriy etish.

FOYDALANILGAN ADABIYOTLAR

1. Abdurahmonova, N. (2021). O‘zbek tili kompyuter lingvistikasi. Toshkent: O‘zbekiston milliy universiteti nashriyoti.
2. Götz, S., & Granger, S. (2024). Learner corpus research for pedagogical purposes. International Journal of Learner Corpus Research, 10(1), 1–25.
3. Granger, S. (2015). Contrastive Interlanguage Analysis sheet. In S. Granger, G. Gilquin, & F. Meunier (Eds.), The Cambridge Handbook of Learner Corpus Research (pp. 45–64). Cambridge University Press.
4. Hunston, S. (2002). Corpora in Applied Linguistics. Cambridge: Cambridge University Press.
5. McEnery, T., & Hardie, A. (2012). Corpus Linguistics: Method, Theory and Application. Cambridge: Cambridge University Press.
6. Sinclair, J. (1991). Corpus, Concordance, Collocation. Oxford: Oxford University Press.
7. Vaswani, A., Shazeer, N., Parmar, J., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30).
8. Mengliev, D. B., Abdurakhmonova, N. Z., Shirinova, R. K., Saparova, M. F., Azimov, I. M., & Ibragimov, B. B. (2024). Automated Detection of Allusions in Uzbek Language: A Computational Approach. In *2024 IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE)* (pp. 1560-1564). IEEE.
9. Abdurakhmonova, N., Ismoilov, A. S., & Xamroev, N. S. O‘g‘li. (2022). Morphological analyzer (morfoAnalyse) Python package for Turkic language. *Science and Education*, 3(9), 146-156.
10. Granger, S. (2003a). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, 20(3), 465-480.

"O'ZBEK VA TURK TILLARIDA AVTOMATIK MORFOLOGIK TAHLIL VA LEMMATIZATSIYA: MASHINA TARJIMASI UCHUN ZAMONAVIY YONDASHUVLAR"

Maxkamova Mohinur Mirodil qizi
O'zbekiston Milliy universiteti,
Jurnalistika va o'zbek filologiyasi fakulteti,
Kompyuter lingvistikasi mutaxassisligi II bosqich magistranti
https://doi.org/10.5281/zenodo.23253273

Abstract. *This article analyzes contemporary approaches to automatic morphological analysis and lemmatization in the Uzbek and Turkish languages, which have an agglutinative structure. It also highlights major linguistic challenges characteristic of these languages, including the large number of word forms (vocabulary explosion), homography, and phonetic changes, as well as their impact on Natural Language Processing (NLP) systems. The study comparatively examines the advantages and disadvantages of rule-based, statistical, and neural network-based approaches. In addition, existing morphological resources and corpora for both languages are analyzed, and recommendations are proposed for the use of highly effective hybrid models in practice and for the development of open morphological test datasets for the Uzbek language.*

Keywords: *Natural Language Processing (NLP), morphological analysis, lemmatization, agglutinative languages, Finite-State Transducers (FST), subword tokenization, Uzbek language, Turkish language.*

Annotatsiya. *Ushbu maqolada agglyutinativ tuzilishga ega bo'lgan o'zbek va turk tillarida avtomatik morfologik tahlil hamda lemmatizatsiya jarayonlaridagi zamonaviy yondashuvlar tahlil qilindi. Shuningdek, bu tillarga xos bo'lgan so'z shakllarining ko'pligi (vocabulary explosion), omoformiya va fonetik o'zgarishlar kabi asosiy lingvistik muammolar hamda ularning tabiiy tilni qayta ishlash (NLP) tizimlariga ta'siri yoritilgan. Tadqiqot ishida qoidaga asoslangan, statistik va neyron tarmoqlariga asoslangan yondashuvlarning afzallik va kamchiliklari qiyosiy tahlil qilinib, ularning natijalari o'rganildi. Shuningdek, ikkala til uchun mavjud ilmiy yondashuvlar va morfologik tahlil usullari ko'rib chiqilib, amaliyotda eng yuqori samara beruvchi gibrid modellardan foydalanish hamda o'zbek tili uchun ochiq morfologik test bazalarini rivojlantirish bo'yicha takliflar ilgari surilgan.*

Kalit so'zlar: *tabiiy tilni qayta ishlash (NLP), morfologik tahlil, lemmatizatsiya, agglutinativ tillar, chekli holatli avtomatlar (FST), subword tokenizatsiya, o'zbek tili, turk tili.*

Аннотация. *В данной статье анализируются современные подходы к автоматическому морфологическому анализу и лемматизации узбекского и турецкого языков, обладающих агглютинативной структурой. Также освещаются основные лингвистические проблемы, характерные для данных языков, включая большое количество словоформ (vocabulary explosion), омографию и фонетические изменения, а также их влияние на системы обработки естественного языка (NLP). В исследовании проводится сравнительный анализ преимуществ и недостатков подходов, основанных на правилах, статистических методах и нейронных сетях. Кроме того, рассматриваются существующие морфологические ресурсы и корпусы обоих языков, а также*

предлагаются рекомендации по использованию эффективных гибридных моделей на практике и развитию открытых морфологических тестовых баз для узбекского языка.

Ключевые слова: *обработка естественного языка (NLP), морфологический анализ, лемматизация, агглютинативные языки, трансдьюсеры с конечным числом состояний (FST), субсловная токенизация, узбекский язык, турецкий язык.*

Tillarning morfologik tuzilishi tabiiy tilni qayta ishlash (NLP) va mashina tarjimasi sohasida modelning samaradorligiga bevosita ta'sir ko'rsatuvchi omillardan biri hisoblanadi. [Abdurahmanova, M., & Rayimjonova, M., 2021]. So'zlarning morfemik tuzilishi, morfemalarning soni va segmentatsiyalanish darajasi bir-biridan sezilarli farq qiladi. Morfologik tipologiyasiga ko'ra tillar izolyatsiyalashgan, fuzional va agglyutinativ xususiyatlarni turli darajada namoyon qilishi mumkin. [Jurafsky D., Martin J. H., 2021]. Masalan, ingliz tili kabi analitik tillarda zamon va sondagi cheklangan turlanishni kuzatishimiz mumkin. O'zbek va turk tillari kabi agglyutinativ morfologik tuzilishga ega bo'lgan tillarda esa grammatik va lug'aviy ma'nolarning katta qismi o'zakka ketma-ket affikslarni qo'shish orqali ifodalanadi. Ushbu tillardagi so'zlarning morfemik tuzilishi va qo'shimchalar tartibi ko'p jihatdan o'xshash (masalan, kitob-lar-imiz — kitap-lar-ımız). Lekin turk tilida unli uyg'unligi (vowel harmony) affikslarning alomorfik variantlarini belgilovchi asosiy omillardan biri hisoblanadi. O'zbek tilida esa unli uyg'unligining kuchsizlanishi bilan birga fonetik almashinishlar (q → g', k → g) va affikslarning grafik variantlari avtomatik tahlil jarayonida alohida hisobga olinishi zarur. Shu sababli, ikki tilning umumiy agglyutinativ tabiatiga qaramay, bu xususiyatlar avtomatik morfologik tahlil va lemmatizatsiya, tokenizatsiya va mashina tarjimasi kabi NLP vazifalarini murakkablashtiradi.

P. Koehn o'zining "Statistical Machine Translation" asarida ta'kidlaganidek, morfologik jihatdan boy tillarda har bir lemma ko'plab so'z shakllarida namoyon bo'lishi mumkin. So'zga asoslangan statistik va neyron tizimlarda ushbu shakllarning har biri alohida birlik sifatida qayta ishlanadi, natijada lug'at hajmining kengayishiga (vocabulary explosion), ma'lumotlar siyrakligi (sparse data) va lug'atda mavjud bo'lmagan so'zlar (out-of-vocabulary–OOV) muammosi yuzaga kelishiga sabab bo'ladi. [Koehn P., 2010: 289]. Ya'ni model bir lemmaga tegishli ayrim shakllarni o'rgangan bo'lsa-da, korpusda uchramagan boshqa shakllarni to'g'ri tahlil qilish yoki qayta ishlashda qiyinchilikka duch kelishi mumkin.

Avtomatik morfologik tahlil va lemmatizatsiya jarayoni morfologik boy tillardagi bu muammolarni hal qilishda muhim ahamiyatga ega. Morfologik tahlil so'zning ichki tuzilishini aniqlab, uni o'zak va affikslarga ajratish va grammatik xususiyatlarini belgilash jarayonidir. So'zning turli grammatik shakllarini uning asosiy lug'aviy shakliga, ya'ni lemmaga keltirishga esa lemmatizatsiya deyiladi.

Ushbu vazifalarni bajarishda keng qo'llanuvchi yondashuvlardan biri qoidaga asoslangan chekli avtomat mashinasi (Finite-State-Transducers–FST) hisoblanadi. "Korpus lingvistikasi" darsligida N. Abduraxmonova qoidaga asoslangan chekli avtomat mashinasi (Finite-State-Transducers–FST) texnologiyasining o'zbek tili uchun samarali va aniq usul ekanini ta'kidlagan. O'zbek tili agglyutinativ tabiatli tillardan biri bo'lganligi sababli bitta fe'lning masalan, ishlamoq fe'lining 100,000dan ortiq morfotaktik shakli va kombinatsiyalari borligini keltiradi. Yuqorida ta'kidlanganidek, o'zbek tilida o'zakka qo'shimcha qo'shilganda, o'zakda tovush tushishi va ortishi, almashishi kabi hodisalar bo'ladi (burun+im–burnim, yurak+im–yuragim, bu+ga–bunga). Bu kabi fonetik o'zgarishlar avtomatik morfologik tahlilda alohida algoritmlar va FST

qoidalarini talab qiladi. FST aniq va qat'iy qoidalar asosida ishlaydi. Biroq tabiiy tilning barcha xususiyatlarini qoidalarga sig'dirish qiyin. Ya'ni ma'no va shakl noaniqliklari yuzaga kelishi mumkin. Masalan, -lar affiksining ko'plik, hurmat, kinoya va taxmin kabi ma'nolari mavjud. Darslikda bu kabi muammolarni bartaraf etishda Yashirin Markov modeli (HMM) kabi statistik va ehtimollik usullarini FST bilan birgalikda qo'llash o'zbek tili uchun eng yaxshi yechimlardan biri sifatida beriladi. [Abduraxmonova N., 2023: 159]. FST modelining o'zbek va turk tillarining morfologik tabiatiga mos bo'lishiga qaramay, bir qancha kamchiliklari ham mavjud. Jumladan, FSTning asosiy kamchililaridan biri uning qo'lda tuzilgan qat'iy qoidalar asosida ishlashidir. Model lug'atda uchramaydigan so'zlarni tahlil qila olmaydi va kontekstni anglay olmagani sababli morfologik ko'p ma'nolikni mustaqil hal qila olmaydi hamda juda katta lingvistik resursni talab etadi. Shu sababli zamonaviy NLPda FSTning yuqori grammatik aniqligi statistik va neyron tarmoqlarning kontekstual imkoniyatlari bilan birlashtirilib, gibrid modellar yaratilmoqda.

Agglyutinativ tillarga avtomatik ishlov berish jarayonida so'z shakllarini yaxlit leksik birlik sifatida emas, balki o'zak va affikslardan iborat kichikroq komponentlarga ajratish eng samarali yondashuvlardan biridir. [Sulevmanov, D., Gatiatullin, A., Prokopyev, N., & Abdurakhmonova, N., 2020]. So'zga asoslangan an'anaviy statistik va neyron modellar agglyutinativ tillardagi murakkab so'zshakllarni alohida token sifatida qabul qiladi. Natijada lug'at hajmi keskin ortib ketadi. K. Oflazer va I. Durgar El-Kahlout turkiy tillar uchun morfemik segmentatsiyalash bo'yicha tadqiqotlar olib borganlar. Analitik tillardagi predlog va yordamchi fe'llar bajaradigan grammatik va sintaktik vazifani turkiy tillarda affikslar bajaradi. Shu sababli morfologik segmentatsiyalash orqali har bir morfemaning alohida token sifatida qayta ishlanishi modelning lingvistik aniqligini oshiradi. O'zbek va turk tillaridagi morfemalar zanjirida tipologik jihatdan umumiylik yuqori bo'lganligi sababli segmentatsiyani parallel amalga oshirish mumkin.

Morfologik qoidalar va FST lug'atlarini qo'lda tuzish katta resurs talab qilgani sababli zamonaviy neyron tarmoqlarida statistik va ehtimollik algoritmlariga asoslangan subword tokenizatsiya usullari (Byte-Pair Encoding–BPE, WordPiece, Unigram Language Model) keng qo'llanilmoqda. R. Sennrich tomonidan taklif etilgan Byte-Pair Encoding (BPE) algoritmining afzallik tomoni shundaki, u lingvistik qoidalarsiz, sof korpus chastotasiga tayangan holda ishlaydi. [Sennrich R., 2016: 1715]. Ushbu usul OOV (out-of-vocabulary) muammosini sezilarli darajada kamaytirgan bo'lsa-da, agglyutinativ tillar uchun ba'zi cheklovlarga ega. Chunki BPE algoritmi har doim ham lingvistik jihatdan to'g'ri bo'lgan morfemik chegaralarni saqlay olmaydi. Masalan, "kitoblarimizdan" so'zini BPE lingvistik o'zak va affikslarga emas, balki quruq chastotaga ko'ra kitob + larim + izdan tarzida segmentatsiyalashi mumkin. Shu sababli, morfologik yo'naltirilgan (morphology-informed) subword tokenizatsiya algoritmlarini yaratish o'zbek va turk NLP tizimlari uchun muhim hisoblanadi.

Shuningdek, P. Koehn o'zining tadqiqotlarida morfologik jihatdan murakkab va boy tillar uchun so'zlarni bir nechta lingvistik omillar (factors) asosida ifodalashni taklif qiladi. Factored modellar yondashuvida so'z yagona birlik sifatida emas, balki uning lemma yoki asosi, morfologik kategoriyalari, grammatik xususiyatlari va boshqa lingvistik belgilari orqali tavsiflanadi. Bu kabi modellar ham lug'at hajmining ortib ketishi muammosini kamaytirishga yordam beradi. Factored modellar ushbu muammoni so'zning ichki lingvistik tuzilishini hisobga olish orqali kamaytirishga imkon beradi. Masalan, so'z lemma va morfologik belgilar

kombinatsiyasi sifatida ifodalanishi mumkin. Bunda model so‘zning lug‘aviy ma’nosi va uning grammatik

xususiyatlarini alohida hamda o‘zaro bog‘liq holda qayta ishlaydi. Bu esa ayniqsa kam uchraydigan yoki ilgari uchramagan so‘z shakllarini tahlil qilishda muhim afzallik beradi. [Koehn P., 2010: 299]. O‘zbek va turk tillarining tipologik o‘xshashligi, ya’ni ikki tilda ham grammatik munosabatlarning katta qismi affikslar bilan ifodalanadi va morfologik kategoriyalar o‘rtasida ham strukturaviy o‘xshashliklar mavjud. Shuning uchun lemma, so‘z turkumi, grammatik shaxs, son, kelishik, zamon va boshqa morfologik belgilarni modelga qo‘shimcha omil sifatida kiritish tillararo mosliklarni aniqroq ifodalash imkonini beradi. Biroq factored modellar yuqori sifatli morfologik tahlil vositalarini talab etadi. Morfologik teglarning noto‘to‘g‘ri qo‘llanishi modelning keyingi bosqichlariga salbiy ta’sir ko‘rsatadi. Shu jihatdan FST asosidagi morfologik tahlil, lemmatizatsiya, morfemik segmentatsiya va zamonaviy subword tokenizatsiyasi factored modellar bilan birgalikda qo‘llanilganda o‘zbek va turk tillari uchun yanada barqaror va lingvistik jihatdan asoslangan tizim yaratish mumkin.

Mazkur maqolada morfologik jihatdan boy tillar, xususan, o‘zbek va turk tillarini mashina tarjimasi jarayonida qayta ishlashning asosiy muammolari va yondashuvlari umumlashtirildi. Tahlil shuni ko‘rsatdiki, ushbu tillardagi so‘z shakllarining ko‘pligi va murakkab morfologik tuzilishi an’anaviy so‘zga asoslangan modellar uchun ma’lumotlar siyrakligi hamda ilgari uchramagan so‘zlar muammosini yuzaga keltiradi. Shu sababli morfologik segmentatsiya, lemmatizatsiya, subword birliklar va factored modellar kabi yondashuvlardan foydalanish muhim ahamiyat kasb etadi.[Zaynobiddin qizi, N. A., & Asliddin qizi, S. G. (2024, November)., 2024]. Ko‘rib chiqilgan usullar bir-birini to‘ldirib, morfologik axborotni saqlash va modelning yangi so‘z shakllarini umumlashtirish imkoniyatini oshiradi. O‘zbek va turk tillarining agglutinativ hamda tipologik jihatdan yaqinligi ushbu xususiyatlarni mashina tarjimasi tizimlarida samarali hisobga olish imkonini beradi. Umuman, lingvistik bilimlarni statistik va neyron yondashuvlar bilan uyg‘unlashtirish mazkur tillar uchun yanada samarali mashina tarjimasi tizimlarini yaratishning istiqbolli yo‘nalishi hisoblanadi.

FOYDALANILGAN ADABIYOTLAR:

1. Koehn, P. Statistical Machine Translation. Cambridge: Cambridge University Press, 2010.
2. Koehn, P. Neural Machine Translation. Cambridge: Cambridge University Press, 2020.
3. Abduraxmonova, N. Korpus lingvistikasi. Toshkent 2023. – 159 b.
4. Jurafsky, D., Martin, J. H. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.
5. Sennrich, R., Haddow, B., Birch, A. Neural Machine Translation of Rare Words with Subword Units // Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Berlin: Association for Computational Linguistics, 2016. – P. 1715–1725.
6. Abdurakhmonova, N., Alisher, I., & Sayfulleyeva, R. (2022, September). MorphUz: Morphological analyzer for the Uzbek language. In *2022 7th International Conference on Computer Science and Engineering (UBMK)* (pp. 61-66). IEEE.
7. Mengliev, D., Abdurakhmonova, N., Zokirova, H., Ibragimov, B., Jurakulova, M., & Abdunazarova, M. (2025, October). From morphological rules to neural networks: A hybrid framework for medical entity extraction in Karakalpak. In *AIP Conference Proceedings* (Vol. 3377, No. 1, p. 070004). AIP Publishing LLC.

8. Sulevmanov, D., Gatiatullin, A., Prokopyev, N., & Abdurakhmonova, N. (2020, November). Turkic morpheme web portal as a platform for turkology research. In *2020 International Conference on Information Science and Communications Technologies (ICISCT)* (pp. 1-5). IEEE.
9. Zaynobiddin qizi, N. A., & Asliddin qizi, S. G. (2024, November). Theoretical Foundations of Corpus-based Uzbek-English Machine Translation. In *2024 IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE)* (pp. 1650-1653). IEEE.
10. Abdurahmanova, M., & Rayimjonova, M. (2021). KORPUS LINGVISTIKASIDA TARJIMASHUNOSLIK MASALASI. *COMPUTER LINGUISTICS: PROBLEMS, SOLUTIONS, PROSPECTS*, *1*(1).

CAT-ВОСИТАЛАР ВА КАТТА ТИЛ МОДЕЛЛАРИ ЁРДАМИДА ТАРЖИМА ТРАНСФОРМАЦИЯЛАРИНИ АВТОМАТИК АНИҚЛАШ: КОМИССАРОВ ТАСНИФИ МИСОЛИДА

Мирмахмудова Ирода Ирисматовна
Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университети, мустақил тадқиқотчиси
https://doi.org/10.5281/zenodo.23253286

Аннотация. *Мақолада В.Н. Комиссаров ва Л.С. Бархударовнинг классик таржима трансформациялари таснифи таржимашунослик нуқтаи назаридан батафсил таҳлил қилиниб, ҳар бир кичик тоифа (транскрипция, калькалаш, конкретлаштириш, генерализация, модуляция, грамматик алмаштириш, антонимик таржима, экспликация, компенсация ва бошқалар) инглиз-ўзбек тил жуфтлиги мисолларида кўрсатилади. Шундан сўнг ушбу лингвистик таснифнинг CAT-воситалар ва катта тил моделлари (LLM) ёрдамида автоматик аниқланиш имконияти корпус-лингвистик тадқиқотлар мисолида баҳоланади. Мақола таржима трансформацияси назарияси билан ҳисоблаш тилшунослиги ўртасидаги методологик боғлиқликни кўрсатишга қаратилган.*

Калит сўзлар: *таржима трансформациялари, Комиссаров таснифи, Бархударов таснифи, таржимашунослик, CAT-воситалар, катта тил моделлари, корпус лингвистикаси.*

Аннотация. *В статье с позиций переводоведения подробно анализируется классическая классификация переводческих трансформаций В.Н. Комиссарова и Л.С. Бархударова: каждая подкатегория (транскрипция, калькирование, конкретизация, генерализация, модуляция, грамматические замены, антонимический перевод, экспликация, компенсация и др.) иллюстрируется примерами на материале англо-узбекской языковой пары. Далее оценивается возможность автоматического выявления данной лингвистической классификации с помощью CAT-инструментов и больших языковых моделей (LLM) на основе корпусно-лингвистических исследований. Статья призвана показать методологическую связь между теорией переводческих трансформаций и компьютерной лингвистикой.*

Ключевые слова: *переводческие трансформации, классификация Комиссарова, классификация Бархударова, переводоведение, CAT-инструменты, большие языковые модели, корпусная лингвистика.*

Abstract. *From a translation-studies perspective, this article provides a detailed analysis of V.N. Komissarov's and L.S. Barkhudarov's classic classification of translation transformations, illustrating each subcategory (transcription, calque, concretization, generalization, modulation, grammatical substitution, antonymic translation, explication, compensation, and others) with English-Uzbek examples. It then evaluates the extent to which this linguistic taxonomy can be automatically detected through CAT tools and large language models (LLMs), drawing on corpus-linguistic research. The article aims to demonstrate the methodological link between translation-transformation theory and computational linguistics.*

Keywords: *translation transformations, Komissarov's classification, Barkhudarov's classification, translation studies, CAT tools, large language models, corpus linguistics.*

КИРИШ

Таржима трансформациялари тушунчаси матннинг манба тилдан мақсад тилга ўтиши жараёнида юзага келадиган тизимли ўзгаришларни тавсифлайди. Бу тушунча В.Н. Комиссаров, Л.С. Бархударов каби рус тилшунослари, шунингдек Ж.-П. Вине ва Ж. Дарбельне каби ғарб олимлари томонидан назарий жиҳатдан асосланган [1; 2; 3]. Сунъий интеллект технологияларининг - хусусан катта тил моделлари (LLM) ва замонавий CAT-воситаларининг - жадал ривожланиши эса табиий равишда янги саволни келтириб чиқаради: мазкур классик лингвистик таснифларда ажратилган трансформация турларини ҳисоблаш воситалари ёрдамида автоматик аниқлаш мумкинми? Тадқиқотнинг мақсади - Комиссаров ва Бархударов таснифларини таржимашунослик нуқтаи назаридан чуқур таҳлил қилиш ва бу таснифни CAT-воситалар ҳамда LLM орқали аниқлаш имкониятларини баҳолашдан иборат.

МАВЗУ БЎЙИЧА ИЛМИЙ АДАБИЁТЛАР ТАҲЛИЛИ

Таржима трансформациялари назарияси борасида олимлар ўртасида умумий келишув мавжуд: барча тадқиқотчилар тилларнинг тизимли фарқлари таржимонни муайян ўзгартиришлар қилишга мажбур этишини тан олади. Бироқ аниқ таснифлашда ягона андоза йўқ. Бархударов барча трансформацияларни тўрт содда амалга - кўчириш, алмаштириш, қўшиш ва тушириб қолдиришга - жамлайди [2]. Комиссаров эса уч поғонали (лексик - грамматик - лексик-грамматик) таснифни таклиф этиб, ҳар бир поғонани яна бир қатор кичик усулларга ажратади [1]. Вине ва Дарбельне тўғридан-тўғри ва билвосита таржима усулларига асосланган етти турдаги моделни таклиф қилган, кейинчалик Молина ва Уртадо Альбир бу моделни ўн саккиз хил техникага кенгайтирган [4]. Ўзбекистонлик тадқиқотчилар Нормуратова ва Курбанова эса Комиссаров таснифининг инглиз-ўзбек бадиий таржима мисолидаги амалий қўлланилишини кўрсатиб, трансформацияларнинг матнда бир-бири билан чамбарчас боғлиқ равишда намоён бўлишини таъкидлайдилар [5]. Демак, назарий адабиётда классификациялар хилма-хил бўлса-да, уларнинг барчаси трансформацияларни тасодифий эмас, балки тилларнинг тизимли фарқидан келиб чиқадиган қонуниятли ҳодиса сифатида тан олади.

Ҳисоблаш тилшунослиги эса бошқа йўналишни ривожлантирган: Baroni M., Bernardini S. ни таржима қилинган матнларни асл матндан ажратиш учун таянч вектор машиналаридан фойдаланиб, функционал сўзлар тарқалиши трансформация изини ошкор қилувчи энг ишончли белги эканини кўрсатган [6]. . Zhai Y., Illouz G., Vilnat A кўп тилли тил моделини "ноаниқ созлаш" орқали сўзма-сўз бўлмаган таржимани 85 фоизгача аниқлик билан ажратишга эришган [7], Amponsah-Kaakyire ва ҳаммуаллифлари эса нейрон таснифловчиларнинг ёлғон корреляцияларга мойиллигини огоҳлантирган [8]. Kocmi T., Federmann C. GPT-асосланган GEMBA-MQM методи орқали таржима хатолари (ва билвосита трансформация зарурати) ўринларини аниқ белгилаш мумкинлигини кўрсатган [11; 12], бу эса Lommel A. ва ҳаммуаллифларининг MQM стандартига асосланади [10]. Демак, лингвистик назария трансформацияларнинг сифат тавсифини яратган, ҳисоблаш тилшунослиги эса уларни статистик аниқлаш усулларини ишлаб чиққан, бироқ бу икки йўналиш ўртасида - яъни Комиссаров тоифаларининг ҳар бирини алоҳида-алоҳида аниқлайдиган махсус модел - ҳали мавжуд эмас. Айнан шу бўшлиқни тўлдириш учун

қуйида ҳар бир Комиссаров тоифаси таржимашунослик нуқтаи назаридан алоҳида таҳлил қилинади.

АСОСИЙ ҚИСМ: КОМИССАРОВ ТАСНИФИНИНГ ТАРЖИМАШУНОСЛИК ТАҲЛИЛИ

1. Лексик трансформациялар

Транскрипция ва транслитерация - сўзнинг товуш ёки ҳарф таркибини мақсад тилда қайта тиклаш усули: "Washington" - "Вашингтон", "impeachment" - "импичмент". Калькалаш - сўз ёки бирикманинг ички тузилишини нусха кўчириш: "White House" - "Оқ уй", "International Monetary Fund" - "Халқаро Валюта Фонди". Конкретлаштириш - кенг маънодаги сўзни торроқ маънодаги сўз билан алмаштириш: инглизча "He is a poor swimmer" жумласидаги умумий "was" феъли ўзбек тилида аниқроқ "яхши сузмайди" шаклида берилади. Генерализация - аксинча, тор маънодаги бирликни кенгроқ тушунча билан алмаштиради: "I saw a man six feet two inches tall" жумласидаги аниқ ўлчов ўзбекча "узун бўйли йигит" тарзида умумлаштирилади, чунки феет-дюйм тизими ўзбек китобхони учун табиий эмас. Модуляция - мантиқий нуқтаи назарни ўзгартириш орқали амалга ошириладиган семантик трансформация: "chased themselves across the grass like swallows" ибораси ўзбек таржимасида "майсаларга соя соларди" тарзида берилиб, ҳаракатнинг субъекти билвосита алмашади.

2. Грамматик трансформациялар

Синтактик ассимиляция (сўзма-сўз таржима) - манба ва мақсад тил тузилиши мос келганда қўлланилади ва алоҳида изоҳ талаб қилмайди. Гапни бўлиш ва бирлаштириш - узун, кўп қаватли инглизча гап ўзбек тилида икки ёки уч содда гапга бўлиниши мумкин (ва аксинча): "The pulse of joy that beats in us at twenty becomes sluggish. Our limbs fail, our senses rot" каби икки инглизча гап ўзбек таржимасида биргаликда, узун бир гап сифатида бирлаштирилиши мумкин. Грамматик алмаштириш - сўз туркуми ёки гап бўлаги ўзгариши: юқоридаги "He is a poor swimmer" мисолида отли бирикма (poor swimmer) феълга (яхши сузмайди) айланади; "A new press conference was held in Washington yesterday" жумласида мажҳул нисбат (was held) ўзбек тилида фаол нисбат (бўлиб ўтди) билан алмаштирилади, чунки ўзбек тилида мажҳул нисбат расмий услубда камроқ қўлланилади.

3. Лексик-грамматик трансформациялар

Антонимик таржима - тасдиқ шаклини инкор шакли орқали (ёки аксинча) етказиш: "I might say I'd never seen her before" жумласи рус/ўзбек таржимасида манфий модаллик сақланган ҳолда "буни биринчи марта кўраётганимни айтишим мумкин" тарзида қайта қурилади. Экспликация (изоҳли таржима) - мақсад тил ўқувчисида етарли фон билим бўлмаган тушунчани батафсил тасвирлаш орқали таржима қилиш: "Society" сўзи контекстда "жамиятнинг юқори қатлами" тарзида кенгайтирилиб таржима қилиниши мумкин. Компенсация - манба матндаги услубий ёки маъновий йўқотишни матннинг бошқа бир жойида қоплаш: масалан, инглизча сўзлашув тилидаги эркин оҳанг ўзбек таржимасида йўқолса, таржимон буни бошқа жумлада қўшимча экспрессив восита (ундов, кучайтирувчи зарралар) орқали тиклаши мумкин. Ушбу уч мисолдан кўринадики, лексик-грамматик тоифа энг мураккаб бўлиб, у соф статистик белгилар орқали эмас, балки чуқур семантик-прагматик тушуниш орқалигина аниқланади.

4. Комиссаров ва Бархударов тоифаларини ҳисоблаш воситалари орқали аниқлаш

Юқорида кўрсатилган мисоллар икки хил аниқланиш даражасига эга. Бархударовнинг тўрт содда тоифаси (кўчириш, алмаштириш, қўшиш, тушириб қолдириш) юзаки статистик белгилар - token сони нисбати, сўз тартиби фарқи, POS-тег тақсимоти - орқали нисбатан осон аниқланади, чунки булар матн узунлиги ва тузилишидаги ҳисобланадиган номувофиқликлар билан боғлиқ. Комиссаровнинг лексик тоифалари (айниқса транскрипция, калькалаш) номлар рўйхати ва тилшунослик луғатлари ёрдамида қоида-асосли (rule-based) алгоритмларда ҳам муваффақиятли аниқланади. Аммо модуляция, антонимик таржима, экспликация ва компенсация каби чуқур семантик тоифалар фақат катта тил моделлари орқали, few-shot prompting стратегиясидан фойдаланилганда аниқланиши мумкин: моделга ҳар бир тоифа таърифи ва 2–3 намуна берилиб, сўнгра янги жумла жуфтлигида қайси тоифа қўлланилгани сўралади [7; 11; 12]. Бироқ Amponsax-Каакьяйре ва ҳаммуаллифлари огоҳлантирганидек, нейрон моделлар кўпинча трансформация турини эмас, балки матн мавзусини "ўрганиб олиши" мумкин [8], шу сабабли модел натижаси ҳамиша таржимашунос-эксперт томонидан тасдиқланиши шарт.

Демак, CAT-воситалар ва LLM таржима трансформацияларини тўлиқ автоматик, инсон иштирокисиз аниқлай олмайди, бироқ улар таржимашунос учун икки муҳим вазифани бажаради: биринчидан, катта ҳажмдаги корпусда трансформация турларини дастлабки (pre-annotation) белгилаб, инсон меҳнатини тезлаштиради; иккинчидан, COMET ва TransQuest каби нейрон баҳолаш тизимлари орқали [9; 13] муайян трансформация тури кўп учраган матнларнинг умумий таржима сифатига таъсирини миқдорий ўлчаш имконини беради - бу эса анъанавий таржимашуносликда деярли имконсиз бўлган янги тадқиқот йўналишидир.

ХУЛОСА

Ушбу тадқиқотда Комиссаров ва Бархударовнинг классик таржима трансформациялари таснифи таржимашунослик нуқтаи назаридан батафсил таҳлил қилиниб, ҳар бир кичик тоифа аниқ инглиз-ўзбек мисоллари орқали кўрсатилди. Таҳлил шуни исботладики, юзаки, ҳисобланадиган трансформациялар (кўчириш, алмаштириш, транскрипция, калькалаш) CAT-воситалар ва оддий корпус-лингвистик белгилар орқали муваффақиятли аниқланади, чуқур семантик трансформациялар (модуляция, антонимик таржима, экспликация, компенсация) эса фақат катта тил моделлари ва инсон эксперти ҳамкорлигида ишончли тарзда аниқланиши мумкин. Келгуси тадқиқотлар доирасида ушбу таҳлил асосида annotatsiya қилинган инглиз-ўзбек пилот корпусини яратиш ва few-shot prompting стратегияларининг самарадорлигини амалий синовдан ўтказиш режалаштирилмоқда.

ФОЙДАЛАНИЛГАН АДАБИЁТЛАР

1. Комиссаров В.Н. Теория перевода (лингвистические аспекты). – М.: Высшая школа, 1990. – 253 с.

2. Бархударов Л.С. Язык и перевод: Вопросы общей и частной теории перевода. – М.: Международные отношения, 1975. – 240 с.

3. Vinay J.-P., Darbelnet J. Stylistique comparée du français et de l'anglais: méthode de traduction. – Paris: Didier, 1958.

4. Molina L., Hurtado Albir A. Translation Techniques Revisited: A Dynamic and Functionalist Approach // Meta. – 2002. – Vol. 47, № 4. – P. 498–512.
5. Normuratova V.I., Kurbanova D.Sh. Transformations in Translation Studies // O'zbekistonda xorijiy tillar. – 2025. – 11-jild, № 3. – B. 253–265.
6. Baroni M., Bernardini S. A New Approach to the Study of Translationese: Machine-learning the Difference between Original and Translated Text // Literary and Linguistic Computing. – 2005. – Vol. 21, № 3. – P. 259–274.
7. Zhai Y., Illouz G., Vilnat A. Detecting Non-literal Translations by Fine-tuning Cross-lingual Pre-trained Language Models // Proceedings of COLING 2020. – Barcelona, 2020. – P. 5944–5956.
8. Amponsah-Kaakyire K., Pylypenko D., van Genabith J., España-Bonet C. Explaining Translationese: why are Neural Classifiers Better and what do they Learn? // Proceedings of the Fifth BlackboxNLP Workshop. – 2022. – P. 281–296.
9. Rei R., Stewart C., Farinha A.C., Lavie A. COMET: A Neural Framework for MT Evaluation // Proceedings of EMNLP 2020. – P. 2685–2702.
10. Lommel A., Burchardt A., Uszkoreit H. Multidimensional Quality Metrics (MQM): A Flexible System for Assessing Translation Quality // Tradumàtica. – 2014. – № 12. – P. 455–463.
11. Kocmi T., Federmann C. Large Language Models Are State-of-the-Art Evaluators of Translation Quality // Proceedings of EAMT 2023. – Tampere, 2023. – P. 193–203.
12. Kocmi T., Federmann C. GEMBA-MQM: Detecting Translation Quality Error Spans with GPT-4 // Proceedings of WMT 2023. – Singapore, 2023. – P. 768–775.
13. Ranasinghe T., Orasan C., Mitkov R. TransQuest: Translation Quality Estimation with Cross-lingual Transformers // Proceedings of COLING 2020. – P. 5070–5081.

RASMIY HUJJATLAR TILIDA QO‘LLANILADIGAN LEKSIK VA GRAMMATIK BIRLIKLARNING O‘ZIGA XOSLIGI

Dildora Mirvaliyeva
O‘zMU Kompyuter lingvistikasi va
amaliy tilshunoslik kafedrasi o‘qituvchisi
https://doi.org/10.5281/zenodo.23253305

Annotatsiya. *Mazkur maqolada rasmiy hujjatlar tilida qo‘llaniladigan leksik va grammatik birliklarning o‘ziga xos xususiyatlari tahlil qilinadi. Rasmiy hujjatlar davlat boshqaruvi, huquqiy munosabatlar va ish yuritish jarayonida muhim kommunikativ vosita sifatida xizmat qilishi sababli, ularda qo‘llaniladigan til birliklarining aniqligi, izchilligi, standartligi va me’yoriyligi alohida ahamiyat kasb etadi. Shuningdek, rasmiy matnlarda fikrning xolis, aniq va mantiqiy ifodalanishini ta’minlovchi leksik-grammatik vositalar tahlil qilinadi.*

Kalit so‘zlar: *rasmiy hujjatlar, rasmiy uslub, rasmiy til, leksik birliklar, grammatik birliklar, ish yuritish tili, terminlar, qoliplashgan birikmalar, grammatik shakllar, til me’yorlari, standartlashuv, hujjat tili, davlat tili.*

Аннотация. *В данной статье анализируются специфические особенности лексических и грамматических единиц, используемых в языке официальных документов. Поскольку официальные документы служат важным коммуникативным средством в процессе государственного управления, правовых отношений и делопроизводства, особое значение имеют точность, последовательность, стандартизация и нормативность используемых в них языковых единиц. Также анализируются лексико-грамматические средства, обеспечивающие объективное, точное и логическое выражение мысли в официальных текстах.*

Ключевые слова: *официальные документы, официальный стиль, официальный язык, лексические единицы, грамматические единицы, язык делопроизводства, термины, шаблонные сочетания, грамматические формы, языковые нормы, стандартизация, язык документов, государственный язык.*

Abstract. *This article analyzes the specific characteristics of lexical and grammatical units used in the language of official documents. Since official documents serve as an important communicative tool in the process of public administration, legal relations, and record-keeping, the accuracy, consistency, standardization, and normality of the language units used in them are of particular importance. It also analyzes the lexico-grammatical means that ensure the objective, clear, and logical expression of thought in official texts.*

Keywords: *official documents, official style, official language, lexical units, grammatical units, clerical language, terms, standardized phrases, grammatical forms, language norms, standardization, document language, state language.*

Til jamiyatdagi ijtimoiy munosabatlarni tartibga solish, axborotni uzatish va turli kommunikativ ehtiyojlarni qondirishning asosiy vositasidir. Jamiyat taraqqiyoti bilan tilning funksional imkoniyatlari ham kengayib, uning turli sohalarda qo‘llanishiga xos bo‘lgan uslubiy tizimlar shakllanadi. Ana shunday funksional ko‘rinishlardan biri rasmiy uslub bo‘lib, u davlat boshqaruvi, huquq, diplomatiya, ma’muriy faoliyat, ish yuritish va hujjatchilik sohalarida keng

qo‘llanadi.

Rasmiy hujjatlar davlat va jamiyat hayotida muayyan huquqiy, tashkiliy yoki ma’muriy vazifani bajaradigan yozma kommunikativ birliklardir. Ariza, bayonnoma, dalolatnoma, ma’lumotnoma, buyruq, farmoyish, qaror, shartnoma, xizmat xati, tavsifnoma kabi hujjatlarning har biri o‘ziga xos kompozitsion va lingvistik tuzilishga ega. Biroq ularning barchasini birlashtiruvchi umumiy jihat — mazmunning aniq, ixcham, xolis va me’yoriy shaklda ifodalanishidir. Rasmiy hujjatlar tilining lingvistik xususiyatlarini o‘rganishda, avvalo, undagi leksik va grammatik birliklarning qo‘llanish qonuniyatlariga e’tibor qaratish zarur. Chunki rasmiy matnning aniqligi va standartligi bevosita unda tanlangan so‘zlar, terminlar, grammatik shakllar va sintaktik konstruksiyalar bilan bog‘liq.

Rasmiy matnda mazmunning turlicha talqin qilinishiga yo‘l qo‘ymaslik muhim hisoblanadi. Shu sababli ko‘p ma’noli, obrazli yoki emotsional bo‘yoqqa ega birliklardan ko‘ra, semantik jihatdan aniq va barqaror birliklarga ustuvorlik beriladi. Masalan, “mazkur”, “ushbu”, “muvofiq”, “asosan”, “belgilangan tartibda”, “amalga oshirish”, “taqdim etish”, “tasdiqlash”, “yuklatilsin”, “ko‘rib chiqilsin” kabi birliklar rasmiy matnlarda keng uchraydi. Ularning muntazam qo‘llanishi hujjat tilining standartlashuviga xizmat qiladi.

Ushbu uslub qat’iy me’yorlar asosida shakllanadi, matnning har bir unsuri mazmunni aniq ifodalash va rasmiy munosabatlarni tartibga solishga yo‘naltiriladi. Rasmiy matnlarning muhim belgilaridan biri andozaviy iboralar va standart sintaktik qoliplarning keng qo‘llanishidir. “Mazkur qaror asosida...”, “O‘zbekiston Respublikasi qonunlariga muvofiq...” kabi ifodalar matnning izchilligini ta’minlaydi va mazmunni bir xil shaklda yetkazishga imkon beradi. Bunday birliklar hujjat matnida yagona uslubni saqlash, fikrni aniq chegaralash hamda turli tomonlar uchun tushunarli shakl yaratishga yordam beradi. Stereotipik iboralar vaqt tejalishiga olib keladi, hujjatning huquqiy aniqligini mustahkamlaydi va tarjima jarayonida noto‘g‘ri talqin ehtimolini kamaytiradi. Rasmiy matnlarda maxsus terminologiya yetakchi o‘rin tutadi. Terminlar matnning sohasi va mazmuniga qarab tanlanadi. Yuridik hujjatlarda “javobgarlik”, “ijro etuvchi organ”, “qonuniy asos” kabi birliklar qo‘llanilsa, iqtisodiy matnlarda “byudjet tasdiqlash”, “moliya rejasi” kabi terminlar uchraydi. Diplomatik yozishmalarda esa “memorandum”, “konvensiya”, “elchi vakolati” kabi atamalar faol ishlatiladi. Terminlarning barqaror ma’noga ega bo‘lishi matnni bir xil talqinda qabul qilishni ta’minlaydi.

Rasmiy matnlar subyektiv baholash, hissiy ifodalar va shaxsiy munosabatni aks ettiruvchi birliklardan xoli bo‘ladi. Masalan, “Mazkur buyruq 2024-yil 1-yanvardan kuchga kiradi” kabi jumla aniq ma’lumot beradi, ortiqcha hissiy ranglardan xoli. Bunday neytrallik hujjat mazmunini xolis shaklda ifodalashga, tomonlar o‘rtasidagi munosabatlarni muvozanatda ushlab turishga imkon beradi.

Rasmiy hujjatlar tilining grammatik tizimida ayrim shakllar boshqa funksional uslublarga qaraganda faolroq qo‘llanadi. Ayniqsa, harakat nomi, majhul nisbat, sifatdosh, ravishdosh va buyruq-istak mayli shakllarining qo‘llanish chastotasi yuqori.

Harakat nomi shakllari. Rasmiy hujjatlarda -ish affiksli harakat nomlari faol qo‘llanadi: *amalga oshirish, ta’minlash, tashkil etish, tasdiqlash, ko‘rib chiqish, nazorat qilish, belgilash, tayinlash.*

Harakat nomining faol qo‘llanishi rasmiy matnda jarayon yoki harakatni shaxsga bog‘langan fe’l shaklidan ko‘ra, umumlashgan va obyektiv tarzda ifodalash imkonini beradi. Masalan: *“Hujjatlarni o‘z vaqtida taqdim etish ta’minlansin.”* Ushbu gapda *“taqdim etish”* shakli muayyan harakatning bajarilishini umumiy talab sifatida ifodalamoqda.

Majhul nisbat shakllari. Rasmiy hujjatlar tilida majhul nisbat ham faol qo‘llanadi: *tasdiqlandi, belgilandi, tayinlandi, ko‘rib chiqildi, ishlab chiqildi, amalga oshirildi.*

Majhul nisbatning muhim funksiyalaridan biri harakat bajaruvchisini ikkinchi darajaga olib, bajarilgan ish yoki amalga oshirilishi lozim bo‘lgan jarayonning o‘ziga e’tiborni qaratishdir. Masalan*: "Yangi nizom tasdiqlandi."*

Bu gapda asosiy kommunikativ urg‘u nizomni kim tasdiqlaganiga emas, balki uning tasdiqlanganligiga qaratilgan.

Buyruq-istak mayli. Buyruq, farmoyish va topshiriq mazmunidagi hujjatlarda buyruq-istak maylining shakllari muhim funksional ahamiyatga ega: *tasdiqlansin, tayinlansin, yuklatilsin, amalga oshirilsin, nazorat qilinsin, bekor qilinsin.*

Masalan: *"Mazkur topshiriqning ijrosi nazoratga olinsin."* Bunday grammatik shakl rasmiy hujjatda topshiriq yoki talabning majburiy xarakterini ifodalashga xizmat qiladi.

Lingvistik xususiyatlar orasida passiv konstruksiyalarning ustunligi alohida ko‘zga tashlanadi: *"Shartnoma tasdiqlandi", "Majburiyat bajarildi"*. Bunday shakllar rasmiylikni kuchaytiradi va diqqatni harakat natijasiga qaratadi. Shuningdek, ko‘plik shakli orqali murojaat qilish keng uchraydi*: "Sizga murojaat qilamiz", "Tomonlar kelishuviga ko‘ra..."*. Bu shakl rasmiy muhitda hurmat va tenglikni ifodalash vositasi sifatida ishlatiladi [Ulashmurodova Sh.,2024:171].

Rasmiy uslub davlat boshqaruvi, huquqiy tizim va xalqaro munosabatlar doirasida asosiy til vositasi sifatida qo‘llanadi. Uning qat’iy tuzilishi, standart birliklari va neytral ifoda vositalari matnning aniqligi hamda tushunarli bo‘lishini ta’minlaydi.

Davlat tilida rasmiy hujjatlarni tuzish jarayoni davlat boshqaruvi faoliyati bilan bevosita bog‘liq bo‘lib, hujjatlar orqali idoralararo munosabatlar tartibga solinadi. Qarorlar, farmonlar, buyruqlar, arizalar va boshqa turdagi hujjatlar mazmun va shakl jihatidan belgilangan me’yorlarga muvofiq rasmiylashtiriladi. Har bir hujjatning to‘g‘ri tuzilishi uning mazmunini aniq yetkazish va amaliy qo‘llanishini ta’minlaydi. Rasmiy hujjatlar tilida aniqlik va tushunarlilik asosiy talablardan biri hisoblanadi. Matnda qo‘llaniladigan so‘z va iboralar ortiqcha murakkablikdan xoli, bir ma’noli va ravshan bo‘lishi zarur. Har bir jumla aniq fikrni ifodalashi, turlicha talqin qilinish ehtimolini kamaytirishi lozim. Bunday yondashuv hujjatni o‘qigan shaxs tomonidan uning mazmunini to‘g‘ri anglashga imkon beradi. Yuridik aniqlik rasmiy hujjatlarning muhim xususiyatlaridan biridir. Matnda qo‘llaniladigan terminlar va formulalar huquqiy me’yorlarga mos kelishi, hujjat mazmunida noaniqlik yuzaga keltirmasligi talab etiladi. Huquqiy terminologiyaning o‘rinli qo‘llanishi qaror va normativ hujjatlarning mazmunini to‘g‘ri ifodalashga yordam beradi.

Rasmiy hujjatlarda til va imlo qoidalariga qat’iy rioya qilish ham muhim omillardan biridir. O‘zbek tilining orfografik va punktuatsion me’yorlari asosida yozilgan hujjatlar aniq va tushunarli bo‘ladi. Terminologiyaning to‘g‘ri tanlanishi va izchil qo‘llanishi matnning mazmuniy yaxlitligini saqlaydi. Rasmiylik darajasi hujjatning tilida va tuzilishida namoyon bo‘ladi. Matnda norasmiy, og‘zaki nutqqa xos yoki mavhum iboralardan foydalanilmaydi. Hujjatlar ko‘pincha mutaxassislar tomonidan ko‘rib chiqiladi, bu jarayon matnning me’yorlarga mosligini ta’minlashga qaratiladi [Bobojonova D.,2025:46].

FOYDALANILGAN ADABIYOTLAR

1. Aminov M va boshq. Davlat tilida ish yuritish. Amaliy qo‘llanma. –T.: “O‘zbekiston nashriyoti” DUK, 2021. – 528 b.
2. Bobojonova D. O. *Rasmiy hujjatlarni tuzishda davlat tilidagi aniq talablar va standartlar* // *Prospects for Innovative Technologies in Science and Education: International online conference*. – 2025-yil 13-yanvar. – 47-b.
3. Mahmudov N. va b. Davlat tilida ish yuritish (9-nashri).– Toshkent, 2020. –528 b.
4. Taxirov Z.T. O‘zbek tilining amaliy uslubiyati. – Toshkent, 2019. – 200 b.
5. Ulashmurodova Sh. D. *Rasmiy uslubdagi matnlarning o‘ziga xos xususiyatlari* // *European science international conference: Studying the progress of science and its shortcomings*. – Toshkent: O‘zbekiston davlat jahon tillari universiteti, 2024. – 171-b.
6. Mengliyev, B., Xoliyorov, O., & Abdurahmonova, N. (2014). O‘zbek tilidan Universal qo ‘llanma (qayta ishlangan 3-nashri). Toshkent: akademnashr, 389.

O'ZBEK TILIDAGI MATNLARDA XATOLIKLARNI ANIQLASHDA DAMERAU–LEVENSHTEYN MASOFASI VA MARKOV ZANJIRI MODELLARINING QIYOSIY TAHLILI

Mirzayeva Faxriya Shavkat qizi

https://doi.org/10.5281/zenodo.23253323

Annotatsiya. *Imlo xatolarini aniqlash va tuzatish masalasida ikki yondashuv bir xil sharoitda qiyoslandi: so'zning yozma shakliga tayanuvchi Damerau–Levenshteyn masofasi va kontekstga tayanuvchi Markov zanjiri modeli. 538 879 tokenli korpus 80 va 20 foiz nisbatida ajratilib, 1200 juftlikdan iborat to'plam shakllantirilib, u sozlash (300) va test (900) qismlariga ajratildi. Tuzatish bosqichida masofa mezoni 93,1 %, Markov modeli 72,1 % natija berdi. Shovqinli kanal tamoyili asosida keltirib chiqarilgan gibrid model λ = 5,0 da 96,1 % ga erishdi, farqning statistik ahamiyatliligi McNemar testi bilan tasdiqlandi (p < 0,001). Aniqlash bosqichida nisbat teskari: lug'atda mavjud so'z hosil qiluvchi xatolarni lug'atga tayanuvchi tekshiruv umuman ko'rmaydi (F1 = 0,0 %), Markov modeli esa 73,2 % beradi.*

Kalit so'zlar: *imlo xatolarini tuzatish, Damerau–Levenshteyn masofasi, Markov zanjiri, shovqinli kanal modeli, transpozitsiya, o'zbek tili.*

Annotation. *Two approaches to spelling error detection and correction were compared under identical conditions: the Damerau–Levenshtein distance, which relies on the written form of the word, and a Markov chain model, which relies on context. A corpus of 538,879 tokens was split in an 80/20 ratio, and a set of 1,200 pairs was formed and divided into tuning (300) and test (900) subsets. At the correction stage the distance criterion achieved 93.1% and the Markov model 72.1%. A hybrid model derived from the noisy channel principle reached 96.1% at λ = 5.0, and the statistical significance of the difference was confirmed by McNemar's test (p < 0.001). At the detection stage the ratio is reversed: dictionary-based checking fails entirely on errors that produce a word present in the dictionary (F1 = 0.0%), whereas the Markov model yields 73.2%.*

Keywords: *spelling correction, Damerau–Levenshtein distance, Markov chain, noisy channel model, transposition, Uzbek language.*

Аннотация. *Два подхода к обнаружению и исправлению орфографических ошибок сопоставлены в одинаковых условиях: расстояние Дамерау–Левенштейна, опирающееся на письменную форму слова, и модель цепи Маркова, опирающаяся на контекст. Корпус объёмом 538 879 токенов разделён в соотношении 80 и 20 процентов; сформирован набор из 1200 пар, разделённый на настроечную (300) и тестовую (900) части. На этапе исправления критерий расстояния дал 93,1 %, модель Маркова — 72,1 %. Гибридная модель, выведенная на основе принципа шумного канала, при λ = 5,0 достигла 96,1 %; статистическая значимость различия подтверждена тестом Макнемара (p < 0,001). На этапе обнаружения соотношение обратное: словарная проверка полностью не выявляет ошибки, порождающие существующее в словаре слово (F1 = 0,0 %), тогда как модель Маркова даёт 73,2 %.*

Ключевые слова: *исправление орфографических ошибок, расстояние Дамерау–Левенштейна, цепь Маркова, модель шумного канала, транспозиция, узбекский язык.*

Kirish

Raqamli hujjatlar soni ortgani sari o‘zbek tilidagi matnlarni avtomatik imloviy tekshirish vositalariga ehtiyoj kuchaymoqda. Boshqa tillar uchun ishlab chiqilgan yechimlarni bevosita ko‘chirish qiyin: o‘zbek tili aglutinativ tuzilishga ega, bu lug‘at hajmini oshiradi va statistik modellarda ma’lumotlar siyrakligini kuchaytiradi [Jurafsky D., ____; Elov B., 2025].

Bu yo‘nalishda [Elov B., 2025] da n-gram til modelini qurish bosqichlari yoritilgan, [Ochilov M.M., 2025] da esa tahrir masofasi nomzod generatori, neyron til modellari tartiblovchi sifatida qo‘llanilgan. Biroq mazkur ishlarda komponentlarning alohida hissasi ajratilmagan, sun’iy xato generatsiyasiga transpozitsiya kiritilmagan, masofa esa o‘zbek yozuvining xususiyatlarini hisobga olmagan holda belgi darajasida hisoblangan.

Tadqiqot maqsadi — ikki yondashuvni bir xil sharoitda qiyosiy baholash va ularning qaysi bosqichda ustunlikka ega ekanini aniqlash.

1. Nazariy asos

Damerau–Levenshteyn masofasi bir satrni ikkinchisiga aylantirish uchun zarur elementar amallarning eng kichik sonini bildiradi. Klassik Levenshteyn ta’rifidagi uch amalga (qo‘shish, o‘chirish, almashtirish) to‘rtinchisi — yonma-yon ikki belgining o‘rin almashishi (transpozitsiya) qo‘shiladi [Damerau F.J., 1964]. Masofa dinamik dasturlash matritsasi orqali hisoblanadi:

$$d(i,j) = min \{ d(i-1,j)+1 , d(i,j-1)+1 , d(i-1,j-1)+c , d(i-2,j-2)+1 \} \quad (1)$$

bu yerda c = 0, agar a(i) = b(j), aks holda c = 1. To‘rtinchi shart faqat a(i) = b(j−1) va a(i−1) = b(j) bo‘lganda qo‘llaniladi. Murakkablik O(nm).

Markov farazi bo‘yicha navbatdagi so‘zning ehtimolligini aniqlash uchun faqat so‘nggi bir necha so‘z hisobga olinadi. Bigram modelida shartli ehtimollik Laplas silliqlashi bilan baholanadi:

$$P(w \mid c) = (C(c, w) + 1) / (C(c) + |V|) \quad (2)$$

Maxrajga |V| qo‘shilishi ehtimolliklar yig‘indisining birga teng bo‘lishini ta’minlaydi.

2. Gibrid modelning keltirib chiqarilishi

Ikki yondashuvni birlashtirish maqsadga muvofiqligini asoslash uchun ularning umumiy nazariy ramkadagi o‘rnini aniqlash lozim. Quyida ko‘rsatilishicha, tahrir masofasi va til modeli bitta ehtimoliy ifodaning ikki turli hadida joylashadi, ya‘ni ular bir-birini almashtirmaydi, balki to‘ldiradi.

Imlo tuzatish masalasi Bayes teoremasi orqali ifodalanadi. Xato yozilgan so‘z x kuzatilganda, asl so‘z w bo‘lish ehtimolligi:

$$P(w \mid x) = P(x \mid w) \cdot P(w) / P(x) \quad (3)$$

Eng ehtimolli nomzodni tanlashda maxrajdagi P(x) barcha nomzodlar uchun bir xil bo‘lgani sababli tartibga ta’sir qilmaydi va tashlab yuboriladi. Natijada shovqinli kanal modelining klassik ko‘rinishi hosil bo‘ladi [Kernighan M., 1990]:

$$\hat{w} = arg\ max\ [P(x \mid w) \cdot P(w)] \quad (4)$$

Ifoda ikki qismdan iborat: xato modeli P(x | w) va til modeli P(w). Bu tadqiqot uchun printsipial ahamiyatga ega. Damerau–Levenshteyn masofasi birinchi hadga, Markov zanjiri modeli esa ikkinchi hadga mos keladi. Boshqacha aytganda, taqqoslanayotgan ikki yondashuv bitta tenglamaning turli o‘rinlarini egallaydi va aynan shu ularning raqobatlashuvchi emas, to‘ldiruvchi ekanini matematik jihatdan asoslaydi. Mazkur tadqiqotda ikkala qismga ham o‘zgartirish kiritildi.

Til modeli. Klassik ko‘rinishda P(w) unigram ehtimollik bo‘lib, kontekstni hisobga olmaydi. Biz uni bigram shartli ehtimollikka almashtirdik, ya’ni P(w) o‘rniga P(w | c) qo‘llanildi, bu yerda c — oldingi so‘z.

Xato modeli. P(x | w) ni baholash annotatsiyalangan xato korpusini talab qiladi, o‘zbek tili uchun bunday resurs mavjud emas. Shu sababli u yaqinlashtirish bilan almashtirildi. Yaqinlashtirish quyidagi farazga asoslanadi: xato so‘z asl so‘zdan qanchalik uzoq bo‘lsa, uning yuzaga kelish ehtimoli eksponensial ravishda kamayadi:

$$P(x \mid w) = (1/Z) \cdot \exp(-\lambda \cdot d(x, w)) \quad (5)$$

bu yerda Z — normallashtiruvchi ko‘paytuvchi bo‘lib, u ifodani ehtimollikka aylantiradi, ya’ni barcha variantlar yig‘indisini birga tenglashtiradi. Nomzodlarni tartiblashda Z tartibga ta’sir qilmaydi, chunki barcha nomzodlarni bir xil songa bo‘lish ularning nisbiy joylashuvini o‘zgartirmaydi. Shu sababli u keyingi hisob-kitoblarda tashlab yuboriladi.

Ifoda (4) ga (5) ni qo‘yib, logarifm olamiz. Logarifmik almashtirish ko‘paytmani yig‘indiga aylantiradi va tartibni saqlaydi:

$$\hat{w} = \arg\max [\log P(w \mid c) - \lambda \cdot d(x, w)] \quad (6)$$

Birinchi had nomzodning kontekstga mosligini, ikkinchisi esa yozma uzoqlikni ifodalaydi. λ koeffitsienti ikki qismning nisbiy og‘irligini belgilaydi va, ifoda (5) dan ko‘rinib turganidek, eksponensial kamayish tezligini bildiradi. λ = 5,0 qiymatida har bir qo‘shimcha tahrir amali xato ehtimolligini exp(5) ≈ 148 marta kamaytiradi, bu Damerau kuzatuviga mos keladi, unga ko‘ra imlo xatolarining asosiy qismi bitta amaldan iborat [Damerau F.J., 1964].

Ta’kidlash lozimki, ifoda (5) yaqinlashtirish hisoblanadi. Unda Z barcha nomzodlar uchun taxminan bir xil deb qabul qilinadi. Qat’iy holda Z har bir w uchun farq qilishi mumkin, biroq bir xil uzunlikdagi nomzodlar uchun bu farq ahamiyatsiz.

3. Metodika

Tadqiqot uchun Uzbek-Corpus-Sample korpusi tanlandi [Kuriyozov E., ____]. Korpus gaplar darajasida 80 va 20 foiz nisbatida ajratildi. Nomzodlar lug‘ati, chastotalar va bigram statistikasi faqat o‘quv qismidan olindi, test so‘zlari esa faqat test qismidagi gaplardan tanlandi. Bu til modelining test materialini oldindan ko‘rmasligini ta’minlaydi. Shakllantirilgan 1200 juftlik o‘z navbatida ikkiga bo‘lindi: 300 juftlik λ koeffitsientini sozlash uchun, qolgan 900 tasi yakuniy baholash uchun ajratildi. Test qismiga bir marta, sozlash yakunlangandan so‘nggina murojaat qilindi.

1-jadval. Korpusning miqdoriy tavsifi

Ko‘rsatkich	Butun korpus	O‘quv qismi
Gaplar soni	99 650	79 720
Tokenlar soni	538 879	430 988
Tiplar soni	47 033	41 964
Noyob bigramlar	—	191 552
Nomzodlar lug‘ati	—	8 854

O‘zbek lotin yozuvida o‘ va g‘ harflari ikki belgi bilan ifodalansa-da, yagona harf hisoblanadi. Masofani belgi darajasida hisoblash ularni bo‘lib yuboradi. Nazorat tekshiruvi buni

tasdiqladi: bog‘dagi va bodg‘agi juftligi uchun belgi darajasidagi masofa 2 ga, harf darajasidagi masofa 1 ga teng chiqdi. Shu sababli so‘zlar harflar ketma-ketligi sifatida ifodalandi.

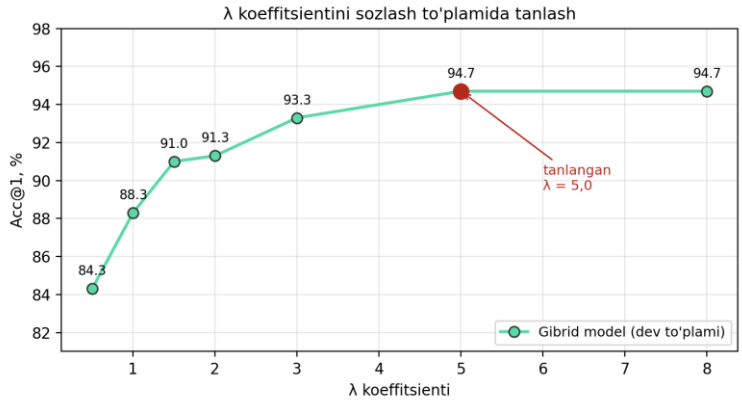
Test to‘plami sun’iy xato generatsiyasi bilan shakllantirildi. Xato modelida alifbodan tasodifiy harf tanlash o‘rniga joriy harfning QWERTY klaviaturasidagi qo‘shnilari hamda o‘zbek tiliga xos fonetik yaqinliklar (x va h, q va k, o va a, u va o‘) asos qilib olindi. To‘rt xato turi qo‘llanildi, test qismidagi taqsimoti quyidagicha: o‘chirish (205), qo‘shish (225), almashtirish (230), transpozitsiya (240). Uchala yondashuv uchun nomzodlar to‘plami bir xil tarzda shakllantirildi ($d \leq 2$), ular faqat tartiblash mezoni bilan farqlanadi.

4. Natijalar

2-jadval. Yondashuvlarning samaradorligi (n = 900)

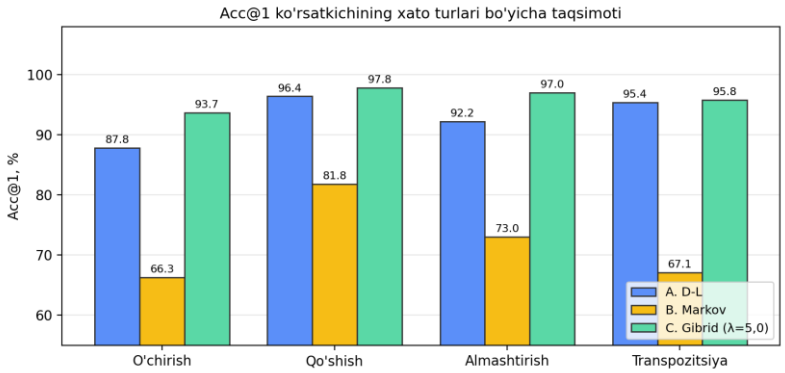
Yondashuv	Acc@1	Acc@5
A. Damerau–Levenshteyn	93,1 %	100 %
B. Markov zanjiri (bigram)	72,1 %	94,7 %
C. Gibrid model (λ = 5,0)	**96,1 %**	**100 %**

Masofa mezoni Acc@5 bo‘yicha 100 % berdi, ya’ni to‘g‘ri variant har doim dastlabki besh nomzod ichiga tushadi, Acc@1 esa 93,1 % ni tashkil etdi. Ular orasidagi 6,9 foiz punktlik farq nomzodni topish emas, uni tartiblash bosqichidagi zaiflikni ko‘rsatadi va kontekstli axborot uchun maydon qoldiradi. Markov zanjiri modeli yakka holda 72,1 % berdi, chunki u nomzodlarni shakllantira olmaydi va faqat tayyor to‘plamni tartiblaydi.



1-rasm. λ koeffitsientini sozlash to‘plamida tanlash

λ koeffitsienti sozlash to‘plamida 0,5 dan 8,0 gacha o‘zgartirib chiqildi. Kichik qiymatlarda gibrid model masofa mezonidan past natija beradi, λ = 5,0 dan boshlab esa egri chiziq to‘yinadi. Bu qonuniyat ifoda (6) tarkibidan kelib chiqadi: silliqlangan ehtimolliklarning logarifmi taxminan −8 dan −12 gacha oraliqda joylashadi, masofa esa 1 yoki 2 qiymatini oladi. Kichik λ da masofa hadi juda kichik bo‘lib qoladi va kontekstli qism hukmronlik qiladi, λ = 5,0 da esa masofa yetakchilikni oladi, kontekstli qism faqat masofalar teng bo‘lgan nomzodlar orasida hal qiluvchi ovoz beradi. Test qismida gibrid model 96,1 % qayd etib, masofa mezonidan 3,0 foiz punktga ustun keldi. Farqning tasodifiy emasligini tekshirish uchun McNemar testi qo‘llanildi: gibrid model 29 ta holatda masofa mezoni xato qilgan javobni to‘g‘riladi va faqat 2 ta holatda teskari natija kuzatildi ($p < 0,001$), ya’ni ustunlik statistik jihatdan ahamiyatlidir.



2-rasm. Acc@1 ko‘rsatkichining xato turlari bo‘yicha taqsimoti

3-jadval. Acc@1 xato turlari bo‘yicha

Xato turi	n	D-L	Markov	Gibrid	Levenshteyn
O‘chirish	205	87,8 %	66,3 %	93,7 %	87,8 %
Almashtirish	230	92,2 %	73,0 %	97,0 %	92,2 %
Qo‘shish	225	96,4 %	81,8 %	97,8 %	96,9 %
Transpozitsiya	240	95,4 %	67,1 %	95,8 %	45,4 %
Jami	**900**	**93,1 %**	**72,1 %**	**96,1 %**	**79,9 %**

Gibrid modelning foydasi turlar bo‘yicha notekis taqsimlangan: o‘chirishda 5,9 va almashtirishda 4,8 foiz punkt, transpozitsiyada esa atigi 0,4 foiz punkt. Buning sababi nomzodlar to‘plamining kengligida. Harf o‘chirilganda yoki almashtirilganda so‘z uzunligi saqlanadi yoki qisqaradi, lug‘atda teng masofada joylashgan variantlar ko‘payadi va kontekst tanlov uchun zarur bo‘ladi. Transpozitsiyada esa masofa mezoni yakka holda 95,4 % beradi, ya’ni qo‘shimcha axborot uchun maydon deyarli qolmaydi.

Oxirgi ustun transpozitsiya amalining ta’sirini ko‘rsatadi. O‘chirish va almashtirish turlarida ikkala metrika bir xil natija beradi, transpozitsiyada esa farq 50,0 foiz punktni tashkil qiladi (95,4 % va 45,4 %). Klassik ta’rifda transpozitsiya ikki mustaqil amal sifatida baholanib, masofa 2 ga teng bo‘ladi, natijada to‘g‘ri variant lug‘atdagi o‘nlab boshqa so‘zlar orasida quyi o‘rinlarga siljiydi. Qo‘shish turida esa Levenshteyn masofasi 0,5 foiz punktga yuqori natija berdi. Bu kutilmagan ko‘rinsa-da, izohi mavjud: Damerau–Levenshteyn masofasi ta’rifi bo‘yicha Levenshteyn masofasidan kichik yoki unga teng, demak $d \leq 2$ sharti bilan u kengroq nomzodlar to‘plamini beradi. Kengroq to‘plam esa to‘g‘ri variant uchun raqobatni kuchaytiradi. Umumiy hisobda bu ta’sir transpozitsiyadagi yutuq bilan qoplanadi (93,1 % va 79,9 %).

4-jadval. Xatolarni aniqlash bosqichi (F1)

Xato sinfi	Lug‘at (D-L)	Markov
Lug‘atda yo‘q shakl	100,0 %	74,8 %
Lug‘atda bor shakl	0,0 %	73,2 %
To‘liq to‘plam	66,4 %	76,7 %

Aniqlash bosqichi tuzatishdan farqli masala hisoblanadi. Unda tizim har bir so‘z uchun ikkilik qaror qabul qiladi: so‘zda xato bormi yoki yo‘q. Shu sababli bu yerda Acc@k emas,

Precision, Recall va F1 ko‘rsatkichlari qo‘llanildi. Test to‘plami 1197 yozuvdan iborat bo‘lib, unga 297 ta lug‘atda mavjud bo‘lmagan shakl, 300 ta lug‘atda mavjud shakl va 600 ta xatosiz so‘z kiritildi.

Xato natijasida lug‘atda mavjud so‘z hosil bo‘lgan sinfda lug‘atga tayanuvchi detektor 0,0 % qayd etdi. Bu o‘lchov xatosi emas, usul ta’rifidan kelib chiqadigan natijadir: hosil bo‘lgan shakl lug‘atda mavjud bo‘lgani uchun tekshiruv mexanizmi umuman ishga tushmaydi. Masalan, otamning so‘zining odamning ko‘rinishiga o‘tishida ikkala shakl ham to‘g‘ri so‘z bo‘lib, ularni faqat kontekst farqlaydi. Natijani lug‘at hajmini oshirish yoki masofa chegarasini kengaytirish orqali yaxshilab bo‘lmaydi.

Xulosa

Damerau–Levenshteyn masofasi va Markov zanjiri modeli tayanch manbasi bilan farq qiladi va shu sababli raqobatlashuvchi emas, bir-birini to‘ldiruvchi vositalar hisoblanadi. Masofa mezoni nomzodlarni shakllantiradi, ammo teng masofadagi variantlar orasidan tanlov qila olmaydi. Statistik model tanlov qiladi, ammo nomzodlarni mustaqil hosil qila olmaydi. Tuzatish bosqichida birinchisi ustun (93,1 % va 72,1 %), aniqlash bosqichida esa nisbat teskari yo‘nalishda o‘zgaradi (0,0 % va 73,2 %).

Gibrid modelni baholashda λ koeffitsientini sozlash majburiy shart hisoblanadi. Kichik qiymatlarda model masofa mezonidan past natija beradi, sozlangandan so‘ng esa undan 3,0 foiz punktga ustun keladi va bu ustunlik statistik jihatdan ahamiyatli ($p < 0,001$). Transpozitsiya amalini hisobga olish o‘zbek tili materialida empirik asoslangan zarurat bo‘lib, bu turda Damerau–Levenshteyn masofasi klassik Levenshteyn masofasidan 50,0 foiz punktga yuqori natija beradi.

Cheklovlar. Test so‘zlari o‘quv lug‘atida mavjud bo‘lish sharti bilan tanlangan, lug‘atdan tashqaridagi so‘zlar sinovdan o‘tkazilmagan. Acc@5 ko‘rsatkichining 100 % ga teng bo‘lishi ham shu shartning oqibati bo‘lib, usulning mutlaq imkoniyatini emas, tajriba tuzilishini aks ettiradi. Sun’iy xatolar to‘rt tur bo‘yicha teng ehtimollik bilan kiritilgan, bu real chastota taqsimotini aks ettirmaydi. Kelgusi ishlarda so‘z chegarasi buzilishidan kelib chiqadigan xatolarni qamrab olish va tajribani kattaroq korpusda takrorlash rejalashtirilgan.

ADABIYOTLAR

1. Jurafsky D., Martin J.H. Speech and Language Processing. 3rd ed. draft. Ch. 3: N-gram Language Models.
2. Kernighan M., Church K., Gale W. A spelling correction program based on a noisy channel model // COLING. – 1990. – P. 205–210.
3. Elov B., Primova M. O‘zbek matnlari uchun n-gram til modeli // Management and Future Technologies. – 2025. – Vol. 2, Issue 3. – B. 166–176.
4. Ochilov M.M., Narzullayev O.O., Xolmatov O.A. Mashinali o‘qitish algoritmlari asosida o‘zbek tili matnlaridagi imlo xatolarini aniqlash va tuzatish // Raqamli texnologiyalarning nazariy va amaliy masalalari xalqaro jurnali. – 2025. – № 8(1). – B. 85–94.
5. Kuriyozov E. Uzbek-Corpus-Sample. – GitHub repozitoriysi.
6. Damerau F.J. A technique for computer detection and correction of spelling errors // Communications of the ACM. – 1964. – Vol. 7, № 3. – P. 171–176.
7. Levenshtein V.I. Binary codes capable of correcting deletions, insertions and reversals // Soviet Physics Doklady. – 1966. – Vol. 10. – P. 707–710.

OYBEKNING "NAVOIY" ROMANIDA PORTRET BIRLIKLARINING LEKSIK-SEMANTIK MUNOSABATLARI

Mirzayeva Maftuna Baxtiyor qizi
Abu Rayhon Beruniy nomidagi Urganch davlat universiteti tayanch doktoranti
https://doi.org/10.5281/zenodo.23253355

Annotatsiya. *Maqolada Oybekning "Navoiy" romanidagi portret birliklarining leksik-semantik munosabatlari parallel korpusning o'zbekcha segmentlari asosida o'rganiladi. Komponent, kontekstual va qiyosiy tahlil orqali ma'noviy yaqinlik, qarama-qarshilik, darajalanish, tur–jins, butun–qism hamda obrazli aloqalar farqlanadi. Tahlil umumiy semaning mavjudligi to'liq sinonimiyani, tana a'zosining tilga olinishi esa sinekdoxani avtomatik belgilamasligini ko'rsatadi. Birliklarning personajni individuallashtirishdagi vazifasi differensial sema, uslubiy bo'yoq va kontekstual baholash bilan izohlanadi. Semantik kategoriya, birliklararo munosabat va badiiy vosita turini alohida qayd etish portret korpusini annotatsiyalash hamda tarjimaviy o'zgarishlarni tekshirishning metodologik asosi sifatida taklif etiladi.*

Kalit so'zlar: *badiiy portret, parallel korpus, sinonimiya, antonimiya, graduonimiya, giponimiya, partonimiya, metafora.*

Аннотация. *В статье исследуются лексико-семантические отношения портретных единиц в романе Айбека «Навои» на материале узбекских сегментов параллельного корпуса. Компонентный, контекстуальный и сопоставительный анализ позволяют разграничить смысловую близость, противопоставление, градацию, родовидовые, партитивные и образные связи. Показано, что общая сема не гарантирует полной синонимии, а упоминание части тела само по себе не образует синекдоху. Функция индивидуализации персонажа определяется дифференциальными семами, стилистической окраской и контекстуальной оценкой. Раздельная фиксация семантической категории, межъединичного отношения и типа художественного средства предлагается как методологическая основа корпусной разметки и последующего анализа переводческих изменений.*

Ключевые слова: *художественный портрет, параллельный корпус, синонимия, антонимия, градация, гипонимия, партитивность, метафора.*

Abstract. *This article examines lexical-semantic relations among portrait units in Oybek's novel Navoi, using the Uzbek segments of a parallel corpus. Componential, contextual and comparative analysis distinguishes semantic similarity, opposition, gradation, taxonomic, part–whole and figurative relations. The analysis shows that a shared semantic feature does not establish complete synonymy, while mentioning a body part does not automatically constitute synecdoche. Character individuation depends on distinctive features, stylistic colouring and contextual evaluation. Recording semantic categories, relations between units and types of figurative devices separately is proposed as a methodological basis for corpus annotation and subsequent analysis of translation shifts.*

Keywords: *literary portrait, parallel corpus, synonymy, antonymy, gradation, hyponymy, meronymy, metaphor.*

Kirish

Badiiy portret personajning tashqi koʻrinishi, harakati va ruhiy holatini birlashtiruvchi mazmuniy tuzilma sifatida shakllanadi. Uning birliklarini yuz, gavda, kiyim, nigoh yoki xarakter belgilariga ajratish tasvirning qaysi sohaga tegishli ekanini koʻrsatadi. Biroq bunday tasnif tavsiflar qanday bogʻlanishi, bir-birini aniqlashtirishi yoki qarama-qarshi qoʻyilishini toʻliq ochmaydi. Portretning semantik yaxlitligini tushuntirish uchun mavzuiy guruhlash bilan birliklararo munosabatlarni farqlash zarur.

Leksik munosabatlar nazariyasida sinonimiya, antonimiya, giponimiya va meronimiya mustaqil semantik aloqalar sifatida oʻrganiladi [2]. Badiiy matnda ushbu aloqalar kontekst ta'sirida murakkablashadi: lugʻatda sinonim boʻlmagan birliklar bir vaziyatni yoritishda yaqinlashadi, antonim boʻlmagan kiyim nomlari esa kontrast portret tarkibida ishlaydi. Oʻzbekcha soʻzlarning semantik oʻxshashligi va aloqadorligini farqlashga qaratilgan SimRelUz tadqiqoti ham bu ikki tushunchani bir xillashtirmaslikning amaliy ahamiyatini koʻrsatadi [3].

Maqolaning maqsadi "Navoiy" romanidagi portret birliklari oʻrtasidagi munosabatlarni umumiy va differensial semalar hamda kontekstual vazifa asosida izohlashdir. Tadqiqot uch savolga javob beradi: 1) ma'no yaqinligi, qarama-qarshilik va darajalanish personajni qanday individuallashtiradi; 2) tur–jins va butun–qism aloqalari portret maydonini qanday tashkil etadi; 3) koʻpma'nolilik va obrazli koʻchimlar tashqi qiyofani ruhiy-ijtimoiy tavsif bilan qanday bogʻlaydi? Tahlilning taklif etilayotgan jihati semantik kategoriya, munosabat turi va badiiy ifoda mexanizmini alohida, lekin oʻzaro bogʻliq qatlamlarda tavsiflashdir.

Material va tadqiqot usullari

Material sifatida roman asosida tuzilgan oʻzbekcha-inglizcha ishchi parallel korpusning original segmentlari olindi [1]. Misollar munosabatni kontekst orqali asoslash imkoniyati va personaj tavsifiga daxldorligi boʻyicha maqsadli tanlandi. Havoladagi T/R korpus jadvalining tartib raqamini bildiradi; takror raqamlar gap boshlanishi orqali farqlanadi. Ushbu koʻrsatkichlar bosma nashr sahifalari emas. Maqola sifat tahliliga asoslanadi; tanlangan misollarning natijalari butun roman boʻyicha chastotaviy ustunlik sifatida umumlashtirilmaydi.

Komponent tahlil birliklarning umumiy va differensial semalarini, kontekstual tahlil esa referent, vaziyat va baholashni aniqlash uchun qoʻllandi. Qiyoslashda til tizimidagi aloqa bilan muayyan epizodda yuzaga kelgan munosabat ajratildi. Darajalanish belgining umumiy shkalasi saqlanishi asosida tekshirildi [4]. Metaforik qoʻllanishlarni izohlashda MIP usulining kontekstual ma'noni asosiyroq ma'no bilan solishtirish tamoyiliga tayanildi [6]. Bu tamoyildan foydalanish butun korpus MIP boʻyicha mustaqil annotatsiyadan oʻtkazilganini anglatmaydi.

Badiiy vosita turi (BVT) leksik-semantik munosabat bilan tenglashtirilmadi. Masalan, oʻxshatish ifodaning badiiy qurilishini, kontekstual sinonimiya uning boshqa birlik bilan aloqasini bildiradi. Neytral muqobillar va umumlashtiruvchi nomlar qiyoslashga jalb qilindi, ammo ular shu sababning oʻzi bilan alohida BVT hisoblanmadi. Koʻp soʻzli ifodalarning ichki tahlili ham yaxlit obrazni statistik hisobda asossiz boʻlib yuborishga olib kelmasligi kerak [7].

Natijalar va muhokama

Birinchi tadqiqot savoliga javob

Sinonimik aloqaning asosiy sharti umumiy semantik yadroning mavjudligidir. Aloiddin Mashhadiyning "asabiy, jizzaki, sezgir" sifatlari bilan tavsiflanishida "asabiy" va "jizzaki" emotsional ta'sirchanlik doirasida yaqinlashadi [1, T/R 17–19]. Biroq birinchisi ruhiy taranglikni, ikkinchisi tez achchiqlanishga moyillikni ajratib koʻrsatadi. "Sezgir" esa ta'sirni

nozik qabul qilishni ifodalaydi. Keyingi ranjishga oid izoh uch belgini bitta ruhiy portretda birlashtiradi, lekin ularni o‘zaro to‘liq almashinadigan birliklarga aylantirmaydi. Bu holat qisman sinonimiya va semantik to‘ldiruvchilikning birga ishlashini ko‘rsatadi.

To‘qli merganning “qo‘ydek yuvosh, mutavoze o‘tiribdi” tarzidagi tasviri kontekstual yaqinlikning boshqa ko‘rinishidir [1, T/R 211, “Majididdin o‘z oldida...”]. “Qo‘ydek yuvosh” osoyishtalik va qarshilik ko‘rsatmaslikni obrazli kuchaytiradi; “mutavoze” o‘zini kamtarona tutishning muomala jihatini bildiradi. Umumiy sema ayni vaziyatdagi tiyilgan tutumdan iborat. Differensial sema birinchisida jonivorga qiyos, ikkinchisida ijtimoiy tavoze bilan bog‘langan. Ularning yonma-yon kelishi tavsifni takrorlamaydi, tashqi holatga munosabat mazmunini qo‘shadi. Shuningdek, bu sahnadagi yuvoshlik personajning doimiy jismoniy ojizligi sifatida talqin qilinmaydi.

Turobiyga nisbatan ishlatilgan “kamtar” va “kamtarin” birliklari semantik jihatdan juda yaqin muqobillarni tashkil etadi [1, T/R 131–137]. Ular o‘zini yuqori qo‘ymaslik belgisini saqlagan holda ifodaning uslubiy ko‘rinishini o‘zgartiradi. Biroq turli epizodlardagi ikki shaklning mavjudligi kamtarlik darajasi oshganini isbotlamaydi. Demak, to‘liq sinonimga yaqinlik, uslubiy variatsiya va graduonimik farq bir xil hodisa emas. Darajalanish xulosasi uchun umumiy belgi bilan birga uning kuchayishini ko‘rsatuvchi dalil ham talab etiladi.

Antonimik qiyosda “oriq” va “semiz” tana to‘laligi o‘qining qarama-qarshi tomonlarini belgilaydi [1, T/R 6, 93]. Lekin ularning turli gaplarda uchrashi tegishli personajlar maxsus kompozitsion juftlikka aylantirilganini anglatmaydi. Bunday kontrastning aniqroq ko‘rinishi mullabachchalar bilan To‘g‘onbekning parallel tasvirida kuzatiladi: “top-toza” to‘n va “qordek oppoq salla”ga “qo‘sqi po‘stin”, katta bo‘rk, chorpaxil gavda hamda “lapanglab” yurish muqobil qo‘yiladi [1, T/R 82–89]. “Salla–bo‘rk” lug‘aviy antonim emas; qarama-qarshilik kiyim, gavda va harakat belgilarining birikishidan hosil bo‘ladi.

Graduallikda esa aynan bir parametr bo‘yicha tartiblanish saqlanadi. Aloiddinning “yarim yumuq ko‘zlarini ... butunlay yumib” olishi ko‘z yumuqligining analitik darajalanishini ifodalaydi [1, T/R 67]. Vaziyatdagi ensa qotishi ushbu jismoniy o‘zgarishga emotsional vazifa yuklaydi. Buni leksik sinonimlar qatori bilan tenglashtirib bo‘lmaydi: daraja ravishlar va predikativ qurilma orqali berilgan. “Qaradi–tikildi–xo‘mraydi” kabi qatorlarda esa yo‘nalish, davomiylik va affektiv belgi almashishi mumkinligi sababli bitta intensivlik shkalasini oldindan qabul qilish to‘g‘ri emas.

Shunday qilib, birinchi savol doirasida individuallashtirish uch xil yo‘l bilan izohlanadi: yaqin birliklar belgining turli tomonlarini aniqlashtiradi, kontrast tavsiflar personajlar farqini ochadi, gradatsiya esa holatning o‘zgarishini ko‘rsatadi. Baholovchi ma’no ushbu munosabatlarning natijasi sifatida kontekstda aniqlanadi; barcha misollarni dastlabdan “ijobiy–salbiy” juftligiga kiritish ularning semantik tarkibini soddalashtiradi.

Ikkinchi tadqiqot savoliga javob

Giperonimiya-giponimiya tushunchalarning umumiylik darajasiga, partonimiya yoki meronimiya esa obyektning tarkibiy tuzilishiga asoslanadi. WordNet terminologiyasida bu farq muayyan narsaning “bir turi” va “bir qismi” bo‘lishi orqali ajratiladi [5]. Korpusdagi “salla” va “bo‘rk” nomlari “bosh kiyim” giperonimi ostidagi kogiponimlardir [1, T/R 84, 87]. Ularning umumiyligi sinonimik tenglikni bildirmaydi. Ayni birliklar epizodda kontrast yarata olishi tur–jins aloqasi bilan kontekstual funksiya turli tahlil qatlamlariga tegishli ekanini ko‘rsatadi.

“Fiziognomik belgi–ko‘z–qora ko‘z” ketma-ketligi esa bir xil giponimik munosabatlardan tuzilmagan. Birinchi bo‘g‘in tadqiqotchining tasnif kategoriyasi, ikkinchisi

tana a’zosi nomi, uchinchisi atributiv tavsifdir. Shuning uchun tasnif daraxtida kategoriyaga mansublik, anatomik tarkib va belgini aniqlashtirish alohida bog‘lanishlar sifatida qayd etilishi lozim. Aks holda tadqiqotning metatili bilan badiiy matndagi leksik birliklar bir sathga tushib qoladi.

Aloiddin portretidagi “chuvak yuz”, “dag‘al qora soqol”, “payvasta quyuq qosh” va “yarim yumuq xira ko‘z” bir personaj qiyofasiga birlashadi [1, T/R 9–12]. Bunda yuz va ko‘z anatomik tarkibni, sifatlovchilar esa shakl, tekstura, rang, zichlik yoki holatni ifodalaydi. Portretning kompozitsion yaxlitligi ana shu farqli belgilar uyg‘unligidan yuzaga keladi. Binobarin, tana qismlarining munosabati bilan ularni birlashtiruvchi badiiy tasvirning yaxlitligi o‘zaro bog‘liq bo‘lsa-da, aynan bir hodisa emas.

Partonimik aloqa har bir qo‘llanishda ma’no ko‘chishini yuzaga keltirmaydi. “Yirik ko‘zli” ifodasida a’zo nomi belgi egasini almashtirmaydi, balki uning tashqi xususiyatini bildiradi. “Ko‘zlari ... quvonib ketdi” tasvirida esa ichki emotsiya tashqi belgi orqali ifodalanib, metonimik talqinga asos paydo bo‘ladi [1, T/R 221]. Sinekdoxani aniqlash uchun qism nomining butun referent o‘rnida ishlatilishi ko‘rsatilishi zarur. Demak, tana a’zosining mavjudligi – partonimik asos, ruhiy holatga o‘tish – kontekstual semantik mexanizmdir.

Shu mezon kiyim va ijtimoiy maqom aloqasini ham aniqlashtiradi. “Xitoyi shoyi to‘nli” davlatmandlar va “juldur kiyimli” dehqonlar tavsifida kiyim materialining sifati hamda holati ijtimoiy farqni ko‘rinadigan qiladi [1, T/R 109–112]. Kiyim tananing anatomik qismi emas; uning maqom bilan bog‘lanishi egalik va ijtimoiy belgi orqali shakllanadi. Ikkinchi savolning javobi shundan iboratki, portret maydoni yagona tur–jins zanjiri bilan ifodalanmaydi: unda tasnifiy ierarxiya, butun–qism, atributiv va ijtimoiy aloqalar tutashadi.

Uchinchi tadqiqot savoliga javob

Bir leksemaning turli portret vazifalarida ishlashi uning ma’nosini kontekstdagi birikishiga qarab aniqlashni talab etadi. “Qaddi-qomati nozik” va “tabiatining nozikligi” ifodalarida “nozik” jismoniy ko‘rinish hamda ruhiy xususiyat sohalariga tegishli [1, T/R 20, 192]. Ikkinchi ifodaning Navoiy g‘ururi va yordam so‘ramasligi bilan bog‘lanishi uni jismoniy ojizlikdan ajratadi. “Dag‘al qora soqol” va “dag‘al iboralar” qiyosida ham tekstura belgisi bilan nutqiy qo‘pollik farqlanadi [1, T/R 9, 8]. Demak, bir leksemaga barcha kontekstlarda bitta portret kategoriyasini biriktirish yetarli emas.

Kontekstual bahoning o‘zgarishi esa doimo yangi lug‘aviy ma’no paydo bo‘lganini bildirmaydi. Navoiy haqidagi “insoniy g‘ururi” birikmasi o‘z qadrini saqlashni oldinga chiqarsa, To‘g‘onbekning “g‘ururning va qo‘pol kuchning go‘yo timsoli” sifatidagi tavsifi namoyishkorona kuchni ajratadi [1, T/R 192; 212, “Qo‘sqi po‘stin kiygan...”]. Baholashdagi farq leksemaning atrofidagi birliklar va personaj tutumi bilan izohlanadi. Ko‘pma’nolilik, konnotatsiya va vaziyatga bog‘liq taassurot shu sababli alohida tekshiriladi.

Obrazli aloqalarda manba sohasi bilan ko‘chiriladigan belgi ham farqlanadi. “Burgutday hurpayib” hamda “qo‘ydek yuvosh” zoonimik modelga kiradi, ammo birinchisi hujumkorlikni, ikkinchisi qarshiliksiz tutumni ifodalaydi [1, T/R 8; 211, “Majididdin o‘z oldida...”]. Umumiy manba sohasi bir xil semantik natijani kafolatlamaydi. “-day”, “-dek” ko‘rsatkichlari mavjudligi sababli bular BVT sathida o‘xshatish sifatida tavsiflanadi. Konseptual qiyosning mavjudligi ularning badiiy shaklini avtomatik metaforaga aylantirmaydi.

“Qordek oppoq salla”da tabiiy hodisaning rangi kiyim detaliga ko‘chiriladi; “oppoq” esa intensivlikni kuchaytiradi [1, T/R 84]. Bu rangning axloqiy poklik ramzi ekanini aytish uchun qo‘shimcha kontekst zarur. “Oq tusha boshlagan soqol”da esa oqarish yosh o‘zgarishining tashqi

belgisi bo‘lib xizmat qiladi [1, T/R 144]. Shunday qilib, rang, yosh va baholash o‘rtasidagi bog‘lanish har safar bir xil emas. Mifologik manbali “devqomat”da esa gavdaning kattaligi ekspressiv qiyos bilan birlashadi; uning neytral “baland bo‘yli” tavsifiga to‘liq teng emasligi shu qo‘shimcha mazmun bilan belgilanadi [1, T/R 117].

“Ko‘zlarida havas, hasad yonib ketardi” ifodasi ikki mexanizmni tutashtiradi [1, T/R 214]. “Yonish” hisning faollashishini olov jarayoniga qiyoslashga, “ko‘z” esa ichki holatning tashqi belgisi sifatida talqin qilinishiga asos bo‘ladi. Birinchi aloqa metaforik, ikkinchisi metonimikdir. Sultonmurodning “ko‘ksi to‘lib-toshardi” tasvirida quvonch ichki makonni to‘ldiradigan mazmun sifatida ifodalanadi [1, T/R 103]. Bu misollarda “HIS – YONISH” va “ICHKI OLAM – SIG‘IM” modellarini qo‘llash kontekstdagi emotsional sabab bilan dalillanadi.

Madaniy assotsiatsiyada ham qiyosning nishoni aniqlashtirilishi lozim. “Saxiylikda Hotamtoy” fazilatni madaniy etalon orqali kuchaytiradi [1, T/R 262]. “Har bir so‘zi bebaho gavhardir” esa Navoiyning yuzini emas, so‘zining qiymatini tavsiflaydi [1, T/R 157]. U keng ma’nodagi intellektual portretga daxldor. Shuningdek, “kapalak gulzorda aylanganidek” qiyosi ijodiy muhitdagi faoliyatni tasvirlaydi; unda jonivor va o‘simlikka oid komponentlarning mavjudligi qomatning o‘simlikka qiyoslanganini anglatmaydi [1, T/R 161]. Demak, manba sohalari tasnifi bilan portret funksiyasi birgalikda tekshiriladi.

Uchinchi savolga ko‘ra, tashqi tasvirdan ruhiy mazmunga o‘tish leksik ma’no variantlari, metaforik o‘xshashlik, metonimik yaqinlik va madaniy baholash orqali yuzaga keladi. Bir birlik bir necha aloqada qatnashishi mumkin. Bunday ko‘p qatlamlilikni saqlash uchun uning shakli, qaysi referentga tegishliligi va kontekstual vazifasi alohida qayd etiladi.

1-jadval. Portret birliklarining semantik munosabatlarini ajratish mezonlari

Munosabat	Korpus misoli	Ajratuvchi mezon
Sinonimik yaqinlik	Asabiy – jizzaki	Umumiy ruhiy belgi va differensial sema
Antonimik qiyos	oriq – semiz	Tana to‘laligining qarama-qarshi qutblari
Darajalanish	yarim yumuq – butunlay yumib	Bitta parametr bo‘yicha o‘zgarish
Giponimiya	bosh kiyim – salla, bo‘rk	Nomlarning tur–jins ierarxiyasi
Partonimiya	tana – ko‘z	Obyektning butun–qism tuzilishi
Ko‘pma’nolilik	nozik qomat – nozik tabiat	Bir leksemaning bog‘liq ma’no variantlari
Metaforik aloqa	havas, hasad yonib ketardi	Hisning olov jarayoniga qiyoslanishi

Jadvalda qisqartirib berilgan misollar yuqorida ko‘rsatilgan korpus kontekstlariga tayanadi; “bosh kiyim” va “tana” tahliliy umumlashtiruvchi nomlardir. Munosabatlarni bu tarzda ajratish ularni o‘zaro istisno qiluvchi guruhlarga aylantirmaydi. Masalan, “salla” bilan “bo‘rk” bir sathda kogiponim, epizodning tasviriy tuzilishida esa kontrastning tarkibiy qismlaridir. “Qo‘ydek yuvosh” o‘xshatishi ham kontekstual yaqinlikda qatnashadi. Shu sababli bitta birlik uchun bir necha munosabatni ko‘rsatish uning mazmuniy aloqalarini aniqroq aks ettiradi.

Xulosa

Tahlil qilingan misollar portret birliklarining semantik tizimi mavzuiy guruhlar bilan cheklanmasligini ko‘rsatadi. Sinonimiya belgining umumiyligini differensial semalar bilan

birlashtiradi; antonimiya va kompozitsion kontrast qiyofalar farqini ochadi; darajalanish esa bir parametrning o‘zgarishini kuzatishga imkon beradi. Giperonimiya-giponimiya nomlarning tasnifiy munosabatini, partonimiya tarkibiy tuzilishni bildiradi. Ko‘pma’nolilik va obrazli assotsiatsiyalar bu aloqalarni personajning ruhiy hamda ijtimoiy tavsifi bilan bog‘laydi.

Maqolaning metodologik natijasi portretni annotatsiyalashda uch mustaqil maydonni ajratishdir: semantik kategoriya, birliklararo munosabat va BVT turi. Har bir munosabat uchun qiyoslanayotgan birlik, umumiy sema, differensial sema, personaj hamda kontekst dalili saqlanishi kerak. Bir necha munosabatning mavjudligi bitta qo‘llanishni takroran hisoblashga asos bo‘lmaydi. Kelgusidagi chastota tahlilida alohida qo‘llanish, takrorlanmas leksik model va munosabatlar soni turli ko‘rsatkichlar sifatida hisoblanadi.

Mazkur xulosalar tarjima tahlili uchun ham mezon beradi: obrazli muqobilning neytrallashuvi, ikki giponimning bitta umumiy nom bilan berilishi yoki intensivlik belgisining tushirilishi originaldagi aloqani o‘zgartirishi mumkin. Bular inglizcha tarjimada allaqachon tasdiqlangan natijalar emas, keyingi qiyosiy tekshiruv yo‘nalishlaridir. Tanlanma sifat tahlilining chegarasini saqlash va bahsli kodlashlarni mustaqil annotator bilan tekshirish xulosalarning tekshiriluvchanligini oshirishga xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR

1. Oybekning “Navoiy” romani asosida tuzilgan o‘zbekcha-inglizcha ishchi parallel korpus. Tuzuvchi: M. Mirzayeva. Nashr qilinmagan elektron tadqiqot materiali, T.docx. Havola birligi: T/R va zarur o‘rinda gap boshlanishi.
2. Cruse D. A. Lexical Semantics. Cambridge: Cambridge University Press, 1986. 310 p.
3. Salaev U., Kuriyozov E., Gómez-Rodríguez C. SimRelUz: Similarity and Relatedness Scores as a Semantic Evaluation Dataset for Uzbek Language // Proceedings of SIGUL. 2022. P. 199–206. https://aclanthology.org/2022.sigul-1.26/
4. Kennedy C., McNally L. Scale Structure, Degree Modification, and the Semantics of Gradable Predicates // Language. 2005. Vol. 81(2). P. 345–381. https://doi.org/10.1353/lan.2005.0071
5. Tohirov Z., Tukhtasinov I., Akhmedova A., Babajanova L., Egamberganova A., Yakubov M. Development and Evaluation of a Combined Model for Genre Classification of Uzbek Media Texts. Scopus konferensiya // 2025 IEEE XVII International Scientific and Technical Conference on Actual Problems of Electronic Instrument Engineering (APEIE). 2025. https://ieeexplore.ieee.org/document/11289361
6. Princeton University. WordNet Reference Manual: wngloss(7WN). https://wordnet.princeton.edu/documentation/wngloss7wn (murojaat sanasi: 12.09.2026).
7. Abduraxmonova, N. Z. (2018). Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) il.dis.aftoref. *2018). Abdurakhmonova N. The bases of automatic morphological analysis for machine translation. Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta.2016; 2 (38): 12-014.-264 c*.
8. Pragglejaz Group. MIP: A Method for Identifying Metaphorically Used Words in Discourse // Metaphor and Symbol. 2007. Vol. 22(1). P. 1–39. https://doi.org/10.1080/10926480709336752

O‘ZBEKCHA-RUSCHA PARALLEL KORPUSLARI ASOSIDA ALIGNER TIZIMINI YARATISHDA NLP TEXNOLOGIYALARI

Muxammadiyeva Kamola Doston qizi
O‘zbekiston Milliy universiteti Jurnalistika va o‘zbek filologiyasi fakulteti,
Kompyuter lingvistikasi mutaxasisligi II bosqich magistranti
https://doi.org/10.5281/zenodo.23253374

Annotatsiya. *Mazkur maqolada o‘zbek va rus tillari o‘rtasidagi parallel korpuslar asosida avtomatik aligner tizimini ishlab chiqish masalasi yoritiladi. Parallel korpuslar mashina tarjimasi, lingvistik tahlil va til modellarini o‘qitishda muhim resurs bo‘lib, o‘zbek tilining resurs jihatdan kam ta’minlanganligi ushbu yo‘nalishda maxsus yondashuvlarni talab etadi. Tadqiqotda parallel matnlarni yig‘ish, preprocessing, tokenizatsiya, gaplarga ajratish va normalizatsiya bosqichlari amalga oshirilgan. Gaplar darajasidagi moslashtirish uchun uzunlikka asoslangan Gale–Church algoritmi, lexical-based yondashuv va statistik modellar elementlarini birlashtirgan gibrid model taklif etilgan. Tajriba natijalari gibrid yondashuvning boshqa alohida metodlarga nisbatan yuqori samaradorlik ko‘rsatganini ko‘rsatadi.*

Kalit so‘zlar: *parallel korpus, o‘zbek tili, rus tili, alignment, NLP, mashina tarjimasi, korpus lingvistikasi, tokenizatsiya, gap segmentatsiyasi, statistik model, gibrid yondashuv.*

Abstract. *This article addresses the development of an automatic alignment system based on Uzbek–Russian parallel corpora. Parallel corpora are an important resource for machine translation, linguistic analysis, and the training of language models. The limited availability of linguistic resources for Uzbek requires specialized approaches in this field. The study involves the stages of parallel text collection, preprocessing, tokenization, sentence segmentation, and normalization. For sentence-level alignment, a hybrid model is proposed that combines the length-based Gale–Church algorithm, a lexical-based approach, and elements of statistical models. The experimental results demonstrate that the hybrid approach achieves higher performance compared with the individual application of the respective methods.*

Keywords: *parallel corpus, Uzbek language, Russian language, alignment, NLP, machine translation, corpus linguistics, tokenization, sentence segmentation, statistical model, hybrid approach.*

Аннотация. *В данной статье рассматривается вопрос разработки системы автоматического выравнивания (aligner) на основе параллельных корпусов узбекского и русского языков. Параллельные корпуса являются важным ресурсом для машинного перевода, лингвистического анализа и обучения языковых моделей. Недостаточная ресурсная обеспеченность узбекского языка требует применения специальных подходов в данном направлении. В исследовании реализованы этапы сбора параллельных текстов, предварительной обработки (preprocessing), токенизации, сегментации текстов на предложения и нормализации. Для выравнивания на уровне предложений предложена гибридная модель, объединяющая алгоритм Gale–Church, основанный на длине текстовых сегментов, лексический подход и элементы статистических моделей. Результаты экспериментов показывают, что гибридный подход продемонстрировал более высокую эффективность по сравнению с отдельным применением каждого из рассматриваемых методов.*

Ключевые слова: *параллельный корпус, узбекский язык, русский язык, выравнивание (alignment), NLP, машинный перевод, корпусная лингвистика, токенизация, сегментация предложений, статистическая модель, гибридный подход.*

Kirish. So‘nggi yillarda tabiiy tilni qayta ishlash (Natural Language Processing – NLP) texnologiyalarining jadal rivojlanishi turli tillar o‘rtasida avtomatik axborot almashinuvi imkoniyatlarini kengaytirmoqda. Bu jarayonda parallel korpuslar muhim lingvistik va hisoblash resursi sifatida namoyon bo‘ladi. Parallel korpus bir xil yoki mazmunan ekvivalent matnlarning ikki yoki undan ortiq tillardagi variantlarini o‘z ichiga oladi. Ular mashina tarjimasi tizimlarini yaratish, matnlarni qiyosiy tahlil qilish, tarjima modellarini o‘qitish va til birliklari o‘rtasidagi munosabatlarni aniqlashda keng qo‘llaniladi. O‘zbek tili NLP uchun resurs jihatdan kam ta’minlangan tillardan biri hisoblanadi. Rus tilida esa lingvistik resurslar, korpuslar va dasturiy vositalar nisbatan ko‘proq shakllangan. Shu sababli o‘zbek-rus tillari o‘rtasida sifatli parallel korpus yaratish va uni avtomatik moslashtirish mashina tarjimasi hamda kompyuter lingvistikasi uchun dolzarb masalalardan biridir. Mavjud alignment algoritmlarining katta qismi resursga boy tillar uchun ishlab chiqilgan bo‘lib, katta hajmdagi ma’lumotlarni talab qiladi. O‘zbek tilida ma’lumotlarning nisbatan cheklanganligi, shuningdek, uning agglutinativ morfologik tuzilishi va erkin so‘z tartibi alignment jarayonini murakkablashtiradi.

Mazkur tadqiqotning asosiy maqsadi o‘zbek va rus tillaridagi parallel matnlar asosida avtomatik aligner tizimini yaratish va uning samaradorligini baholashdan iborat. Buning uchun parallel korpus shakllantiriladi, matnlar oldindan qayta ishlanadi, turli alignment yondashuvlari sinovdan o‘tkaziladi va ularning natijalari precision, recall, F1-score hamda alignment accuracy mezonlari orqali baholanadi.

Mavzuga oid adabiyotlarning tahlili. Parallel korpuslar NLP sohasidagi asosiy lingvistik resurslardan biri bo‘lib, ular yordamida turli tillardagi ekvivalent birliklarni aniqlash va tarjima modellarini o‘qitish mumkin. Alignment masalasi dastlab sodda, tezkor va hisoblash jihatdan tejamkor usullar yordamida hal qilingan. Shunday yondashuvlardan biri gap uzunligiga asoslangan alignment bo‘lib, unda tarjima qilingan segmentlarning uzunlik ko‘rsatkichlari o‘rtasidagi munosabatdan foydalaniladi. Bunday usul katta tezlikda ishlashi bilan afzal bo‘lsa-da, semantik moslikni yetarli darajada hisobga olmaydi. Keyinchalik statistik alignment modellarining rivojlanishi bilan moslik ehtimollik asosida aniqlana boshlandi. IBM modellariga asoslangan yondashuvlar so‘zlar va tarjima birliklari o‘rtasidagi ehtimoliy munosabatlarni hisoblash imkonini berdi. Biroq statistik modellar samarali ishlashi uchun katta hajmdagi parallel ma’lumot talab qilishi mumkin. So‘nggi yillarda neyron tarmoqlar va embeddinglarga asoslangan usullar ham keng qo‘llanmoqda. Ular semantik yaqinlikni chuqurroq aniqlashi mumkin, ammo katta dataset va hisoblash resurslarini talab qiladi. Shu sababli kam resursli tillar uchun statistik, lingvistik va lexical belgilarni birlashtiruvchi gibrid yondashuvlar muhim hisoblanadi. Tizim samaradorligini baholashda precision topilgan mosliklarning qanchasi to‘g‘ri ekanini, recall mavjud mosliklarning qanchasi aniqlanganini ko‘rsatadi. F1-score esa ushbu ikki ko‘rsatkichning muvozanatli umumlashtirilgan qiymatidir. Alignment accuracy umumiy moslashtirishning to‘g‘rilik darajasini ifodalaydi. BLEU va NIST kabi metrikalar ham mavjud bo‘lsa-da, ular asosan mashina tarjimasi sifatini baholashga yo‘naltirilgan.

Tadqiqot metodologiyasi. Tadqiqotda o‘zbek-rus parallel matnlari asosida avtomatik alignment tizimini ishlab chiqish uchun kompleks yondashuv qo‘llanildi. Metodologiya ma’lumotlarni yig‘ish, korpus sifatini nazorat qilish, oldindan qayta ishlash, alignment

algoritmlarini qo‘llash va natijalarni baholash bosqichlarini qamrab oladi. Parallel korpus rasmiy hujjatlar, tarjima qilingan badiiy va ilmiy adabiyotlar hamda yangiliklar saytlaridan olingan matnlar asosida shakllantirildi. Turli janr va uslubdagi materiallardan foydalanish modelning turli matnlarga moslashish imkoniyatini oshirishga xizmat qiladi. Noto‘g‘ri moslashtirilgan yoki to‘liq bo‘lmagan segmentlar korpusdan chiqarib tashlandi. Korpusni alignmentga tayyorlash uchun avvalo sentence segmentation amalga oshirilib, matnlar gaplarga ajratildi. Keyin tokenization yordamida gaplar so‘zlar va tinish belgilaridan tashkil topgan tokenlarga bo‘lindi. Normalizatsiya jarayonida katta-kichik harflar bir xil ko‘rinishga keltirildi, ortiqcha belgilar va formatlash elementlari olib tashlandi. Zarur hollarda semantik yuklamasi past bo‘lgan yuqori chastotali birliklar bilan ishlash uchun stop-wordlar qo‘llanildi. Ushbu preprocessing bosqichlari keyingi alignment jarayonining barqaror va aniq ishlashi uchun zarur.

Alignment jarayonida bir nechta yondashuvlar ketma-ket qo‘llanildi. Birinchi bosqichda sentence-length based Gale–Church algoritmi yordamida dastlabki gap mosliklari aniqlandi. U tezkor bo‘lib, boshlang‘ich alignmentni shakllantirish uchun qulay. Ikkinchi bosqichda lexical-based alignment yordamida gaplar tarkibidagi so‘zlar o‘rtasidagi mosliklar hisobga olindi. Bu tarjima juftliklaridagi lexical va semantik yaqinlikni aniqlashni yaxshilaydi. Uchinchi bosqichda IBM modellariga yaqin statistik yondashuv elementlari qo‘llanib, ehtimollik asosida mosliklar aniqlashtirildi. Yakunda ushbu usullarni birlashtirgan gibrid model ishlab chiqildi. Tizimni amalga oshirishda Python dasturlash tilidan foydalanildi. NLP operatsiyalarini bajarish uchun NLTK va SpaCy kutubxonalari, zarur hollarda esa neyron modellar va embeddinglar bilan ishlash uchun HuggingFace texnologiyalaridan foydalanish imkoniyati ko‘zda tutildi. Tizim samaradorligi precision, recall, F1-score va alignment accuracy mezonlari bilan baholandi.

Tahlil va natijalar. Tajriba natijalari alohida alignment yondashuvlarining kuchli va zaif tomonlarini ko‘rsatdi. Gale–Church algoritmi tezkor ishlagan bo‘lsa-da, ayrim hollarda uzunligi o‘xshash, ammo mazmunan mos kelmaydigan gaplarni bir-biriga moslashtirdi. Bunga o‘zbek tilidagi erkin so‘z tartibi va rus tilidagi sintaktik farqlar ham ta’sir ko‘rsatdi. Demak, faqat gap uzunligiga tayanish semantik moslikni to‘liq ta’minlamaydi. Lexical-based yondashuv qo‘llanganda alignment sifati sezilarli yaxshilandi. Ayniqsa terminologik jihatdan yaqin matnlarda aniqlik yuqori bo‘ldi. Biroq sinonimiya va kontekstual farqlar ayrim hollarda noto‘g‘ri mosliklarga sabab bo‘ldi. Statistik yondashuv esa ehtimollik asosida mosliklarni aniqlash imkonini berdi, lekin mavjud korpus hajmi cheklanganligi sababli uning natijalari barcha holatlarda bir xil barqarorlikni ko‘rsatmadi. Eng yaxshi natija uzunlikka asoslangan, lexical va statistik usullarni birlashtirgan gibrid modelda kuzatildi. Dastlabki mosliklar Gale–Church orqali aniqlanib, keyinchalik lexical va statistik filtrlar yordamida aniqlashtirildi. Natijada alohida metodlarning kamchiliklari qisman bartaraf etilib, ularning kuchli tomonlari bir-birini to‘ldirdi. Bu yondashuv kam resursli o‘zbek tili uchun ayniqsa muhimdir.

Tadqiqotda olingan ko‘rsatkichlar quyidagicha:

Algoritm turi	Precision (%)	Recall (%)	F1-score (%)	Alignment accuracy (%)
Gale-Church	72	68	70	69
Lexical-based	81	76	78	77
Statistik model	84	79	81	80
Gibrid model	89	85	87	86

Natijalardan ko‘rinadiki, gibrid model barcha asosiy ko‘rsatkichlar bo‘yicha eng yuqori natijani qayd etgan. Uning F1-score ko‘rsatkichi 87% bo‘lib, Gale–Church algoritmidagi 70%, lexical-based yondashuvdagi 78% va statistik modeldagi 81% natijalardan yuqoridir. Alignment accuracy ham gibrid modelda 86% ni tashkil etib, boshqa usullarga nisbatan ustunlikni ko‘rsatdi. Bu natija o‘zbek-rus parallel matnlari uchun bitta metodga tayanishdan ko‘ra, bir nechta metodni ketma-ket va o‘zaro to‘ldiruvchi tarzda qo‘llash maqsadga muvofiqligini ko‘rsatadi.

Tahlil jarayonida o‘zbek va rus tillari o‘rtasidagi strukturaviy farqlar ham muhim omil ekani aniqlandi. O‘zbek tilining agglutinativ xususiyati sababli grammatik ma’nolar ko‘pincha qo‘shimchalar orqali ifodalanadi, rus tilida esa flektiv tuzilma ustun. Bundan tashqari, o‘zbek tilidagi nisbatan erkin so‘z tartibi va rus tilining sintaktik qurilishi so‘z darajasidagi moslikni murakkablashtiradi. Shuning uchun kelgusidagi tizimlarda morfologik va kontekstual belgilarni chuqurroq hisobga olish zarur.

Xulosa va takliflar. Tadqiqot natijasida o‘zbek-rus parallel matnlari asosida gibrid yondashuvga tayangan avtomatik alignment tizimi ishlab chiqildi va uning samaradorligi baholandi. Tizimda Gale–Church, lexical va statistik metodlarning kombinatsiyasi qo‘llanilib, har bir usulning alignment jarayonidagi vazifasi belgilandi. Tajriba natijalari gibrid model alohida yondashuvlarga nisbatan yuqori samaradorlik ko‘rsatishini tasdiqladi. Bu natija o‘zbek tili kabi kam resursli tillar uchun maxsus moslashtirilgan gibrid yondashuvlarning amaliy ahamiyatini ko‘rsatadi. Ishlab chiqilgan aligner tizimi mashina tarjimasi tizimlari uchun sifatli trening ma’lumotlarini shakllantirishda, shuningdek tarjima strategiyalarini o‘rganish, terminologik bazalar yaratish va ikki tilli lug‘atlar tuzishda qo‘llanishi mumkin. Shu bilan birga, tizimning hozirgi imkoniyatlari mavjud korpus hajmi va o‘zbek tilining murakkab lingvistik xususiyatlari bilan cheklanadi. Kelgusidagi ishlar uchun bir nechta muhim yo‘nalishlar taklif etiladi. Birinchidan, parallel korpus hajmini kengaytirish va turli janr hamda uslubdagi matnlarni ko‘paytirish zarur. Ikkinchidan, neyron tarmoqlar va embeddinglarga asoslangan alignment modellarini joriy etish orqali sinonimiya va kontekstual mosliklarni yaxshiroq aniqlash mumkin. Uchinchidan, o‘zbek tilining agglutinativ xususiyatlarini hisobga oluvchi maxsus morfologik analizator yaratish alignment aniqligini oshirishi mumkin. To‘rtinchidan, ishlab chiqilgan tizimni real vaqt rejimida ishlovchi tarjima platformalariga integratsiya qilish uning amaliy qo‘llanish doirasini kengaytiradi.

Xulosa qilib aytganda, o‘zbek-rus parallel korpuslari asosida avtomatik aligner yaratish o‘zbek tilidagi NLP resurslarini rivojlantirish, mashina tarjimasini takomillashtirish va ikki tilli lingvistik tadqiqotlarni qo‘llab-quvvatlash uchun muhim yo‘nalishdir. Tadqiqotda olingan natijalar gibrid modelning samaradorligini ko‘rsatdi va kelgusida korpusni kengaytirish, morfologik tahlil hamda neyron texnologiyalarni birlashtirish orqali tizim sifatini yanada oshirish imkonini beradi.

FOYDALANILGAN ADABIYOTLAR

1. Abdurahmonova N. Mashina tarjimasining lingivistik asoslari. Akademnashr. Toshkent, 2012. – 30 b.
2. Xamroyeva Sh., Matyakubova N. Korpus lingvistikasida “Aligner” dasturlarning
3. Jurafsky D., Martin J. H. Speech and Language Processing. – 3-nashr. – Pearson, 2023. – 700 b.
4. Koehn P. Statistical Machine Translation. – Cambridge: Cambridge University Press, 2010. – 433 p.

5. Brown P. F., Della Pietra S. A., Della Pietra V. J., Mercer R. L. The mathematics of statistical machine translation: Parameter estimation // Computational Linguistics. – 1993. – Vol. 19, No. 2. – P. 263–311.
6. Gale W. A., Church K. W. A program for aligning sentences in bilingual corpora // Computational Linguistics. – 1993. – Vol. 19, No. 1. – P. 75–102.
7. Och F. J., Ney H. A systematic comparison of various statistical alignment models // Computational Linguistics. – 2003. – Vol. 29, No. 1. – P. 19–51.
8. Tiedemann J. Bitext Alignment. – Morgan & Claypool Publishers, 2011. – 120 p.
9. Proceedings of Machine Translation Summit IX. – New Orleans, USA, 2003. – (ilmiy maqolalar to‘plami).
10. Abdurakhmonova, N. Z. (2019). A two-level morphological analysis of the uzbek corpus. In Иностранная филология. Социальная и национальная вариативность языка и литературы (pp. 425-430).
11. Mengliev, D., Abdurakhmonova, N., Kholmurodova, I., Ibragimov, B., Latipova, G., & Kadirova, Z. (2025, October). Bridging dialectal variations in Uzbek texts: A comparative evaluation of modern approaches. In AIP Conference Proceedings (Vol. 3377, No. 1, p. 070001). AIP Publishing LLC.
12. Abdurakhmonova, N., Mengliev, D., Mardonova, S., Asadov, T., Xalilov, F., & Maxsudova, F. (2025, September). A Hybrid Named-Entity Recognition Algorithm for Ecological Documentation in Uzbekistan. In 2025 10th International Conference on Computer Science and Engineering (UBMK) (pp. 1778-1781). IEEE.
13. Zaynobiddin qizi, N. A., & Asliddin qizi, S. G. (2024, November). Theoretical Foundations of Corpus-based Uzbek-English Machine Translation. In 2024 IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE) (pp. 1650-1653). IEEE.
14. Abdurakhmonova, N., Alisher, I., & Sayfulleyeva, R. (2022, September). MorphUz: Morphological analyzer for the Uzbek language. In 2022 7th International Conference on Computer Science and Engineering (UBMK) (pp. 61-66). IEEE.
15. Abdurakhmonova, N., Alisher, I., & Toirova, G. (2022, September). Applying web crawler technologies for compiling parallel corpora as one stage of natural language processing. In 2022 7th International Conference on Computer Science and Engineering (UBMK) (pp. 73-75). IEEE.

IBORALARNI MATNDAN ANIQLASH VA IKKI TILLI EKVIVALENTLIGINI BAHOLASHDA ONTOLOGIYAGA ASOSLANGAN MATEMATIK MODEL

Elmira Nazirova[1], Nilufar Madetbayeva[2]
[1]Professor, t.f.d. (DSc),
Toshkent axborot texnologiyalari universiteti.
[2]tayanch doktorant, Raqamli texnologiyalar va sun'iy intellektni rivojlantirish ilmiy-tadqiqot institute.
https://doi.org/10.5281/zenodo.23253388

Annotatsiya. *Maqolada o'zbek va turk iboralarini matndan avtomatik aniqlash hamda ularning o'zaro ekvivalentlik darajasini baholash uchun ontologiyaga asoslangan matematik model taklif etiladi. Model ikki tilli bilimlar bazasiga tayanadi. Unda o'zbek tilining 2 098 ta va turk tilining 2 496 ta iborasi rasmiylashtirilgan, ular o'rtasida 3 193 ta tekshirilgan ekvivalentlik munosabati o'rnatilgan va to'rt darajali shkala qo'llanilgan. Ontologik bilim mashinaviy o'rganish usullariga ikki yo'l bilan kiritilgan, ya'ni belgilar to'plamini kengaytirish va aprior taqsimotni shakllantirish orqali. Haqiqiy korpusda o'tkazilgan tajribalar ontologik bilim kiritilganda aniqlash sifati barqaror o'sishini va bashorat noaniqligi kamayishini ko'rsatdi.*

Kalit so'zlar. *ibora, frazeologik birlik, ontologiya, ekvivalentlik darajasi, mashinaviy o'rganish, aprior taqsimot, kam resursli til.*

Annotation. *The paper proposes an ontology-based mathematical model for the automatic detection of Uzbek and Turkish idioms in running text and for grading the equivalence between them. The model relies on a bilingual knowledge base that formalises 2 098 Uzbek and 2 496 Turkish phraseological units, links them through 3 193 expert-verified equivalence relations and applies a four-level equivalence scale. Ontological knowledge is injected into machine learning methods in two ways, namely by extending the feature space and by shaping the prior distribution. Experiments on a real corpus show that the injection of ontological knowledge yields a stable improvement in detection quality and reduces predictive uncertainty.*

Keywords. *idiom, phraseological unit, ontology, equivalence degree, machine learning, prior distribution, low-resource language.*

Аннотация. *В статье предлагается онтологическая математическая модель для автоматического распознавания узбекских и турецких фразеологизмов в тексте и для оценки степени их эквивалентности. Модель опирается на двуязычную базу знаний, в которой формализованы 2 098 узбекских и 2 496 турецких фразеологических единиц, установлены 3 193 проверенных отношения эквивалентности и применена четырёхуровневая шкала. Онтологическое знание вводится в методы машинного обучения двумя способами, а именно через расширение признакового пространства и через формирование априорного распределения. Эксперименты на реальном корпусе показали устойчивый рост качества распознавания и снижение неопределённости прогноза.*

Ключевые слова. *фразеологизм, фразеологическая единица, онтология, степень эквивалентности, машинное обучение, априорное распределение, малоресурсный язык.*

So‘nggi yillarda mamlakatimizda ona tilimizni asrab-avaylash va uni zamonaviy texnologiyalar bilan uyg‘unlashtirish davlat siyosati darajasiga ko‘tarildi. O‘zbekiston Respublikasi Prezidentining 2019-yil 21-oktabrdagi PF-5850-son Farmonida o‘zbek tilining davlat tili sifatidagi nufuzi va mavqeyini tubdan oshirish vazifasi qo‘yilgan [1]. 2020-yil 20-oktabrdagi PF-6084-son Farmoni bilan esa 2020-2030-yillarda o‘zbek tilini rivojlantirish va til siyosatini takomillashtirish konsepsiyasi tasdiqlandi [2]. Ushbu hujjatlarda ona tilining ilm-fan va zamonaviy texnologiyalar tili sifatidagi nufuzini oshirish, uning boy imkoniyatlarini raqamli vositalar orqali namoyon etish alohida vazifa etib belgilangan.

Ikkinchi tomondan, Prezidentimizning 2021-yil 17-fevraldagi PQ-4996-son qarori sun’iy intellekt texnologiyalarini jadal joriy etish uchun keng shart-sharoitlar yaratdi va bu yo‘nalishdagi ilmiy tadqiqotlarni ustuvor darajaga olib chiqdi [3]. Mazkur ikki yo‘nalish kesishgan nuqtada o‘zbek tilini sun’iy intellekt vositalariga to‘laqonli o‘rgatish masalasi turadi va tilimizning eng nozik qatlami bo‘lgan iboralar bu masalaning ayni tugunini tashkil etadi.

Iboralar xalqimizning asrlar davomida shakllangan tafakkuri, urf-odatlari va milliy qadriyatlarining tildagi jonli ifodasidir. “Qo‘li ochiq”, “yuragi keng”, “tili shirin” kabi birikmalarda millatimizning dunyoqarashi, odob-axloq mezonlari va o‘zaro munosabat madaniyati mujassam [4, 5]. Shu bois ularni raqamli tizimlarga to‘g‘ri o‘rgatish faqat texnik vazifa emas, balki madaniy merosimizni raqamli davrga bekamu ko‘st olib o‘tish vazifasi hamdir.

Ayni paytda ibora, ya’ni frazeologik birlik, avtomatik tahlil uchun tilning eng murakkab birliklaridan biri sanaladi. Uning ma’nosi tarkibidagi so‘zlar ma’nolari yig‘indisiga teng emas, shu sababli matnni avtomatik qayta ishlash tizimlari iborani odatiy so‘z birikmasi sifatida talqin qiladi va uni so‘zma-so‘z tarjima qilib yuboradi [7]. Kam resursli tillarda muammo yanada keskin tus oladi, chunki iborani aniqlashga o‘rgatish uchun zarur bo‘lgan katta hajmdagi belgilangan to‘plamlar deyarli mavjud emas.

O‘zbek va turk tillari qarindosh bo‘lgani uchun ularning iboralarida o‘xshashlik ko‘p uchraydi, ammo bu o‘xshashlik ba’zan aldamchi bo‘lib chiqadi. Shakli deyarli bir xil, ma’nosi esa boshqacha bo‘lgan juftliklar tarjimada eng xavfli xatolarni keltirib chiqaradi. Shuning uchun ikki til iboralari orasidagi munosabatni oddiy lug‘aviy moslik sifatida emas, balki darajalangan ekvivalentlik sifatida rasmiylashtirish talab etiladi.

Mazkur ishning maqsadi iboralarni matndan aniqlash va ularning o‘zbek-turk ekvivalentligini baholash uchun ontologiyaga asoslangan matematik modelni taklif etish hamda bu modelning turli mashinaviy o‘rganish usullariga qo‘shgan haqiqiy hissasini o‘lchashdan iborat. Tadqiqotda tayyor statistik usullar qayta ixtiro qilinmaydi. Aksincha, ularga lingvistik bilim ontologiya ko‘rinishida kiritiladi va bu kiritishning samarasi bir xil sharoitda, juftlashgan sinovlar yordamida tekshiriladi.

Modelning asosini ikki tilli bilimlar bazasi tashkil etadi. U quyidagi ko‘rinishda rasmiylashtirilgan.

$$B = \langle I_{uz}, I_{tr}, A, E, \tau \rangle \qquad (1)$$

Bu yerda:

– I_{uz} va I_{tr} – o‘zbek va turk iboralari to‘plamlari;

– A – iboraning ma’noviy, tuzilishiy, uslubiy va mavzuviy belgilari to‘plami;

– E – ekvivalentlik munosabatlari to‘plami;

– τ – har bir juftlikka ekvivalentlik darajasini biriktiruvchi akslantirish.

O‘zbek ontologiyasi OWL formatida qurilgan va unda 2 098 ta ibora, 1 104 ta lemma hamda 28 310 ga yaqin uchlik mavjud. Iboralar hasComponentLemma, hasAnchorLemma, hasSemanticCategory, hasStructuralCategory, hasStylisticCategory, hasThematicCategory va hasTranslationEquivalence munosabatlari orqali bog‘langan. Turk ontologiyasi 2 496 ta iborani qamrab oladi va besh qirrali tasnifga ega. Ikkala ontologiya OntoLex-Lemon uslubidagi ekvivalentlik ko‘prigi orqali birlashtirilgan [14].

Ekvivalentlik shkalasi to‘rt darajadan iborat. E1 to‘liq ekvivalentlik, E2 qisman ekvivalentlik, E3 analog, E0 esa bir tilda muqobili topilmagan holatni bildiradi. Ekspertlar tomonidan tekshirilgan baza 3 193 ta juftlikni o‘z ichiga oladi. Baholash ishonchliligi 1 477 ta juftlikda o‘lchandi va Krippendorff kelishuv ko‘rsatkichi 0.967 ga teng bo‘ldi [15]. Solishtirish jarayonida shakli o‘xshash, ma’nosi mos kelmaydigan 14 ta juftlik aniq tasdiqlandi.

Ontologiyaga qo‘shimcha ravishda 4 408 qatorli sintaktik barqarorlik korpusi qurildi. U iboralarning matnda o‘zgarishga qanchalik moyilligini o‘lchaydi. Olingan moslashuvchanlik gradienti Vinogradov tasnifini empirik tarzda tasdiqladi, ya’ni chatishma 0.202, birikma 0.246 va qo‘shilma 0.290 qiymatlarini berdi [6]. Bu tartib iborani matndan qidirishda qanday darajadagi o‘zgarishga yo‘l qo‘yish kerakligini belgilab beradi.

Ekvivalentlik darajasini baholash masalasi tartiblangan tasniflash masalasi sifatida qo‘yildi. Har bir juftlik uchun uchta signal hisoblanadi va ular chiziqli birlashtiriladi.

$$\eta(u,t) = w_1\varphi_{sem}(u,t) + w_2\varphi_{ont}(u,t) + w_3\varphi_{form}(u,t) \quad (2)$$

Bu yerda:

– φ_{sem} – LaBSE ko‘p tilli modeli bergan vektorlar orasidagi ma’noviy yaqinlik [12];

– φ_{ont} – ontologiya grafidagi masofaga asoslangan tushunchaviy yaqinlik;

– φ_{form} – transliteratsiyani hisobga oluvchi shakliy o‘xshashlik;

– w_1, w_2, w_3 – o‘rgatish jarayonida topiladigan vazn koeffitsiyentlari;

– u va t – o‘zbek va turk iboralaridan tuzilgan juftlik.

Daraja ehtimolligi tartibli logistik regressiya orqali baholanadi.

$$P(E \le k \mid u,t) = \frac{1}{1 + exp(-(\theta_k - \eta(u,t)))} \quad (3)$$

Bu yerda:

– θ_k – k darajaga mos keluvchi chegara parametri;

– k – E1, E2 va E3 darajalari bo‘yicha tartib ko‘rsatkichi.

Iboralarni matndan aniqlash masalasida ontologik bilim ikki xil yo‘l bilan kiritildi. Birinchi yo‘lda ontologiyadan olingan belgilar chastota va o‘zaro axborot kabi statistik belgilarga qo‘shildi [8]. Bunda ontologiya tugunlari, ya’ni iboralar, lemmalar va ma’noviy turkumlar uchun TransE usulida vektor tasvirlar hisoblandi va ular belgilar qatoriga kiritildi [13]. Ikkinchi yo‘lda ontologiya modelning aprior taqsimotini shakllantirdi, ya’ni bilim o‘rgatish maqsad funksiyasining bir qismiga aylandi.

Bayescha logistik regressiya uchun aprior kovariatsiya matritsasi ontologik o‘xshashlik matritsasidan qurildi.

$$\Sigma_{onto} = K_{onto} + 10^{-5} \cdot I \quad (4)$$

Bu yerda:

– K_{onto} – iboralarning ontologik belgilari bo‘yicha hisoblangan o‘xshashlik matritsasi;

– I – birlik matritsa, qo‘shimcha had matritsani teskarilanuvchi holga keltiradi.

Gradient bosting oilasidagi ehtimollik modeli uchun esa ontologik aprior maqsad funksiyasiga jarima hadi sifatida kiritildi [11].

$$L_{MAP} = \frac{1}{2}\sum\left[ln(2\pi\sigma^2) + \frac{(y-\mu)^2}{\sigma^2}\right] + \lambda\sum\frac{(\mu-\mu_0)^2}{\sigma_0^2} \quad (5)$$

Bu yerda:

– μ va σ^2 – model bergan o‘rtacha qiymat va dispersiya;

– μ_0 va σ_0^2 – ontologiya bergan aprior o‘rtacha va dispersiya;

– λ – apriorning ta’sir kuchini boshqaruvchi giperparametr;

– y – haqiqiy belgi, ya’ni nomzodning chindan ibora ekanligi.

Aprior qiymatlar ontologiyadagi hasAnchorLemma munosabati va affiks qoidalari mos kelgan hollarda kuchli, qolgan hollarda betaraf tanlanadi. Shu tariqa lingvistik bilim modelga qattiq qoida sifatida emas, balki ehtimollik ko‘rinishida beriladi va u ma’lumot bilan raqobatlasha oladi.

Tajribalar uchun sakkizta badiiy asardan yig‘ilgan korpus ishlatildi. Uning hajmi 154 502 gap va 1.07 million lemmaga yaqin, lemmalarga ajratish UzMorphAnalyser morfologik tahlil vositasi yordamida bajarildi [18]. Etalon to‘plam ikki bosqichda tayyorlandi. Dastlab 400 ta gap qo‘lda belgilandi, keyinchalik to‘plam 860 ta gapga kengaytirildi. Belgilash chog‘ida takliflarning bir qismi qayta ko‘rib chiqildi, ya’ni 86 ta taklif musbatdan manfiyga va 21 ta taklif manfiydan musbatga o‘zgartirildi. Barcha solishtirishlar bir xil bo‘linishlarda va oqib chiqish xavfini yo‘qotgan holda o‘tkazildi, chunki belgilar faqat gap matnidan avtomatik tarzda hosil qilindi.

Natijalar quyidagi jadvalda keltirilgan.

1-jadval. Ontologik bilim kiritilishining aniqlash sifatiga ta’siri

Tajriba	Baholash sharoiti	Ontologiyasiz model	Ontologiyali model
Bayescha logistik regressiya	400 gap, 5 karrali nazorat	AUC 0.621, F1 0.531	AUC 0.738 ± 0.053, F1 0.737 ± 0.063
NGBoost va MAP apriori	860 gap, 5 boshlang‘ich qiymat o‘rtachasi	F1 0.210 ± 0.010, AUC 0.727 ± 0.021	F1 0.263 ± 0.018, AUC 0.744 ± 0.014

Bayescha modelda ontologik aprior barcha beshta bo‘linishda ijobiy o‘sish berdi, o‘rtacha AUC farqi 0.118 ni tashkil etdi. Kam uchraydigan iboralar guruhida vaznlarning o‘rtacha moduli 0.320 dan 0.845 gacha ko‘tarildi, ya’ni ontologiya aynan ma’lumot yetishmayotgan joyda kuchli ishladi. NGBoost tajribasida esa ontologik aprior oltita ko‘rsatkichning o‘rtachasida ustunlik qildi va bashorat noaniqligini umumiy holda 27.6 foizga, haqiqiy iboralar guruhida 15.9 foizga kamaytirdi.

Gradient bosting kaskadi alohida tajribada sinaldi. Ontologik belgilar bilan ishlaydigan kaskad aniqlik 0.710, qamrov 0.884 va F1 0.787 ko‘rsatkichlarini berdi. Ontologiyasiz gradient bosting bilan o‘tkazilgan juftlashgan McNemar sinovida farq 37 ga 13 nisbatda va p = 0.0011 darajasida statistik ahamiyatli bo‘ldi [10]. Tayanch vektorlar usuli asosidagi aniqlash modelida esa F1 ko‘rsatkichi 0.795 ga yetdi [9], ekvivalentlik darajasini tasniflashda 1 215 ta juftlikda tasodifiy o‘rmon usuli 0.586 makro-F1 natijasini berdi.

Shuni alohida ta’kidlash kerakki, jadvaldagi va matndagi raqamlar bir-biri bilan bevosita solishtirilmaydi, chunki tajribalar turli etalon to‘plamlarda va masalaning turli qo‘yilishida

bajarilgan. Solishtirish faqat har bir tajriba ichida, bir xil sharoitdagi ikki model orasida o‘tkazilgan.

Tadqiqotning asosiy cheklovi ontologiya qamrovi bilan bog‘liq. Etalon to‘plamdagi haqiqiy iboralarning 44 foizi ontologiyadagi qolip bilan mos keldi, kuchli aprior esa ularning 28 foiziga yetdi. Shunga qaramay ontologiya bergan signal ajratuvchi kuchga ega bo‘lib chiqdi, ya’ni kuchli apriorli guruhda musbat nomzodlar ulushi 39.0 foizni, betaraf guruhda esa 2.2 foizni tashkil etdi. Bu taxminan o‘n sakkiz baravar farq demakdir. Ikkinchi cheklov morfologik va sintaktik o‘zgarishlarga taalluqli. Ibora tarkibiga begona so‘z kirgan hollarda noaniqlikning kamayishi deyarli sezilmadi, demak bu guruh uchun qidiruv qoliplari yanada moslashuvchan bo‘lishi lozim.

Modelning amaliy tekshiruvi uchun ikki tilli tarjima ilovasi ishga tushirildi. Unda iboralar avval ontologiya orqali qidiriladi, mos ekvivalent topilmagan hollarda esa umumiy tarjima yo‘liga o‘tiladi. Ilova ustidagi kuzatuvlar nazariy modelda aniqlangan cheklovlarni takrorladi, ya’ni xatolarning asosiy qismi ontologiyada qayd etilmagan iboralar hisobiga yuzaga keldi.

Xulosa qilib aytganda, iboralarni avtomatik aniqlash va ikki tilli ekvivalentligini baholash masalasida lingvistik bilimni ontologiya ko‘rinishida rasmiylashtirish va uni statistik modelga aprior sifatida kiritish samarali yondashuv bo‘lib chiqdi. Bu yondashuv kam resursli tillar uchun ayniqsa qimmatli, chunki u katta hajmdagi belgilangan ma’lumot o‘rniga tuzilgan bilimni ishlatadi. Keyingi ishlarda ontologiya qamrovini kengaytirish va ibora tarkibiga begona so‘z kirgan holatlar uchun maxsus qoliplar ishlab chiqish rejalashtirilgan. Ontologiyaga asoslangan bunday modellar davlat siyosati darajasida qo‘yilgan vazifalarning, ya’ni ona tilimizni raqamli davr texnologiyalari bilan uyg‘unlashtirish va milliy qadriyatlarimiz xazinasi bo‘lgan iboralarni raqamli muhitda asrab qolish vazifalarining amaliy yechimiga qo‘shilgan hissa sifatida qaraladi [16, 17].

FOYDALANILGAN ADABIYOTLAR

1. O‘zbekiston Respublikasi Prezidentining 2019-yil 21-oktabrdagi PF-5850-son Farmoni. O‘zbek tilining davlat tili sifatidagi nufuzi va mavqeyini tubdan oshirish chora-tadbirlari to‘g‘risida // Qonunchilik ma’lumotlari milliy bazasi, lex.uz.
2. O‘zbekiston Respublikasi Prezidentining 2020-yil 20-oktabrdagi PF-6084-son Farmoni. Mamlakatimizda o‘zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish chora-tadbirlari to‘g‘risida // Qonunchilik ma’lumotlari milliy bazasi, lex.uz.
3. O‘zbekiston Respublikasi Prezidentining 2021-yil 17-fevraldagi PQ-4996-son qarori. Sun’iy intellekt texnologiyalarini jadal joriy etish uchun shart-sharoitlar yaratish chora-tadbirlari to‘g‘risida // Qonunchilik ma’lumotlari milliy bazasi, lex.uz.
4. Rahmatullayev Sh. O‘zbek tilining izohli frazeologik lug‘ati. Toshkent, O‘qituvchi, 1978.
5. Aksoy Ö. A. Atasözleri ve Deyimler Sözlüğü. İstanbul, İnkılâp Kitabevi, 1988.
6. Виноградов В. В. Об основных типах фразеологических единиц в русском языке. Москва, Издательство АН СССР, 1947.
7. Baldwin T., Kim S. N. Multiword Expressions // Handbook of Natural Language Processing, 2nd ed. Boca Raton, CRC Press, 2010. P. 267–292.
8. Church K., Hanks P. Word association norms, mutual information, and lexicography // Computational Linguistics, 1990. Vol. 16, No. 1. P. 22–29.
9. Cortes C., Vapnik V. Support-vector networks // Machine Learning, 1995. Vol. 20, No. 3. P. 273–297.

10. Chen T., Guestrin C. XGBoost. A scalable tree boosting system // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, ACM, 2016. P. 785–794.
11. Duan T., Anand A., Ding D. Y. et al. NGBoost. Natural gradient boosting for probabilistic prediction // Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020. P. 2690–2700.
12. Feng F., Yang Y., Cer D., Arivazhagan N., Wang W. Language-agnostic BERT sentence embedding // Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Dublin, ACL, 2022. P. 878–891.
13. Bordes A., Usunier N., Garcia-Duran A., Weston J., Yakhnenko O. Translating embeddings for modeling multi-relational data // Advances in Neural Information Processing Systems 26. Red Hook, Curran Associates, 2013. P. 2787–2795.
14. Cimiano P., McCrae J. P., Buitelaar P. Lexicon Model for Ontologies. Community Report. W3C, 2016.
15. Krippendorff K. Content Analysis. An Introduction to Its Methodology. Thousand Oaks, Sage Publications, 2004.
16. Abduraxmonova N. Kompyuter lingvistikasi. Darslik. Toshkent, Nodirabegim, 2021. 400 b.
17. Ismailov A. et al. A comparative study of stemming algorithms for use with the Uzbek language // 2016 3rd International Conference on Computer and Information Sciences (ICCOINS). IEEE, 2016.
18. Salaev U., Kuriyozov E., Gómez-Rodríguez C. UzMorphAnalyser. A morphological analysis model for the Uzbek language using inflectional endings // AIP Conference Proceedings, 2024. DOI 10.1063/5.0241461.

MUHAMMAD YUSUF MUALLIFLIK KORPUSI: SHAKLLANTIRISH TAMOYILLARI VA LINGVOSTATISTIK TAHLIL

Normamatova Mashhura
O‘zbekiston Milliy Universiteti
magistratura 1-bosqich talabasi.
https://doi.org/10.5281/zenodo.23253399

Annotatsiya: *Ushbu maqolada Muhammad Yusuf ijodiga oid mualliflik korpusini yaratishning nazariy va amaliy asoslari ko‘rib chiqiladi. Tadqiqot korpus lingvistikasi yondashuvi asosida amalga oshirilib, shoir she’riy matnlarini elektron formatda tizimlashtirish, ularni metama’lumotlar bilan boyitish hamda lingvostatistik tahlilga mos holga keltirish masalalari yoritilgan. Shuningdek, korpus asosida yuqori chastotali leksik birliklarning statistik va semantik tahlili amalga oshirilib, shoir ijodiga xos emotsional, falsafiy va milliy-madaniy konseptlar aniqlangan.*
Kalit so‘zlar: *korpus lingvistikasi, mualliflik korpusi, elektron matnlar, lingvostatistika, leksik tahlil, semantik maydon, chastotali tahlil, metama’lumotlar, o‘zbek she’riyati, lingvopoetika*
Abstract. *This article examines the theoretical and practical principles for creating an authorial corpus of Muhammad Yusuf ’s work. The research is based on a corpus linguistics approach and covers the systematization of the poet’s texts in electronic format, their enrichment with metadata, and their adaptation for linguostatistical analysis. Furthermore, a statistical and semantic analysis of high-frequency lexical units was conducted using the corpus, which identified the emotional, philosophical, and national-cultural concepts characteristic of the poet’s work.*
Keywords: *corpus linguistics, author corpus, electronic texts, linguostatistics, lexical analysis, semantic field, frequency analysis, metadata, Uzbek poetry, linguopoetics*

Jahon tilshunosligida kompyuter texnologiyalari imkoniyatlarining oshib borishi bilan korpus lingvistikasi muammolarini o‘rganish jarayoni ham jadal sur’atda rivojlanmoqda. Korpus lingvistikasi tadqiqotlari rivojlanishi bilan til korpuslarining bir nechta turlari ajratildi. Jumladan, o‘zbek tilshunosligida ham milliy, parallel va mualliflik korpuslari borasida bir qancha muhim tadqiqotlar va amaliy ishlar amalga oshirildi. Shunday bo‘lsada, korpus lingvistikasida jahon e’tirof etgan allomalar hamda milliy adabiyotimizning yirik namoyondalari mualliflik korpusi birliklarini tanlash, ularning lingvistik va ekstralingvistik ma’lumotlarini shakllantirish, lisoniy birliklarni semantik teglash, mualliflik korpusi uchun lingvistik ta’minot yaratish yechimini kutayotgan muhim masalalardan bo‘lib turibdi. Shu bois ushbu tadqiqotda Muhammad Yusuf milliy korpusi va mazkur korpus uchun shoir qalamiga mansub she’rlarning bazasini yaratish ko‘zda tutilgan.

Jahon tilshunosligida korpus lingvistikasi jadal rivojlanish bosqichidagi soha bo‘lib, ushbu sohada turli maqsad va yo‘nalishga bag‘ishlangan ko‘plab ilmiy tadqiqotlar mavjuddir. Xususan, Braun korpusi yaratuvchilari N. Frensis va G. Kucheralarning korpus tuzish tamoyillariga oid qarashlari [1] muhim nazariy asos bo‘lib xizmat qiladi. Rus tilshunosligida

esa V. P. Zaxarov, A. B. Kutuzov, Ye. V. Nedoshivina, V. V.Rikov hamda V.Plungyanlarning ishlari korpus lingvistikasi rivojiga kata hissa qo‘shgan.

O‘zbek korpusshunosligida katta hajmdagi matnlar bilan ishlash, maxsus elektron matnlar bazasining ilmiy tadqiqotlar va ta’lim jarayonidagi ahamiyati masalalari dastlab A. Po‘latov [2] va A. Rahimov [3] tomonidan ilgari surilgan. Keyinchalik Sh. Xamroyeva [4], A. Eshmo‘minov [5], D. Axmedova [6], O‘. Xoliyorov [7], N. Abduraxmonova [8], O. Abdullayeva [9], A. Raxmanova [10] hamda D. Elova [11] tomonidan til korpuslari va ularning turlari kengroq o‘rganildi. Ushbu tadqiqotlar natijasida mualliflik korpuslarini shakllantirishning lingvistik asoslari, tabiiy tilni qayta ishlash, sinonim birliklarni semantik teglash, leksik–semantik belgilash, o‘zbek tilining ta’limiy korpusini yaratish, elektron korpus modellarini ishlab chiqish va kompyuter texnologiyalari asosida korpus tuzish masalalari yoritildi.

Korpus materiali va manbalar bazasi – bu til korpusini yaratishda foydalaniladigan barcha matnlar majmui hamda ularning manbalarini o‘z ichiga olgan tizimdir [14]. Korpus materiali deganda turli janr, uslub va sohalarga oid tanlab olingan matnlar yig‘indisi tushuniladi. Ushbu tadqiqotda Muhammad Yusuf ijodiga oid matnlar asosida mualliflik korpusi shakllantirildi. Mualliflik korpusining yaratilishi tadqiqotning ilmiy aniqligi va ishonchliligini ta’minlashga xizmat qiladi, chunki unda bir muallifga tegishli bo‘lgan matnlar jamlanib, yagona tizim asosida saralanadi va qayta ishlanadi. Bu esa shoirning leksik boyligi, so‘z tanlash mahorati, uslubiy birliklardan foydalanish darajasi hamda semantik qatlamdagi o‘ziga xosliklarini chuqurroq yoritish imkonini beradi. Mualliflik korpusi, nafaqat badiiy asarlar tilini o‘rganish balki muallif tilining uslubiyatini o‘rganishda ham muhim ahamiyat kasb etadi.

Shu jihatdan qaraganda, mualliflik korpusining samarali ishlashi va ilmiy tahlillarda ishonchli natija berishi uning ma’lumotlar bazasini to‘g‘ri va tizimli shakllantirishga bevosita bog‘liqdir. Chunki, lingvopoetik va lingvokulturologik tahlillar aynan korpusdagi matnlarning raqamli bazasiga tayangan holda amalga oshiriladi [12: 45].

Shuningdek, mazkur korpus asosida olib boriladigan tahlillar orqali Muhammad Yusuf ijodida uchraydigan takrorlanuvchi birliklar, obrazli ifodalar, ramziy ma’nolar va milliy–madaniy komponentlarning qo‘llanish xususiyatlari aniqlanadi. Bu esa nafaqat shoirning individual uslubini tavsiflash, balki zamonaviy o‘zbek badiiy tilining rivojlanish jarayonidagi o‘rnini belgilashda ham muhim ahamiyat kasb etadi. Shu jihatdan, shakllantirilgan mualliflik korpusi tadqiqotning nazariy va amaliy natijalarini asoslashda tayanch manba sifatida xizmat qiladi. Mazkur mualliflik korpusining material bazasi Muhammad Yusuf ijodiga oid turli janr va formatdagi manbalarni qamrab olgan holda tizimli ravishda shakllantirildi. Xususan, korpus tarkibiga shoirning 327 ta she’riy asari, 2 ta dostoni hamda uning bevosita ishtirokida yozib olingan 20 ta videomaterial kiritildi. Ushbu materiallar majmuasi shoir ijodini nafaqat matn darajasida, balki multimodal yondashuv asosida – ya’ni audio va vizual komponentlar bilan uyg‘un holda o‘rganish uchun imkoniyat yaratadi.

Korpus materiallarini shakllantirishda asosiy manbalar sifatida shoirning “Quyoshga qarab oqqan suv”, “Saylanma”, “Alyorim qolur” hamda “Erka kiyik” nomli to‘plamlaridan foydalanildi. Shu bilan bir qatorda, mazkur to‘plamlarda mavjud bo‘lmagan yoki qo‘shimcha variant sifatida uchraydigan ayrim she’rlar ishonchli elektron manbalar orqali to‘ldirildi. Keltirilgan ilmiy yondashuvlarga muvofiq ravishda, mazkur tadqiqot doirasida tuzilgan mualliflik korpusida ham metama’lumotlarni kiritish va tizimlashtirishga alohida e’tibor berildi.

Korpusga kiritilgan she'riy matnlar yozuv tizimi jihatidan ham tartibga solindi: jami 327 ta she'rning 84 tasi kirill yozuvidagi manbalardan olinib, qolgan qismi lotin yozuvidagi matnlar asosida shakllantirildi. Bu holat o'zbek yozuv tizimining ikki grafik shaklda qo'llanishini hisobga olgan holda, matnlarni qiyosiy va tizimli o'rganish imkonini yaratadi. Manbalar bazasining umumiy holati va statistik taqsimoti quyida keltirilgan jadvallar orqali aks ettiriladi.

N	JANRI	SONI
1	She'r	**327**
2	Doston	**2**
3	Maqola	**20**
4	She'rlarning video va audio ko'rinishi	**100**
5	Muhammad Yusuf ishtirokidagi videolar	**20**

1-jadval. Korpus ma'lumotlar bazasi tarkibi.

She'riy materiallarni korpusga joylashtirishdan avval ularni tanlash va qayta ishlash jarayoni qat'iy ilmiy mezonlar asosida amalga oshirildi. Xususan, barcha matnlar ishonchli manbalar asosida saralab olinib, dastlab yagona formatga keltirildi hamda TXT (UTF–8 kodlash tizimida) fayl ko'rinishida saqlandi. Metama'lumotlar tizimi korpusning ajralmas tarkibiy qismi bo'lib, u nafaqat materiallarni tartibga solish, balki ularni turli mezonlar asosida saralash, guruhlash va qiyosiy tahlil qilish imkonini beradi. Xususan, metama'lumotlar yordamida matnlarni vaqt (nashr yili), yozuv tizimi (kirill/lotin), janr xususiyati yoki manbaiga ko'ra ajratib o'rganish mumkin. Bu esa korpus lingvistikasi doirasida olib boriladigan tadqiqotlarning aniqligi va tizimliligini ta'minlaydi. Korpusning asosiy vazifasi Muhammad Yusuf asarlarining lingvistik birliklarini statistik va kontekstual jihatdan tahlil qilish, muallifning individual uslubiy xususiyatlarini aniqlash hamda matnlar ustida avtomatlashtirilgan qidiruv tizimini amalga oshirishdan iboratdir. Korpus materiali elektron matnlar bazasi shaklida tashkil etildi. Bunda ijodkorning tanlab olingan to'plamlaridagi har bir she'ri alohida fayllar ko'rinishida saqlanib, korpus dasturlarida ishlash uchun mos format (txt UTF-8)ga keltirildi. Matnlarning standartlashtirilgan shaklda joylashtirilishi keyingi bosqichlarda konkordans, chastotaviy tahlil, n-gram va kalit so'zlar tahlilini amalga oshirish imkonini berdi.

Statistik birlik	Miqdori
Korpusga kiritilgan she'rlar	327
Umumiy tokenlar soni	41850
So'z shakllari	11823

2- jadval Muhammad Yusuf mualliflik korpusining umumiy statistik ko'rsatkichlari

Jadvalda Muhammad Yusuf mualliflik korpusining asosiy lingvostatistik ko‘rsatkichlari keltirildi. Korpus tarkibidagi har bir poetik matn alohida elektron fayl shaklida joylashtirildi hamda korpus dasturlarida lingvistik tahlilni amalga oshirish uchun moslashtirildi. Muhammad Yusuf asarlari asosida yaratilgan mualliflik korpusi ichki tuzilish jihatidan tizimli va izchil tashkil etilgan lingvistik baza hisoblanadi [13]. Korpus tarkibidagi barcha matnlar shoir ijodiga tegishli bo‘lib, har bir she’r alohida elektron matn sifatida ajratilgan va mustaqil lingvistik birlik sifatida qayta ishlanadi.

Muhammad Yusuf mualliflik korpusi interfeysi va bo‘limlar sonini belgilashda ijodiy merosidan kelib chiqqan holda ishlab chiqilgan. Muhammad Yusufning she’riy to‘plam va dostonlari va bundan tashqari internet manbalaridagi sara she’rlari tanlab olindi. Bular quyidagilar: “Quyoshga qarab oqqan suv”, “Saylanma”, “Alyorim qolur” hamda “Erka kiyik”. Korpusga qolgan asarlarini ham kiritishni kelgusi ilmiy ishlarimizda davom ettirish maqsad qilingan.

Muhammad Yusuf mualliflik korpusi faqat lingvistik tahlil uchun mo‘ljallangan matnlar majmuasi emas, balki foydalanuvchi uchun qulay navigatsion tuzilishga ega raqamli axborot tizimi sifatida ham ishlab chiqilgan. Korpus interfeysi o‘zaro bog‘langan bir nechta funksional bo‘limlardan iborat bo‘lib, u foydalanuvchiga shoir ijodini tizimli va bosqichma-bosqich o‘rganish imkonini beradi.

Korpus interfeysi quyidagi 7 asosiy bo‘lim asosida tashkil etilgan:

1. Bosh sahifa. Bosh sahifa korpusning kirish nuqtasi hisoblanadi. Unda Muhammad Yusuf ijodiga umumiy kirish, korpusning maqsadi va qisqacha tavsifi beriladi. Foydalanuvchi ushbu sahifa orqali tizimning boshqa bo‘limlariga yo‘naltiriladi.

2. Muallif haqida. Ushbu bo‘limda Muhammad Yusuf hayoti, ijodiy faoliyati, adabiy merosi va uning o‘zbek adabiyotidagi o‘rni haqida ma’lumotlar keltiriladi. Bu qism korpusning kontekstual va biografik asosini tashkil etadi.

3. Asarlari. Bu bo‘limda shoirning she’riy asarlari tizimli ravishda joylashtirilgan. Har bir she’r alohida matn sifatida berilib, ularni o‘qish va tahlil qilish imkoniyati yaratilgan.

4. Dostonlari. Muhammad Yusuf ijodidagi dostonlar alohida bo‘lim sifatida ajratilgan. Bu yerda yirik hajmli asarlar strukturaviy jihatdan tartiblangan holda taqdim etiladi.

5. Videolar. Ushbu bo‘limda shoir hayoti, ijodi va uning asarlari bilan bog‘liq video materiallar joylashtiriladi. Bu korpusning multimediaviy komponentini tashkil etadi.

6. Audiolar. Audiolar bo‘limida Muhammad Yusuf she’rlarining o‘qilishi, audio yozuvlar va ijro variantlari beriladi. Bu foydalanuvchiga poetik matnni eshitish orqali qabul qilish imkonini yaratadi.[13]

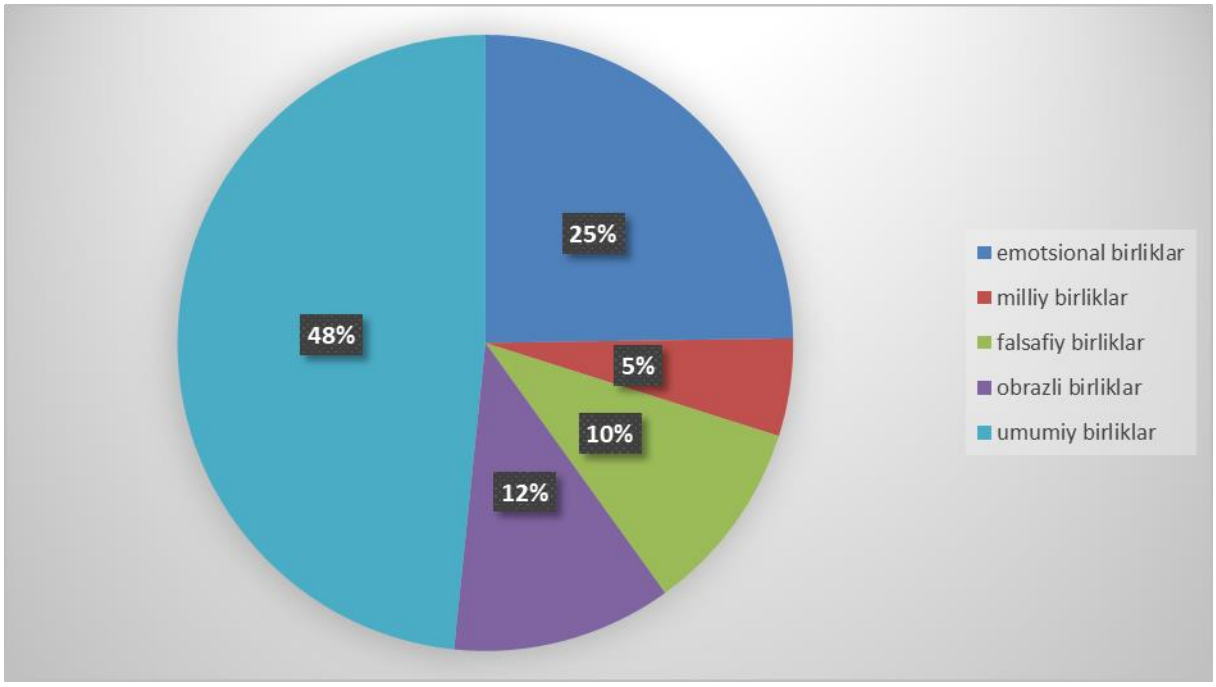
Muhammad Yusuf asarlari korpusini shakllantirishda lingvostatistik yondashuv muhim metodologik asoslardan biri hisoblanadi. Lingvostatistika matnni o‘rganishning o‘ziga xos ilmiy usuli bo‘lib, u til birliklarining miqdoriy xususiyatlarini aniqlash orqali ularning matndagi funksional va tizimli munosabatlarini ochib berishga xizmat qiladi.

Shu nuqtayi nazardan, Muhammad Yusuf she’rlari korpusi asosida olib borilgan tahlilda ham til birliklarining miqdoriy ko‘rsatkichlarini aniqlash asosiy bosqichlardan biri sifatida belgilandi. Xususan, ushbu bo‘limda korpus materiallari asosida so‘zlarning takrorlanish darajasi, ularning matn ichidagi taqsimlanishi hamda eng faol leksik birliklar aniqlanib, shoir ijodiga xos lingvopoetik xususiyatlar statistik asosda yoritildi. Muhammad Yusuf she’rlari korpusida eng yuqori chastotaga ega bo‘lgan so‘zlar quyidagi jadvalda keltiriladi:

Muhammad Yusuf she'rlari korpusi tahlili jarayonida eng yuqori chastotali o'nta leksik birlik ajratib olindi. Ushbu birliklar matnda takrorlanish darajasi yuqoriligi bilan birga, shoirning asosiy g'oyaviy va obrazli maydonlarini ham belgilab beradi.

T/r	So'z	Chastota	Semantik yo'nalish
1.	Bir	556	Umumiy/miqdoriy birlik
2.	Yig'lamoq	154	Emotsional holat
3.	Yor	122	Muhabbat obrazi
4.	Kun	114	Vaqt/borliq
5.	Qiz	106	Inson obrazi
6.	Dunyo	80	Falsafiy konsept
7.	Vatan	71	Milliy konsept
8.	Jon	67	Hissiy markaz
9.	Yolg'on	61	Axloqiy-falsafiy tushuncha
10.	Gul	54	Estetik obraz

3- jadval . Muhammad Yusuf she'rlarida yuqori chastotali takroriy so'zlar semantik tahlili.



2 - diagramma. Muhammad Yusuf she'rlarida yuqori chastotali so'zlarning semantik taqsimoti (foizlarda).

Olingan natijalar shuni ko'rsatadiki, yuqori chastotali leksik birliklar Muhammad Yusuf she'riyatida muayyan semantik markazlarni shakllantiradi. Xususan, "*yor*", "*jon*", "*yig'lamoq*" kabi birliklar emotsional-ekspressiv qatlamni tashkil etib, lirik subyektning ichki kechinmalarini ifodalashda faol ishtirok etadi.

"*Dunyo*", "*kun*", "*yolg'on*" kabi birliklar esa falsafiy-mavhum tasvirlarni shakllantirib, borliq va inson munosabatiga oid poetik mushohadalarni aks ettiradi. "Vatan" leksemasi milliy-ma'naviy konsept sifatida shoir ijodining markaziy g'oyaviy yo'nalishlaridan birini ifodalaydi.

Shu bilan birga, "*gul*" va "*qiz*" kabi obrazli birliklar estetik tasvir vositasi sifatida poetik matnning badiiy qatlamini kuchaytiradi. "*Bir*" va "*kun*" kabi umumiy leksik birliklarning takrorlanishi esa matnning tabiiy nutqqa yaqinligini va sintaktik soddaligini ta'minlaydi.

Umuman olganda, takroriy birliklar Muhammad Yusuf she'riyatida tasodifiy hodisa emas, balki poetik tizimning markaziy elementlaridan biri bo'lib, emotsional, falsafiy va estetik qatlamlarning uyg'unlashuvini ta'minlaydi.

Tadqiqot natijalari Muhammad Yusuf mualliflik korpusi o'zbek korpus lingvistikasini rivojlantirish, milliy adabiy merosni raqamlashtirish hamda kelgusidagi lingvopoetik, va kompyuter lingvistikasi yo'nalishidagi ilmiy izlanishlar uchun mustahkam nazariy-amaliy manba vazifasini bajarishini tasdiqlaydi.

FOYDALANILGAN ADABIYOTLAR

1. Френсис Н., Кучера Г. Вычислительный анализ современного американского варианта английского языка. – М., 1967; Синклер Д. Предисловие к книге "Как использовать корпуса в преподавании иностранного языка"/ Д. Синклер [Электронный ресурс]. – Режим доступа: http://www/ruscorpora.ru/corpora-infro.html.
2. Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011. – 520 б
3. Rahimov A. Kompyuter lingvistikasi asoslari. – Toshkent: Akademnashr. 2011. – 160 b.
4. Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фал. док. (PhD) дис. автореф. – Қарши, 2018. – 53 б.
5. Эшмўминов А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол. фан. бўйича фал. док. (PhD) дисс. – Қарши, 2019. – 140 б.
6. Ахмедова Д. Атов бирликларини ўзбек тили корпуслари учун лексик-семантик теглашнинг лингвистик асос ва моделлари: Филол. фан. бўйича фал. док. (PhD) дисс. – Бухоро, 2020. – 156 б
7. Холиёров Ў. Ўзбек тили таълимий корпусини тузишнинг лингвистик асослари. Филол. фан. бўйича фал. док. (PhD) дисс. автореф. – Термиз, 2021. – 53 б.
8. Абдурахмонова Н. Электрон корпусларнинг компьютер моделлари: Филол. фан. док. дисс. автореф. – Тошкент, 2021. – 74 б.
9. Абдуллаева О. Ўзбек тилининг интернет ахборот матнлари корпусини шакллантиришнинг назарий ва амалий асослари. Филол. фан. бўйича фал. док. (PhD) дисс. афтореф. – Тошкент, 2022. – 54 б.
10. Raxmanova A. O'zbek tili Milliy korpusini yaratishda kompyuter usullari. Filol. fan. bo'yicha fal. dok. (PhD) diss. – Farg'ona, 2022. – 52 b.
11. Элова Д. Ўзбек тили корпуси бирликларининг услубий тегларини яратиш тамойиллари ва лингвистик таъминот. Филол. фан. бўйича фал. док. (PhD) дисс. афтореф. – Тошкент, 2022. – 65 б.
12. Abduraxmonova N.Z. Kompyuter lingvistikasi (darslik). – Toshkent: Globeedit, 2021. – 260 b.
13. Abdurakhmonova, N., & Rajabova, K. (2025, September). Stages of Creating Uzbek Language Multimedia Corpus. In *2025 10th International Conference on Computer Science and Engineering (UBMK)* (pp. 1628-1633). IEEE.

14. Abduraxmonova, N., & Shamsiyeva, G. (2026). KO 'P TILLI KORPUSLAR ASOSIDA MASHINA TARJIMASI TIZIMLARINI TAKOMILLASHTIRISH BOSQICHLARI. *COMPUTER LINGUISTICS: PROBLEMS, SOLUTIONS, PROSPECTS*, *1*(1).
15. Zaynobiddin qizi, N. A., & Asliddin qizi, S. G. (2024, November). Theoretical Foundations of Corpus-based Uzbek-English Machine Translation. In *2024 IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE)* (pp. 1650-1653). IEEE.

TABIIY TILNI QAYTA ISHLASH VA BILIMLAR GRAFLARI: INTEGRATSIYALASHGAN YONDASHUV VA ZAMONAVIY TENDENSIYALAR

Nazokat O'rinboyeva
O'zbekiston Milliy universiteti tadqiqotchisi
nnazokat71@gmail.com
https://doi.org/10.5281/zenodo.23253432

Tabiiy tilni qayta ishlash (NLP - Natural Language Processing) — bu sun'iy intellekt va tilshunoslikning kesishmasidagi soha bo'lib, kompyuterlarga inson tilini (matn va nutq ko'rinishida) tushunish, talqin qilish va tahlil qilish imkonini beradi. NLP kompyuter yangiliklar matnini shunchaki "shovqin" (noise) yoki harflar to'plami deb qabul qiladi.

Tabiiy tilni qayta ishlash texnologiyasi yagona algoritmik yechim emas, balki kompyuter lingvistikasi, matematik modellashtirish va chuqur o'qitish (Deep Learning) metodologiyalarining o'zaro integratsiyasiga asoslangan ko'p komponentli tizimdir[1]. Ushbu texnologiyaning amaliy tatbiqi kundalik hayotda ovozli navigatsiya tizimlari, raqamli asistentlar va chatbotlar kabi intellektual interfeyslar shaklida namoyon bo'lmoqda.

NLP texnologiyalarining funksional imkoniyatlari faqatgina servis xizmatlari bilan cheklanib qolmaydi. Hozirgi kunda ushbu soha ulkan hajmdagi axborot oqimlarini qisqa vaqt ichida tahlil qilish, ularning semantik mazmunini interpretatsiya qilish va eng muhim ma'lumotlarni ajratib olish (abstractive/extractive summarization) kabi strategik vazifalarni bajarishda hal qiluvchi rol o'ynamoqda.

Ayniqsa, matnli ma'lumotlardan bilimlarni ajratib olish (Knowledge Extraction) va ularni ontologik tuzilmalarga keltirish jarayonida yuzaga kelgan global muammolarni bartaraf etishning eng samarali texnologik yechimi bo'lib xizmat qiladi. Bu esa yirik biznes tizimlarini boshqarish va ilmiy-tadqiqot jarayonlarini avtomatlashtirishda yuqori samaradorlikni ta'minlaydi.

Tabiiy tilni qayta ishlash texnologiyalari bugungi kunda axborot almashinuvi va foydalanuvchi interfeyslarining ajralmas qismiga aylandi. Ularning asosiy amaliy yo'nalishlarini quyidagicha guruhlash mumkin[2]:

Neyron mashina tarjimasi (NMT) modellariga asoslangan real vaqtdagi tarjima texnologiyasi (Simultaneous Machine Translation) nutq yoki matnni manba tilidan maqsadli tilga minimal kechikish (latency) bilan o'girish jarayoni bo'lib, texnologiyaning an'anaviy tarjimadan farqi shundaki, u gapning to'liq yakunlanishini kutmasdan, semantik birliklarni segmentlar (chunks) asosida qayta ishlaydi[3]. Hozirgi kunda bu jarayon asosan Transformer arxitekturasiga tayanadi. Undagi "diqqat mexanizmi" (attention mechanism) matnning uzoq masofali bog'liqliklarini soniyalar ichida tahlil qilib, tarjimaning aniqligini ta'minlaydi[4]. Bu usul nafaqat so'zma-so'z o'girish, balki kontekstual ma'noni saqlab qolish imkonini beradi.

[1] What is Natural Lab nguage Processing? And it's applications | LinkedIn
[2] Natural Language Processing (NLP): How It Works, Benefits & Real-World Examples
[3] Vaswani A. et al. Attention is all you need // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – Pp. 5998-6008.
[4] Cho K. et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation // arXiv preprint arXiv:1406.1078. – 2014. – Pp. 1-15.

Tabiiy tilni qayta ishlashning fundamental yo'nalishlaridan biri bo'lgan matn klassifikatsiyasida berilgan matn fragmentini oldindan belgilangan toifalardan biriga avtomatik biriktirish jarayonidir. Ushbu texnologiyaning eng keng tarqalgan amaliy namunasi elektron pochta xizmatlaridagi spam-filtrlash [5] tizimlaridir.

Spam-filtrlarni yaratishda xabarlarning lingvistik va statistik xususiyatlari tahlil qilinib, klassifikatsiya jarayonida an'anaviy ravishda quyidagi modellardan foydalaniladi:

Naive Bayes (Nayiv Beyes) modeli ehtimollar nazariyasiga asoslangan bo'lib, matndagi so'zlarning paydo bo'lish chastotasi bo'yicha xabarning "spam" yoki "ham" (foydali) ekanligini hisoblaydi. Bu model katta hajmdagi ma'lumotlar bilan tezkor ishlashi tufayli yuqori samaradorlik ko'rsatadi [6].

Zamonaviy NLP tizimlarining eng murakkab va ko'p qirrali yo'nalishlaridan biri intellektual muloqot interfeyslari hisoblanadi. Ovozli yordamchilar va chatbotlar foydalanuvchi bilan tabiiy tilda muloqot qilish uchun bir nechta yuqori texnologik jarayonlarning sintezidan foydalanadi. Ushbu tizimlarning ishlash mexanizmini quyidagi asosiy bosqichlarga ajratish mumkin[7] :

- Nutqni avtomatik aniqlash (Automatic Speech Recognition - ASR) bosqichi akustik signallarni (ovozni) raqamli matn ko'rinishiga o'tkazadi.
- Tabiiy tilni tushunish (Natural Language Understanding - NLU) matnning semantik mazmunini tahlil qilish, foydalanuvchi niyatini (intent recognition) aniqlash va muhim birliklarni (entity extraction) ajratib olish bosqich bo'lib bevosita NER (Named Entity Recognition) texnologiyalari bilan uzviy bog'liqdir [8].
- Muloqotni boshqarish (Dialogue Management) jarayoni kontekstdan kelib chiqib mantiqiy javob algoritmini shakllantirishga xizmat qiladi.
- Nutq generatsiyasi (Text-to-Speech - TTS) matnli javobni sun'iy intellekt yordamida tabiiy ovozga aylantirish vazifasini bajaradi.

Ushbu tizimlarda foydalanuvchi niyatini aniqlash (intent recognition) muloqotning samaradorligini belgilovchi asosiy omil hisoblanadi. Bunda chuqur o'qitish (Deep Learning) modellaridan foydalanib, foydalanuvchi so'rovidagi yashirin semantik bog'liqliklar tahlil qilinadi.

FOYDALANILGAN ADABIYOTLAR

1. 1.Abualigah L. et al. Text Summarization: A Brief Review // Recent Advances in NLP and AI. – 2020. – Vol. 2. – Pp. 1-15.
2. 2.Aggarwal C. C., Zhai C. A Survey of Text Classification Algorithms // Mining Text Data. – Springer, Boston, MA, 2012. – Pp. 163-222.
3. 3.Allahyari M. et al. Text Summarization Techniques: A Brief Survey // arXiv preprint arXiv:1707.02268. – 2017. – Pp. 1-9.

[5] Aggarwal C. C., Zhai C. A Survey of Text Classification Algorithms // Mining Text Data. – Springer, Boston, MA, 2012. – Pp. 163-222.

[6] Metsis V., Androutsopoulos I., Paliouras G. Spam Filtering with Naive Bayes – Which Naive Bayes? // CEAS. – 2006. – Vol. 17. – Pp. 28-69.

[7] Jurafsky D., Martin J. H. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. – Pearson, 2023.

[8] McTear M. Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots. – Morgan & Claypool Publishers, 2020. – Pp. 45-112.

4. 4.Bahdanau D., Cho K., Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate // International Conference on Learning Representations (ICLR). – San Diego, 2015. – Pp. 1-15.
5. 5.Bao Y. et al. The Effect of Text Preprocessing on Machine Learning Algorithms // Proceedings of the IEEE International Conference on Big Data. – 2019. – Pp. 4621-4629.
6. 6.Cho K. et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation // arXiv preprint arXiv:1406.1078. – 2014. – Pp. 1-15.
7. 7.Fox C. Lexical Analysis and Stoplists // Information Retrieval: Data Structures and Algorithms. – Prentice-Hall, Inc., 1992. – Pp. 102-130.
8. 8.Gao J., Galley M., Li L. Neural Approaches to Conversational AI // Foundations and Trends in Information Retrieval. – 2019. – Vol. 13, No. 2-3. – Pp. 127-298.

NUTQ KORPUSINI ANNOTATSIYALASHNING LINGVISTIK ASOSLARI: SAMARQAND VILOYATI SHEVASI MISOLIDA

O'rinboyeva Roza Toliboy qizi
O'zMU tayanch doktoranti
orinboyevaroza@gmail.com
https://doi.org/10.5281/zenodo.23253534

Annotatsiya. *Mazkur maqolada o'zbek tili shevalariga, xususan, Samarqand shevasiga asoslangan nutq texnologiyalari (STT va TTS) uchun mo'ljallangan audio korpuslarni annotatsiyalashning lingvistik asoslari tahlil qilinadi. Maqolada annotatsiyalash jarayonining ketma-ket bosqichlari – audio sifatini nazorat qilish va segmentatsiya, so'zma-so'z transkripsiya tamoyili, orfografik me'yorlash, dialektal xususiyatlarni normallashtirmaslik tamoyili, kod almashinish hodisasining belgilanishi – nazariy jihatdan asoslanadi va amaliy misollar bilan yoritiladi. Tadqiqot natijalari shevaga xos fonetik-leksik xususiyatlarni yo'qotmagan holda sifatli va takrorlanuvchan annotatsiya tizimini yaratish mumkinligini ko'rsatadi.*

Kalit so'zlar: *nutq korpusi, annotatsiyalash, Samarqand shevasi, nutqni matnga aylantirish (STT), nutq sintezi (TTS), kod almashinish.*

Abstract. *This article analyzes the linguistic foundations of annotating audio corpora designed for speech technologies (STT and TTS) based on Uzbek dialects, specifically the Samarkand dialect. The study provides a theoretical rationale and practical examples for sequential annotation stages, including audio quality control and segmentation, word-for-word transcription principles, orthographic normalization, the principle of non-normalization of dialectal features, and the tagging of code-switching phenomena. The findings demonstrate that a high-quality and reproducible annotation framework can be developed without compromising dialect-specific phonetic and lexical nuances.*

Keywords: *speech corpus, annotation, Samarkand dialect, speech-to-text (STT), text-to-speech (TTS), code-switching.*

Аннотация. *В данной статье анализируются лингвистические основы аннотирования аудиокорпусов, предназначенных для речевых технологий (STT и TTS) на основе диалектов узбекского языка, в частности Самаркандского говора. В работе теоретически обосновываются и иллюстрируются практическими примерами последовательные этапы процесса аннотирования: контроль качества и сегментация аудио, принцип дословного транскрибирования, орфографическая нормализация, принцип ненормализации диалектных особенностей, а также разметка явления переключения кода. Результаты исследования показывают возможность создания качественной и воспроизводимой системы аннотирования без утраты специфических фонетико-лексических особенностей диалекта.*

Ключевые слова: *речевой корпус, аннотирование, Самаркандский говор, распознавание речи (STT), синтез речи (TTS), переключение кода.*

KIRISH

So'nggi yillarda sun'iy intellektga asoslangan nutq texnologiyalarining rivojlanishi ko'p jihatdan sifatli, izchil va lingvistik asoslangan hamda annotatsiyalangan korpuslarning

mavjudligiga bog‘liq. Bunday korpuslarni yaratish inson qo‘l mehnatini talab qiladi. O‘zbek tili boshqa tillar singari, ko‘plab hududiy shevalar tizimiga ega bo‘lib, ular fonetik, leksik va qisman morfologik jihatdan adabiy tildan farqlanadi. Bugunga qadar o‘zbek adabiy tili asosida shakllangan nutq texnologiyalari sifatida UzbekVoiceAI, Uznutq kabi platformalarni keltirish mumkin [Abduraxmonova N., O‘rinboyeva R., 2026:63]. Ular to‘liq adabiy tilga moslashgani uchun sheva vakillari nutqini yetarli aniqlikda sintez qila olmasligi mumkin. Jahon korpus lingvistikasida ham dialektal ma’lumotlarning yetishmasligi asosiy ochiq muammo sifatida belgilanadi. Xuddi shunday, kam resursli tillar va shevalar uchun TTS tizimlarini yaratish yo‘nalishida ham qator amaliy loyihalar mavjud. Jumladan, tunis arab shevasi uchun ishlab chiqilgan va qo‘lda diakritikalangan TunArTTS korpusi [Laouirine I. va b., 2024], Xitoyning Xaynan shevalari bo‘yicha olib borilgan tadqiqotlardagi kundalik so‘zlashuv nutqini yozib olish va to‘liq annotatsiyalash orqali yaratilgan maxsus korpus [Qi Z. va b., 2025], shuningdek, barcha dunyo tillari uchun yaratilgan Common Voice [Abduraxmonova N., O‘rinboyeva R., 2026:57] hamda Hindiston shevalari uchun yaratilgan ULCA loyihasi [https://ulca.bhashini.gov.in/] bugungi kunda dunyo shevalarini o‘rganishga qaratilgan muhim tadqiqotlardan hisoblanadi.

O‘zbek tili materialida olib borilgan so‘nggi tadqiqotlar ham ushbu yo‘nalishning jadal rivojlanayotganini ko‘rsatadi. Xususan, o‘zbek tilida imlo xatolarini aniqlash va tuzatishga qaratilgan Uzbekcorrector.uz platformasi, o‘zbek tilida yaratilgan audio korpus [Urazaliyeva M., 2026] hamda bolalar nutqi korpusi [Umirzaqova M.,2026] shular jumlasidandir.

Nutq korpusi annotatsiyasi – akustik signal bilan lingvistik birliklar o‘rtasidagi muvofiqlikni ta’minlovchi ko‘p qatlamli jarayon bo‘lib, u fonetika, dialektologiya, kompyuter lingvistikasi va korpus lingvistikasi tutashgan nuqtada yotadi. Annotatsiyalash uchun muhim jihatlardan biri audio yozuvda mavjud bo‘lgan nutqiy hodisalarning aniq ifodalanishidir. Masalan, og‘zaki nutqda ikkilanish yoki hayratlanishni ifodalash uchun m-m-m-, a-a-a-a, hmmm kabi tovushlarni cho‘zib talaffuz qilish kuzatiladi. Shuningdek, yuqori aniqlikka erishish uchun bir xil hodisani korpus bo‘ylab bir xil qoida asosida belgilash talab etiladi. Bu bazani dasturga yuklashda kelib chiqadigan muammolarning oldini olishga xizmat qiladi. Keyingi bosqichda foydalaniladigan model (STT yoki TTS) uchun zarur bo‘lgan barcha lingvistik ma’lumotlar annotatsiyada aks ettirilishi lozim.

Dialektologik nuqtai nazardan, Samarqand viloyati Qo‘shrabot tumani shevasi, asosan, o‘zbek tilining qarluq lahjasiga mansub bo‘lib, adabiy tilning shakllanishida ishtirok etgan markaziy shevalar guruhiga kiradi. Shu bilan birga, unda fors-tojik tili bilan uzoq davom etgan hududiy aloqadorlik natijasida yuzaga kelgan fonetik va leksik o‘ziga xosliklar, jumladan, unlilar tizimidagi farqlar va bir qator dialektizmlar kuzatiladi.

Nutq korpusini annotatsiyalash yaxlit, ammo shartli ravishda bir necha izchil bosqichga ajratiladigan jarayondir. Quyida Samarqand shevasi materiallari asosida ushbu jarayonning tizimlashtirilgan bosqichli tahlili keltiriladi.

1. Audio sifatini nazorat qilish va segmentatsiya.

Annotatsiyalashning dastlabki bosqichi audio materialning akustik sifatini baholashdan iborat. Yengil fon shovqini va past darajadagi qaytariqlar qabul qilinishi mumkin bo‘lsa-da, signalning kesilishi, kuchli darajadagi nutqlar ustma-ustligi va elektron artefaktlar korpus sifatiga jiddiy putur yetkazadi, shu sababli bunday yozuvlar rad etiladi. Segmentatsiya, ya’ni audioni kichik bo‘laklarga (odatda 5–15 soniya) bo‘lish bosqichida lingvistik yaxlitlik tamoyiliga rioya qilinadi: segment chegarasi so‘z ichida emas, balki tabiiy pauza yoki nafas olish oralig‘ida

belgilanadi. Bu tamoyil prosodik butunlikni saqlash va model o‘qitishda kontekst yo‘qotilishining oldini olish uchun zarur.

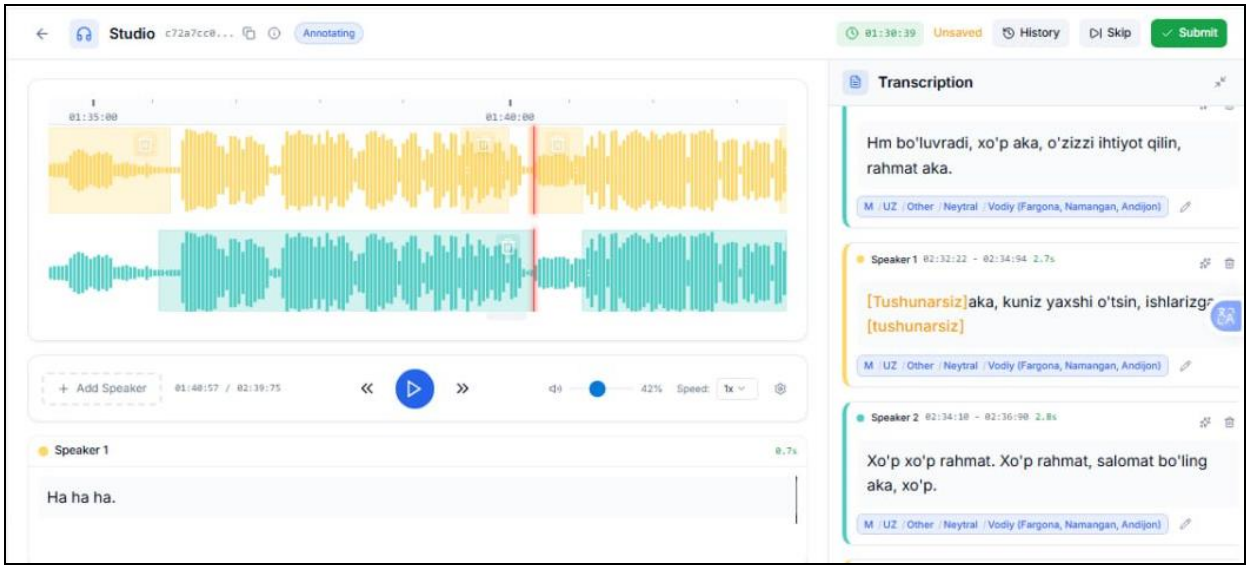
1-rasm. Audioni segmentlarga ajratish jarayoni.

2. So‘zma-so‘z transkripsiya qilish.

Annotatsiyaning markaziy metodologik tamoyillaridan biri akustik-grafik moslik tamoyilidir. Ushbu tamoyilga ko‘ra, annotator matnni sheva variantida aynan eshitilgan shaklga



muvofiq yozadi. Grammatik xatolar, ikkilanish tovushlari (mm, uh, ee), takrorlashlar va duduqlanish hodisalari ham to‘liq qayd etiladi. Masalan, so‘zlovchi *“Hmm, aligiii, manziliz qaydaydi?”* desa, annotatsiyada aynan shu shakl aks ettirilishi lozim. Bu yondashuv akustik model uchun haqiqiy nutq signaliga mos keluvchi, tabiiy nutqiy hodisalarni o‘zida mujassam etgan ma’lumotlar bazasini shakllantirish imkonini beradi.



2-rasm. So‘zma-so‘z transkripsiya qilish jarayoni.

3. Orfografik me’yor: talaffuz va imlo o‘rtasidagi muvozanat.

O‘zbek tilida ayrim fonemalar (masalan, x va h) so‘zlashuv nutqida, odatda, farqlanmay talaffuz qilinadi, biroq annotatsiya jarayonida bunday holatlarda so‘zning imlo lug‘atidagi me’yoriy shakli tanlanadi. Masalan, so‘zlovchi bo‘g‘iz undoshi bilan “*habar*” yoki “*harakter*” tarzida talaffuz qilsa ham aynan shu imloviy shakl qayd etiladi. Bunday yondashuv akustik

model (talaffuzga asoslangan) bilan til modeli (leksik-semantik to‘g‘rilikka asoslangan) o‘rtasidagi muvozanatni ta’minlaydi. Shuningdek, gap boshi, atoqli otlar (Toshkent, Samarqand) va qisqartmalar (MIB, MCHJ) katta harf bilan yozilishi, raqamlar esa to‘liq so‘z shaklida ifodalanishi (masalan, 2026 – ikki ming yigirma olti) belgilangan me’yoriy qoidalar sirasiga kiradi.

4. Dialektal xususiyatlarning annotatsiyada aks ettirilishi.

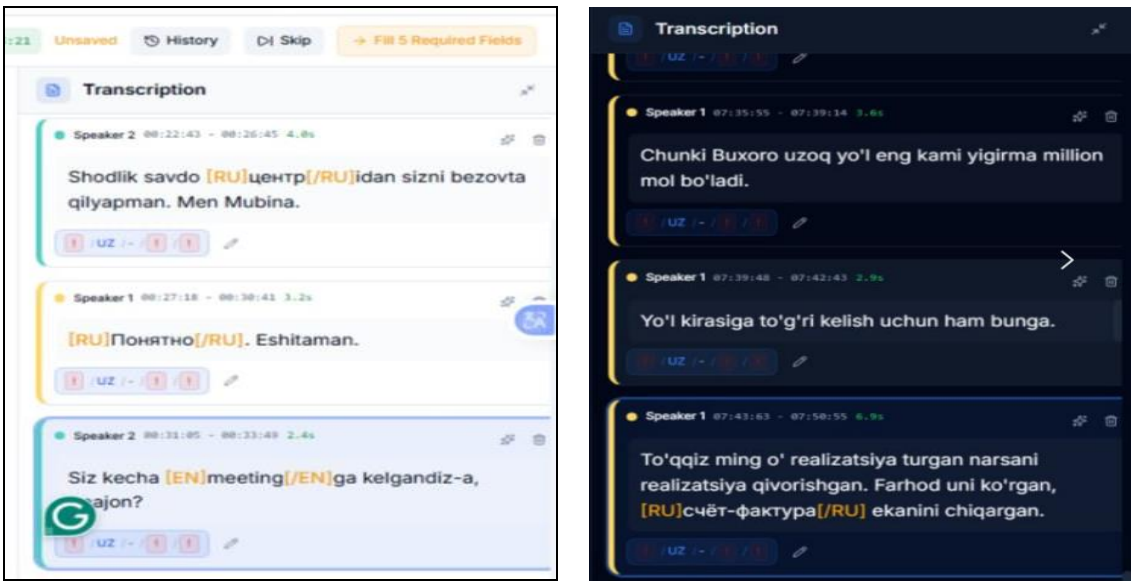
Sheva materialini annotatsiyalashdagi eng muhim metodologik tamoyil – dialektizmlarni adabiy tilga “to‘g‘rilamaslik”dir. Bu tadqiqotning asosiy maqsadi hisoblanib, STT tizimi uchun yaratilayotgan bazada variativlikni ta’minlab beradi. Samarqand shevasi misolida bunday yondashuv, xususan, unli tovushlarning talaffuzdagi qisqarishi, ba’zi fe’l shakllarining o‘ziga xos qo‘shimchalanishi hamda mahalliy leksik birliklarning saqlanishi kabi hodisalarni to‘liq qayd etish imkonini beradi. Masalan, Samarqand viloyatining Qo‘shrabot tumani shevasida so‘z boshida kelgan “*o‘*” unlisidan oldin “*v*” undoshi qo‘shib talaffuz qilinadi: *o‘tirgandan => vo‘tirgandan, o‘ttiz => vo‘ttis, o‘rtahol => vo‘rtahol.*

Kelishik qo‘shimchalarining *-di, -ti (-ni), -din, -tin (-ning), -g‘a (-ga), -ta (-da), -tan, -nan (-dan)* shaklida qo‘llanishi kuzatiladi. Masalan, *xabardi, maktaptin, ukangnan, qishta* kabi. Ayrim leksemalarda ham fonetik o‘zgarishlar kuzatiladi. Masalan, *katata* (buva), *kasa* (kosa), *taraba* (taroveh) kabi. Ushbu hodisalarning annotatsiyada ayni holda berilishi ishning sifatini oshirishga xizmat qiladi.

5. Kod almashinish hodisasi.

O‘zbekiston hududida o‘zbek-rus va qisman o‘zbek-ingliz kod almashinishi keng tarqalgan bo‘lib, bu hodisa annotatsiya nuqtai nazaridan alohida murakkablik tug‘diradi. Annotatsiya tizimida boshqa tildagi birliklar maxsus teglar ([RU]...[/RU], [EN]...[/EN], [KZ]...[/KZ], [TJ]...[/TJ]) yordamida belgilanadi. Bunda muhim metodologik farqlash – o‘zlashgan so‘zlar (masalan, kompyuter, bank, telefon) bilan haqiqiy kod almashinish o‘rtasidagi chegaraning aniqlanishidir: agar so‘z o‘zbek tilining izohli lug‘atiga kirgan bo‘lsa, unga teg qo‘yilmaydi, aks holda u kod almashinish sifatida teglanadi. O‘zbek tilining agglyutinativ tabiati kod almashinish annotatsiyasida qo‘shimcha qiyinchilik yaratadi: chet tili o‘zagiga o‘zbekcha qo‘shimcha qo‘shilganda, faqat o‘zak teg ichiga olinadi, qo‘shimcha esa tegdan tashqarida qoladi. Masalan, *“Shodlik savdo [RU]центр[/RU]idan sizni bezovta qilyapman. Ismim Mubina”* jumlasida *центр* so‘zi kirill alifbosida, rus tilida hamda [RU]…[/RU] tegi ostida berilgan bo‘lib, “*markaz*” ma’nosida qo‘llanilgan. Bu og‘zaki nutqda ko‘p qo‘llaniladigan holatlardan biri.

Bundan tashqari, rus tilidagi birliklar faqat kirill alifbosida yozilishi shart – lotin transliteratsiyasi (masalan, “xarasho” yoki “privet”) annotatsiya me’yorlariga zid hisoblanib, ishning sifatiga ta’sir ko‘rsatadi.



3-rasm. Kod almashinish hodisasining teglanishi.

Bundan tashqari, annotatsiyaning sifatli amalga oshirilishi annotatorlar o‘rtasidagi izchillikka ham bog‘liq omil hisoblanadi. Amaliyot shuni ko‘rsatadiki, murakkab holatlarda (masalan, tushunarsiz nutq parchalari yoki bir necha til aralashgan gaplar) qaror qabul qilishdagi farqlar, faqat tushunarsiz hollarda [tushunarsiz] tegidan foydalanish, imloviy va punktuatsion xatoliklarga yo‘l qo‘ymaslik, aynan eshitilgan holda transkripsiya qilish ishning muhim qismlaridan biri hisoblanadi. Buning uchun til modellaridan foydalanish maqsadga muvofiq [12].

6. Korpus hajmi va annotatorlar o‘rtasidagi izchillik.

Sheva materiallariga asoslangan nutq korpusini yaratishda dastlab 100 soatlik audio materialdan foydalanildi. Taxminiy hisob-kitoblarga ko‘ra, 100 soatlik audio material 700–900 ming so‘z birligini tashkil etadi (o‘rtacha nutq tezligi 120–150 so‘z/daqiqa asosida); bu ko‘rsatkich nutqning tezroq yoki sekinroq ekanligiga ko‘ra o‘zgarishi mumkin. Ushbu taxminni tekshirish maqsadida annotatsiyalangan 3 daqiqalik audio namunasi hisoblab chiqildi: nutq tezligi o‘rtacha bo‘lgan ushbu segmentda so‘zlar soni 432 tani tashkil etdi (bu ko‘rsatkichda takror kelgan so‘zlar ham alohida hisoblangan), ya’ni ~144 so‘z/daqiqa – bu natija yuqorida keltirilgan diapazon ichida joylashib, taxminiy hisob-kitobning ishonchliligini tasdiqlaydi.

Korpus hajmi bilan bir qatorda, annotatsiya sifatini ta’minlashning muhim mezonlaridan biri annotatorlar o‘rtasidagi izchillikni (inter-annotator agreement) miqdoriy baholashdir. Loyihada 12 nafar annotator ishtirok etganligi sababli, ikki annotator uchun mo‘ljallangan Cohen’s kappa koeffitsienti qo‘llanilmaydi; buning o‘rniga ikkitadan ortiq baholovchi ishtirokida qo‘llash mumkin bo‘lgan Krippendorff’s alpha koeffitsientidan foydalaniladi. Bunga o‘xshash yondashuv, jumladan, ko‘p annotatorli nutq korpuslarida (THAI-SER) ham qo‘llanilgan, bunda annotatorlar soni birlikdan-birlikka farqlanishi sababli aynan Krippendorff’s alpha tanlab olingan [Wongpithayadisai J. va b. 2025].

Annotatorlar oldida turgan qarorlarning tabiatiga ko‘ra izchillik quyidagi yo‘nalishlarda baholanadi. Birinchidan, o‘zlashgan so‘zlar bilan haqiqiy kod almashinish hodisasi o‘rtasidagi farqlash – nominal (kategorial) xususiyatga ega qaror bo‘lib, bunda annotator har bir birlikni ikki sinfning biriga (“o‘zlashgan” yoki “kod almashinish”) kiritadi; bu turdagi qarorlar uchun Krippendorff’s alphaning nominal ma’lumot uchun mo‘ljallangan varianti qo‘llaniladi. Ikkinchidan, x va h fonemalarining orfografik me’yorga muvofiqlashtirilishi ham nominal tabiatga ega bo‘lib, annotator izohli lug‘atga tayangan holda ikki variantdan birini (“x” yoki “h”

shakli) tanlaydi; bu qaror ham xuddi shu usul bilan baholanadi. Uchinchidan, segmentatsiya chegaralarining belgilanishi boshqa tabiatga ega – bu nominal tasniflash emas, balki vaqt o‘qi bo‘yicha chegara belgilash masalasidir, shu sababli unga standart kappa yoki alpha koeffitsientlarini to‘g‘ridan-to‘g‘ri qo‘llash metodologik jihatdan noto‘g‘ri bo‘ladi. Bunday hollarda annotatorlar tomonidan belgilangan chegaralar orasidagi vaqt farqi (masalan, ±200 millisoniya oralig‘ida mos kelish "kelishilgan" deb hisoblanadi) yoki chegaralar bilan chegaralangan segmentlar kesishmasiga asoslangan overlap-based F1 ko‘rsatkichidan foydalanish maqsadga muvofiqdir.

Ushbu ko‘rsatkichlar – korpus hajmining empirik tasdig‘i va uch yo‘nalishli izchillik baholash tizimi – annotatsiya jarayonining ishonchliligini miqdoriy asoslash va kelgusi tadqiqotlar uchun taqqoslanadigan bazaviy ko‘rsatkich yaratish imkonini beradi.

Xulosa qilib aytganda, o‘zbek tili shevalariga, xususan, Samarqand shevasiga asoslangan nutq korpuslarini annotatsiyalash – bir qator o‘zaro bog‘liq lingvistik va texnik bosqichlarni o‘z ichiga olgan murakkab, ko‘p qatlamli jarayondir. Audio sifatini nazorat qilish va segmentatsiyadan boshlab, so‘zma-so‘z transkripsiya, orfografik me'yorlash, dialektal xususiyatlarni normallashtirmaslik, kod almashinishni belgilashgacha bo‘lgan barcha bosqichlar yagona metodologik tizim asosida amalga oshirilganda, natijada shevaning fonetik-leksik o‘ziga xosligini saqlagan, lingvistik jihatdan izchil va nutq texnologiyalari uchun funksional jihatdan yetarli korpus yaratiladi. Kelgusi tadqiqotlarda ushbu metodologik modelni o‘zbek tilining boshqa shevalari (Xorazm, Farg‘ona vodiysi, Qashqadaryo-Surxondaryo) materialida qo‘llash va annotatorlar o‘rtasidagi izchillik ko‘rsatkichlarini miqdoriy baholash, shuningdek, sheva materialida nomlangan birliklarni (NER) belgilash masalasi alohida taraqqiyot yo‘nalishi sifatida belgilanadi.

FOYDALANILGAN ADABIYOTLAR

1. Abduraxmonova N. Korpus lingvistikasi: darslik. – Toshkent: Globeedit, 2023. – 400 b.
2. Abduraxmonova N., O‘rinboyeva R. O‘zbek tili nutqiy sintezatori uchun lingvistik ta'minot yaratish tamoyillari: monografiya. – Globe Edit, 2026. – 106 b.
3. Abduraxmonova N.Z. Kompyuter lingvistikasi: darslik. – Toshkent: Nodirabegim, 2021. – 394 b.
4. Abduraxmonova N.Z. O‘zbek tili elektron korpusining kompyuter modellari: monografiya. – GlobeEdit, 2021.
5. Gatiatullin A., Suleymanov D., Prokopyev N., Abdurakhmonova N. "Turkic Morpheme": From the Portal to the Linguistic Platform // Aliev R.A. et al. 12th World Conference "Intelligent System for Industrial Automation" (WCIS-2022): Lecture Notes in Networks and Systems. – Cham: Springer, 2024. – Vol. 912. – DOI: https://doi.org/10.1007/978-3-031-53488-1_22
6. Laouirine I., Kammoun R., Bougares F. TunArTTS: Tunisian Arabic Text-To-Speech Corpus // *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024).* – 2024. – Pp. 16879–16889.
7. Qi Z., Li F., Long H. Research on optimal deep learning modeling in HaiNan dialect recognition // *Scientific Reports*. – 2025. – Vol. *15*, № 1.
8. Umirzaqova M. O‘zbek bolalar nutqi korpusi va uni teglash masalalari: magistrlik dissertatsiyasi. – Toshkent, 2026. – 90 b.

9. Urazaliyeva M. O'zbek tili audiokorpusining lingvistik ta'minotini yaratish: filol. fan. fals. dok. (PhD) diss. avtoreferati. – Toshkent, 2026. – 58 b.
10. Wongpithayadisai J. et al. THAI Speech Emotion Recognition (THAI-SER) corpus [Elektron resurs]. – arXiv:2507.09618, 2025. – URL: https://arxiv.org/abs/2507.09618
11. ULCA loyihasi [Elektron resurs]. – URL: https://ulca.bhashini.gov.in/
12. Abdurakhmonova, N., Alisher, I., & Sayfulleyeva, R. (2022, September). MorphUz: Morphological analyzer for the Uzbek language. In 2022 7th International Conference on Computer Science and Engineering (UBMK) (pp. 61-66). IEEE.

FOMA DASTURIY MUHITIDA QARAQALPOQ TILIDAGI SON SO‘Z TURKUMINI MODELLASHTIRISH

Aziz O‘temisov
PhD, dots. Qoraqalpoq davlat universiteti, Amaliy filologiya kafedrasi
https://doi.org/10.5281/zenodo.23253568

Annotatsiya. *Ushbu maqolada qoraqalpoq tilidagi son so‘z turkumini Foma muhitida semantik-struktur xususiyatlari asosida modellashtirish masalasi ko‘rib chiqilgan. Sonlarning til sistemasida tutgan o‘rni, ularning turlari, grammatik belgilari va so‘z turkumi sifatida o‘zaro munosabati, son so‘z turkumining fst, xfst fayllarida tuzilishi, ularga qo‘yiladigan talablar, morfoanalizator fayllarining o‘zaro bog‘liqliklari ilmiy asosda tahlil qilinadi.*

Kalit so‘zlar: *qoraqalpoq tili, foma, xfst, fst, lexc, ma’lumotlar bazasi, til modellari, morfoanalizator, son so‘z turkumi.*

I. KIRISH. Foma — bu chekli avtomatlar va erkin chekli avtomatlar ifodalari bilan ishlashga hamda ularni yaratishga mo‘ljallangan maxsus dasturiy muhit bo‘lib, asosan tabiiy tillarning morfologik strukturalarini modellashtirishda keng qo‘llaniladi [1]. Ushbu dasturiy qurol tilshunoslikda morfoanalizatorlarni ishlab chiqish, so‘z turkumlarining o‘zgarish qoidalarini aniqlash va ularni dasturiy darajada qayta ishlash jarayonlarini avtomatlashtirishda muhim o‘rin egallaydi. Foma muhitining asosiy afzalliklaridan biri shundaki, u leksik birliklarni qoidalar asosida samarali tahlil qilish, fst va xfst, lexc formatidagi fayllar bilan ishlash imkoniyatini beradi [2].

Qoraqalpoq tilidagi so‘z turkumlari, jumladan, morfologik o‘zgarishlarga kam uchraydigan, o‘ziga xos forma yasovchi affikslarga ega bo‘lgan son so‘z turkumining semantik-struktur xususiyatlarini raqamli muhitda to‘g‘ri modellashtirishda ushbu qurol ulkan ahamiyatga ega.

II. MORFOANALIZATORNING SISTEMAVIY KODI VA FONETIK QOIDALARINING BERILISHI.

Avtomatlashtirilgan sistemamiz dastur ishga tushish paytidagi barcha jihatlarni hisobga olgan holda ishlab chiqildi. Biz bu ishda so‘z turkumlari bazasini bir fayl (lexicon.lexc) ichida joylashtirdik. Fonetik qoidalarni o‘z oldiga alohida bir faylga (phonology.foma) joylashtirdik:

KARmorph.foma kodi:

```
# 1. Tokenizatorni yuklash
source tokenizer.xfst

# 2. Fonologiya qoidalarini yuklash
source phonology.foma

# 3. Leksikondan yuklash
read lexc lexicon.lexc
define Lexicon;

# 5. Tarmoqlarni birlashtirish
```

```
# Faqat so‘z darajasidagi tahlil qiluvchi
define MorphoAnalyzer Lexicon .o. Phonology;

# Punktuatsiyaning morfoanalizer bilan parallel ishlashini ta’minlaymiz
# Ya’ni: sistema morfologik tahlil qilsin, yoki punktuatsiyani o‘zi tanisin
define MorphoWithPunct [MorphoAnalyzer | Punct];

# Butun matn darajasidagi tahlil qiluvchi (Tokenizator + Morfologiya + Punct)
define FullSystem QqTokenizer .o. MorphoWithPunct;

# 6. Natijani saqlash
clear stack
push FullSystem
push MorphoWithPunct
save stack KARmorph.bin
```

Phonology.foma fayli qoidalar to‘plami kodi:

```
# 1. Alifbo
define Vowel [ a | á | ı | i | o | ó | u | ú | e ];
define BackVowel [ a | o | u | ı ];
define FrontVowel [ e | i | ó | ú | á ];
define Consonant [ b | d | f | g | ǵ | h | j | k | l | m | n | p | q | r | s | t | v | w | x | y | z | c | ń | sh | ch ];

# 2. Qoidalar
define PtoB p -> b || _ %^ Vowel;
define QtoG q -> ǵ || _ %^ Vowel;
define KtoG k -> g || _ %^ Vowel;

# Garmoniya (-ler, -ge)
define VowelHarmony a -> e || FrontVowel [Consonant | %^ ]* %^ [ l | ǵ | q | n ] _;

# Markerni o‘chirish
define Cleanup %^ -> 0;

# 3. Kompilyatsiya
read lexc lexicon.lexc
define Lexicon;
define Phonology PtoB .o. QtoG .o. KtoG .o. VowelHarmony .o. Cleanup;

# Leksikon va Fonologiyani birlashtirish
regex Lexicon .o. Phonology;

# Natijani saqlash
save stack KARmorph.bin
```

III. SONLARNING MORFOLOGIK XUSUSIYATI VA MA’LUMOTLAR BAZASI.

Qoraqalpoq adabiy tilidagi son so‘z turkumi boshqa so‘z turkumlari singari o‘ziga xos

xususiyatlarga ega. Eng avvalo, tilimizda ularning soni boshqa so‘z turkumlariga solishtirganda o‘ta cheklangan. Masalan, boshqa turkiy tillardagidek umumiy o‘rta son so‘zlarning soni 20 ta so‘zdan iborat: *bir, eki, úsh, tórt, bes, altı, jeti, segiz, toǵız, on, jigirma, otız, qırq, eliw, alpıs, jetpis, seksen, toqsan, júz, mıń* [3]; shuningdek, bugungi kunda internatsional sonlar quyidagilardan iborat: *million, milliard, trillion* h.t.b. Mana, shu ko‘rsatilgan son so‘zlar orqali boshqa barcha qo‘shma son so‘zlar yasaladi va ular cheksizlikkacha davom etaveradi.

Boshqa so‘z turkumlariga solishtirganda son morfologik o‘zgarishga ko‘p uchramaydigan so‘z turkumidir. U ko‘pincha ot so‘zlarning oldida turib (jamlovchi, taqsim sonlardan va kasr sonlarning ayrim turlaridan tashqari), bosh kelishikda qo‘llaniladi. Har xil so‘z o‘zgartiruvchi affikslarni ular faqat substantivlashganda va boshqa hamda ayrim sharoitlarda qabul qila oladi.

- Sonlardan keyingi ot so‘zlar ko‘plik affiksini qabul qilmaydi, hamma vaqt birlik sonda qo‘llaniladi.
- Bu so‘z turkumi hech vaqt boshqa so‘z turkumidan yasalmaydi, aksincha, o‘zi boshqa so‘z turkumlariga o‘ta oladi.
- Son faqat o‘zigagina xos bo‘lgan forma yasovchi affikslarga ega: -(ı)nshı // -inshi, -aw // -ew va h.k.
- O‘zining morfologik sostavi, shuningdek, har xil forma yasovchi va so‘z o‘zgartiruvchi affikslarni qabul qilishi jihatidan ham son bir qator o‘ziga xos xususiyatlarga ega [4]. Shu o‘rinda son so‘z turkumiga tegishli affikslar haqida ham ma’lumot berib o‘tamiz. Asosan, ular:

№	Sonning forma yasovchi affikslari	Misollar
1.	-ınshı // -inshi	*Birinshi, ekinshi, úshinshi, tórtinshi, besinshi...*
2.	-aw // -ew	*Birew, ekew, úshew, tórtew, besew, altaw, jetew.*
3.	-lap // -lep	*Onlap, júzlep, mıńlap, millionlap, milliardlap.*
4.	-laǵan // -legen	*Onlaǵan, jigirmalaǵan, otızlaǵan, júzlegen.*
5.	-day // -dey (-tay // -tey)	*Onday, qırqtay, jetpistey, júzdey, mıńday.*
6.	-dan // -den (-tan // -ten, -nan // -nen)	*Birden, ekiden, jigirmadan, qırqtan, mıńnan*

1-jadval. Sonning forma yasovchi affikslari

Foma muhitida sonlarni so‘z turkumlari bazasiga kiritishda ayrim xususiyatlarni hisobga olish kerak. Bular:

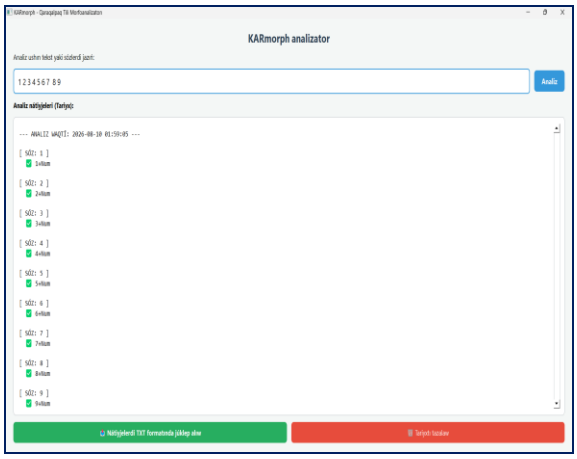
- so‘z bilan ifodalanadigan sonlar bazasining shakllantirilishi, ularning soni biz yuqorida ta’kidlagandek barcha turkiy tillarda umumiy soni 20 tadan oshmaydi, biroq internatsional sonlarni ham e’tibordan chetda qoldirib bo‘lmaydi: *bir, eki, úsh, tórt, bes, altı, jeti, segiz, toǵız, on, jigirma, otız, qırq, eliw, alpıs, jetpis, seksen, toqsan, júz, mıń, million, milliard, trillion, trilliard, kvadrilion, kvadrilliard*;
- arab raqamlarining ham matn tarkibida kelib aniq son ma’nosini anglatishi, ularning soni bor-yo‘g‘i o‘nta bo‘lsa ham matn tarkibida bir-biri bilan birikib, cheksizlikkacha davom etaverishi. Morfoanalizator bu sonlarni ham matn deb qabul qilishi, ularning 0=nol, 1=bir, 2=eki, 3=úsh, 4=tórt, 5=bes, 6=altı, 7=jeti, 8=segiz, 9=toǵız deb qabul etilishi tahlil jarayoni uchun o‘ta ahamiyatli;
- morfoanalizatorlarni ishlab chiqishda arab raqamlari bilan bir qatorda rim raqamlarining ham matn tarkibida son ma’nosini bildirib kelishini hisobga olish va ularni tegishli leksik ma’nolarga (I=bir=1, II=eki=2, III=úsh=3 v.b.) umumlashtirib olib borish tahlil jarayonining aniqligini ta’minlaydi.

Yuqoridagi xususiyatlarni hisobga olgan holda sonlarning ma'lumotlar bazasi quyidagi tartibda berildi [5]:

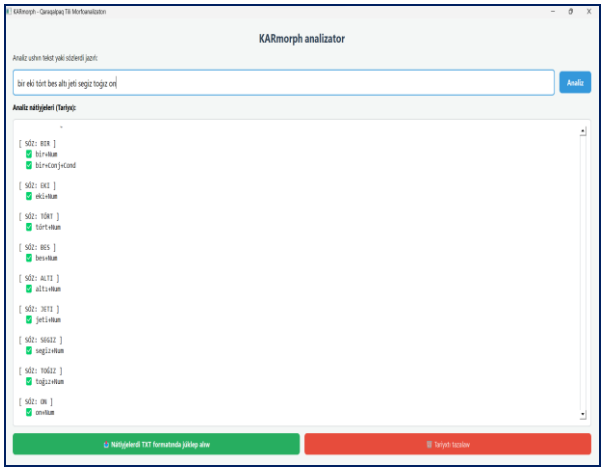
So'z bilan ifodalanadigan sonlar	Arab raqamlari bilan ifodalanadigan sonlar	Rim raqamlari bilan ifodalanadigan sonlar
bir Num_INFL;	1 Num_Ar;	I Num_Rum;
eki Num_INFL;	2 Num_Ar;	II Num_Rum;
úsh Num_INFL;	3 Num_Ar;	III Num_Rum;
tórt Num_INFL;	4 Num_Ar;	IV Num_Rum;
bes Num_INFL;	5 Num_Ar;	V Num_Rum;
altı Num_INFL;	6 Num_Ar;	VI Num_Rum;
jeti Num_INFL;	7 Num_Ar;	VII Num_Rum;
segiz Num_INFL;	8 Num_Ar;	VIII Num_Rum;
toǵız Num_INFL;	9 Num_Ar;	IX Num_Rum;
on Num_INFL;	0 Num_Ar;	X Num_Rum; ! 10

2-jadval. Son so'z turkumining lexc ma'lumotlar bazasi

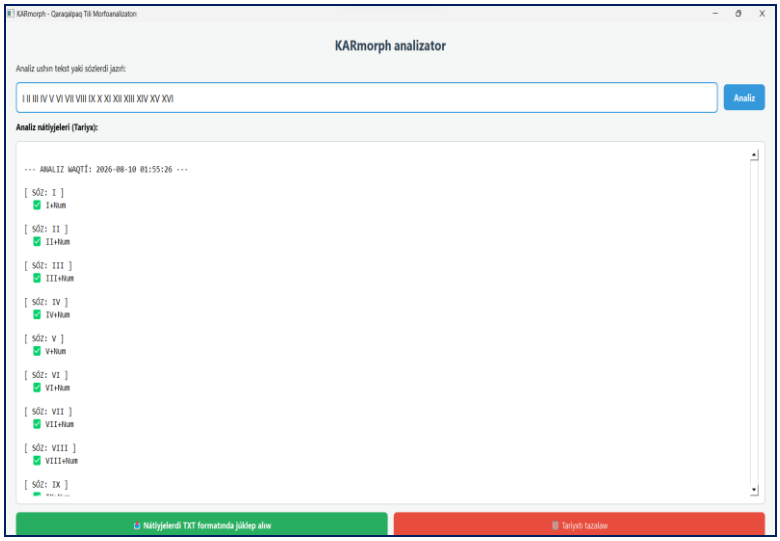
IV. NATIJA. Tuzilayotgan KARmorph.foma analizatorida yuqorida ko'rsatilgan uchinchi jadvaldagi ma'lumotlar asosida tub sonlarni test qilib ko'ramiz:



1-rasm. *So'zlar bilan ifodalanadigan sonlar*



2-rasm. *Rim raqamlari bilan kelgan sonlar*



3-rasm. *Arab raqamlari bilan ifodalanadigan sonlar*

Keltirilgan har uchala holatda ham analizator ma'lumotlar bazasidagi sonlarni chiqarib bera oldi. Ularni tanidi. Biroq, matn tarkibida o'ziga har xil qo'shimchalarni ulab kelish holatlarini tahlil qilish bosqichlarini modellashtirish hali to'liq yakuniga yetgani yo'q. Shu sababli bunday misollarning tahlil natijasi ko'rsatilmadi.

V. XULOSA. Ushbu tadqiqot qoraqalpoq tilidagi son so'z turkumini Foma dasturiy muhitida modellashtirishning nazariy va amaliy ahamiyatini ko'rsatib berdi. Olingan natijalar quyidagi xulosalarga asoslanadi:

- **Modellashtirishning samaradorligi:** foma muhiti qoraqalpoq tili morfologiyasini, jumladan, sonlarning forma yasovchi affikslarini sistemali tarzda aks ettirishda yuqori aniqlikni ta'minlaydi.
- **Texnik xulosalar:** fst va xfst, lexc formatidagi fayllarni ishlab chiqish morfoanalizatorlarning avtomatlashtirilgan tizimlarda ishlashini optimallashtiradi hamda til darajasidagi murakkab morfologik jarayonlarni qayta ishlash imkoniyatini beradi.
- **Perspektiv yo'nalishlar:** son so'z turkumi uchun yaratilgan bu lingvistik modellar kelajakda kompyuter lingvistikasi, avtomatik tarjima tizimlari va korpus lingvistikasini rivojlantirishda asosiy resurs sifatida xizmat qiladi.

Xulosa qilib aytganda, sonlarning semantik va struktur xususiyatlarini dasturiy muhitda to'g'ri modellashtirish qoraqalpoq tilining raqamli resurslarini yanada boyitishga va lingvistik texnologiyalarni yangi bosqichga olib chiqishga xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR

1. A. Bekbergenov *Qaraqalpoq tilinde sanlıqlar.* Nókis – 1976 (8-11-betlar)
2. A. Otemisov and J. Sharbaev, "Development of a Karakalpak Morphological Analyzer in the Foma Environment," *2026 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, Yekaterinburg, Russian Federation, 2026, pp. 1-4, doi: 10.1109/USBEREIT70063.2026.11580433.
3. S. Gebeyehu, W. Wolde, and Z. S. Shibeshi, "Information extraction model from Ge'ez texts," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 30, no. 2, pp. 787–795, May 2023.
4. M. Hulden, "Finite-state morphology: A foma tutorial," 2024. [Online]. Available: https://fomafst.github.io/morphtut.html [Accessed: Jan. 15, 2026].
5. M. Hulden, "Foma: An open-source finite-state toolkit and library," 2024. [Online]. Available: https://fomafst.github.io/ [Accessed: Jan. 15, 2026].

HUQUQIY HUJJATLAR ASOSIDA NER TIZIMINI YARATISH VA MATNLARNI TASNIFLASH

Odilova Olmosxon Dilshodjon qizi
2-kurs magistratura talabasi
Mirzo Ulugbek nomidagi O'zbekiston Milliy Universiteti
https://doi.org/10.5281/zenodo.23253590

Annotatsiya. *Ushbu maqolada huquqiy hujjatlar asosida Named Entity Recognition (NER) tizimini yaratish va huquqiy matnlarni avtomatik tasniflash masalalari tahlil qilinadi. Huquqiy hujjatlar katta hajmdagi strukturaga ega bo'lmagan ma'lumotlarni o'z ichiga olgani sababli ulardan shaxs, tashkilot, hudud, sana, qonun, sud, hujjat raqami va boshqa muhim birliklarni avtomatik ajratib olish zamonaviy tabiiy tilni qayta ishlash texnologiyalarining muhim vazifalaridan biridir. Maqolada NER tushunchasi, huquqiy matnlarda nomlangan birliklarni aniqlashning o'ziga xos jihatlari, huquqiy korpusni shakllantirish va annotatsiyalash, mashinaviy o'qitish hamda Transformer modellaridan foydalanish imkoniyatlari yoritiladi. Shuningdek, huquqiy hujjatlarni mavzu va mazmuniga ko'ra avtomatik tasniflash usullari ko'rib chiqiladi. O'zbek tilidagi huquqiy NER uchun 2026-yilda yaratilgan inson tomonidan tekshirilgan ma'lumotlar to'plami asosida mavjud kategoriyalar va istiqbolli yo'nalishlar tahlil qilinadi. Tadqiqot natijasida huquqiy NER va matn tasniflashni yagona tizimga birlashtirish huquqiy hujjatlarni qidirish, indekslash, saralash va avtomatik tahlil qilish samaradorligini oshirishi mumkinligi asoslanadi.*

Kalit so'zlar: *NER, Named Entity Recognition, tabiiy tilni qayta ishlash, huquqiy NLP, huquqiy hujjatlar, matn tasnifi, mashinaviy o'qitish, BERT, Transformer, huquqiy korpus.*

Аннотация. *В данной статье рассматриваются вопросы создания системы Named Entity Recognition (NER) на основе юридических документов и автоматической классификации юридических текстов. Юридические документы содержат значительный объём неструктурированной информации, поэтому автоматическое извлечение таких сущностей, как имена лиц, организации, географические объекты, даты, законы, суды и номера документов, является важной задачей современных технологий обработки естественного языка. В статье анализируются понятие NER, особенности распознавания именованных сущностей в юридических текстах, этапы формирования и аннотирования юридического корпуса, а также возможности применения методов машинного обучения и моделей Transformer. Особое внимание уделяется классификации юридических документов по тематике и содержанию. На основе опубликованного в 2026 году набора данных для распознавания именованных сущностей в узбекских юридических текстах рассматриваются существующие категории сущностей и перспективы дальнейшего развития данного направления. Обосновывается, что интеграция NER и автоматической классификации текстов может повысить эффективность поиска, индексации, систематизации и анализа юридических документов.*

Ключевые слова: *NER, Named Entity Recognition, обработка естественного языка, юридический NLP, юридические документы, классификация текста, машинное обучение, BERT, Transformer, юридический корпус.*

Annotation. *This article examines the development of a Named Entity Recognition (NER) system based on legal documents and the automatic classification of legal texts. Since legal documents contain a large amount of unstructured information, the automatic extraction of entities such as persons, organizations, locations, dates, laws, courts, and document identifiers represents an important task in modern Natural Language Processing. The article discusses the concept of NER, the specific characteristics of named entity recognition in legal texts, the construction and annotation of legal corpora, and the application of machine learning and Transformer-based models. Particular attention is paid to the automatic classification of legal documents according to their topics and content. The study also examines the entity categories included in a human-reviewed Uzbek legal NER dataset published in 2026. The analysis demonstrates that integrating NER with legal text classification can improve the* [9]*efficiency of searching, indexing, organizing, and automatically analyzing legal documents.*

Keywords: *NER, Named Entity Recognition, Natural Language Processing, legal NLP, legal documents, text classification, machine learning, BERT, Transformer, legal corpus.*

Raqamli texnologiyalarning jadal rivojlanishi natijasida davlat va huquqiy tashkilotlarda elektron hujjatlar hajmi sezilarli darajada ortib bormoqda. Qonunlar, farmonlar, qarorlar, sud hujjatlari, shartnomalar va boshqa huquqiy materiallarda katta miqdordagi strukturaga ega bo‘lmagan ma’lumotlar mavjud. Bunday hujjatlarni inson tomonidan qo‘lda izlash, guruhlash va tahlil qilish ko‘p vaqt talab qiladi. Shu sababli tabiiy tilni qayta ishlash (Natural Language Processing - NLP) texnologiyalaridan foydalanish huquqiy axborotni avtomatik qayta ishlashning muhim yo‘nalishiga aylanmoqda. Named Entity Recognition (NER), ya’ni nomlangan birliklarni aniqlash, NLP va axborot ajratib olishning asosiy vazifalaridan biri hisoblanadi. David Nadeau va Satoshi Sekine NERni matndagi shaxs, tashkilot va joy nomlari, shuningdek vaqt, sana va pul miqdori kabi muhim birliklarni aniqlash va tasniflash bilan bog‘liq vazifa sifatida tavsiflaydi. Mualliflarning ta’kidlashicha, NER tizimlari dastlab qo‘lda yaratilgan qoidalarga tayangan bo‘lsa, keyinchalik mashinaviy o‘qitish usullari keng qo‘llanila boshlagan. Huquqiy sohada NERning ahamiyati yanada yuqori bo‘lib, huquqiy hujjatlar tarkibida oddiy shaxs va tashkilot nomlaridan tashqari qonunlar, normativ-huquqiy hujjatlar, sudlar, ish raqamlari, moddalar va boshqa maxsus birliklar mavjud.[10] Elena Leitner, Georg Rehm va Julian Moreno-Schneider nemis tilidagi sud hujjatlari uchun maxsus huquqiy NER tizimini tadqiq qilib, 19 ta semantik kategoriyani belgilagan. Ularning tadqiqoti huquqiy sohada umumiy NER kategoriyalaridan tashqari, domenning o‘ziga xos birliklarini ham hisobga olish zarurligini ko‘rsatadi.[11] O‘zbek tilida esa huquqiy NER bo‘yicha resurslar hali shakllanish bosqichida. 2026-yilda Bobur Saidov, Zarnigor Fayzullaeva, Umida Bazarova va hammualliflar tomonidan inson tomonidan tekshirilgan o‘zbek huquqiy NER ma’lumotlar to‘plami e’lon qilindi. Mazkur dataset 12 ta asosiy kategoriya - PER, ORG, LOC, DATE, MONEY, POSITION, DOCNO, LAW, COURT, BANK, TIN va CADASTRE ni qamrab oladi. Shu sababli o‘zbek tilidagi huquqiy hujjatlar asosida NER tizimini yaratish va matnlarni avtomatik tasniflash dolzarb ilmiy-amaliy masalalardan biri hisoblanadi.

[9] Nadeau D., Sekine S. A Survey of Named Entity Recognition and Classification. Lingvisticae Investigationes, 2007.
[10] Abdurakhmonova, N., Esanov, A., Abidjanov, A., Baxtiyarova, Z., Koymuradova, G., & Tilavova, M., 2020.
[11] Leitner E., Rehm G., Moreno-Schneider J. Fine-Grained Named Entity Recognition in Legal Documents. Springer, 2019.

NER tizimining asosiy vazifasi matndagi muhim nomlangan birliklarni aniqlash va ularni tegishli semantik kategoriyalarga ajratishdan iborat. Nadeau va Sekine NER bo‘yicha tadqiqotlarni tahlil qilib, ushbu sohaning rivojlanishida qoida asosidagi, statistik va mashinaviy o‘qitishga asoslangan yondashuvlar muhim rol o‘ynaganini ko‘rsatadi.[12] Huquqiy hujjatlar uchun esa NER kategoriyalari oddiy yangilik yoki umumiy matnlar uchun ishlatiladigan kategoriyalardan kengroq bo‘lishi kerak. Masalan, huquqiy hujjatlarda quyidagi birliklarni ajratish mumkin:

PER - shaxs nomlari;
ORG - tashkilot va davlat organlari;
LOC - geografik obyektlar;
DATE - sanalar;
MONEY - pul miqdorlari;
POSITION - lavozimlar;
DOCNO - hujjat raqamlari;
LAW - qonunlar;
COURT - sudlar;
BANK - bank tashkilotlari;
TIN - soliq identifikatsiya raqamlari;
CADASTRE - kadastrga oid identifikatorlar.

Aynan shu 12 ta kategoriya 2026-yilda e’lon qilingan o‘zbek huquqiy NER datasetida ham qo‘llangan.[13]

NER tizimini yaratish uchun birinchi navbatda yetarli hajmdagi va sifatli huquqiy korpus shakllantirilishi lozim. Korpus tarkibiga qonunlar, Prezident farmonlari va qarorlari, Vazirlar Mahkamasi qarorlari, sud hujjatlari, shartnomalar va boshqa huquqiy materiallarni kiritish mumkin. Korpus yaratish jarayonida matnlarni raqamlashtirish, takroriy qismlarni olib tashlash, texnik belgilarni standartlashtirish va hujjatlarni tegishli kategoriyalar bo‘yicha guruhlash talab etiladi[14]. Leitner va hammualliflari tomonidan yaratilgan nemis huquqiy NER datasetida sud qarorlaridan foydalanilib, taxminan 67 000 ta gap va 54 000 ta annotatsiyalangan birlik shakllantirilgan. Bu tajriba huquqiy NER tizimlari uchun katta hajmdagi domen korpusining muhimligini ko‘rsatadi. O‘zbek huquqiy NER datasetida esa ma’lumotlar ikki qatlamga ajratilgan: manbaga asoslangan, inson tomonidan tekshiriladigan asosiy qism va kam uchraydigan kategoriyalarni o‘qitishni qo‘llab-quvvatlash uchun kengaytirilgan qism.

NER tizimlarining rivojlanishida neyron tarmoqlarga asoslangan modellar muhim bosqichni tashkil etadi. Guillaume Lample va hammualliflari BiLSTM asosidagi neyron arxitekturani taklif qilib, so‘z va belgilar darajasidagi reprezentatsiyalarni birgalikda qo‘llash NER vazifasida samarali natija berishini ko‘rsatgan.[15] Bunday yondashuv huquqiy matnlar uchun ham foydali bo‘lishi mumkin. Chunki huquqiy terminlarning ayrimlari murakkab morfologik tuzilishga ega bo‘lib, ularni faqat tayyor lug‘atlar orqali aniqlash yetarli bo‘lmasligi mumkin. Keyingi bosqichda Transformer arxitekturasi asosidagi modellar NER va boshqa NLP vazifalarida keng qo‘llanila boshlandi. Jacob Devlin va hammualliflari tomonidan ishlab

[12] Abdurakhmonova, N., Adizova, N., Jamolitdinova, D., Begimov, A., Sultanova, A., & Juraeva, G. 2020.
[13] Lample G., Ballesteros M., Subramanian S., Kawakami K., Dyer C. Neural Architectures for Named Entity Recognition. ACL, 2016.
[14] Abdurakhmonova, N., Mengliev, D., Mardonova, S., Asadov, T., Xalilov, F., & Maxsudova, F. 2025.
[15] Chalkidis I., Fergadiotis M., Malakasiotis P., Aletras N., Androutsopoulos I. LEGAL-BERT: The Muppets Straight Out of Law School. ACL Findings, 2020.

chiqilgan BERT modeli matnni chap va o‘ng kontekstni birgalikda hisobga olgan holda ifodalash imkonini beradi. [16]BERTning muhim xususiyatlaridan biri shundaki, u oldindan katta hajmdagi matnlarda o‘qitilib, keyinchalik muayyan vazifa uchun moslashtirilishi mumkin. Bu yondashuv huquqiy matnlar uchun ham samarali bo‘lishi mumkin.

Huquqiy matnlarning terminologik xususiyatlari umumiy matnlardan farq qilgani sababli umumiy til modeli har doim ham optimal natija bermasligi mumkin. Ilias Chalkidis va hammualliflari huquqiy domen uchun BERT modellarini o‘rganib, modelni maxsus huquqiy korpusda qo‘shimcha o‘qitish va huquqiy domen uchun alohida model yaratish kabi strategiyalarni taqqoslagan. [17]Mualliflar LEGAL-BERT oilasidagi modellarni taklif qilgan bo‘lib, ushbu modellar huquqiy NLP va computational law vazifalariga moslashtirilgan. Bu natija o‘zbek tilidagi huquqiy NER uchun ham huquqiy korpus asosida maxsus Transformer modelini moslashtirish istiqbolli ekanini ko‘rsatadi. Shunday qilib, o‘zbek huquqiy NER tizimining nazariy modeli quyidagi ketma-ketlikda tashkil etilishi mumkin: Huquqiy hujjat → Matnni tozalash → Tokenizatsiya → Annotatsiya → Transformer/BERT modeli → NER klassifikatori → Nomlangan birliklar.

NER bilan bir qatorda huquqiy matnlarni avtomatik tasniflash ham muhim vazifalardan hisoblanadi. Matn tasnifi hujjatlarni mazmuni, mavzusi yoki huquqiy sohasi bo‘yicha tegishli kategoriyalarga avtomatik ajratishni anglatadi.

Huquqiy hujjatlarni quyidagi guruhlarga ajratish mumkin:

fuqarolik huquqi;
jinoyat huquqi;
mehnat huquqi;
ma’muriy huquq;
soliq va moliya huquqi;
ta’lim sohasi;
sog‘liqni saqlash;
sud hujjatlari;
normativ-huquqiy hujjatlar.

Haihua Chen va hammualliflarining tadqiqotida 30 000 ta AQSh sud hujjati 50 ta kategoriya bo‘yicha tasniflangan va turli mashinaviy o‘qitish hamda chuqur o‘rganish usullari taqqoslangan. Tadqiqotda domen tushunchalariga asoslangan Random Forest yondashuvi ayrim tajribalarda oldindan o‘qitilgan word embeddinglar va chuqur neyron tarmoqlarga asoslangan tizimdan yuqori natija ko‘rsatgani qayd etilgan.[18] Bu holat huquqiy matnlarni tasniflashda faqat murakkab neyron modeldan foydalanish emas, balki huquqiy domenning o‘ziga xos terminlari va tushunchalarini hisobga olish ham muhim ekanini ko‘rsatadi. NER va matn tasniflashni yagona tizimga birlashtirish huquqiy axborot bilan ishlash samaradorligini oshirishi mumkin. Masalan, huquqiy hujjat tizimga kiritilganda NER moduli undan shaxs, tashkilot, sana, qonun va hujjat raqamini ajratib oladi. Keyingi modul esa hujjatning umumiy mazmunini tahlil qilib, uni muayyan huquqiy sohaga tegishli kategoriya bilan belgilaydi. Bunday tizimning umumiy ishlash prinsipi quyidagicha bo‘lishi mumkin:

[16] Chen H., Wu L., Chen J., Lu W., Ding J. A Comparative Study of Automated Legal Text Classification Using Random Forests and Deep Learning. Information Processing & Management, 2022.
[17] Saidov B., Fayzullaeva Z., Bazarova U. et al. Human-Reviewed Uzbek Legal Named Entity Recognition Dataset. F1000Research, 2026.
[18] O‘zbekiston Respublikasi qonunchilik ma’lumotlari milliy bazasi - LexUZ.

1-bosqich: hujjatni qabul qilish;
2-bosqich: matnni tozalash;
3-bosqich: tokenizatsiya;
4-bosqich: nomlangan birliklarni aniqlash;
5-bosqich: hujjat mavzusini tasniflash;
6-bosqich: natijalarni ma'lumotlar bazasiga saqlash;
7-bosqich: qidiruv va avtomatik tahlil.

Natijada foydalanuvchi nafaqat hujjatning o‘zini, balki uning asosiy shaxslari, tashkilotlari, sanalari, qonunlari va huquqiy kategoriyasi orqali ham kerakli ma'lumotni topishi mumkin.

NER tizimini baholashda asosan Precision, Recall va F1-score ko‘rsatkichlaridan foydalaniladi. Nadeau va Sekine NER tizimlarini baholash metodlarini ushbu sohaning muhim tarkibiy qismlaridan biri sifatida ko‘rsatadi

Precision - model aniqlagan birliklarning qanchasi to‘g‘ri ekanini ko‘rsatadi.
Recall - mavjud to‘g‘ri birliklarning qanchasi model tomonidan aniqlanganini bildiradi.
F1-score - Precision va Recall ko‘rsatkichlarining garmonik o‘rtacha qiymati:
$F1 = 2 \times (Precision \times Recall) / (Precision + Recall)$
Huquqiy NER tizimida umumiy F1-score bilan birga har bir kategoriya bo‘yicha alohida natijalarni ham hisoblash maqsadga muvofiq. Masalan, PER, ORG, LAW, COURT va DOCNO kategoriyalarining aniqligi alohida baholanishi kerak.

O‘zbek tilidagi huquqiy NER uchun eng muhim masalalardan biri - sifatli va keng qamrovli annotatsiyalangan korpus yaratishdir. 2026-yilda Saidov va hammualliflar tomonidan e'lon qilingan dataset ushbu yo‘nalishda muhim amaliy qadam bo‘ldi. Unda 12 turdagi nomlangan birliklar belgilangan bo‘lib, dataset XLSX, CSV, JSON va JSONL formatlarida taqdim etilgan. Mazkur tadqiqotda o‘zbek tili huquqiy NER nuqtai nazaridan hali resurslari cheklangan til sifatida ko‘riladi. Shu sababli kelajakda yanada katta hajmdagi huquqiy korpus yaratish, annotatsiya standartlarini ishlab chiqish va turli Transformer modellarini o‘zbek huquqiy matnlari asosida moslashtirish muhim vazifalardan bo‘lib qoladi. Bundan tashqari, huquqiy NER tizimini LexUZ va boshqa elektron huquqiy resurslar bilan integratsiya qilish kelajakda huquqiy hujjatlarni avtomatik indekslash va semantik qidiruv tizimlarini yaratish imkonini berishi mumkin.

Xulosa qilib aytganda, huquqiy hujjatlar asosida NER tizimini yaratish huquqiy axborotni avtomatik qayta ishlashning istiqbolli yo‘nalishlaridan biridir. NER yordamida huquqiy matnlardan shaxslar, tashkilotlar, joylar, sanalar, qonunlar, sudlar, hujjat raqamlari va boshqa muhim birliklarni avtomatik ajratib olish mumkin. Ilmiy tadqiqotlar NER tizimlarining qoida asosidagi usullardan mashinaviy o‘qitish va Transformer modellarigacha rivojlanganini ko‘rsatadi. Huquqiy sohada esa modelni maxsus domen korpusi asosida moslashtirish alohida ahamiyatga ega. O‘zbek tilidagi huquqiy NER bo‘yicha 2026-yilda e'lon qilingan inson tomonidan tekshirilgan dataset ushbu yo‘nalish uchun muhim ilmiy resurs hisoblanadi. Unda PER, ORG, LOC, DATE, MONEY, POSITION, DOCNO, LAW, COURT, BANK, TIN va CADASTRE kabi 12 ta kategoriya mavjud. Shunday qilib, o‘zbek tilidagi huquqiy NER tizimini yaratishning asosiy bosqichlari huquqiy korpusni shakllantirish, ma'lumotlarni annotatsiyalash, NER modelini o‘qitish, huquqiy hujjatlarni avtomatik tasniflash va tizimni Precision, Recall hamda F1-score orqali baholashdan iborat. Bunday tizim kelajakda huquqiy hujjatlarni tezkor

qidirish, saralash, indekslash va tahlil qilish qilish jarayonlarini avtomatlashtirishga xizmat qilishi mumkin.

FOYDALANILGAN ADABIYOTLAR

1. Nadeau D., Sekine S. A Survey of Named Entity Recognition and Classification. Lingvisticae Investigationes, 2007.
2. Leitner E., Rehm G., Moreno-Schneider J. Fine-Grained Named Entity Recognition in Legal Documents. Springer, 2019.
3. Lample G., Ballesteros M., Subramanian S., Kawakami K., Dyer C. Neural Architectures for Named Entity Recognition. ACL, 2016.
4. Chalkidis I., Fergadiotis M., Malakasiotis P., Aletras N., Androutsopoulos I. LEGAL-BERT: The Muppets Straight Out of Law School. ACL Findings, 2020.
5. Chen H., Wu L., Chen J., Lu W., Ding J. A Comparative Study of Automated Legal Text Classification Using Random Forests and Deep Learning. Information Processing & Management, 2022.
6. Saidov B., Fayzullaeva Z., Bazarova U. et al. Human-Reviewed Uzbek Legal Named Entity Recognition Dataset. F1000Research, 2026.
7. O‘zbekiston Respublikasi qonunchilik ma’lumotlari milliy bazasi - LexUZ.
8. Mengliev, D., Abdurakhmonova, N., Shirinova, R., Ibragimov, B., Khudayberganova, D., & Suyunova, M. (2025, October). A comparative study of annotation schemes for named entity recognition in Uzbek legal texts. In *AIP Conference Proceedings* (Vol. 3377, No. 1, p. 070005). AIP Publishing LLC.
9. Abdurakhmonova, N., Esanov, A., Abidjanov, A., Baxtiyarova, Z., Koymuradova, G., & Tilavova, M. (2026, May). Named Entity Recognition for Uzbek Economic and Financial Text. In *2026 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)* (pp. 1-4). IEEE.
10. Abdurakhmonova, N., Adizova, N., Jamolitdinova, D., Begimov, A., Sultanova, A., & Juraeva, G. (2025, November). Development of Hybrid Algorithm for Named Entity Recognition in Uzbek Public Administration Texts. In *2025 IEEE XVII International Scientific and Technical Conference on Actual Problems of Electronic Instrument Engineering (APEIE)* (pp. 1-6). IEEE.
11. Abdurakhmonova, N., Mengliev, D., Mardonova, S., Asadov, T., Xalilov, F., & Maxsudova, F. (2025, September). A Hybrid Named-Entity Recognition Algorithm for Ecological Documentation in Uzbekistan. In *2025 10th International Conference on Computer Science and Engineering (UBMK)* (pp. 1778-1781). IEEE.

RAQAMLI MAKONDA SEVGI FENOMENI: ER-XOTIN MUNOSABATLARIDA HISSIY YAQINLIK VA ISHONCH DINAMIKASI

Olimjonova Ziynat Bobomurod qizi
ISFT instituti katta o'qituvchisi, Oila va gender ilmiy tadqiqot instituti mustaqil izlanuvchisi
https://doi.org/10.5281/zenodo.23253604

Annotatsiya. *Mazkur maqolada raqamli transformatsiya sharoitida er-xotin munosabatlarida hissiy yaqinlik va o'zaro ishonchning namoyon bo'lish xususiyatlari ijtimoiy-psixologik nuqtai nazardan tahlil qilinadi. Raqamli kommunikatsiya vositalarining oilaviy muloqot, hissiy aloqadorlik va munosabatlar sifatiga ta'siri zamonaviy psixologik tadqiqotlar asosida yoritiladi. Sevgi tuyg'usi, hissiy yaqinlik va ishonchning raqamli hamda bevosita muloqot shakllari o'rtasidagi muvozanat bilan bog'liqligi asoslanadi.*

Kalit so'zlar*: raqamli makon, sevgi fenomeni, er-xotin munosabatlari, hissiy yaqinlik, o'zaro ishonch, raqamli kommunikatsiya, rashk, oilaviy barqarorlik, munosabatlardan qoniqish.*

Аннотация. *В данной статье с социально-психологической точки зрения анализируются особенности проявления эмоциональной близости и взаимного доверия в супружеских отношениях в условиях цифровой трансформации. На основе современных психологических исследований рассматривается влияние средств цифровой коммуникации на семейное общение, эмоциональную взаимосвязь и качество супружеских отношений. Обосновывается взаимосвязь чувства любви, эмоциональной близости и доверия с балансом между цифровыми и непосредственными формами общения.*

Ключевые слова: *цифровое пространство, феномен любви, супружеские отношения, эмоциональная близость, взаимное доверие, цифровая коммуникация, ревность, стабильность семьи, удовлетворённость отношениями.*

Abstract. *This article analyzes the manifestations of emotional intimacy and mutual trust in marital relationships in the context of digital transformation from a socio-psychological perspective. Drawing on contemporary psychological research, it examines the influence of digital communication tools on family communication, emotional connectedness, and the quality of marital relationships. The relationship between feelings of love, emotional intimacy, and trust and the balance between digital and face-to-face forms of communication is substantiated.*

Keywords: *digital space, phenomenon of love, marital relationships, emotional intimacy, mutual trust, digital communication, jealousy, family stability, relationship satisfaction.*

Jamiyatning raqamli transformatsiyasi inson hayotining deyarli barcha sohalari qatori shaxslararo va oilaviy munosabatlarning mazmuni, shakli hamda kommunikativ mexanizmlariga ham sezilarli o'zgarishlar olib kirmoqda. Smartfonlar, messenjerlar, ijtimoiy tarmoqlar va boshqa raqamli kommunikatsiya vositalari zamonaviy insonning kundalik muloqot makoniga chuqur integratsiyalashib, er-xotin o'rtasidagi o'zaro aloqaning vaqt va makon chegaralarini kengaytirmoqda. Natijada oilaviy muloqot endilikda faqat bevosita yuzma-yuz kommunikatsiya bilan chegaralanmay, raqamli va real kommunikativ jarayonlarning o'zaro birikmasi sifatida namoyon bo'lmoqda.

Mazkur o‘zgarish sevgi fenomenini ham zamonaviy ijtimoiy-psixologik sharoitlar kontekstida qayta ko‘rib chiqish zaruratini yuzaga keltiradi. Chunki sevgi er-xotin munosabatlarida faqat hissiy kechinma emas, balki o‘zaro ishonch, yaqinlik, g‘amxo‘rlik, mas’uliyat, sherikning ehtiyojlariga javob bera olish hamda munosabatlarni saqlashga qaratilgan xulq-atvor komponentlarini birlashtiruvchi murakkab psixologik hodisadir. Raqamli kommunikatsiyaning kundalik oilaviy hayotga kirib kelishi mazkur komponentlarning namoyon bo‘lish shakllariga ham ta’sir etishi mumkin.

Raqamli makon er-xotin munosabatlari uchun bir vaqtning o‘zida ham “Kommunikativ resurs”, ham “Psixologik xavf omili” bo‘lishi mumkin. Bir tomondan, masofadan turib muntazam aloqa qilish, hissiy qo‘llab-quvvatlash, kundalik tajribalarni tezkor almashish va sherikning hayotida ishtirok etish imkoniyati kengayadi. Ikkinchi tomondan esa raqamli qurilmalarning bevosita muloqot jarayoniga haddan tashqari kirib kelishi sherikka qaratilgan e’tiborning bo‘linishi, muloqot sifatining pasayishi, hissiy masofalanish, rashk, nazorat va ishonch bilan bog‘liq ziddiyatlarni yuzaga keltirishi mumkin.

Zamonaviy ilmiy adabiyotlarda raqamli qurilmalar sababli yuzma-yuz muloqotning uzilishi yoki bo‘linishi texnologik aralashuv (technoference) tushunchasi bilan ifodalanadi. Er-xotin muloqoti jarayonida sherikka emas, balki smartfon yoki boshqa raqamli qurilmaga ustuvor e’tibor qaratish holati esa sherikka nisbatan raqamli e’tiborsizlik (partner phubbing) sifatida tavsiflanadi. Mazkur hodisalar raqamli texnologiyalarning mavjudligidan ko‘ra, ulardan foydalanishning er-xotin o‘rtasidagi bevosita muloqotga qay darajada kirib borishi, sherikka qaratilgan e’tiborni bo‘lishi hamda hissiy yaqinlik va o‘zaro munosabatlar sifatida qanday aks etishini tahlil qilish imkonini beradi. Shuningdek, so‘nggi ma’lumotlarga ko‘ra er-xotin juftiga nisbatan romantik munosabatlar sifati, hissiy yaqinlik, sherikning javobchanligini his etish va munosabatlardan qoniqish kabi ko‘rsatkichlar bilan salbiy bog‘liqligi, konflikt va rashk bilan esa ijobiy bog‘liqligi qayd etilgan. Bu natijalar raqamli makonning er-xotin munosabatlariga ta’sirini “texnologiya foydali yoki zararli” kabi sodda dixotomiya asosida emas, balki kommunikativ kontekst, foydalanish xususiyati va juftlikning psixologik resurslari bilan birgalikda o‘rganish zarurligini ko‘rsatadi.

Zamonaviy sharoitda sevgi tuyg‘usining namoyon bo‘lishini faqat bevosita muloqot doirasida o‘rganish yetarli emas. Er-xotinning kun davomida yozishmalari, ovozli va videoaloqasi, ijtimoiy tarmoqlardagi o‘zaro munosabatlari, bir-birining raqamli faoliyatiga munosabati hamda raqamli shaxsiy chegaralarni qabul qilishi munosabatlarning yangi psixologik kontekstini shakllantirmoqda. Bunda asosiy masala raqamli muloqotning miqdorida emas, balki uning psixologik sifatida namoyon bo‘ladi. Tez-tez yozishish har doim ham yuqori hissiy yaqinlikni anglatmaydi. Xuddi shuningdek, sherikning onlayn faolligini muntazam kuzatish ham yuqori darajadagi g‘amxo‘rlik yoki sevgi belgisi sifatida baholanishi mumkin emas. Bunday xatti-harakatlarning psixologik mazmuni er-xotin o‘rtasidagi ishonch, bog‘lanish xususiyati, shaxsiy chegaralar va mavjud kommunikativ tajriba bilan bog‘liq holda talqin qilinishi lozim.

Hissiy yaqinlik er-xotinning bir-biriga o‘z kechinmalari, fikrlari va ehtiyojlarini ochiq ifodalashi, sherik tomonidan tushunilgan va qabul qilinganligini his etishi hamda psixologik xavfsizlikni boshdan kechirishi bilan bog‘liq murakkab munosabat komponentidir. Shu sababli yaqinlikni faqat birgalikda o‘tkazilgan vaqt miqdori bilan baholash yetarli emas. Raqamli texnologiyalar hissiy yaqinlikni qo‘llab-quvvatlash uchun muhim kommunikativ imkoniyat yaratadi. Ayniqsa, er-xotin vaqtincha bir-biridan uzoqda bo‘lganida raqamli aloqa kundalik voqealarni baham ko‘rish, sherikning holatidan xabardor bo‘lish va hissiy qo‘llab-quvvatlashni

davom ettirish imkonini beradi. Biroq sheriklar jismonan bir makonda bo‘lgan paytda texnologik qurilmalarning ularning bevosita muloqotini muntazam bo‘lib turishi qarama-qarshi psixologik natijaga olib kelishi mumkin.

O‘zaro ishonch er-xotin munosabatlari barqarorligining asosiy psixologik resurslaridan biridir. Ishonch sherikning xatti-harakatlarini doimiy tekshirish zaruratining kamayishi, munosabatlarda psixologik xavfsizlikning mavjudligi va o‘zaro ochiqlik bilan bog‘liq. Raqamli makon ushbu jarayonga yangi o‘lchovlarni olib kiradi. Ijtimoiy tarmoqlardagi faollik, “layk” va izohlar, shaxsiy yozishmalar, onlayn holat, parollar, kontaktlar hamda oldingi munosabatlarga oid raqamli izlar er-xotin o‘rtasida ishonch va shaxsiy chegaralar masalasini yanada murakkablashtirishi mumkin. Ijtimoiy tarmoqlar va romantik rashk bo‘yicha sistematik tahlillarda raqamli platformalarning ayrim xususiyatlari rashkni kuchaytirishi mumkinligi, ayni paytda ishonchning bunday jarayonlarda himoya qiluvchi psixologik omil sifatida namoyon bo‘lishi qayd etilgan. Bu holat raqamli makonda ishonchni faqat “sir saqlamaslik” yoki sherikning telefoniga kirish imkoniyatiga ega bo‘lish bilan tenglashtirish noto‘g‘ri ekanligini ko‘rsatadi. Aksincha, sog‘lom raqamli ishonch er-xotinning o‘zaro kelishilgan chegaralari, shaxsiy makonga hurmat, ochiq kommunikatsiya va raqamli xatti-harakatlar bo‘yicha o‘zaro maqbul me’yorlarning shakllanganligi bilan tavsiflanishi mumkin. Shu sababli zamonaviy er-xotin munosabatlarida psixologik ishonch tushunchasi bevosita munosabatlar bilan bir qatorda “raqamli chegaralar va raqamli ishonch” komponentlarini ham qamrab olmoqda.

Olingan empirik natijalar hamda zamonaviy tadqiqotlarning nazariy tahlili raqamli makonda er-xotin munosabatlarini tushuntirishda uchta asosiy psixologik komponentni ajratish imkonini beradi: sevgi tuyg‘usi, hissiy yaqinlik va o‘zaro ishonch. Mazkur komponentlar o‘zaro chambarchas bog‘liq bo‘lsa-da, ularning har biri raqamli kommunikatsiya sharoitida o‘ziga xos tarzda namoyon bo‘ladi. Sevgi er-xotin munosabatlarining umumiy hissiy-qadriyatli asosini tashkil etsa, hissiy yaqinlik sheriklarning o‘z kechinmalarini ochiq ifodalashi va o‘zaro tushunilganlik hissi bilan bog‘liq. Ishonch esa munosabatlarda psixologik xavfsizlik, sherikning xatti-harakatlarini doimiy nazorat qilish ehtiyojining kamayishi va o‘zaro chegaralarning tan olinishida namoyon bo‘ladi.

Raqamli kommunikatsiya ushbu uch komponentning o‘zaro munosabatini murakkablashtirishi mumkin. Masalan, er-xotinning kun davomida raqamli aloqada bo‘lishi kommunikativ yaqinlikni qo‘llab-quvvatlashi mumkin. Biroq bir makonda bo‘lgan paytda smartfonning muntazam ravishda suhbatni bo‘lishi sherikda e’tiborsizlik hissini yuzaga keltirishi ehtimoli mavjud. Demak, raqamli kontaktlarning ko‘pligi hissiy yaqinlikning yuqoriligini avtomatik ravishda anglatmaydi. Shundan kelib chiqib, raqamli kommunikativ yaqinlik va hissiy-psixologik yaqinlik tushunchalarini farqlash maqsadga muvofiq. Birinchisi er-xotin o‘rtasidagi raqamli kontaktlarning mavjudligi va muntazamligini anglatsa, ikkinchisi ushbu muloqotning mazmuni, sherikning javobchanligi, tushunilganlik hissi va psixologik xavfsizlik bilan belgilanadi.

Raqamli transformatsiya sharoitida oilalar bilan psixologik ishlashda faqat er-xotinning an’anaviy kommunikativ muammolariga emas, ularning raqamli xulq-atvoriga ham e’tibor qaratish zarur. Bunda asosiy vazifa raqamli vositalardan foydalanishni cheklash emas, balki juftlikning o‘ziga mos sog‘lom raqamli chegaralarini shakllantirishdan iborat.

- Birinchidan, er-xotin o‘rtasida raqamli kommunikatsiyaga oid o‘zaro maqbul me’yorlarni kelishib olish muhim. Telefonning shaxsiy daxlsizligi, parollardan foydalanish, ijtimoiy tarmoqlardagi muloqot, boshqa shaxslar bilan yozishmalar hamda birgalikdagi vaqt

davomida smartfondan foydalanish masalalarida ochiq kelishuv mavjudligi tushunmovchiliklarning oldini olishga xizmat qilishi mumkin.

- Ikkinchidan, "raqamli mavjudlik" va "hissiy mavjudlik" o'rtasidagi muvozanatni ta'minlash zarur. Er-xotinning doimiy onlayn aloqasi bevosita muloqotning o'rnini to'liq bosa olmaydi. Shu sababli birgalikdagi vaqt davomida sherikka bo'linmagan e'tibor qaratish, faol tinglash va hissiy tajribalarni bevosita baham ko'rish oilaviy yaqinlikning muhim sharti bo'lib qoladi.
- Uchinchidan, raqamli rashk va nazoratni sevgi yoki g'amxo'rlikning bevosita belgisi sifatida talqin qilishdan saqlanish lozim. Sherikning telefonini muntazam tekshirish, onlayn faolligini nazorat qilish yoki parollarni talab qilish ayrim hollarda ishonchsizlik va psixologik xavfsizlik bilan bog'liq muammolarning tashqi ifodasi bo'lishi mumkin.

FOYDALANILGAN ADABIYOTLAR

1. Shoumarov G'.B. Oila psixologiyasi: darslik. – Toshkent, 2011. – B. 85.
2. Karimova V. Oila psixologiyasi: pedagogika oliy o'quv yurtlari talabalari uchun darslik. – Toshkent, 2008. – 152 b.
3. Андреева Т. В. Семейная психология: учеб. пособие. – СПб.: Речь, 2004.
4. Карабанова О. А. Психология семейных отношений и основы семейного консультирования: учеб. пособие. – М.: Гардарики, 2005.
5. Шнейдер Л.Б. Семейная психология: учеб. пособие для вузов. – М.: Академический проект, 2006.
6. McDaniel B.T., Coyne S.M. "Technoference": The interference of technology in couple relationships and implications for women's personal and relational well-being // Psychology of Popular Media Culture. – 2016.

O‘ZBEK TILI TEZAURUSIDA SIFAT LEKSEMALARINING LINGVISTIK VA METODOLOGIK XUSUSIYATLARI

Po‘latova Gulxayyo G‘ofurjon qizi
Namangan davlat pedagogika instituti o‘qituvchisi, PhD.
https://doi.org/10.5281/zenodo.23253629

Annotatsiya. *Ushbu maqolada o‘zbek tili tezaurusi uchun sifat leksemalarining semantik munosabatlari va sinsetlarini shakllantirish tamoyillari tadqiq etiladi. WordNet modeli asosida sifat sinonimiyasi, antonimiyasi va graduonimiyasi (darajalanishi) tizimlashtirilgan. Shuningdek, sifat komponentli frazeologizmlar va an’anaviy lug‘atlarda cheklangan sifat leksemalarida ma’no ko‘chishi (semantik kengayish va torayish) jarayonlari tahlil qilinib, ularning tezaurus modeli ishlab chiqilgan.*

Аннотация*. В данных тезисах исследуются семантические отношения и принципы формирования синсетов прилагательных для тезауруса узбекского языка. На основе модели WordNet систематизированы синонимия, антонимия и градуонимия прилагательных. Проанализированы процессы семантического переноса (расширения и сужения значения) и фразеологизмы с адъективным компонентом, разработана их тезаурусная модель.*

Abstract. *This thesis examines the semantic relations and principles of synset formation for adjective lexemes in the Uzbek language thesaurus. Based on the WordNet model, synonymy, antonymy, and graduonymy (gradation) of adjectives are systematically organized. Processes of semantic transfer (semantic expansion and narrowing) and adjective-component idioms, which are often limited in traditional dictionaries, are analyzed and their thesaurus models are constructed.*

Kalit so‘zlar: *tezaurus, sifat leksemalari, Wordnet modeli, semantik munosabatlar, graduonimiya.*

Ключевые слова: *Тезаурус, лексемы прилагательных, модель WordNet, семантические отношения, градуонимия.*

Keywords: *thesaurus, adjective lexemes, WordNet model, semantic relations, graduonymy.*

KIRISH

Globallashuv va axborot texnologiyalarining jadal rivojlanishi natijasida jahon tilshunosligida kompyuter lingvistikasi va amaliy leksikografiya sohasi alohida ahamiyat kasb etmoqda. Tabiiy tilni qayta ishlash (NLP) va sun’iy intellekt tizimlarining muloqot samaradorligini oshirishda so‘zlarning o‘zaro munosabatlarini aks ettiruvchi ontologik va tezaurus tipidagi lug‘atlar muhim leksik baza hisoblanadi. Xususan, WordNet kabi leksik-semantik ma’lumotlar bazasi inson va kompyuter muloqotini yangi bosqichga olib chiqdi.

O‘zbek kompyuter tilshunosligi sohasida milliy korpus, mashina tarjimasi va sentiment tahlil bo‘yicha qator tadqiqotlar olib borilayotgan bo‘lsa-da, o‘zbek tilida tezaurus tuzishning lingvistik va metodologik modellari hali to‘liq shakllanmagan. Ayniqsa, o‘zbek tili tezaurusida sifat so‘z turkumining semantik munosabatlarini tizimli va monografik tarzda tadqiq etish ehtiyoji mavjud. Sifat leksemalari predmet va hodisalarning o‘lchov, rang, maza, hid, xossa va

xarakter-xususiyatlarini ifodalovchi, uslubiy moslashuvchanlikka ega bo‘lgan turkum sifatida alohida modellashtirish yondashuvini talab etadi.

ASOSIY QISM

Sifat leksemalari ot va fe’l turkumlaridan farqli o‘laroq qat’iy iyerarxik (giponimik) tuzilishga ega emas. Shu sababli, WordNet tipidagi elektron lug‘atlarda sifat sinsetlarini (sinonimlar to‘plamini) yaratish murakkab metodologik tamoyillar tizimini talab qiladi. Tadqiqot doirasida sifat tezaurusi modelini barpo etishning yetti ustuvor lingvistik tamoyili ishlab chiqildi va asoslandi:

1. Semantik yadroni aniqlash tamoyili: Har bir sifat sinseti bitta markaziy semantik yadro atrofida joylashgan yaqin ma’noli so‘zlardan tashkil topadi. Masalan, *“katta, ulkan, yirik”* sinseti hajm va o‘lchamning me’yordan kattaligi yadrosiga birlashadi.

2. Kontekstual almashuvchanlik tamoyili: Sinonim sifatlar gapda bir-birining o‘rnini qay darajada bosa olishiga qarab guruhlanadi. Masalan, *“go‘zal qiz”* va *“chiroyli qiz”* sinonimik almashina olsa-da, *“go‘zal vaziyat”* va *“chiroyli vaziyat”* birikmalarida uslubiy nomuvofiqlik vujudga keladi.

3. Stilistik va uslubiy moslik tamoyili: Leksemalarning rasmiy, so‘zlashuv, badiiy yoki poetik nutqqa mansubligi va ularning ekspressivlik darajasi inobatga olinadi. Masalan, *“katta” (neytral), “bahaybat” (ekspressiv), “azamat” (og‘zaki/sheva).*

4. Graduatsiya (darajalanish) tamoyili: Sinonimik va yaqin ma’noli sifatlar orasidagi belgi intensivligining ortib yoki kamayib borish tartibi belgilanadi (*kichik* ← *o‘rta* ← *katta* ← *ulkan* ← *bahaybat*).

5. Antonimik o‘q bo‘yicha muvozanat tamoyili: Sifatlar antonimiya o‘qi atrofida model shaklida tashkil qilinadi. Har bir ijobiy/neytral sinset o‘zining qarama-qarshi sinset juftiga ega bo‘lishi va bog‘lanishi shart (masalan*, “chiroyli” – “xunuk”, “go‘zal” – “badbashara”*).

6. Ma’no ko‘chishi, ya’ni semantik kengayish/torayish tamoyili: Sifatlarning vaqt o‘tishi bilan ijtimoiy-madaniy omillar ta’sirida kasb etgan ko‘chma va konnotativ ma’nolari tezaurus tizimida aks ettiriladi (masalan, *“qattiq” jism → “qat’iy” qaror).*

7. Sintaktik rol va kollokatsiya (birikuvchanlik) tamoyili: Sifatlarning ma’lum otlar bilan barqaror birika olish qobiliyati tezaurusda teglash tizimi bilan belgilanadi. Masalan, *“yashil”* sifatining *“chiroq”, “burchak”, “makon”, “iqtisodiyot”* kabi otlar bilan birikuvchanligi turlicha semantik maydon hosil qiladi.

Graduonimiya (darajalanish) sifat leksemalariga xos bo‘lgan eng muhim xususiyatlardan biri bo‘lib, til tizimida belgi intensivligining ortishi yoki kamayishini mantiqiy-semantik tarzda tartiblaydi. WordNet modellarida graduonimik munosabatlarni raqamli ko‘rsatkichlar (teglari) yordamida ifodalash taklif qilingan. Sifat antonimiyasi esa ikki xil ko‘rinishga ajratiladi: darajalanadigan (gradual) va darajalanmaydigan (komplementar) antonimlar. Darajalanadigan antonimlar o‘rtasida o‘rta/neytral nuqta mavjud bo‘lsa (masalan, *“sayoz”* ↔ *“chuqur”*), darajalanmaydigan antonimlar bir-birini mutlaqo inkor etadi va oraliq ma’noga ega bo‘lmaydi (masalan, *“tirik”* ↔ *“o‘lik”*). Quyidagi 1-jadvalda o‘zbek tili tezaurusi uchun ishlab chiqilgan va amaliyotda uzbekcorpus.uz saytida qo‘llanilgan eng yirik graduonimik oppozitsiyalardan biri – “katta-kichik” graduonimik ketma-ketligi berilgan:

1-jadval

№	Daraja Ko‘rsatkichi	Leksema va ifodalangan ma’no	Kontekstdagi namuna
1	+6 (O‘ta yuqori)	haybatli – ulug‘vor, salobatli, bahaybat	haybatli tog‘lar, bahaybat maxluq
2	+5 (G‘oyat yuqori)	azim – ko‘rinish jihatidan g‘oyat ulkan	azim chinor, azim daryo
3	+4 (Juda yuqori)	ulkan – egallagan o‘rni va hajmi yirik	ulkan kashfiyot, ulkan bino
4	+3 (Yuqori)	azamat – ko‘rinish va salobatga ega bo‘lgan	azamat yigit, azamat imorat
5	+2 (O‘rtacha ortiq)	zabardast – jismoniy baquvvat, kuchli	zabardast olim, zabardast yozuvchi
6	+1 (Me'yordan ortiq)	yirik – hajmi va o‘lchamlari nisbatan katta	yirik tadbirkor, yirik korxona
7	0 (Me'yor)	kelishgan / chiroyli – me’yoriy o‘lcham	kelishgan qaddi-qomat, chiroyli dizayn
8	-1 (Kamayish)	kichik – hajmi nisbatan kichik	kichik hovli, kichik maktab
9	-2 (O‘rtacha kam)	kichikroq – biroz kamayish	kichikroq xona, kichikroq mablag‘
10	-3 (Juda kam)	juda kichik – o‘ta cheklangan hajm	juda kichik zarracha, juda kichik xato
11	-4 (O‘ta kam)	jimit / mitti – eng kichik o‘lchov, jajji	jimitdek bola, mitti qushcha

Katta-kichik antonimlarining darajalanishi.

Ushbu graduonimik tizim sinonim sifatlarning til korpusida va mashina tarjimasida ma’noviy nozikliklar va konnotatsiyalarini aniq tarjima qilish imkonini beradi. Har bir leksemaning o‘zaro darajaviy masofasi raqamlash orqali belgilanishi axborot qidiruv tizimlarida semantik moslikni yaxshilaydi.

XULOSA

O‘zbek tili tezaurusi uchun sifat leksemalarining semantik munosabatlarini tadqiq etish kompyuter lingvistikasi va milliy leksikografiyani yangi bosqichga olib chiqadi. WordNet modeliga moslashtirilgan sifatlar tezaurusi tabiiy tilni qayta ishlash, matnlarni avtomatik klassifikatsiya qilish, sentiment tahlili va sun’iy intellekt tizimlariga o‘zbek adabiy tilining boy semantik imkoniyatlarini muvaffaqiyatli tatbiq etish imkonini beradi.

FOYDALANILGAN ADABIYOTLAR

1. O‘zbekiston Respublikasi Prezidentining 2020-yil 20-oktabrdagi “Mamlakatimizda o‘zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish chora-tadbirlari to‘g‘risida”gi PF-6084-son Farmoni.
2. Abduraxmonova N. O‘zbek tili elektron korpusining kompyuter modellari. Monografiya. – Germany: Globeedit, 2021. – 200 b.

3. Abdullayeva R. O‘zbek tilida xarakter ifodalovchi birliklar semantikasi. Monografiya. – Toshkent: Poytaxt exclusive, 2022. – 120 b.
4. Amanlikova N. O‘zbek tilida sifatga oid leksemalarning lingvokulturologik tadqiqi masalalari. Monografiya. – Toshkent: Transport, 2021. – 136 b.
5. Fellbaum C. WordNet: an elektronic lexical database. – Cambridge: MIT Press, 1998. – 422 p.
6. O‘zbek tili sinonimlarining izohli lug‘ati: N. Mahmudov tahriri ostida. – Toshkent: G‘afur G‘ulom nomidagi nashriyot-matbaa ijodiy uyi, 2023. – 567 b.
7. Rahmatullayev Sh., Mahmudov N. va boshqalar. O‘zbek tili frazeologik lug‘ati. – Toshkent: G‘afur G‘ulom nomidagi nashriyot-matbaa ijodiy uyi, 2022. – 636 b.
8. Saparova M. O‘zbek tili tezaurusi uchun ot so‘z turkumining semantik-struktur tadqiqi: Filol. fan. bo‘yicha fals. dok. … diss. avtoref. – Toshkent, 2025. – 51 b.

THE DISTINCTIVE FEATURES OF THESAURUS DICTIONARIES

Qadirova Mokhira Bakhodir qizi, Saparova Mokhira Fayzullayevana*
e-mail: Mohirabahodirovna18@gmail.com
*Associative professor of Mamun university
E-mail: saparovamokhira93@gmail.com
https://doi.org/10.5281/zenodo.23253648

Abstract. *This article examines the linguistic nature and structural organization of thesaurus dictionaries, as well as the distinctive features that differentiate them from other types of dictionaries. It argues that the defining characteristic of thesaurus dictionaries lies not merely in explaining lexical units based on their dictionary meanings, but also in systematically representing the semantic relationships between them, including synonymy, antonymy, hyperonym, hyponymy, meronymy, and other types of semantic relations. The article also considers the development of modern electronic thesaurus and their relationship with semantic networks and natural language processing technologies. In particular, it analyzes the issues involved in the semantic systematization of the Uzbek lexicon, drawing on examples of UZWORDNET and emerging electronic thesaurus and corpus resources developed for the Uzbek language.*

Keywords: *thesaurus, lexicography, semantics, lexical-semantic relations, synonymy, antonymy, hyperonym, hyponymy, meronymy, electronic thesaurus, UZWORDNET, computational linguistics.*

Introduction

The lexical inventory of a language constitutes a complex system that reflects the historical, cultural, and cognitive experience of a society. The lexemes within a language's vocabulary do not exist independently of one another; rather, they enter into various semantic, paradigmatic, syntagmatic, and associative relationships. Therefore, a comprehensive description of a language's lexical system requires not only determining the meanings of individual words but also examining their relationships with other linguistic units. Unlike traditional explanatory or bilingual dictionaries, thesaurus dictionaries do not merely provide definitions of individual lexemes. In her study of the semantic and structural features of thesaurus dictionaries, M. Saparova emphasizes the importance of systematically representing the semantic relationships between lexemes.

From the perspective of modern computational linguistics, semantic resources such as WordNet are particularly important for understanding the concept of a thesaurus. In WordNet, words are organized into synonym sets, or *synsets*, which are interconnected through conceptual and semantic relationships. The aim of this study is to identify the principal structural, semantic, and functional characteristics of thesaurus dictionaries.

Research Objectives

The study addresses the following objectives:

1. to determine the linguistic essence of the concept of a thesaurus dictionary;
2. to describe the principal structural features of thesaurus dictionaries;

3. to identify the lexical-semantic relationships represented in thesaurus dictionaries;
4. to compare thesaurus dictionaries with traditional types of dictionaries;
5. to demonstrate the significance of thesaurus for modern linguistics and computational linguistics.

Methods

This study employs a comprehensive linguistic approach to identify the distinctive features of thesaurus dictionaries. The research methodology is based on lexicographic, semantic, and comparative-analytical approaches.

The object of the study comprises the lexical units represented in thesaurus dictionaries and the semantic relationships between them. The subject of the study encompasses the structure of thesaurus dictionaries, the principles governing the grouping of lexical units, and the ways in which synonymic, antonymic, hyperonymic-hyponymic, meronymic, and associative relationships are represented.

Through theoretical analysis, the study examines scholarly perspectives related to thesaurus, lexicography, and lexical-semantic relations. Lexicographic analysis is employed to investigate the structural features of thesaurus dictionaries, as well as the principles underlying the organization and semantic grouping of dictionary entries.

Semantic analysis is used to examine synonymic, antonymic, hyponymic-hyperonymic, and meronymic relationships between lexemes. In addition, the comparative-analytical method is applied to identify the features that distinguish thesaurus dictionaries from explanatory and other types of dictionaries.

Results

The results of the study demonstrate that, unlike traditional types of dictionaries, particularly explanatory dictionaries, thesaurus dictionaries are not limited to defining the lexical meanings of individual units. Rather, they are designed to systematically represent the various semantic relationships between lexical units. The principal distinctive feature of thesaurus dictionaries is their ability to present words not as isolated units but as elements of a semantically interconnected system. [3; 67-68] The analysis shows that the main characteristics of thesaurus dictionaries can be classified into several groups: semantic systematicity, the diversity of lexical-semantic relations, hierarchical structure, associative connections, and extensive information-retrieval capabilities.

1. Semantic Systematicity of the Thesaurus

The study established that semantic relationships between lexemes occupy a central position in thesaurus dictionaries. M. Saparova's research likewise pays particular attention to identifying the semantic and structural features of noun lexemes in the Uzbek thesaurus and systematizing them on the basis of their semantic relationships. [6; 120] In conventional dictionaries, users generally search for the meaning of a particular word. In a thesaurus, however, users can also identify other lexical units that are semantically related to that word. For example, units associated with the lexeme *uy* ("house") may be represented within the following semantic sequence:

uy → bino → turar joy → xona → oshxona → yotoqxona
house → building → dwelling → room → kitchen → bedroom

In this case, the connections between the lexemes are organized on the basis of conceptual and semantic relationships.

2. Representation of Synonymic Relations

According to the results of the study, synonymic relations constitute one of the important components of a thesaurus. Lexemes expressing the same or closely related meanings are grouped together. For example:

chiroyli → go‘zal → ko‘rkam → latofatli → suluv
beautiful → lovely → attractive → graceful → beautiful/elegant

Such an arrangement enables users to identify synonymous alternatives for a particular lexeme, reduce repetition in speech, and select the most appropriate lexical unit for a given context.

3. Representation of Antonymic Relations

Another important feature of thesaurus dictionaries is the systematic representation of antonymic relationships between lexemes. For example:

katta ↔ kichik, yaxshi ↔ yomon, issiq ↔ sovuq, uzoq ↔ yaqin
big ↔ small, good ↔ bad, hot ↔ cold, far ↔ near

Antonymic relations help reveal the semantic structure of lexical meanings through relationships of opposition and contrast.

4. Hyperonymic-Hyponymic Relations

The analysis demonstrates that hierarchical relationships between general concepts and their specific or subordinate categories are of particular importance in thesaurus dictionaries. For example:

meva → olma, anor, shaftoli *fruit → apple, pomegranate, peach*

Here, *meva* (“fruit”) represents the more general concept, whereas *olma* (“apple”), *anor* (“pomegranate”), and *shaftoli* (“peach”) represent more specific concepts subordinate to it. [2; 57-58] Such hierarchical relationships demonstrate the ability of a thesaurus to represent systems of knowledge and concepts. The representation of hierarchical relationships in Uzbek thesaurus has also been examined by M. Saparova. The researcher demonstrates how the relationships between hyperonymic and hyponymic units make it possible to organize concepts along a continuum from the general to the specific.

5. Presence of Meronymic Relations

The study also identifies relationships between a whole and its constituent parts as an important semantic feature of thesaurus dictionaries. For example:

avtomobil → dvigatel → g‘ildirak → rul *automobile → engine → wheel → steering wheel*

In this type of relationship, one lexeme is connected to a lexical unit denoting a constituent part of the object represented by another lexeme. As a result, a thesaurus can represent the internal semantic structure of objects and concepts.

6. Associative Relations

The results further indicate that thesaurus dictionaries encompass not only strict paradigmatic relationships but also associative connections. For example:

maktab → o‘qituvchi → o‘quvchi → dars → sinf → ta’lim *school → teacher → student → lesson → classroom → education*

Although these units are not directly synonymous or antonymous, they are interconnected through a shared conceptual and communicative domain. In this respect, a thesaurus offers broader semantic capabilities than a conventional synonym dictionary.

Overall, the findings demonstrate that thesaurus dictionaries are distinguished by their ability to systematize semantic relationships, represent connections between lexical units, and

enable users to identify lexical items associated with a particular concept in a comprehensive manner. These characteristics make thesaurus dictionaries important linguistic resources for modern lexicography and computational linguistics.

Discussion

The findings of the study demonstrate that the defining characteristic of thesaurus dictionaries is their relation-based nature. While the conventional model of an explanatory dictionary can be represented as **WORD → DEFINITION**, a thesaurus is better described by the model **WORD → WORD → SEMANTIC RELATIONSHIP**. In the case of an electronic thesaurus, this model can be further expanded as **LEXEME → SYNSET → SEMANTIC RELATIONSHIP → CONTEXT → CORPUS**.

The development of such a system for the Uzbek language would make it possible to digitize its rich lexical-semantic resources and provide a high-quality linguistic resource for natural language processing (NLP).

Conclusion

Based on the findings of the study, thesaurus dictionaries can be regarded as one of the important areas of modern lexicography. Their principal distinctive feature lies not simply in presenting or defining lexical units in alphabetical order, but in organizing them systematically according to their semantic relationships.

The representation of synonymic, antonymic, hyponymic, hyperonymic, meronymic, and associative relationships in thesaurus dictionaries makes it possible to provide a more comprehensive account of the interconnections among lexical units within the vocabulary.

The study demonstrates that when thesaurus dictionaries are structured according to the principles of semantic networks, users can obtain not only information about the meaning of a particular word but also about other concepts and lexemes associated with it.

These features enhance the practical value of thesaurus dictionaries and make them important linguistic resources for electronic dictionaries, semantic search systems, natural language processing, and computational linguistics.

In the development of a national thesaurus for the Uzbek language, UZWORDNET represents an important foundational resource. Further improving the identification of semantic relationships between lexemes on the basis of corpus data and modeling these relationships as an electronic semantic network remain important areas for future research.

REFERENCES

1. Agostini, A., Usmanov, T., Khamdamov, U., Abdurakhmonova, N., & Mamasaidov, M. (2021). UZWORDNET: A lexical-semantic database for the Uzbek language. *Proceedings of the 11th Global WordNet Conference*, 8–19.
2. Abduraxmonova, N., & Xoliyorova, G. (2024). Standard norms for creating a thesaurus of the Uzbek language. *Contemporary Technologies of Computational Linguistics*, 2(22.04), 464–469.
3. Xoliyorova, G. (2024). Modeling the linguistic database of the Uzbek thesaurus. *Acta NUUz*.
4. Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. MIT Press.
5. Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM, 38*(11), 39–41.

DIALEKTAL MATERIALLARNI YIG'ISHNING ZAMONAVIY METODLARI VA ULARNI SAMARQAND SHEVASI KORPUSINI YARATISHDA QO'LLASH

Qurbonnazarova Aisla O'ktamjonovna
O'zbekiston Milliy universiteti, Jurnalistika va o'zbek filologiyasi fakulteti,
Kompyuter lingvistikasi mutaxassisligi II bosqich magistranti
asilaqurbonnazarova74@gmail.com
https://doi.org/10.5281/zenodo.23253662

Annotatsiya. *Maqolada dialektal materiallarni yig'ishning zamonaviy metodlari va ularni Samarqand shevasi korpusini yaratishda qo'llash imkoniyatlari qiyosiy tahlil qilinadi. Jahon tajribasi asosida respondentlarni tanlash, tabiiy va yo'naltirilgan nutqni qayd etish, audio-video hujjatlashtirish, metadata va transkripsiyani tashkil etish tamoyillari umumlashtirilib, Samarqand shevasi uchun kompleks metodologik model taklif etiladi. Model materiallarni keyingi lingvistik annotatsiya va kompyuter lingvistik tahlilga tayyorlashga yo'naltirilgan.*

Kalit so'zlar: *dialektal korpus, Samarqand shevasi, dialektal material, dala tadqiqoti, spontan nutq, intervyu, audio yozuv, video yozuv, metadata, transkripsiya, lingvistik annotatsiya, kompyuter lingvistikasi.*

Abstract. *This article examines modern methods of collecting dialectal materials and their applicability to the development of a dialect corpus of the Samarkand variety of Uzbek. The study comparatively analyzes international practices of dialect corpus construction, with particular attention to geographical and socio-demographic sampling, field interviews, spontaneous speech recording, audio-video documentation, metadata construction, and multimedia transcription. Existing research on Uzbek dialect corpora is also reviewed in order to identify possibilities for integrating international methodological practices with the Uzbek linguistic context. As a result, a methodological model is proposed for collecting Samarkand dialect materials by combining geographically and socio-demographically balanced participant selection, natural and guided speech recording, audio-video documentation with metadata, transcription, and preparation for linguistic annotation. The proposed model can serve as a practical basis for developing an audio-multimedia corpus of the Samarkand dialect and for its subsequent application in computational linguistics research.*

Keywords: *dialect corpus, Samarkand dialect, dialectal data, fieldwork, spontaneous speech, interview, audio recording, video recording, metadata, transcription, linguistic annotation, computational linguistics.*

Аннотация. *В статье рассматриваются современные методы сбора диалектных материалов и возможности их применения при создании корпуса самаркандского говора узбекского языка. Сопоставляются методы полевого исследования, отбора информантов, записи спонтанной и направленной речи, аудио-видеодокументирования, формирования метаданных и транскрибирования. На основе международного опыта и существующих разработок узбекской диалектной корпусологии предлагается комплексная методологическая модель сбора материалов, ориентированная на последующую компьютерную обработку и лингвистическую аннотацию.*

Ключевые слова: *диалектный корпус, самаркандский говор, полевое исследование, спонтанная речь, интервью, аудиозапись, видеозапись, метаданные, транскрипция, компьютерная лингвистика.*

KIRISH

Raqamli lingvistikaning rivojlanishi dialektal nutqni faqat tavsifiy material sifatida emas, balki qayta ishlanadigan, qidiriladigan va turli lingvistik parametrlar bo‘yicha tahlil qilinadigan korpus resursi sifatida tashkil etish zaruratini kuchaytirmoqda. Dialektal korpusning ilmiy qiymati uning hajmi bilangina emas, balki tanlanmaning reprezentativligi, nutqning tabiiyligi, so‘zlovchi haqidagi ekstralingvistik ma’lumotlarning to‘liqligi va audio-video materiallarning saqlanishi bilan ham belgilanadi.

Dala tadqiqotida material yig‘ish usuli tadqiqotda olinishi ko‘zda tutilgan dalil turiga, tekshiruv shakliga, foydalaniladigan vositalarga va suhbat strategiyasiga mos bo‘lishi lozim. Shu bilan birga, tadqiqotchining ishtiroki nutqning tabiiyligiga ta’sir qilishi mumkinligi sababli intervyu va yozuv jarayonini imkon qadar kam intruziv tashkil etish muhimdir [1, 284–299-b.].

Jahon tajribasi dialektal korpuslar uchun turli janrdagi tabiiy nutqni, shaxsiy hikoyalarni, folklor materiallarini va intervyularni birgalikda yig‘ish, ularni audio yoki video yozuvlar bilan bog‘lash hamda so‘zlovchining ijtimoiy-demografik ma’lumotlarini metadata sifatida saqlash keng qo‘llanishini ko‘rsatadi [8, 633–644-b.]. Masalan, Russian National Corpus Dialect Corpus hozirgi holatida 3067 matn va 790 310 so‘zni qamrab oladi; unda spontan nutq, shaxsiy hikoyalar va folklor materiallari mavjud bo‘lib, matnlarning taxminan uchdan bir qismi to‘liq audio-video yozuvlar bilan ta’minlangan.

O‘zbek tilida ham dialektal korpusni shakllantirishning nazariy va kompyuter-texnologik asoslari ishlab chiqilmoqda. O‘zbekiston Respublikasi Oliy attestatsiya komissiyasining ma’lumotlariga ko‘ra, S. Abdurahmanovaning “O‘zbek tili dialektal korpusi ma’lumotlar bazasini shakllantirish” mavzusidagi PhD tadqiqotida dialektal korpusning ekstralingvistik va intralingvistik ma’lumotlar bazasini tizimlashtirish, lingvistik annotatsiyalash tamoyillarini ishlab chiqish va lakunar birliklar subkorpusini shakllantirish masalalari tadqiq etilgan. O‘zbek tili elektron korpusi platformasida esa dialekt, audio va multimedia korpuslari alohida resurslar sifatida ko‘rsatilgan.

Shu bilan birga, Samarqand shevasi uchun tabiiy og‘zaki nutqni dala sharoitida yig‘ish, respondentlarni hududiy va ijtimoiy-demografik mezonlar asosida tanlash, nutq janrlarini muvozanatlash, audio-video yozuvlarni standartlashtirish va ularni keyingi kompyuter lingvistik qayta ishlashga tayyorlash bosqichlarini yagona metodologik modelda birlashtirish masalasi alohida e’tibor talab qiladi.

Tadqiqotning maqsadi — jahon dialektal korpuslari tajribasi va o‘zbek dialektal korpusshunosligidagi mavjud ishlanmalarni qiyosiy tahlil qilish hamda Samarqand shevasi materiallarini yig‘ish uchun amaliy-metodologik model taklif etish.

Tadqiqot vazifalari quyidagilardan iborat: 1) jahon dialektal korpuslarida material yig‘ish usullarini aniqlash; 2) respondent tanlash, intervyu, audio-video yozuv, metadata va transkripsiya tamoyillarini qiyoslash; 3) mavjud o‘zbek dialektal korpus tajribasini tahlil qilish; 4) Samarqand shevasi uchun hududiy va ijtimoiy-demografik tanlanma modelini taklif etish; 5) yig‘ilgan materiallarni keyingi kompyuter lingvistik tahlilga moslashtirish tamoyillarini belgilash.

MUAMMONING O'RGANILGANLIK DARAJASI VA ADABIYOTLAR TAHLILI

Dialektologiyada dala intervyusi material olishning asosiy vositalaridan biri bo'lib, uning metodikasi XX asr davomida dialektologik atlaslar, so'rovnomalar va sotsiolingvistik tadqiqotlar bilan birga takomillashgan. G.Bailey dala intervyusida izlanayotgan dalil turini oldindan aniqlash, tegishli tadqiqot rejimi va vositalarini tanlash hamda tadqiqotchi ta'sirini kamaytirish zarurligini ta'kidlaydi [1, 284–299-b.]. R.Macaulay esa dialektologik tanlanmani shakllantirishda respondentlarni tanlash va namunani tashkil etish masalalarini alohida metodologik muammo sifatida ko'rib chiqadi. C.Llamas dialekt so'rovnomasining maqsadga muvofiq tuzilishi [3, 253–267-b.], J.K.Chambers esa yozma dialekt so'rovlari imkoniyatlarini tahlil qiladi [4, 268–283-b.].

Korpus lingvistikasi nuqtayi nazaridan dialekt materialini tizimli, mashinada qayta ishlanadigan ma'lumotlar majmuasiga aylantirish zarur. B.Szmrecsanyi va L.Anderwald korpusga asoslangan dialekt tadqiqotlarida tabiiy nutq materiallaridan sifat va miqdoriy tahlilni birgalikda foydalanish imkoniyatlarini ko'rsatadi [5, 300–313-b.]. S.Zufferey korpus qurishda sampling, reprezentativlik, kodlash, spoken corpus uchun kontekstual ma'lumotlar va etik-huquqiy masalalarni muhim metodologik komponentlar sifatida ko'rsatadi.

Xalqaro korpuslar ushbu tamoyillarning turli modellarini namoyish etadi. Nationwide Speech Project AQShning olti dialektal hududidan 60 nafar so'zlovchini qamrab olgan bo'lib, materiallar so'zlar, gaplar, parchalar va intervyu nutqini o'z ichiga oladi [8, 633–644-b.]. CORAAL esa mintaqaviy va ijtimoiy variativlikni o'rganishga mo'ljallangan audio, ELAN fayllari, metadata va matnlarni birgalikda taqdim etadi. Russian National Corpus Dialect Corpus spontan nutq, shaxsiy hikoyalar, folklor va multimedia yozuvlarini birlashtiradi.

Beserman Multimedia Corpus tajribasida matnlarning audio-video bilan vaqt bo'yicha bog'lanishi, ELAN transkripsiyalari va alohida metadata fayllari qo'llanadi. Mazkur tajriba dialektal korpusni oddiy matnlar to'plami emas, balki multimedia va strukturaviy metadata bilan bog'langan raqamli resurs sifatida qurish mumkinligini ko'rsatadi.

O'zbek tilida dialektal korpusni yaratishga qaratilgan ilmiy ishlarda esa ma'lumotlar bazasi, annotatsiya, dialektal birliklarning xususiyatlarini tizimlashtirish va milliy korpus bilan integratsiya masalalari ishlab chiqilmoqda. Demak, mavjud ilmiy baza Samarqand shevasi uchun yangi korpus konsepsiyasini mutlaqo noldan emas, balki mavjud nazariy va texnologik tajribani dala ma'lumotlarini yig'ish metodikasi bilan chuqurroq integratsiyalash asosida qurishni taqozo etadi.

TADQIQOT METODOLOGIYASI

Tadqiqotda qiyosiy-tahliliy, tavsifiy, korpusologik va metodologik modellashtirish yondashuvlari uyg'unlashtirildi. Jahon dialektal korpuslari bo'yicha ilmiy adabiyotlar va rasmiy korpus resurslari tahlil qilinib, material yig'ishning takrorlanadigan metodik komponentlari ajratib olindi. O'zbek tilining dialektal korpusiga oid rasmiy ilmiy ma'lumotlar bilan xalqaro tajriba qiyoslandi [11, 305–320-b.].

Tadqiqotning obyekti — dialektal korpus tuzish texnologiyalarida qo'llanadigan dialektal materiallar. Predmeti — dala tadqiqoti, respondent tanlash, intervyu, spontan nutq, audio-video yozuv, metadata va transkripsiya orqali dialektal materiallarni yig'ish usullari. Tadqiqotning amaliy yo'nalishi Samarqand shevasi materiallarini keyinchalik kompyuterda qayta ishlashga tayyorlash bilan belgilanadi.

Muhim metodologik cheklov shundaki, maqolada taklif qilinayotgan respondentlar tarkibi va materiallar foizlari empirik jihatdan sinovdan o'tkazilgan universal norma sifatida

emas, balki pilot tadqiqot uchun dastlabki model sifatida qaraladi. Ularning maqbulligi real dala tadqiqoti natijalari bilan tekshirilishi lozim.

ASOSIY QISM VA NATIJALAR

Jahon dialektal korpuslarini shakllantirish tajribasi

Jahon tajribasining qiyosiy tahlili besh asosiy tamoyilni ajratish imkonini berdi: 1) hududiy va ijtimoiy jihatdan asoslangan tanlanma; 2) tabiiy og‘zaki nutqqa ustuvorlik; 3) audio-video materiallarni saqlash; 4) to‘liq metadata; 5) materiallarni keyingi kompyuter lingvistik qayta ishlashga mos holda yig‘ish.

Russian National Corpus Dialect Corpus keng hududiy qamrovga ega bo‘lib, spontan nutq, shaxsiy hikoyalar va folklor materiallarini o‘z ichiga oladi. Rasmiy korpus sahifasida 3067 matn va 790 310 so‘z qayd etilgan; materiallarning taxminan uchdan biri to‘liq audio-video yozuvlar bilan ta’minlangan. Korpus ma’lumotlarining bunday tuzilishi dialektal materialni yozma transkripsiya bilangina cheklamaslik zarurligini ko‘rsatadi.

Nationwide Speech Project modeli hududlar bo‘yicha muvozanatli so‘zlovchilar tarkibi va bir necha turdagi nutq materiallarini birlashtirishga asoslangan [8, 633–644-b.]. CORAAL modeli esa audio, ELAN, metadata va matn komponentlarini yagona tadqiqot ekotizimida saqlash imkonini beradi. Beserman Multimedia Corpusda esa dialog, monolog va eksperimental materiallar audio-video bilan bog‘langan holda saqlanadi.

Korpus / model	Asosiy material	Texnologik xususiyat	Samarqand shevasi korpusini yig‘ishga tatbiq etiladigan jihatlari
Russian National Corpus Dialect Corpus	Spontan nutq, hikoya, folklor	Metadata, qidiruv, audio-video	Janr va multimedia xilma-xilligini saqlash
Nationwide Speech Project	So‘z, gap, matn, intervyu	Hududlar bo‘yicha balans	Hududiy va ijtimoiy tanlanmani rejalash
CORAAL	Og‘zaki nutq va intervyular	Audio, ELAN, metadata, matn	Multimedia va metadata integratsiyasi
Beserman Multimedia Corpus	Dialog, monolog, intervyu	ELAN, audio-video, metadata	Nutq turini va fayllarni strukturaviy bog‘lash

1-jadval. Jahon dialektal korpuslari tajribasining Samarqand shevasi korpusini yig‘ishga tatbiq etiladigan jihatlari.

O‘zbek dialektal korpusi tajribasi va metodologik bo‘shliq

O‘zbek dialektal korpusini yaratish bo‘yicha mavjud tadqiqotlar milliy korpus tarkibida dialektal subkorpuslarning o‘rnini belgilash, ma’lumotlar bazasini shakllantirish va lingvistik annotatsiya tamoyillarini ishlab chiqish kabi yo‘nalishlarda muhim natijalarni taqdim etadi. Bu holat maqolada taklif qilinayotgan modelni mavjud ishlanmalarga qarshi qo‘yish emas, balki ularni dala materiallarini yig‘ish metodikasi bilan to‘ldirish zarurligini ko‘rsatadi.

Metodologik jihatdan dolzarb masala Samarqand shevasi misolida respondentlarni hududiy va ijtimoiy-demografik mezonlar asosida tanlash, tabiiy va yo‘naltirilgan nutqni birgalikda qayd etish, turli janrdagi materiallarni muvozanatlash, audio-video yozuvlarni

standartlashtirish va metadata bilan hujjatlashtirish jarayonlarini yagona protokolga keltirishdir. Mazkur masala alohida metodologik model sifatida yanada tizimlashtirilishi mumkin.

Samarqand shevasi uchun taklif etilayotgan metodologik model

Samarqand shevasi dialektal korpusini shakllantirish uchun quyidagi ketma-ketlik taklif etiladi:

RESPONDENT TANLASH → DALA TADQIQOTI → ERKIN SUHBAT + YARIM TUZILGAN INTERVYU + MAXSUS TOPSHIRIQLAR → AUDIO/VIDEO YOZUV → METADATA → TRANSKRIPSIYA → NORMALIZATSIYA → LINGVISTIK ANNOTATSIYA → SIFAT NAZORATI → KORPUS → QIDIRUV VA KOMPYUTER LINGVISTIK TAHLIL.

Hududiy va ijtimoiy-demografik tanlanma

Korpusning reprezentativligi faqat geografik hudud bilan belgilanmaydi. Samarqand shevasi bo‘yicha material yig‘ishda tuman, shahar/qishloq, mahalla yoki qishloq, so‘zlovchining tug‘ilgan joyi va yashash tarixi, yoshi, jinsi, ta’limi, kasbi hamda boshqa tillar bilan aloqasi kabi ko‘rsatkichlarni metadata sifatida qayd etish taklif qilinadi.

Respondentlar uchun mahalliy dialekt bilan aloqasi barqaror bo‘lgan so‘zlovchilarni jalb qilish, uzoq muddat boshqa hududda yashagan shaxslar bo‘yicha esa migratsion tarixni alohida qayd etish maqsadga muvofiq. Bunday ma’lumotlar keyinchalik hududiy va ijtimoiy variativlikni tahlil qilish imkonini beradi.

Dialektal material turlari

Taklif etilayotgan dastlabki material tarkibi quyidagicha: erkin suhbat — 40%; yarim tuzilgan intervyu — 25%; tematik suhbat — 15%; monolog, hikoya va shaxsiy narrativ — 10%; folklor — 5%; maqsadli lingvistik topshiriqlar — 5%. Ushbu nisbatlar Beserman kabi multimedia korpuslaridagi material xilma-xilligidan ilhomlangan metodologik taklif bo‘lib, universal standart yoki empirik natija sifatida talqin qilinmasligi kerak [10, 57-62-b.].

Erkin suhbat tabiiy nutqni maksimal darajada saqlashga xizmat qiladi. Yarim tuzilgan intervyu esa barcha respondentlarda taqqoslanadigan mavzular bo‘yicha material olish imkonini beradi. Maxsus lingvistik topshiriqlar fonetik yoki grammatik hodisalarni maqsadli tekshirish uchun qo‘llanadi. Shu tariqa tabiiylik va nazorat qilinadigan taqqoslash o‘rtasidagi muvozanat ta’minlanadi.

Audio-video yozuv va metadata

Audio yozuvlarda nutqning qayta ishlanishi va arxivlanishi uchun barqaror, siqilmagan yoki sifatni maksimal saqlaydigan formatlardan foydalanish maqsadga muvofiq. Video yozuvlar esa nutq jarayonidagi kommunikativ vaziyat va zarur hollarda noverbal kontekstni saqlash imkonini beradi. Yozuv sifati, qurilma, sana, joy va davomiylik metadata tarkibida qayd etilishi kerak.

Har bir materialga yagona identifikator berish tavsiya etiladi. Masalan: SAM-F-55-001-A.wav — audio; SAM-F-55-001-V.mp4 — video; SAM-F-55-001.eaf — ELAN annotatsiyasi. Bu nomlash tizimi yakuniy korpus arxitekturasining bir qismi sifatida qaralishi lozim.

Transkripsiya va annotatsiya

Dialektal korpusda transkripsiyaning ikki qatlamli modeli taklif qilinadi: birinchi qatlamda so‘zlovchining real dialektal nutqi maksimal darajada saqlanadi, ikkinchi qatlamda esa adabiy tilga yaqinlashtirilgan yoki normalizatsiyalangan shakl beriladi. Bu ikki qatlam dialektal shaklni yo‘qotmagan holda standart til bilan qiyoslash va qidiruv imkonini kengaytiradi.

Multimedia annotatsiyasi uchun ELAN kabi tier-based muhitdan foydalanish maqsadga muvofiq. Dastlabki qatlamlar quyidagicha tashkil etilishi mumkin: Speaker, Dialect transcription, Standard/normalized Uzbek, Morpheme, Gloss, Translation, Phonetic transcription va Context/Comment. ELAN vaqtga bog‘langan annotatsiyalarni audio-video bilan sinxronlashtirish imkonini beradi.

Bu bosqichda transkripsiya qidiruv uchun, original audio-video esa ilmiy tekshiruv va fonetik-kontekstual tahlil uchun xizmat qiladi. Shuning uchun korpusning matnli qismi media fayllarning o‘rnini bosmasligi kerak.

Metadata modeli va sifat nazorati

Har bir yozuv uchun Speaker ID, hudud, tuman, mahalla/qishloq, jins, yosh, tug‘ilgan joy, yashash tarixi, ta’lim, kasb, ona tili, boshqa tillar, ota-onaning kelib chiqishi, yozuv sanasi, yozuv turi, davomiyligi, audio/video formati, transkripsiya tizimi, rozilik holati va foydalanish darajasi kabi maydonlarni kiritish taklif etiladi.

Sifat nazoratida kamida to‘rt bosqich nazarda tutiladi: yozuv sifatini tekshirish; metadata to‘liqligini tekshirish; transkripsiya va audio/video mosligini tekshirish; annotatsiya qatlamlari o‘rtasidagi muvofiqlikni tekshirish. Ikki tadqiqotchi tomonidan tanlangan namunani mustaqil qayta ko‘rib chiqish kelajakdagi pilot bosqichda foydali nazorat vositasi bo‘lishi mumkin.

Etik va huquqiy talablar

Og‘zaki nutq yozuvlari shaxsiy ma’lumotlarni o‘z ichiga olishi mumkinligi sababli respondentning ongli roziligi olinishi, materialdan foydalanish darajasi oldindan tushuntirilishi va kerak bo‘lganda shaxsni aniqlash imkonini beruvchi ma’lumotlar cheklanishi kerak. Ochiq foydalanish, cheklangan foydalanish va faqat tadqiqot guruhi uchun saqlash kabi kirish darajalarini belgilash multimedia korpus uchun maqbul tashkiliy yechim bo‘lishi mumkin.

MUHOKAMA

Tahlil natijalari dialektal korpus uchun tabiiy nutq va nazorat qilinadigan nutq materiallarini qarama-qarshi qo‘yish emas, balki ularni funksional jihatdan to‘ldiruvchi qatlamlar sifatida ko‘rish zarurligini ko‘rsatadi. Erkin suhbat korpusning tabiiyligini ta’minlasa, yarim tuzilgan intervyu va maxsus topshiriqlar ma’lum lingvistik hodisalarni turli respondentlar kesimida taqqoslash imkonini beradi.

Ikkinchi muhim masala reprezentativlikdir. Hududiy samplingning o‘zi yetarli emas: yosh, jins, ta’lim, kasb, migratsion tarix kabi ijtimoiy ko‘rsatkichlar ham hisobga olinishi lozim. Aks holda korpusda ma’lum bir guruhning nutqi haddan tashqari ko‘p ifodalanishi va keyingi statistik xulosalar tor tanlanmaga bog‘lanib qolishi mumkin [2, 241–252-b.].

Uchinchi jihat — multimedia ma’lumotlarining ilmiy qiymati. Transkripsiya qidiruv va avtomatlashtirilgan tahlilni yengillashtiradi, biroq fonetik xususiyatlar, pauza, intonatsiya va kontekstni qayta tekshirish uchun original audio-video yozuv zarur. Shu sababli Samarqand korpusida transkripsiya va media fayllarni alohida emas, o‘zaro bog‘langan resurslar sifatida tashkil etish taklif qilinadi [11, 305–320-b.].

To‘rtinchi jihat metadata bilan bog‘liq. Dialektal materialni faqat matn sifatida saqlash so‘zlovchining hududiy va ijtimoiy xususiyatlarini keyingi tahlilda hisobga olish imkonini kamaytiradi. Metadata korpusning yordamchi tavsifi emas, balki uning ilmiy interpretatsiyasi uchun zarur tarkibiy qismdir [8, 633–644-b.].

Beshinchi jihat — xalqaro modellarni mexanik ko‘chirmaslik. Har bir dialektning hududiy tuzilishi, so‘zlovchilar tarkibi, til kontaktlari, etik sharoitlari va mavjud texnik

infratuzilmasi farq qiladi. Shu sababli Samarqand modeli xalqaro tajribadagi tamoyillarni saqlagan holda mahalliy dala sharoitiga moslashtirilishi kerak.

Taklif etilgan modelning cheklovi shundaki, u hozircha konseptual-metodologik bosqichda. Respondentlar soni, hududlar bo‘yicha taqsimot va material janrlarining foiz nisbati pilot tadqiqot orqali tekshirilmagan. Shuning uchun ushbu ko‘rsatkichlar yakuniy standart emas, keyingi empirik tadqiqot uchun boshlang‘ich parametrlar sifatida qaraladi.

XULOSA

Tadqiqotda dialektal materiallarni yig‘ishning zamonaviy usullari jahon dialektal korpuslari va mavjud o‘zbek korpusshunosligi tajribasi asosida qiyosiy tahlil qilindi. Tahlil natijasida dialektal korpusning ilmiy qiymatini belgilovchi asosiy omillar — reprezentativ tanlanma, tabiiy og‘zaki nutq, multimedia hujjatlashtirish, to‘liq metadata va keyingi kompyuter lingvistik qayta ishlashga mos strukturaviy tashkil etish — sifatida ajratildi.

Samarqand shevasi uchun respondentlarni hududiy va ijtimoiy-demografik mezonlar asosida tanlash, erkin suhbatni yarim tuzilgan intervyu, tematik suhbat, narrativ, folklor va maqsadli topshiriqlar bilan to‘ldirish, audio-video materiallarni metadata bilan hujjatlashtirish, ikki qatlamli transkripsiya va ELAN asosidagi annotatsiyaga tayyorlashni birlashtiruvchi metodologik model taklif etildi.

Mazkur modelning ilmiy yangiligi mavjud xalqaro va milliy tajribalarni Samarqand shevasi materiallarini yig‘ishning yagona metodik ketma-ketligiga integratsiyalash bilan bog‘liq. Biroq modelning yakuniy samaradorligi pilot dala tadqiqotlari bilan tekshirilishi kerak. Shu jihatdan maqolada taklif etilgan material foizlari va sampling parametrlari qat’iy norma emas.

Kelgusida model asosida Samarqand shevasining real audio-multimedia korpusini shakllantirish, hududiy va ijtimoiy variativlikni empirik o‘rganish, fonetik, leksik, morfologik va sintaktik annotatsiya qatlamlarini ishlab chiqish hamda korpusni O‘zbek tili milliy korpusi infratuzilmasi bilan integratsiya qilish mumkin.

AMALIY VA ILMIY TAKLIFLAR

- Samarqand shevasi bo‘yicha pilot dala tadqiqotini bir necha hududda sinov tariqasida o‘tkazish va sampling parametrlarini empirik ma’lumotlar asosida aniqlashtirish.
- Har bir respondent uchun standart metadata formasini ishlab chiqish va barcha audio-video materiallarni yagona identifikator bilan bog‘lash.
- Tabiiy nutqni asosiy material sifatida saqlagan holda yarim tuzilgan intervyu va maqsadli lingvistik topshiriqlarni qo‘shimcha qatlam sifatida kiritish.
- Original audio/video fayllarni saqlash, transkripsiyani esa ular bilan vaqt bo‘yicha bog‘lash.
- Dialektal va normalizatsiyalangan transkripsiyani parallel qatlamlarda tashkil etish.
- ELAN yoki mos multimedia annotatsiya vositasida tier-based annotatsiya modelini ishlab chiqish.
- Ochiq, cheklangan va yopiq foydalanish darajalarini respondent roziligi asosida belgilash.
- Pilot korpus asosida keyinchalik fonetik, leksik, morfologik va sintaktik annotatsiya standartlarini takomillashtirish.

FOYDALANILGAN ADABIYOTLAR

1. G.Bailey. Field Interviews in Dialectology // The Handbook of Dialectology / eds. C.Boberg, J.Nerbonne, D.Watt. — Wiley-Blackwell, 2017. — P. 284–299. — DOI: 10.1002/9781118827628.ch16.

2. R.Macaulay. Dialect Sampling Methods // The Handbook of Dialectology / eds. C.Boberg, J.Nerbonne, D.Watt. — Wiley-Blackwell, 2017. — P. 241–252. — DOI: 10.1002/9781118827628.ch13.
3. C.Llamas. The Dialect Questionnaire // The Handbook of Dialectology / eds. C.Boberg, J.Nerbonne, D.Watt. — Wiley-Blackwell, 2017. — P. 253–267. — DOI: 10.1002/9781118827628.ch14.
4. J.K.Chambers. Written Dialect Surveys // The Handbook of Dialectology / eds. C.Boberg, J.Nerbonne, D.Watt. — Wiley-Blackwell, 2017. — P. 268–283. — DOI: 10.1002/9781118827628.ch15.
5. B.Szmrecsanyi, L.Anderwald. Corpus-Based Approaches to Dialect Study // The Handbook of Dialectology / eds. C.Boberg, J.Nerbonne, D.Watt. — Wiley-Blackwell, 2017. — P. 300–313. — DOI: 10.1002/9781118827628.ch17.
6. S.Zufferey. How to Build a Corpus // Introduction to Corpus Linguistics. — Wiley, 2020. — DOI: 10.1002/9781119779728.ch6.
7. Russian National Corpus. Dialect Corpus [Elektron resurs]. — URL: https://ruscorpora.ru/en/corpus/dialect.
8. C.G.Clopper, D.B.Pisoni. The Nationwide Speech Project: A New Corpus of American English Dialects // Speech Communication. — 2006. — Vol. 48, No. 6. — P. 633–644. — DOI: 10.1016/j.specom.2005.09.010.
9. T.Kendall, C.Farrington. The Corpus of Regional African American Language (CORAAL). Version 2023.06. — 2023. — DOI: 10.7264/1ad5-6t35.
10. T.Arkhangelskiy, M.Usacheva. Beserman Multimedia Corpus [Dataset]. — 2025. — DOI: 10.25592/uhhfdm.16991.
11. H.Sloetjes. ELAN: Multimedia Annotation Application // The Oxford Handbook of Corpus Phonology. — Oxford University Press, 2013. — P. 305–320. — DOI: 10.1093/oxfordhb/9780199571932.013.019.
12. N.Thieberger, A.L.Berez-Kroeker. Linguistic Data Management // The Oxford Handbook of Linguistic Fieldwork / ed. N.Thieberger. — Oxford University Press, 2011. — P. 90–118. — DOI: 10.1093/oxfordhb/9780199571888.013.0005.
13. S.B.Abdurahmanova. O‘zbek tili dialektal korpusi ma’lumotlar bazasini shakllantirish: filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. — Toshkent, 2025.
14. O‘zbek tili korpusi [Elektron resurs]. — Dialekt korpusi, audio korpus va multimedia korpusi. — URL: https://uzbekcorpus.uz/newIndex.

BADIIY DISKURSDA FRAZEOLOGIK BIRLIKLARNING LINGVOTERAPEVTIK TA'SIRI (CHO‘LPONNING "KECHA VA KUNDUZ" ROMANI MISOLIDA)

Rahimov Usmonjon Ergashevich[19], Sirojiddinova Muxarram Raximjon qizi[20]

https://doi.org/10.5281/zenodo.23253687

Annotatsiya. *Mazkur maqolada badiiy diskursda qo‘llanuvchi frazeologik birliklarning adresat ongi va ruhiyatiga ko‘rsatadigan ta'siri lingvoterapevtik nuqtai nazardan tahlil qilinadi. Diskursiv yondashuv frazeologik birlikni faqat matn ichidagi til birligi sifatida emas, balki muallif–matn–adresat kommunikativ zanjirining faol elementi sifatida ko‘rib chiqishga imkon beradi. Frazeologik birliklarning obrazliligi, ekspressivligi va emotsional-baholovchi semantikasi adresatda hissiy tanish (identifikatsiya), katarsis va ruhiy yengillashuv hamda shifo potensialini yuzaga keltirish salohiyatiga ega ekanligi asoslanadi. Tadqiqot materializ sifatida Cho‘lponning "Kecha va kunduz" romanidagi frazeologik birliklar tahlilga tortiladi.*

Kalit so‘zlar: *frazeologik birlik, badiiy diskurs, lingvoterapiya, biblioterapiya, emotsional-ekspressiv semantika, obrazlilik, katarsis, adresat retseptsiyasi, Cho‘lpon, "Kecha va kunduz".*

Ushbu annotatsiya hamda kalit so‘zlarning akademik darajadagi ingliz va rus tillaridagi tarjimasi:

Abstract. *This article analyzes the impact of phraseological units used in literary discourse on the recipient's consciousness and psyche from a linguotherapeutic perspective. The discursive approach allows for viewing phraseological units not merely as textual linguistic elements, but as active components within the author–text–recipient communicative chain. It is established that the imagery, expressiveness, and emotional-evaluative semantics of phraseological units have the potential to evoke emotional identification, catharsis, psychological relief, and therapeutic effects in the recipient. Phraseological units from Cho‘lpon's novel Night and Day (Kecha va kunduz) serve as the primary research material for analysis.*

Keywords: *phraseological unit, literary discourse, linguotherapy, bibliotherapy, emotional-expressive semantics, imagery, catharsis, recipient reception, Cho‘lpon," Night and Day".*

Аннотация. *В данной статье с лингвотерапевтической точки зрения анализируется влияние фразеологических единиц, используемых в художественном дискурсе, на сознание и психику адресата. Дискурсивный подход позволяет рассматривать фразеологическую единицу не просто как языковую единицу внутри текста, а как активный элемент коммуникативной цепи «автор–текст–адресат». Обосновывается, что образность, экспрессивность и эмоционально-оценочная семантика фразеологических единиц обладают потенциалом вызывать у адресата эмоциональное тождество (идентификацию), катарсис, психологическое разряжение и*

[19] Andijon davlat universiteti O‘zbek tilshunosligi kafedrasi dotsenti, filologiya fanlari nomzodi
[20] Andijon davlat universiteti O‘zbek tilshunosligi kafedrasi o‘qituvchisi

терапевтический эффект. В качестве материала исследования анализируются фразеологические единицы из романа Чулпана «Ночь и день» (Kecha va kunduz).

Ключевые слова: *фразеологическая единица, художественный дискурс, лингвотерапия, библиотерапия, эмоционально-экспрессивная семантика, образность, катарсис, рецепция адресата, Чулпан, «Ночь и день».*

Til nafaqat axborot uzatish quroli, balki inson ruhiy holatini ifodalash orqali boshqalarga bevosita ta'sir ko‘rsatuvchi vosita hamdir. Bu haqiqat, ayniqsa, badiiy diskurs misolida yaqqol namoyon bo‘ladi: yozuvchi so‘z va so‘z birikmalari orqali nafaqat voqelikni tasvirlaydi, balki adresat bilan bevosita hissiy-kommunikativ aloqaga kirishadi, yozuvchi – matn – o‘quvci(adresat) munosabatdorligi kommunikativ kanal – badiiy matn orqali yuzaga chiqadi va tinglovchida muayyan ruhiy kechinmalarni uyg‘otadi.

Frazeologik birliklar — ikki yoki undan ortiq leksik birliklarning birikuvidan hosil bo‘lgan, ko‘chma ma’noli hamda yaxlit obrazli ma'noga ega turg‘un birikmalar — bu jarayonda alohida o‘rin tutadi, chunki ular tilning eng "hissiy, mantiqiy, ma‘naviy zichlashgan" qatlamini tashkil etadi.

Diskurs tushunchasi lotincha “discursus” – “mulohaza yuritish”, “oldinga va orqaga yugurish”, “harakat”, “aylanish”, “suhbat” yoki fransuzcha “Discours” – “nutq” degan ma’nolarga borib taqaladi. Gapdan uzun bo‘lgan, izchil ma’no va aniq maqsadni ko‘zlagan har qanday bog‘langan nutq yoki yozuv birligi diskurs deyiladi[21]. Diskurs tushunchasi matn tushunchasidan farqli o‘laroq, til birligini nafaqat lingvistik struktura, balki uning yuzaga kelish sharoiti, muallif niyati, adresat retseptsiyasi va kommunikativ-pragmatik vazifasi bilan birgalikda o‘rganishni talab qiladi. Shu nuqtai nazardan, Cho‘lponning "Kecha va kunduz" romanidagi frazeologik birliklar shunchaki ko‘chma ma’noli til vositasi emas, balki muallif–qahramon–adresat o‘rtasidagi ruhiy-kommunikativ ko‘prik vazifasini bajaradi: ular orqali qahramonning ichki kechinmasi adresatga uzatiladi va adresatda hissiy javob (rezonans) uyg‘otadi. Aynan shu diskursiv-kommunikativ ta'sir mazkur maqolada "lingvoterapevtik ta'sir" deb nomlanadi.

So‘nggi yillarda tilshunoslikda antroposentrik yo‘nalishdagi tadqiqotlarning ko‘payishi natijasida til hodisalarini nafaqat struktur-semantik, balki psixologik va terapevtik nuqtai nazardan o‘rganishga ehtiyoj tug‘ildi. O‘zbek tilshunosligida bu yo‘nalish tadqiqotchilar tomonidan "lingvoterapiya" deb nomlanib, lingvistika, psixologiya, tibbiyot va neyrolingvistikaning integratsion kesishmasida shakllangan gibrid makrosistema sifatida talqin etiladi[22]. Ushbu maqolaning maqsadi — badiiy diskursdagi, xususan Cho‘lponning "Kecha va kunduz" romanidagi frazeologik birliklarning aynan shu — lingvoterapevtik — funksiyasini badiiy matnlar asosida nazariy asoslab berishdir.

Lingvoterapiya tushunchasi o‘zbek tilshunosligida yangi yo‘nalish sifatida talqin etilmoqda. Lingvoterapiya — antroposentrik va pragmatik paradigmalar kesishuvida shakllangan, lingvistika, psixologiya, tibbiyot va neyrolingvistikaning integratsion tutashuvida

[21] M.U. Saidova , N.A.Nadirxanova . Diskurs Tushunchasi Tahlili Va Tadqiqi // Miasto Przyszłości Kielce 2023 Impact // Factor: 9.2 ISSN-L: 2544-980X. 136-138 bet

[22] U. Rahimov. Lingvoterapiya va uning integratsion yo‘nalishlari tadqiqi//— Andijon davlat universiteti

yuzaga kelgan gibrid yo‘nalish bo‘lib, "shaxs ruhiy holatini barqarorlashtirish va kognitiv tuzilmalarni ijobiy tomonga siljishiga yo‘naltirilgan konstruktiv ta'sir" sifatida talqin etiladi [23].

Lingvoterapiya makrosistema sifatida bir qator mikrosistemalarga bo‘linadi[24]: fonoterapiya, leksikoterapiya, morfoterapiya, sintaksoterapiya, poetoterapiya, teolingvistikoterapiya, germenevtikoterapiya, stilistikoterapiya, kognitoterapiya, sinkretizmoterapiya, sinergoterapiya. Muhimi shundaki, biz tadqiq qilayotgan frazeoterapiya mikrosistema sifatida leksikoterapiya tarkibiga kirib, gibrid sistemaning bir bo‘lagi sifatida muhim ahamiyatga ega. Shu sababli ushbu kichik tadqiqot “frazeoterapiyani” — Cho‘lponning "Kecha va kunduz" romanidan olingan frazeolgik birliklari asosida konkretlashtirishga qaratilgan. Shu ma'noda mazkur tadqiqot lingvoterapiya nazariyasi bilan frazeologiya nazariyasi o‘rtasidagi amaliy ko‘prik vazifasini bajaradi.

Frazeologik birliklarning boshqa til birliklaridan farqi shundaki, ular denotativ ma'no bilan bir qatorda kuchli konnotativ — emotsional, baholovchi, ekspressiv — ma'no qatlamiga ega. Masalan, "yuragi orziqmoq", "ko‘ngli bo‘sh", "boshi qotmoq" kabi birliklar shunchaki holatni nomlamaydi, balki adresatda shu holatni ichdan his qilish, unga hamdardlik bildirish imkonini yaratadi. Shu sababli frazeologik birliklar adresat ruhiyatida bir qancha lingvoterapevtik funksiyaga ega bo‘ladi:

1. Obrazli identifikatsiya funksiyasi — adresat frazeologik obraz orqali qahramon kechinmasini o‘zinikidek his qiladi.

2. Emotsional zichlashtirish funksiyasi— murakkab ruhiy holat bir necha so‘zdan iborat turg‘un birikma orqali ixcham va ta'sirchan ifodalanadi;

3. Katarsis funksiyasi — og‘ir, og‘riqli kechinmalarning badiiy-obrazli shaklda ifodalanishi adresatda ruhiy yengillashuvni keltirib chiqaradi.

4. Milliy-madaniy tasalli funksiyasi— xalq donishmandligidan kelib chiqqan frazemalar (masalan, maqol-frazemalar) adresatga tanish, "o‘zinikidek" tuyuluvchi ruhiy tayanch beradi.

Bilamizki, lingvoterapevtik ta'sir bir necha bosqichlar orqali izohlaniladi:

- Idrok bosqichi — adresat frazeologik birlikni matn kontekstida o‘qiydi va uning obrazli ma'nosini tiklaydi (dekodlaydi);
- hissiy rezonans bosqichi— obraz adresatning shaxsiy tajribasi bilan bog‘lanib, hissiy javob (rezonans) uyg‘otadi;
- ruhiy ishlov berish bosqichi — bu hissiy javob orqali adresat o‘z ichki kechinmalarini "chetdan turib" ko‘rish, ularni anglash imkoniga ega bo‘ladi;
- yengillashuv (katarsis) bosqichi — matn yakunida yoki muayyan epizodda hissiy zo‘riqish pasayadi, adresat ruhiy muvozanatga qaytadi;
- yozuvchi syujet pragmatikasi orqali o‘quvchini o‘z asari ichiga olib kirib, ruhiy boshqaradi.

Bu jarayonda frazeologik birlikning obrazliligi muhim rol o‘ynaydi: mavhum tushunchalar (qayg‘u, umid, pushaymonlik, sog‘inch) konkret, ko‘rgazmali obrazlar (yurak, ko‘ngil, bag‘ir, ko‘z yoshi kabi somatik komponentlar orqali) vositasida beriladi, bu esa insonga perlokutiv ta’sirini osonlashtiradi.

Tadqiqot materiali sifatida Cho‘lponning "Kecha va kunduz" romani asosida tuzilgan frazeologik birliklar bazasi (456 FB, MORF/SEM/THM/TR/TDI teglari bilan belgilangan)dan

[23] U. Rahimov, "Diskursda manipulyatsiya va lingvoterapevtik ta'sirning psixolingvistik tadqiqi//— Andijon davlat universiteti

[24] U. Rahimov. Lingvoterapiya va uning integratsion yo‘nalishlari tadqiqi//— Andijon davlat universiteti

foydalanildi. Bilamizki, frazeologik birliklar tematik belgisiga ko'ra bir qancha guruhlarga (NAT (tabiat), SOM (somatik), SOC (ijtimoiy), PSY (psixologik), TIME (vaqt)) bo'linadi. Ular orasida psixologik (PSY) yoki somatik (SOM) guruhga mansub frazeologizmlar insonning ichki kechinmalari va ruhiy holatini ifodalashda ustunlik qilishi aniqlandi. Diskursiv tahlil nuqtai nazaridan yana bir muhim jihat shundaki, tanlangan misollarning aksariyati qahramonlar nutqida — dialog, ichki monolog yoki xitob shaklida — namoyon bo'ladi (jadvaldagi kontekst gaplarining tirnoq ichida berilishi buni ko'rsatadi). Bu tasodifiy emas: aynan bevosita nutq vaziyati frazeologik birlikning kommunikativ-pragmatik quvvatini(zaryadini) kuchaytiradi va uni adresatga "jonli" ruhiy signal sifatida yetkazadi:

№	Frazema	Bet	Matndagi konteksti	Ma'nosi	Lingvoterapevtik funksiyasi
1	ko'ngli bo'shamoq	107	— Шошманг! Эшонбувамга бориб дод дейман! Ори рост, бир бўлган ишни энди буздириш мумкин эмаску-я, ишқилиб сиздан дод деб бир кўнглимни бўшатмасам бўлмайди...	Ruhiy kechinmalar ta'sirida yig'lab yuborish holatiga kelmoq [25]	**Katarsis funksiyasi** — frazema ichki og'riqni tashqariga chiqarish, "bo'shatish" harakatini tom ma'noda obrazlashtiradi; bu bibliyoterapiyadagi "emotsional bo'shalish" mexanizmining til vositasidagi to'g'ridan-to'g'ri aksidir.
2	Yurag (i)ni ezmoq	115	— Зебихоннинг аҳволи қалай, айланай? Жиндак гапириб беринг,ҳеч нарса билолмай, юрагимиз эзилиб кетди. Кўзимнинг қорачўғидай яхши кўрардим.	Ruhan azoblanmoq	**Hissiy identifikatsiya funksiyasi**— somatik obraz (yurak "ezilishi") adresatda boshqa qahramonga hamdardlik va empatiyani jismoniy sezgi darajasida uyg'otadi
3	ko'nglini to'q tutmoq	162	« — «Кўнглингни тўқ тут, азизим: Мирёқуб билан Акбарали мингбоши ўртасида ер билан осмон қадар фарқ борлигига биз қаттиқишонамиз! Акбарали бошқа одам. Мирёқуб бутунлай бошқача одам...»	Ishongani holda xotirjam bo'lmoq (o'z-o'zicha)	**Tasalli-tinchlantirish funksiyasi** — frazema muloqot vositasi sifatida qahramonning o'ziga (va adresatga) ishonch va xotirjamlik bag'ishlaydi, tashvishni pasaytiradi
4	taqdirga tan	108	«Фақат кампирнинг бу сўзларида энди	peshonasiga	**Moslashuv (adaptatsiya) funksiyasi** — og'ir voqelikni qabul

[25] Sh. Rahmatullayev. O'zbek tilining izohli frazeologik lug'ati. – Toshkent: O'qituvchi, 1978. 155-bet.

	bermoq		тақдирга тан бериш унлари бор эди»	rozi bo‘lmoq	qilish orqali ichki qarshilikni yumshatadi; psixologik "acceptance" mexanizmining badiiy ifodasi
5	og‘ir tog‘ ag‘anamoq	58	«...қишлоққа қайтганида, мингбошининг устидан ҳар қандай оғир тоғ ҳам ағанаган бўларди»	og‘ir tashvishdan qutulmoq	**Yengillashuv funksiyasi** — og‘ir yuk obrazi orqali tashvishning cho‘kishi/ko‘tarilishi jismoniy miqyosda tasvirlanadi, bu adresatda ham "yengil tortish" hissini uyg‘otadi
6	ko‘ngli g‘ashidan yorilmoq	104	« Онанинг кўнгли сир беркитиш орқасида ғашидан ёрилай деганда, қизнинг кўнгли дам Ўлмасжонни, дам Салтанатхонни эркалатиб ўйнарди.	Qattiq tashvishda bo‘lmoq	**Ichki taranglikni ifodalash funksiyasi** — bostirilgan hissiyot keskin obraz ("yorilish") orqali beriladi, bu adresatda taniqli — bosilgan tuyg‘u — holatini aks ettiradi
7	imoni uchmoq	214	«— Нимага имонинг ўчди бирданига?» деб қичқирди мингбоши ҳайбатли овоз билан	qattiq qo‘rqmoq	**Kuchli affekt ifodasi** — qo‘rquv holatini diniy-ma'naviy obraz orqali kuchaytirib beradi, adresatda zudlik bilan tushuniladigan keskin hissiy signal yaratadi
8	yurakdan urmoq	34	Бу оҳанг жуда очиқ эшитилган бўлса керакки, Холмат дарҳол фаҳмига борди ва «а, ҳароми-ей!» дегандай қилиб унга қарагач: «— Юракдан урган экан-а!..» — деди.	sevib qolmoq	**Ijobiy hissiy zichlashtirish funksiyasi** —frazeologik obraz ijobiy kechinmani (sevgi) ruhiy kuch bilan ifodagan

Shuningdek, asarda keltirilgan frazeologik birliklarning lingvoterapevtik funksiyasiga ko‘ra tahlil qilib quyidagi tipologiyani shakllantirdik:

1. **Katarsis - hissiy yengillik** — ko‘nglini bo‘shatmoq, alangasi osmonga chiqmoq, ko‘ngli g‘ashidan yorilmoq;
2. **Moslashuv (acceptance)** — taqdirga tan bermoq;
3. **Ichki tiyilish - o‘z-o‘zini nazora**t — damini ichiga solmoq, dami chiqmaslik, quruq «surat» shakliga kirmoq, itlikni o‘ldirmoq;
4. **Hissiy identifikatsiya - empatiya** — yuragi ezilmoq, ko‘ngliga og‘ir olmoq;
5. **Tasalli-tinchlantirish** — ko‘nglini to‘q tutmoq;
6. **Kuchli affekt ifodasi (qo‘rquv, g‘azab, aql yo‘qotish**) — imoni uchmoq, to‘nini teskari kiymoq, esini yemoq, sil qilib yubormoq;
7. **Ijobiy hissiy zichlashtirish (sevgi, quvonch)** — yurakdan urmoq, jirtak chalmoq;
8. **Himoya-distansiya** — parvoyi falak, tirnog‘i ham tegmaslik.

Ushbu tipologiya shuni ko‘rsatadiki, Cho‘lpon romanidagi frazeologik birliklar faqat og‘riqli kechinmalarni emas, balki ulardan chiqish yo‘llarini — bostirish, qabul qilish, bo‘shatish, tasalli topish kabi psixologik strategiyalarni ham til vositasida obrazlashtiradi. Aynan shu xususiyat ularni lingvoterapevtik tahlil uchun boy material ekanligini ko‘rsatadi.

Badiiy diskursda frazeologik birliklar shunchaki stilistik bezak emas, balki muallif–qahramon–adresat kommunikativ zanjirida muhim funksiyaga ega hamda adresat ruhiyatiga bevosita ta'sir ko‘rsatuvchi kuchli vositadir. Shuningdek, uning yuqorida keltirilgan funksiyalari badiiy diskursda takrorlanib, yaxlit "hissiy-somatik kod" hosil qiladi, natijada adresat ongida barqaror kognitiv freym va ijobiy perlokutiv samarani yuzaga keltiradi.

FOYDALANILGAN ADABIYOTLAR

1. Rahimov U.E. Diskursda manipulyatsiya va lingvoterapevtik ta'sirning psixolingvistik tadqiqi //— Andijon davlat universiteti.
2. Rahimov U.E. Lingvoterapiya va uning integratsion yo‘nalishlari tadqiqi //— Andijon davlat universiteti.
3. Rahimov U. O‘zbek tilida pragmatika presuppozitsiya hamda matnning konseptologik tadqiqi. —Andijon, 2026.
4. Sh. Rahmatullayev. O‘zbek tilining izohli frazeologik lug‘ati. – Toshkent: O‘qituvchi, 1978.
5. M.U. Saidova , N.A.Nadirxanova . Diskurs Tushunchasi Tahlili Va Tadqiqi // Miasto Przyszłości Kielce 2023 Impact // Factor: 9.2 ISSN-L: 2544-980X. 136-138 bet.

O'ZBEK VA TURK RAQAMLİ NUTQİDA YANGİ LEKSİK BİRLİKLAR: QİYOSİY TAHLİL

Safiye Kırlangıç
PhD Candidate, Department of Modern Turkic Dialects and Literatures, Institute of Social Sciences, Ankara University Visiting Researcher, Department of Computer Linguistics and Applied Linguistics, National University of Uzbekistan Named after Mirzo Ulugbek Email: sfykirlangic@gmail.com Phone: +998509955423
https://doi.org/10.5281/zenodo.23253707

Annotatsiya. *Raqamlashuv va ijtimoiy tarmoq platformalarining keng tarqalishi turkiy yozuv tillarining lug'at tarkibini jadal ravishda o'zgartirmoqda. Ushbu tadqiqot o'zbek va turk ijtimoiy tarmoq nutqida (Instagram, Telegram, TikTok va YouTube) paydo bo'layotgan ingliz tilidan olingan raqamli neologizmlarning shakllanish jarayonlari hamda lisoniy moslashuv mexanizmlarini qiyosiy jihatdan o'rganadi. Tadqiqotning ma'lumotlar bazasini 2025–2026-yillar oralig'ida to'plangan 150 ta post va taxminan 1 000 ta foydalanuvchi sharhlari asosida aniqlangan, takroriy qo'llanishga ega 60 ta raqamli neologizm tashkil etadi. Aniqlangan birliklar leksik-semantik, shakliy (morfologik) va morfofonologik jihatdan tahlil qilindi. Natijalar har ikkala tilning agglutinativ tuzilishi neologizmlarni morfologik negiz (stem) sifatida qabul qilishda moslashuvchan ekanligini ko'rsatadi. Biroq fe'llashtirish strategiyalarida o'zbek tili analitik/yordamchi fe'lli qurilmalardan (post qilmoq, strim qilmoq) foydalansa, turk tilida qo'shimchalar orqali yasalish mexanizmlari (postlamak, likelamak) ko'proq ko'zga tashlanadi. Yozuv darajasida esa o'zbek tilida talaffuzga yaqin shakllar (strim, kringe), turk tilida esa manba tildagi imloning saqlanib, turkcha qo'shimchalar bilan qo'llanishi diqqatni tortadi. Ushbu gibrid tuzilmalar tabiiy tilni qayta ishlash tizimlarida (NLP) o'zaklashtirish (stemming) va so'z turkumlarini belgilash (POS tagging) jarayonlarida morfologik noaniqliklarga olib kelishi mumkin. Xulosa qilib aytganda, raqamli lug'at tarkibi global internet jargoni bilan birga tillarning o'z morfologik imkoniyatlari negizida ham shakllanayotgani ko'rinadi.*

Kalit so'zlar: *Raqamli neologizm, ijtimoiy tarmoq nutqi, o'zbek tili, turk tili, lug'at tarkibi, morfologik moslashuv, lisoniy qiyoslash.*

1. KIRISH

XXI asrda raqamli kommunikatsiya texnologiyalari va ijtimoiy tarmoq platformalarining keng tarqalishi nafaqat muloqot va o'zaro ta'sir shakllarini, balki tillarning lug'at tarkibi va qo'llanish sohalarini ham o'zgartirdi. Instagram, Telegram, TikTok va YouTube kabi platformalar yangi tushunchalar, muloqot shakllari va raqamli xatti-harakatlar yuzaga keladigan dinamik kommunikatsiya muhitlariga aylandi. Ushbu muhitlarning tezkor va global xususiyati yangi leksik birliklarning qisqa vaqt ichida paydo bo'lishi, tarqalishi va turli tillarga o'tishiga zamin yaratmoqda (Shakirov, 2022: 173; Sobirova, 2022: 1442). Raqamli kommunikatsiya muhitlarida yuzaga keladigan yangi leksik birliklar, umuman olganda, raqamli neologizmlar doirasida baholanishi mumkin. Ular yangi so'zlar bilan bir qatorda, xorijiy tillardan o'zlashtirilgan birliklarni, mavjud so'zlarning yangi ma'nolar kasb etishini hamda xorijiy manbali birliklarning maqsad tilining morfologik va fonologik xususiyatlariga moslashtirilishini ham qamrab oladi. Ijtimoiy tarmoq nutqida, ayniqsa, ingliz tilining ta'siri yaqqol namoyon bo'ladi; texnologiya,

raqamli kontent yaratish, onlayn muloqot, marketing va ijtimoiy media madaniyatiga oid ko‘plab tushunchalar ingliz tilidan olingan birliklar orqali turli tillarga o‘tmoqda.

Turk va o‘zbek tillari strukturaviy jihatdan agglutinativ xususiyatga ega bo‘lgan hamda xorijiy manbali leksik birliklarni o‘z grammatik tizimlariga kiritish imkoniyatiga ega bo‘lgan ikki turkiy tildir. Biroq ikki tilning raqamli kommunikatsiya muhitida uchraydigan xorijiy manbali birliklarni o‘zlashtirish va moslashtirish shakllari bir xil bo‘lishi kutilmaydi. Xususan, xorijiy manbali negizlarga so‘z yasovchi yoki shakl yasovchi qo‘shimchalarning qo‘shilishi, yordamchi fe’llar bilan qo‘llanishi, yangi fe’llarning hosil qilinishi hamda o‘zlashgan birliklarning fonetik va yozuv tizimiga moslashtirilishi jihatidan turli tendensiyalar yuzaga kelishi mumkin (Ismailov va boshq., 2016: 7; Jurafsky & Martin, 2007: 40–41). Bu holat ijtimoiy tarmoq lug‘at tarkibini nafaqat zamonaviy so‘z qo‘llanilishi, balki yangi leksik birliklarning maqsad tilining strukturaviy tizimiga qanday kiritilishi nuqtayi nazaridan ham muhim tadqiqot sohasi sifatida namoyon etadi. Raqamli neologizmlarga bag‘ishlangan tadqiqotlar xorijiy manbali birliklarning turli tillarga o‘tishi, ma’no o‘zgarishlari va leksik moslashuv jarayonlariga e’tibor qaratadi. Biroq o‘zbek va turk tillarining bir xil raqamli kommunikatsiya muhitida duch keladigan yangi leksik birliklarni qanday o‘xshash va farqli strategiyalar asosida shakllantirishi masalasini qiyosiy jihatdan o‘rganish ushbu ikki tilning zamonaviy lug‘at tarkibi va lisoniy moslashuv mexanizmlarini aniqlash nuqtayi nazaridan muhim ahamiyatga ega (Primova, 2025: 207; Sobirova, 2022: 1448). Mazkur maqolaning maqsadi o‘zbek va turk ijtimoiy tarmoq nutqida qo‘llaniladigan raqamli neologizmlarni aniqlash hamda ushbu birliklarni leksik, semantik va morfologik xususiyatlari jihatidan qiyosiy o‘rganishdan iborat. Shu maqsadda raqamli neologizmlarning qo‘llanish sohalari va semantik kategoriyalari aniqlanadi; xorijiy manbali birliklarning ikki tilda qo‘shimcha qabul qilishi, fe’llashishi hamda fonetik va yozuv jihatidan moslashish shakllari qiyoslanadi.

2. METOD VA MA’LUMOTLAR BAZASI

Mazkur tadqiqotda tavsifiy va qiyosiy lingvistik tahlil usullari sifat tahlili yondashuvi bilan birgalikda qo‘llanildi. Tadqiqotning ma’lumotlar bazasini 2025–2026-yillar davomida Instagram, Telegram, TikTok va YouTube kabi ijtimoiy tarmoq platformalarida o‘zbek va turk tillarida e’lon qilingan kontentlarda aniqlangan raqamli neologizmlar tashkil etadi. Ma’lumotlarni to‘plash jarayonida ommaga ochiq kontent yaratuvchi akkauntlar, kanallar va materiallar ko‘rib chiqildi; jami 150 ta ijtimoiy tarmoq posti hamda taxminan 1 000 ta izoh va boshqa foydalanuvchi o‘zaro ta’sirlari o‘rganildi. Ushbu kuzatuv natijasida takroriy qo‘llanishga va yaqqol foydalanish tendensiyasiga ega bo‘lgan 60 ta raqamli neologizm tahlilga kiritildi. Ma’lumotlarni tanlashda raqamli kommunikatsiya va ijtimoiy tarmoqdan foydalanishga xos bo‘lgan yangi yoki keng tarqalish jarayonidagi leksik birliklarga e’tibor qaratildi. Xususan, ingliz tili, avvalo, xorijiy manbali birliklar bilan bir qatorda, mavjud leksik birliklarning raqamli kommunikatsiya muhitida kasb etgan yangi ma’no va qo‘llanish shakllari ham baholandi. Birlikning faqat bitta kontentda uchrashi yetarli mezon sifatida qabul qilinmadi; turli kontentlarda yoki foydalanuvchilar tomonidan takroran qo‘llangan birliklarga ustuvorlik berildi. Shu tariqa vaqtinchalik va individual qo‘llanishlar bilan ijtimoiy tarmoq nutqida keng tarqalish tendensiyasini namoyon etayotgan leksik birliklar o‘rtasida farq qilindi. Aniqlangan raqamli neologizmlar uch asosiy jihatdan o‘rganildi:

- **Leksik-semantik kategorizatsiya:** Birliklarning qo‘llanish sohalari va ma’no xususiyatlari aniqlanib, tematik kategoriyalar shakllantirildi.

- **Shakliy (morfologik) moslashuv:** Xorijiy manbali birliklarning o‘zbek va turk tillarining morfologik tizimlari bilan integratsiyalashuvi; ko‘plik, egalik va kelishik qo‘shimchalari bilan qo‘llanishi, so‘z yasovchi qo‘shimchalar orqali yangi shakllarning hosil qilinishi hamda fe’llashtirish mexanizmlari o‘rganildi.
- **Morfofonologik va yozuv jihatidan moslashuv:** Birliklarning maqsad tilining fonetik, imlo va alifbo xususiyatlariga moslashtirilish shakllari qiyoslandi.

Qiyosiy tahlilda har ikki tilda bir xil yoki o‘xshash raqamli tushunchalarni ifodalovchi birliklar bilan bir qatorda, faqat bir tilda uchraydigan o‘ziga xos qo‘llanishlar ham birgalikda baholandi. Tahlil natijasida ma’lumotlar; (i) kelib chiqishi, (ii) semantik kategoriyasi, (iii) so‘z turkumi, (iv) morfologik moslashuvi, (v) fe’llashishi, (vi) fonetik va yozuv jihatidan moslashuvi hamda (vii) turk yoki o‘zbek tilining ichki manbalari asosida hosil qilingan muqobillariga ko‘ra tasniflandi. Shu tariqa ma’lumotlar bazasi raqamli neologizmlarning shakliy va semantik xususiyatlarini qiyoslash imkonini beruvchi tizimlashtirilgan ma’lumotlar matritsasiga aylantirildi. Mazkur tasnifning kelgusida tabiiy tilni qayta ishlash (NLP), hissiyot tahlili va so‘z turkumlarini belgilash (POS tagging) kabi kompyuter lingvistikasi sohalarida qo‘llanishi mumkin bo‘lgan resurslarni ishlab chiqishga hissa qo‘shishi kutiladi (Abduraxmonova va boshq., 2022: 13; Primova, 2025: 208).

3. NATIJALAR VA MUHOKAMA

Tadqiqot doirasida to‘plangan 60 ta raqamli neologizm leksik-semantik kategoriyalar, morfologik integratsiya, fe’llashtirish mexanizmlari hamda morfofonologik va yozuv jihatidan moslashuv nuqtayi nazaridan o‘zbek va turk tillari ma’lumotlari asosida qiyosiy tahlil qilindi. Natijalar raqamli neologizmlar har ikki tilda o‘xshash kommunikativ ehtiyojlar asosida yuzaga kelishini, biroq maqsad tillarning lug‘at tarkibi va morfologik tizimlariga kiritilishida turli strategiyalardan foydalanilishi mumkinligini ko‘rsatadi.

3.1. Leksik-semantik kategorizatsiya

Ijtimoiy tarmoq nutqida aniqlangan raqamli neologizmlar semantik sohalari va qo‘llanish kontekstlariga ko‘ra uch asosiy tematik guruhga tasniflandi. Birliklarning ikki tildagi qo‘llanishi quyidagi jadvalda qiyosiy ravishda keltirilgan.

1-jadval. O‘zbek va turk tillarida raqamli neologizmlarning semantik kategoriyalari

Semantik kategoriya	**O‘zbek tilidagi ijtimoiy tarmoq qo‘llanilishi**	**Turk tilidagi ijtimoiy tarmoq qo‘llanilishi**	**Konseptual ma’no / funksiya**
Raqamli kontent va media	kontent, strim, kover, post, storis, tresh, rils	içerik/kontent, stream/yayın, cover, post, story, trash, reels	Platforma ichidagi materiallar, efir shakllari va raqamli kontent turlari
Ijtimoiy o‘zaro ta’sir va xatti-harakat	hayp, kringe, hater, follower, cancelar, heyt	hype, cringe, hater, takipçi/follower, linç/cancel, hate	Foydalanuvchilar o‘rtasidagi reaksiyalar, raqamli munosabatlar va ijtimoiy baholashlar
Raqamli iqtisodiyot va mehnat madaniyati	freelance, deadline, support, target, donat	freelance, deadline, support, target, donate/bağış	Masofaviy mehnat, raqamli marketing, qo‘llab-quvvatlash va onlayn iqtisodiy

			faoliyat

Jadvaldagi birliklar ijtimoiy tarmoq lug‘at tarkibidagi yangilanish turli kommunikativ va qo‘llanish sohalariga tarqalganini ko‘rsatadi. Raqamli kontent va media kategoriyasida post, story/storis, stream/strim, cover/kover va content/kontent kabi birliklar kontent yaratish, ulashish va e’lon qilish kabi funksiyalar bilan ajralib turadi. Ijtimoiy o‘zaro ta’sir va xatti-harakat kategoriyasidagi hype/hayp, cringe/kringe, hater va follower kabi birliklar esa onlayn reaksiyalar, baholashlar va ijtimoiy xatti-harakatlarni ifodalaydi. Raqamli iqtisodiyot va mehnat madaniyati kategoriyasidagi freelance, deadline, support va target kabi birliklar esa raqamli mehnat, marketing va onlayn iqtisodiy faoliyat bilan bog‘liq. Bir xil raqamli tushunchaning ikki tilda turli shakllarda ifodalanishi ham kuzatiladi. Turk tilida içerik, takipçi va yayın kabi o‘zlashib ketgan muqobillar bilan bir qatorda xorijiy manbali shakllar ham qo‘llanilsa, o‘zbek tilidagi ijtimoiy tarmoq nutqida xorijiy manbali shakllarning bevosita yoki maqsad tilining fonetik va yozuv xususiyatlariga moslashtirilgan ko‘rinishlari ko‘proq namoyon bo‘lishi mumkin. Bu holat raqamli lug‘at tarkibidagi yangilanishni faqat o‘zlashma so‘zlarning mavjudligi bilan emas, balki ularning maqsad tilidagi qo‘llanish va moslashuv shakllari bilan birgalikda baholash zarurligini ko‘rsatadi (Primova, 2025: 214–215).

3.2. Morfologik moslashuv va shakl yasovchi qo‘shimchalarning integratsiyasi

Xorijiy manbali raqamli birliklarning maqsad tillarning agglutinativ tuzilishiga kiritilishi jarayonida ham o‘zbek, ham turk tillari o‘z morfologik tizimlariga xos shakl yasovchi qo‘shimchalarni xorijiy manbali negizlarga qo‘shishi mumkin. Ayniqsa, ko‘plik, egalik va kelishik qo‘shimchalarining qo‘llanishi ushbu birliklarning maqsad tilida grammatik kontekst doirasida funksional xususiyat kasb etganini ko‘rsatadi.

Egalik va ko‘plik qo‘shimchalarining qo‘llanishi:

- **O‘zbek tili:** kontent-im “icheriyim”, storis-i “hikoyasi”, follower-lar-imiz “izdoshlarmiz”, post-lar-da “postlarda”.
- **Turk tili:** kontent-im “icheriyim”, story-si “hikoyasi”, follower-lar-ımız “izdoshlarmiz”, post-lar-da “postlarda”.

Ushbu misollar xorijiy manbali raqamli birliklarning har ikki tilda ham morfologik negiz sifatida faoliyat ko‘rsatishi va ularga maqsad tilining shakl yasovchi qo‘shimchalari qo‘shilishi mumkinligini ko‘rsatadi. Bu jarayon mazkur birliklarning maqsad tilining sintaktik va grammatik munosabatlari doirasida qo‘llanish imkonini beradi (Ismailov va boshq., 2016: 8; Jurafsky & Martin, 2007: 77).

3.3. Fe’llashtirish mexanizmlaridagi strukturaviy farqlar

Ingliz tilidan kelib chiqqan raqamli birliklarning harakatni ifodalovchi shakllarga aylantirilishida o‘zbek va turk tillari o‘rtasida turli tendensiyalar kuzatiladi. Umuman olganda, xorijiy manbali ot negizlarining o‘zbek tilida yordamchi fe’llar, turk tilida esa so‘z yasovchi qo‘shimchalar orqali fe’llashtirilishi kuzatiladi.

Ushbu tuzilmani quyidagicha ko‘rsatish mumkin:

- **[Xorijiy manbali negiz] ⟶ O‘zbek tili:** [Negiz] + [Yordamchi fe’l]
- **[Xorijiy manbali negiz] ⟶ Turk tili:** [Negiz] + [So‘z yasovchi qo‘shimcha]

O‘zbek tilida xorijiy manbali ot negizli raqamli birliklarning qilmoq “yapmoq” va bo‘lmoq “bo‘lmoq” kabi yordamchi fe’llar bilan birikib, harakatni ifodalovchi shakl hosil qilishi diqqatga sazovordir. Post qilmoq “post ulashmoq”, strim qilmoq “jonli efir qilish”, kringe bo‘lmoq va hayp ko‘tarmoq kabi qo‘llanishlar xorijiy manbali birliklarning o‘zbek tilidagi mavjud fe’l hosil qilish imkoniyatlari doirasida ishlatilishiga misol bo‘la oladi. Turk tilida esa -

la/-le so‘z yasovchi qo‘shimchasi orqali fe’l yasash muhim hosil qilish usulidir. Postlamak, favlamak, likelamak va stalklamak kabi misollarda xorijiy manbali negiz turk tilining fe’l yasash mexanizmiga kiritiladi (Qiyos. Ismailov va boshq., 2016: 7). Shu bilan birga, cringe olmak kabi yordamchi fe’lli tuzilmalar ham qo‘llaniladi. Demak, har ikki tilda turli fe’llashtirish strategiyalari mavjud bo‘lsa-da, o‘zbek tilida yordamchi fe’lli tuzilmalar, turk tilida esa qo‘shimchalar orqali yasalish ko‘proq namoyon bo‘ladi, deyish mumkin.

3.4. Morfofonologik va yozuv jihatidan (ortografik) moslashuv

Raqamli neologizmlarning maqsad tilining fonetik, imlo va alifbo xususiyatlariga moslashtirilishi jihatidan ham ikki til o‘rtasida turli tendensiyalar kuzatiladi. O‘zbek ijtimoiy tarmoq nutqida ayrim xorijiy manbali birliklarning talaffuzga yaqin shakllarda yozilishi diqqatga sazovordir. Strim (stream), kringe (cringe) va hayp (hype) kabi shakllar xorijiy manbali birliklarning o‘zbek lotin alifbosi va mahalliy qo‘llanish odatlarining ta’siri ostida qayta shakllantirilishi mumkinligini ko‘rsatadi. Turk tilida esa ayrim raqamli birliklarning manba tildagi yozuv shakliga yaqin holda saqlanib, turkcha qo‘shimchalar bilan qo‘llanishi diqqatga sazovordir. Biroq ijtimoiy tarmoq muhitida yozuv shakllari standart til qoidalaridan chetga chiqishi mumkinligini hisobga olish lozim. Shu sababli bunday qo‘llanishlarni standart turk tili imlo qoidasi sifatida emas, balki ijtimoiy tarmoq nutqida kuzatiladigan foydalanish tendensiyalari sifatida baholash kerak. Morfofonologik va yozuv jihatidan moslashuv shu tariqa raqamli neologizmlarning maqsad tilining fonetik va yozuv tizimlari asosida qayta shakllantirilish jarayonini namoyon etadi (Sobirova, 2022: 1448).

3.5. Tabiiy tilni qayta ishlash (NLP) va kompyuter lingvistikasi nuqtayi nazaridan baholash

Tadqiqotda aniqlangan raqamli neologizmlar ijtimoiy tarmoq ma’lumotlarining tabiiy tilni qayta ishlash nuqtayi nazaridan duch keladigan ayrim muammolarini ham namoyon etadi. Xususan, xorijiy manbali negizlarning turk yoki o‘zbek tilining shakl yasovchi va so‘z yasovchi qo‘shimchalari bilan birikishi avtomatik til tahlili tizimlarida morfologik noaniqlikni yuzaga keltirishi mumkin. Tabiiy tilni qayta ishlash jarayonlarida tokenizatsiya, lemmatizatsiya va o‘zaklashtirish (stemming) kabi usullardan foydalaniladi (Sobirova, 2022: 1447). Ijtimoiy tarmoq ma’lumotlarida uchraydigan gibrid tuzilmalar, ayniqsa, o‘zaklashtirish va morfologik tahlil jarayonlarida qiyinchilik tug‘dirishi mumkin. Turk tilidagi like'lamak, o‘zbek tilidagi strimda kabi shakllarda xorijiy manbali negiz bilan mahalliy qo‘shimchaning birikishi avtomatik analizatorlarning mazkur birliklarni to‘g‘ri tahlil qilishini qiyinlashtiradi. Xuddi shuningdek, cringe/kringe kabi birliklar kontekstga qarab turli so‘z turkumi yoki funksiyani ifodalashi mumkinligi sababli, so‘z turkumlarini belgilashda kontekstga asoslangan yondashuvlarga ehtiyoj tug‘iladi (Jurafsky & Martin, 2007: 40–41). Hype/hayp, cringe/kringe va hate/heyt kabi birliklar foydalanuvchilarning ijobiy yoki salbiy baholarini ifodalashi mumkinligi sababli, hissiyot tahlili tizimlarida hisobga olinishi lozim bo‘lgan leksik unsurlar hisoblanadi. Mazkur natijalar raqamli neologizmlarni nafaqat lingvistik, balki kompyuter lingvistikasi va tabiiy tilni qayta ishlash nuqtayi nazaridan ham hisobga olish zarurligini ko‘rsatadi. Ayniqsa, xorijiy manbali negizlarning maqsad tilining qo‘shimchalari bilan birikishidan hosil bo‘lgan gibrid shakllar turk va o‘zbek tillari uchun ijtimoiy tarmoqqa yo‘naltirilgan morfologik tahlil va lug‘at resurslarini ishlab chiqishda e’tiborga olinishi lozim (Abduraxmonova va boshq., 2022: 14; Primova, 2025: 214–215).

4. XULOSA

Mazkur tadqiqotda o‘zbek va turk ijtimoiy tarmoq nutqida (Instagram, Telegram, TikTok va YouTube) keng tarqalgan raqamli neologizmlar leksik-semantik, morfologik, morfofonologik va kompyuter lingvistikasi nuqtayi nazaridan qiyosiy o‘rganildi. Natijalar raqamli neologizmlar har ikki tilda Raqamli kontent va media, Ijtimoiy o‘zaro ta’sir va xatti-harakat hamda Raqamli iqtisodiyot va mehnat madaniyati sohalarida ko‘proq uchrashini ko‘rsatdi. O‘zbek va turk tillarining agglutinativ tuzilishi xorijiy manbali birliklarning morfologik negiz sifatida qo‘llanishi va ularga egalik, ko‘plik hamda kelishik qo‘shimchalarining qo‘shilishiga imkon beradi. Fe’llashtirishda esa o‘zbek tilida analitik va qo‘shma fe’l qoliplari (post qilmoq, strim qilmoq), turk tilida esa qo‘shimchalar orqali yasalish (postlamak, likelamak) ko‘proq namoyon bo‘ladi. Yozuv jihatidan o‘zbek tilida talaffuzga yaqin shakllar (strim, kringe), turk tilida esa manba tildagi yozuvga yaqin shakllarning turkcha qo‘shimchalar bilan qo‘llanishi kuzatiladi. Xorijiy manbali negiz va mahalliy qo‘shimchalardan tashkil topgan gibrid birliklar tabiiy tilni qayta ishlashda o‘zaklashtirish, so‘z turkumlarini belgilash va hissiyot tahlili jarayonlari uchun maxsus yondashuvlar hamda kengaytirilgan lug‘at resurslariga ehtiyoj tug‘diradi. Xulosa qilib aytganda, raqamli neologizmlar o‘zbek va turk tillarining lug‘at tarkibini dinamik ravishda boyitmoqda. Ikki tilning agglutinativ tuzilishi va moslashuvchan mexanizmlari global kommunikatsiya unsurlarining mahalliy grammatik tizimlarga kiritilishini osonlashtiradi; bunda o‘zbek tilida yordamchi fe’lli, turk tilida esa qo‘shimchalar orqali fe’llashtirish strategiyalari ko‘proq namoyon bo‘ladi.

FOYDALANILGAN ADABIYOTLAR

1. Abduraxmonova, N., Tuliyev, U., & Abduvahobov, G‘. (2022). O‘zbek tili korpus lingvistikasining nazariy va amaliy masalalari. *Kompyuter Lingvistikasi: Muammo va Yechimlar* (s. 11–15). Taşkent: O‘zbekiston Milliy Universiteti.
2. Ismailov, A., Abdul Jalil, M. M., Abdullah, Z., & Abd Rahim, N. H. (2016). *A Comparative Study of Stemming Algorithms for use with the Uzbek Language*. IEEE 3rd International Conference on Computer and Information Sciences (ICCOINS) (s. 7–12). Kuala Terengganu: IEEE. https://doi.org/10.1109/iccoins.2016.7783180
3. Jurafsky, D., & Martin, J. H. (2007). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (2. Baskı). Draft, Pearson Prentice Hall.
4. Primova, M. H. (2025). O‘zbek tili milliy korpusi tuzilishi: Asosiy komponentlar va axborot turlari. *V Xalqaro ilmiy-amaliy konferensiya: Kompyuter Lingvistikasi: Muammolar, Yechim, Istiqbollar* (s. 207–221). Taşkent: Alisher Navoiy nomidagi ToshDO‘TAU.
5. Shakirov, R. S. (2022). Korpus lingvistikasi va parallel korpuslar xususida. *Tarjimashunoslik va Lingvodidaktika* (s. 173–176). Buhara: Buxoro Davlat Universiteti.
6. Sobirova, N. G‘. (2022). Korpus lingvistikasi va parallel korpuslar tavsifi. *Academic Research in Educational Sciences*, 3(5), 1442–1450. https://doi.org/10.24412/2181-1385-2022-5-1442-1450

ERNEST HEMENGUEY VA OYBEK PERSONAJLARIDA "SADOQAT" KONSEPTINI TARKIBLANISHI

Sattorova Nafisa Isomidinovna
KDPI o'qituvchisi, SamDChTI mustaqil izlanuvchisi
https://doi.org/10.5281/zenodo.23253740

Annotatsiya. *Ushbu maqolada Ernest Hemingway va Oybek nasriy merosida "sadoqat" (loyalty) konseptining kognitiv-semantik tuzilishi va badiiy matnda namoyon bo'lish xususiyatlari qiyosiy tahlil qilingan. Tadqiqot davomida Hemingway qahramonlaridagi individualistik, ekzistensial sadoqat hamda Oybek ijodidagi ijtimoiy-tarixiy, milliy "sadoqat" konsepti ichki monologlar va uslubiy vositalar orqali ochib berilgan.*

Kalit so'zlar: *konsept, sadoqat, Hemingway, Oybek, monolog, Aysberg nazariyasi, antiteza, axloqiy kodeks, kognitiv, ishonch.*

Abstract. *This article provides a comparative analysis of the cognitive-semantic structure and literary representation of the concept of "loyalty" as an ethnic and universal value in the prose heritage of Ernest Hemingway and Oybek. The study explores the individualistic, existential loyalty in Hemingway's characters alongside the socio-historical and national loyalty in Oybek's works, examined through textual markers, internal monologues, and stylistic devices.*

Keywords: *concept, loyalty, Hemingway, Oybek, monologue, Iceberg theory, ontological metaphor, antithesis, moral code.*

Аннотация. *В данной статье проводится сравнительный анализ когнитивно-семантической структуры и художественной репрезентации концепта «верность» (loyalty) как этнической и универсальной ценности в прозаическом наследии Эрнеста Хемингуэя и Ойбека. В ходе исследования на основе текстовых маркеров, внутренних монологов и стилистических приемов раскрываются индивидуалистическая, экзистенциальная верность героев Хемингуэя и социально-историческое, национальное понимание верности в творчестве Ойбека.*

Ключевые слова: *концепт, верность, Хемингуэй, Ойбек, монолог, теория айсберга, онтологическая метафора, антитеза, моральный кодекс.*

1.KIRISH(INTRODUCTION)

Avvalo, sadoqat shunchaki axloqiy kategoriya emas, balki muayyan lisoniy tizimda o'zining kognitiv xaritasiga ega bo'lgan murakkab konseptosfera ekanligini ta'kidlash lozim. Ma'lumki, Heminguey asarlarida sadoqat ko'proq individualistik xarakterga ega. U qahramonning shaxsiy prinsiplariga, burch tuyg'usiga va "qiyinchilikda matonat ko'rsatish" orqali namoyon bo'ladi. Oybek ijodi uchun sadoqat milliy va ijtimoiy-madaniy qadriyatlar, xalqning milliy ozodlik ruhi, oilaviy va g'oyaviy bog'liqlik bilan chambarchas bog'liq hisoblanadi. Badiiy muloqot tizimida makrokonseptlar tahlili har bir ijodkorning individual-mualliflik lisoniy manzarasini ochib beruvchi birlamchi omillardan hisoblanadi. Shu ma'noda, E.Heminguey va Oybek nasriy merosida etnik va universal qadriyat hisoblangan "sadoqat" konseptining qahramonlar nutqi va xarakteristikasidagi kognitiv tuzilishi tahlili o'ziga xos ilmiy ahamiyat kasb etadi. Har ikki adib asarlarida ushbu konseptual maydon semantik jihatdan turlicha mikromaydonga (shaxsiy sadoqat, g'oyaviy sadoqat, milliy burch va his-tuyg'ular barqarorligi)

bo‘linsa-da, ular funksional-stilistik jihatdan personajlarning ma'naviy evrilishlarini ko‘rsatishga xizmat qilgan. E.Heminguey va Oybek qahramonlari tizimida “sadoqat” konseptining tarkiblanishi faqatgina leksik-frazeologik birliklargina ifodalanmay, balki qahramonlarning psixologik-situativ hatti-harakatlari orqali ham talqin qilinadi.

2. TADQIQOT METODLARI (METHODS)

Heminguey qahramonlarida (masalan, "Alvido, qurol!", "Chol va dengiz" asarlarida) sadoqat shaxsiy burch konsepti sifatida namoyon bo‘lsa, Oybek ijodida ("Qutlug‘ qon", "Navoiy" romanlari) ushbu tushuncha ijtimoiy-tarixiy omillar va jamiyat hayoti bilandeterminantlar va mental etnosotsium bilan uzviy bog‘langan holda voqelanadi.

Ernest Hemingway asarlaridagi bosh qahramonlar monologlarida "sadoqat" konsepti nafaqat bevosita tushuncha sifatida, balki qahramonning ichki axloqiy kodeksi va hayotga nisbatan qat'iy munosabati sifatida konseptuallashtiriladi. Monologlar bu jarayonning eng nozik va yashirin qatlami hisoblanadi. Heminguey asarlarida monologlar odatda qisqa, aniq va chuqur ichki kechinmani ifodalovchi lakonik nutq xususiyatiga ega. "Sadoqat" (loyalty) konsepti bu monologlarda, ayniqsa, "Aysberg nazariyasi" ga muvofiq ifodalanadi.

Adabiyotshunos olim Karlos Beyker Ernest Hemenguey qahramonlarining sadoqatini "birodarlik" va "vazifaga sodiqlik" prizmasidan ko‘radi."Heminguey qahramonlari uchun sadoqat — bu o‘z-o‘zini anglashning eng yuqori nuqtasi. 'Alvido, qurol!' asaridagi Frederik Genri armiyadan ketishi mumkin, lekin u o‘zining ichki axloqiy prinsiplariga va sevgisiga oxirigacha sadoqatli qoladi."

3. TADQIQOT NATIJALARI (RESULTS)

Ernest Hemengueyning eng mashhur asarlaridan biri bo‘lgan "Chol va dengiz" (The Old Man and the Sea) asarida Santyagoning monologlari sadoqatning eng yorqin ko‘rinishlaridan biri – "kurashga sodiqlik" va "o‘zligiga sodiqlik"ni namoyish etadi. Santyago marskonini tutishga 84 kunlik omadsizlikdan keyin ham sadoqat bilan kirishadi. Uning monologlaridagi love, respect, hope kabi nutq vositalari “sadoqat” konseptini ifodalab keladi. Ushbu birliklar mag‘lubiyatni qabul qilmaslikka , kurashishga undaydi va o‘z kasbiga bo‘lgan sadoqatni ifodalaydi. Adibning “Chol va dengiz” asaridan olingan quyidagi parchada cholning o‘z kasbiga qanchalik sodiq ekanligini ko‘rishimiz mumkin.

“Fish,” he said, “I love you and respect you very much. But I will kill you dead before this day ends." . Let us hope so, he thought [“Chol va dengiz”, 19]

Bu yerda chol (Santyago) va baliq (Marlin) o‘rtasidagi munosabat shunchaki ovchi va o‘lja o‘rtasidagi to‘qnashuv emas, balki chuqur ma'naviy bog‘liqlikdir.

Monologning tuzilishi va unda qo‘llanilgan so‘zlar cholning ichki dunyosini ochib beradi. Yuqorida berilgan parchada antiteza (zidlash) qo‘llanilgan bo‘lib, "I love you and respect you" va "I will kill you" jumlalari bir-biriga qarama-qarshi qo‘yilgan. Bu lisoniy zidlik Santyagoning qalbida kechayotgan hayot uchun kurash va tabiatga bo‘lgan muhabbat o‘rtasidagi muvozanatni ko‘rsatadi. Shuni ham ta’kidlab o‘tish kerakki, Santyago baliqchi bo‘lib tug‘ilgan. Uning "o‘ldirish" haqidagi va'dasi yovuzlik emas, balki o‘z tabiatiga va vazifasiga bo‘lgan sadoqatidir. U baliqni o‘ldirish orqali o‘zining "men"ligini va hayotdagi maqsadini tasdiqlaydi.

"I will kill you " iborasidagi "will" yordamchi fe'li shunchaki kelasi zamonni emas, balki qat'iy iroda va muqarrar taqdirni anglatadi. Baliqqa "Fish" (Baliq) deb murojaat qilishi orqali uni muloqotga qodir tengdosh subyekt sifatida ko‘radi. "I respect you" jumlasi orqali Santyago baliqni dushman deb emas, "birodar" deb bilishini ko‘rishimiz mumkin. U baliqqa sadoqat bilan oxirigacha kurashishni va'da bermoqda. Bu yerda sadoqat raqibni past ko‘rmaslik va unga

munosib kurash taqdim etishdir. "Let us hope so" jumlasida cholning o‘z kuchiga bo‘lgan shubhasi emas, balki taqdirga tan berishi ko‘rinadi. U o‘z so‘zining ustidan chiqishga, ya'ni oxirigacha sadoqatli qolishga intiladi.

Hemengueyga zamondosh bo‘lgan o‘zbek adabiyotining zabardast vakillaridan biri hisoblanmish Oybek asarlarida sadoqat asosan uch xil konsept ko‘rinishga: Vatanga sadoqat, G‘oyaga sadoqat va Insoniy (sevgidagi) sadoqat ega sanaladi.

Oybekning “Navoiy” romanidagi Navoiy monologlarida sadoqat konsepti majburiyat leksik-semantik maydoni orqali ifodalanadi.

4.MUHOKAMA (DISCUSSION)

Dramaturg va olim Izzat Sulton Oybek asarlaridagi qahramonlarning ichki dunyosi haqida gapirganda, ularning og‘ir sinovlarda o‘z asliga sodiq qolishini ta’kidlaydi va Oybek qahramonlari (ayniqsa Navoiy va uning shogirdlari) hayotning eng murakkab chorrahalarida ham o‘z insoniy qadr-qimmatiga xiyonat qilmasliklarini, bu sadoqat — Oybekning estetik idealining asosi ekanligini aytib o‘tadi.

Oybekning "Navoiy" romanida bosh qahramon Alisher Navoiyning ichki monologlari va o‘y-xayollari uning davlatga, xalqqa, do‘stlikka (Husayn Boyqoroga) va ijodga bo‘lgan mutlaq sadoqatini ko‘rsatuvchi lingvistik maydon hisoblanadi.

Quyida Oybek qalamiga mansub “Navoiy” romanidan olingan parchada ham shuni ko‘rishimiz mumkin:

“…Aqlli kengash shulki, — javob berdi Navoiy, —Badiuzzamon Mirzoni aqlga chaqirasiz, insof va vijdonga da’vat qilasiz. Uning ko‘nglida mehru muhabbat uyg‘otishga harakat etasiz.— Navoiy bir nafas sukut qilib davom etdi.— O‘rtada bo‘lib o‘tgan ginalarni bartaraf etib, samimiy sadoqat, vafo, muhabbatni barpo qilmoq kerak… [Navoiy, 427]

Oybekning «Navoiy» romanidan olingan ushbu parcha sadoqat konseptining nafaqat axloqiy, balki ijtimoiy-siyosiy va kognitiv-semantik ko‘lamini ochib beruvchi muhim lisoniy materialdir. Navoiy nutqida sadoqat tushunchasi mavhum hissiyot darajasidan mantiqiy va zaruriy harakat darajasiga ko‘tariladi.
Sadoqat konsepti matnda yakkalanib kelmaydi, balki bir-birini taqozo etuvchi bir qancha leksemalar zanjiri orqali quriladi:
"Samimiy sadoqat", "vafo", "muhabbat": Bu yerda sadoqat leksemasi „vafo“ va „muhabbat“ so‘zlari bilan bir qatorda (gradatsiya usulida) qo‘llangan. Tilshunoslikda bu "konseptual maydon" deyiladi. Ya’ni, Navoiy diskursida sadoqat — muhabbat va vafosiz mavjud bo‘lmaydigan yaxlit bir butunlikdir. "Insof" va "Vijdon" Sadoqatning asosi sifatida bu ikki leksema keltirilgan. Bu — sadoqatning kognitiv poydevori vijdon ekanligini ko‘rsatuvchi lisoniy markerlardir. Matndagi eng muhim lisoniy birliklardan biri "barpo qilmoq" fe’lidir. Odatda "barpo qilmoq" fe’li moddiy obyektlarga (bino, shahar, imorat) nisbatan ishlatiladi. Sadoqat, vafo va muhabbatga nisbatan ushbu fe’lning qo‘llanilishi ontologik metaforani yuzaga keltiradi. Sadoqat tayyor narsa emas, u "buzilgan" (gina-kudurat tufayli) holatdan qayta "quriladigan", mehnat va iroda talab qiladigan lisoniy-ma’naviy konstruksiya sifatida modellashtirilgan. "Aqlga chaqirasiz", "da’vat qilasiz", "harakat etasiz".
Bu yerdagi buyruq-istak ma’nosidagi fe’llar sadoqatning passiv holat emas, balki faol kommunikativ harakat ekanligini bildiradi. Sadoqatga erishish uchun avvalo "aql" va "insof" kabi kognitiv bosqichlardan o‘tish kerakligi lisoniy uqtirilmoqda.
Samimiy sifati bu yerda sadoqat konseptining "soxtalik" yoki "majburiyat" semalaridan xoli ekanligini ko‘rsatadi. Bu sadoqatning intensivlik darajasini oshiruvchi lisoniy vositadir. Navoiy

nutqidagi ushbu birliklar sadoqatning ijtimoiy barqarorlik va insoniy komillik ramzi ekanligini ilmiy-lisoniy jihatdan asoslab beradi. Bu monologlarda Navoiy o‘z sadoqatini deklarativ va imperativ gaplar orqali tasdiqlaydi. "Burchdir", "ahdimdir" kabi tasdiqlovchi sintaktik konstruksiyalar sadoqatning qat’iy xarakterga egaligini ko‘rsatadi.

Adabiyotshunos Ozod Sharafiddinov Oybekning "Navoiy" romanidagi Alisher Navoiy obrazini tahlil qilar ekan, sadoqatni do‘stlik va el-yurt manfaati kontekstida ko‘radi: "Oybek Navoiyni chizganda, uning sulton Husayn Boyqaro bilan munosabatlarida faqat do‘stlik sadoqatini emas, balki adolatga va xalq farovonligiga bo‘lgan sadoqatini birinchi o‘ringa qo‘yadi. Navoiy uchun sadoqat — bu o‘z e’tiqodidan voz kechmaslikdir."

Akademik Matyoqub Qo‘shjonov esa Oybek ijodining eng mohir tadqiqotchilaridan biri hisoblanadi. U Yo‘lchi va Gulnor ("Qutlug‘ qon") obrazlari orqali sadoqatni quyidagicha ta’riflaydi:"Yo‘lchi va Gulnor sevgisi shunchaki ehtiros emas, bu — ijtimoiy zulmga qarshi turuvchi ma’naviy qalqondir. Ularning bir-biriga bo‘lgan sadoqati insoniy erkinlik va poklikning oliy ko‘rinishidir."

Oybek o‘z qahramonlari monologlarida sadoqat tushunchasini tilshunoslik jihatdan turli uslubiy vositalar orqali konseptuallashtiradi: Ya’ni Oybek asarlarida “sadoqat” konseptini ifodalashga hizmat qiladigan "burch", "ahd", "fido", "vafo" kabi o‘zbekcha-forscha so‘zlarni faol ishlatadi. Shuningdek, Qat’iy tasdiqlovchi ("...dir") yoki buyruq-undov gaplar boshqa so‘z bilan aytganda, sintaktik konstruksiyalar orqali g‘oyaviy qat’iylikni ifodalaydi. Bu usullar Oybekka qahramonning sadoqatini shunchaki xislat emas, balki ichki jang maydoni va hayotiy falsafa sifatida ko‘rsatishga imkon beradi.

Hemingwayning bosh qahramonlari monologlarida "sadoqat" konsepti to‘g‘ridan-to‘g‘ri nomlanishi kamdan-kam uchraydi. Aksincha, u qahramonning ichki tanlovlari, qiyin vaziyatdagi harakatlari va matonatli yondashuvlari orqali, asosan ichki monologlar va lakonik nutq shaklida ifodalanadi. Bu uning minimalistik uslubining sadoqat konseptini konseptuallashtirishdagi asosiy funksiyasidir.

5.CONCLUSION

Ikkala yozuvchi Ernest Hemenguey va Oybek ham sadoqatni insonning hayotdagi eng muhim va deyarli yagona qiymati sifatida ko‘rsatadi, ammo bu ularning o‘ziga xos sharoitlarda namoyon bo‘ladi.

“Sadoqat” konsepti ikkala yozuvchi ijodida barcha qadriyatlarning markazi hisoblanadi. Oybek personajlarida Sadoqat vatan, xalq va g‘oya uchun yashashning yagona asosini tashkil etsa, Hemenguey asarlarida Sadoqatsizlik fojia degani. Sadoqat, ayniqsa Do‘stlik va Sevgidagi sadoqat, yozuvchining dunyosida mavjud bo‘lgan kam sonli ma'naviy qoidalardan biridir.

Oybek va Hemingway sadoqatni ifodalashda turli xil badiiy vositalardan foydalanganlar. Oybek asarlaridagi sadoqat "ijtimoiy inson" va "Tarixiy shaxs" ning majburiy burchi bo‘lsa, Hemingway asarlaridagi sadoqat "Yo‘qolgan avlod" vakilining yagona najot yo‘li va shaxsiy kodeksi hisoblanadi. Oybek sadoqatni Xalqqa va G‘oyaga xizmat qilish orqali ulug‘lasa, Hemingway sadoqatni Hukumat va jamiyatdan uzilib, yaqin insoniga sodiq qolish orqali ulug‘laydi.

FOYDALANILGAN ADABIYOTLAR

1. Hemingway, E. (1952). *The Old Man and the Sea*. Charles Scribner's Sons.
2. Hemingway, E. (1929). *A Farewell to Arms*. Charles Scribner's Sons.
3. Oybek. (1944). *Navoiy* (roman). Toshkent.

4. Oybek. (1940). *Qutlug‘ qon* (roman). Toshkent.
5. Baker, C. (1972). *Hemingway: The Writer as Artist*. Princeton University Press.
6. Sulton, I. (1980). *Adabiyot nazariyasi*. Toshkent: O‘qituvchi.
7. Sharafiddinov, O. (2004). *Ijodkorning baxti*. Toshkent: Sharq.
8. Qo‘shjonov, M. (1983). *Oybek mahorati*. Toshkent: G‘afur G‘ulom nomidagi Adabiyot va san’at nashriyoti.
9. Karimov, N. (1985). *Oybek: Hayoti va ijodi*. Toshkent: Yosh gvardiya.
10. Boldick, C. (2001). *The Oxford Dictionary of Literary Terms*. Oxford University Press.

KORXONALARDA YASHIL INVESTITSIYALARNI MOLIYALASHTIRISH SAMARADORLIGINI BAHOLASHNING EKONOMETRIK MODELI

Sharipov Qudratbek Quraltayevich
Ma'mun universiteti o`qituvchisi
sharipovqudrat@gmail.com
https://orcid.org/0009-0002-8062-2450
https://doi.org/10.5281/zenodo.23253752

Annotatsiya. *Ushbu maqolada korxonalarda yashil investitsiyalarni moliyalashtirish samaradorligini baholashning ekonometrik modeli ishlab chiqilgan. Tadqiqotda panel ma'lumotlar regressiyasi, sobit (Fixed Effects) va tasodifiy (Random Effects) effektlar modellari, shuningdek dinamik GMM (Arellano–Bond) yondashuvi asosida yashil investitsiya samaradorligiga ta'sir etuvchi omillar tahlil qilingan. Samaradorlik indeksini shakllantirishda ma'lumotlarni qamrab olish tahlili (DEA) usulidan foydalanilgan. Model natijalari korxonalarda moliyalashtirish manbalarini optimallashtirish va yashil iqtisodiyot siyosatini takomillashtirish uchun amaliy ahamiyatga ega.*

Kalit so'zlar: *yashil investitsiyalar, moliyalashtirish samaradorligi, ekonometrik model, panel ma'lumotlar, DEA, GMM, barqaror rivojlanish.*

Аннотация. *В данной статье разработана эконометрическая модель оценки эффективности финансирования зелёных инвестиций на предприятиях. В исследовании проанализированы факторы, влияющие на эффективность зелёных инвестиций, на основе регрессии панельных данных, моделей с фиксированными и случайными эффектами, а также динамического подхода GMM (Ареллано–Бонда). Для формирования индекса эффективности использован метод анализа среды функционирования (DEA). Результаты модели имеют практическое значение для оптимизации источников финансирования предприятий и совершенствования политики зелёной экономики.*

Ключевые слова: *зелёные инвестиции, эффективность финансирования, эконометрическая модель, панельные данные, DEA, GMM, устойчивое развитие.*

Abstract. *This article develops an econometric model for evaluating the efficiency of financing green investments in enterprises. The study analyzes the factors affecting green investment efficiency based on panel data regression, Fixed Effects and Random Effects models, as well as the dynamic GMM (Arellano–Bond) approach. Data Envelopment Analysis (DEA) is used to construct the efficiency index. The model's results are of practical significance for optimizing enterprises' financing sources and improving green economy policy.*

Keywords: *green investments, financing efficiency, econometric model, panel data, DEA, GMM, sustainable development.*

So'nggi yillarda barqaror rivojlanish kontseptsiyasi doirasida korxonalarning yashil (ekologik) investitsiyalarga yo'naltirgan mablag'lari hajmi barqaror ravishda o'sib bormoqda. Iqlim o'zgarishi, resurslardan foydalanish samaradorligini oshirish va uglerodsizlantirish siyosatining kuchayishi korxonalarni energiya tejamkor texnologiyalarga, qayta tiklanuvchi energiya manbalariga, chiqindilarni qayta ishlash tizimlariga va ekologik toza ishlab chiqarish

jarayonlariga sarmoya kiritishga undamoqda. Biroq bunday investitsiyalarning moliyalashtirish manbalari (o‘z mablag‘lari, bank kreditlari, yashil obligatsiyalar, davlat subsidiyalari, xalqaro grantlar) xilma-xilligi va ularning har birining qiymati hamda risk darajasi turlicha bo‘lganligi sababli, yashil investitsiyalarni moliyalashtirish samaradorligini miqdoriy jihatdan aniq baholash dolzarb ilmiy-amaliy muammo hisoblanadi.

Ushbu tezisning maqsadi — korxonalarda yashil investitsiyalarni moliyalashtirish samaradorligini baholash uchun ekonometrik modelni ishlab chiqish, uning nazariy asoslarini, tenglama spetsifikatsiyasini, baholash usullarini (OLS, Fixed Effects, Random Effects, GMM) hamda diagnostik testlarini aniq formulalar asosida ko‘rsatib berishdan iborat.

Yashil investitsiyalarni moliyalashtirish samaradorligi deganda, korxona tomonidan ekologik loyihalarga yo‘naltirilgan har bir moliyaviy birlik hisobiga erishilgan iqtisodiy, ekologik va moliyaviy natijalar nisbati tushuniladi. Bunda samaradorlik ko‘p o‘lchovli kategoriya bo‘lib, u quyidagi asosiy komponentlarni o‘z ichiga oladi: (1) moliyaviy samaradorlik — investitsiya rentabelligi va sof joriy qiymat; (2) ekologik samaradorlik — chiqindilar va zararli emissiyalarning kamayishi; (3) institutsional samaradorlik — moliyalashtirish manbalarining barqarorligi va qiymati. Yashil investitsiya loyihasining moliyaviy jozibadorligini baholashda sof joriy qiymat (NPV) va daromadlilik indeksi keng qo‘llaniladi:

$$NPV = \Sigma_{t=1} n\ CF_t / (1 + r)^t - I_0 \quad (1)$$

Bunda: CF_t — t-davrdagi yashil loyihadan tushgan pul oqimi (energiya tejash, soliq imtiyozlari va h.k. hisobiga); r — diskont stavkasi (moliyalashtirish manbaining o‘rtacha tortilgan qiymati, WACC); I0 — boshlang‘ich investitsiya hajmi; n — loyiha muddati.

Yashil investitsiyalarni moliyalashtirish samaradorligiga ta'sir etuvchi omillarni aniqlash va ularning statistik ahamiyatini baholash uchun panel ma'lumotlar (panel data) asosida qurilgan ekonometrik model qo‘llaniladi. Panel model bir necha korxona (kesim, i = 1,...,N) bo‘yicha bir necha yil (vaqt, t = 1,...,T) davomidagi kuzatuvlarni birlashtirib, individual va vaqt bo‘yicha o‘zgarmas omillarni nazorat qilish imkonini beradi.

Yashil investitsiyalar moliyalashtirish samaradorligini bog‘liq o‘zgaruvchi sifatida quyidagi ko‘p faktorli chiziqli regressiya tenglamasi orqali ifodalaymiz:

$$GIE_{it} = \beta_0 + \beta_1 FIN_{it} + \beta_2 SIZE_{it} + \beta_3 LEV_{it} + \beta_4 ROA_{it} + \beta_5 RND_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (2)$$

GIE_it — i-korxonaning t-yildagi yashil investitsiya samaradorlik indeksi (chiqim/kirim nisbati); FIN_it — yashil moliyalashtirish hajmi (jami aktivlarga nisbatan ulush); SIZE_it — korxona hajmi (aktivlar logarifmi); LEV_it — moliyaviy richag (qarz/aktiv nisbati); ROA_it — aktivlar rentabelligi; RND_it — ilmiy-tadqiqot xarajatlari ulushi; μ_i — korxonaga xos individual effekt; λ_t — vaqt effekti; ε_it — tasodifiy xatolik hadi; β0...β5 — baholanadigan koeffitsientlar.

Bog‘liq o‘zgaruvchi GIE_it ni shakllantirish uchun ma'lumotlarni qamrab olish tahlili (Data Envelopment Analysis, DEA) usulidan foydalanish tavsiya etiladi, chunki u bir nechta kirim va chiqim ko‘rsatkichlarini bitta nisbiy samaradorlik ballida birlashtirish imkonini beradi:

$$E_i = (\Sigma_{r=1} s\ u_r y_{ri}) / (\Sigma_{m=1} k\ v_m x_{mi}) \quad (3)$$

Bunda: E_i — i-korxonaning nisbiy samaradorlik koeffitsienti (0 ≤ E_i ≤ 1); y_ri — r-chiqim ko‘rsatkichi (masalan, CO2 emissiyasining kamayishi, energiya tejash hajmi, qo‘shimcha daromad); x_mi — m-kirim ko‘rsatkichi (yashil investitsiya hajmi, moliyalashtirish xarajatlari); u_r, v_m — chiqim va kirim uchun optimallashtirilgan tortiq koeffitsientlari, chiziqli dasturlash yo‘li bilan topiladi.

Individual korxonalarga xos kuzatilmaydigan xususiyatlarni hisobga olish uchun sobit effektlar (Fixed Effects, FE) modeli quyidagicha ifodalanadi:

$$Y_{it} = \alpha_i + \beta' X_{it} + \varepsilon_{it} \quad (4)$$

Tasodifiy effektlar (Random Effects, RE) modelida esa individual effekt tasodifiy komponenta sifatida kiritiladi:

$$Y_{it} = \beta_0 + \beta' X_{it} + u_i + \varepsilon_{it} \quad (5)$$

Ikki model o‘rtasida to‘g‘ri tanlovni amalga oshirish uchun Xausman testi qo‘llaniladi. Nol gipoteza (H0: RE model izchil va samarali) rad etilsa, FE model afzal ko‘riladi:

$$H = (\beta_{FE} - \beta_{RE})' \, [Var(\beta_{FE}) - Var(\beta_{RE})]^{-1} \, (\beta_{FE} - \beta_{RE}) \sim \chi^2(k) \quad (6)$$

Agar hisoblangan H statistikasi tanqidiy qiymatdan katta bo‘lsa ($p < 0,05$), H0 rad etiladi va FE model qo‘llaniladi.

Dinamik panel model va umumlashtirilgan momentlar usuli:

Yashil investitsiya samaradorligining oldingi davrdagi qiymatiga bog‘liqligini (inersiya effekti) hisobga olish uchun Arellano–Bond dinamik panel modeli qo‘llaniladi:

$$Y_{it} = \rho Y_{i,t-1} + \beta' X_{it} + \mu_i + \varepsilon_{it} \quad (7)$$

Individual effektlarni yo‘qotish uchun birinchi tafovutlar (first differencing) usuli qo‘llaniladi:

$$\Delta Y_{it} = \rho \Delta Y_{i,t-1} + \beta' \Delta X_{it} + \Delta \varepsilon_{it} \quad (8)$$

Endogenlik muammosini bartaraf etish uchun instrumental o‘zgaruvchilar sifatida bog‘liq o‘zgaruvchining kechikkan qiymatlaridan foydalaniladi, moment sharti quyidagicha yoziladi:

$$E[Y_{i,t-s} \cdot \Delta \varepsilon_{it}] = 0, \quad s \geq 2 \quad (9)$$

Model diagnostikasi va statistik baholash mezonlari:

Model sifatini baholash va uning statistik ishonchliligini tekshirish uchun quyidagi ko‘rsatkichlar hisoblanadi.

Model tomonidan tushuntirilgan dispersiya ulushi (aniqlash koeffitsienti):

$$R^2 = 1 - (\Sigma(Y_i - \hat{Y}_i)^2) / (\Sigma(Y_i - \bar{Y})^2) \quad (10)$$

Ko‘p qo‘llidoshlik (multikollinearlik) darajasini aniqlash uchun dispersiyani oshirish koeffitsienti (VIF) hisoblanadi:

$$VIF_j = 1 / (1 - R_j^2) \quad (11)$$

VIF_j > 10 bo‘lsa, mos j-omil bo‘yicha multikollinearlik muammosi mavjud deb hisoblanadi va model qayta spetsifikatsiya qilinishi lozim.

Qoldiqlarning avtokorrelyatsiyasini tekshirish uchun Durbin–Uotson statistikasi qo‘llaniladi:

$$DW = \Sigma_{t=2}n \, (\varepsilon_t - \varepsilon_{t-1})^2 / \Sigma_{t=1}n \, \varepsilon_t^2 \quad (12)$$

$DW \approx 2$ qiymati qoldiqlarda avtokorrelyatsiya yo‘qligini bildiradi; $DW < 1,5$ yoki $DW > 2,5$ bo‘lsa, model qoldiqlarida seriya bog‘liqligi mavjudligi ehtimoli yuqori.

Gomoskedastiklik farazini tekshirish uchun Breusch–Pagan testi qo‘llaniladi, uning test statistikasi:

$$LM = n \cdot R_{aux}^2 \sim \chi^2(k) \quad (13)$$

Amaliy tatbiq va kutilayotgan natijalar: Ishlab chiqilgan ekonometrik modelni amalda qo‘llash uchun O‘zbekiston sanoat korxonalarining kamida 5 yillik ($T \geq 5$) moliyaviy hisobotlari va ekologik ko‘rsatkichlari (energiya sarfi, chiqindilar hajmi, CO2 emissiyasi) asosida balanslashtirilgan panel ma'lumotlar to‘plami (balanced panel dataset) shakllantirilishi lozim. Model orqali quyidagi gipotezalar tekshiriladi: (1) yashil moliyalashtirish hajmining o‘sishi samaradorlik indeksiga ijobiy ta'sir ko‘rsatadi; (2) korxona hajmi kattalashgan sari moliyalashtirish samaradorligi masshtab effekti hisobiga oshadi; (3) yuqori moliyaviy richag (qarz yuki) samaradorlikka salbiy ta'sir ko‘rsatishi mumkin.

Bir birlik moliyalashtirishga to‘g‘ri keladigan ekologik samara (uglerod chiqindisining kamayishi) quyidagi ko‘rsatkich orqali baholanadi:

$$CE_i = \Delta CO_{2,i} / FIN_i \quad (14)$$

CE_i — i-korxonaning bir so‘m yashil investitsiyaga to‘g‘ri keladigan uglerod chiqindisi kamayish darajasi (tonna/mln so‘m).

Model natijalari asosida korxona rahbariyati va davlat boshqaruv organlari uchun moliyalashtirish manbalarini optimallashtirish, yashil obligatsiyalar chiqarish hajmini asoslash hamda subsidiyalash siyosatini takomillashtirish bo‘yicha aniq tavsiyalar ishlab chiqish mumkin bo‘ladi.

Xulosa.

Xulosa qilib aytganda, korxonalarda yashil investitsiyalarni moliyalashtirish samaradorligini baholash murakkab, ko‘p omilli jarayon bo‘lib, uni to‘g‘ri baholash uchun DEA usulida shakllantirilgan samaradorlik indeksini panel ma'lumotlar regressiyasi (FE/RE), zarur holda dinamik GMM modeli bilan birlashtirgan kompleks ekonometrik yondashuv talab etiladi. Taklif etilgan model (1)–(14) formulalar tizimi orqali nafaqat moliyaviy, balki ekologik samaradorlikni ham miqdoriy baholash imkonini beradi, bu esa korxonalarning barqaror rivojlanish strategiyasini ilmiy asoslangan tarzda shakllantirishga xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR

1. Arellano, M., Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies, 58(2), 277–297.
2. Charnes, A., Cooper, W.W., Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. European Journal of Operational Research, 2(6), 429–444.
3. Wooldridge, J.M. (2010). Econometric Analysis of Cross Section and Panel Data. MIT Press.
4. O‘zbekiston Respublikasi "Yashil iqtisodiyot" konsepsiyasi va tegishli me'yoriy-huquqiy hujjatlar.
5. Mamayusupova Sh.U. O‘zbekistonda yashil moliyalashtirishning rivojlanish bosqichlari va istiqbollari. Journal of Innovation in Education and Social Research, 2026, 4(5), 75–82.
6. Tursunxodjayeva Sh.Z. "Yashil" moliyalashtirish moliyaviy risklarni kamaytirish instrumenti sifatida. Moliya tizimini rivojlantirishning zamonaviy tendensiyalari va istiqbollari, 2024, 1(4), 295–302.
7. Bozorboyev D.P. "Yashil moliyalashtirish" tizimining tarkibi va amal qilish xususiyatlari. International Scientific and Practical Conference (ISAPR) materiallari, 2024, 1-sessiya, 2-qism.

QOIDAVIY KKT FORMALIZMI VA YIRIK TIL MODELLARI INTEGRATSIYASI: ARXITEKTURA VARIANTLARI TASNIFI VA BAHOLASH MEZONLARI

Sirojiddinov Ziyoviddin Sharofiddin o'g'li
«Kapital Sug'urta» AJ, Toshkent, O'zbekiston
ORCID: 0000-0002-6249-032X · ziyo07070@gmail.com
Hakimov Muftax Hamidovich
Professor, Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti
ORCID: 0000-0001-6937-800X · muftah@mail.ru
https://doi.org/10.5281/zenodo.23253809

Annotatsiya. *Tezisda ijtimoiy ahamiyatli sohalar raqamlashtirilishi sharoitida kam resursli o'zbek tili uchun qoidaviy mashina tarjimasi bilan yirik til modellarini (YTM) birlashtirish masalasi ko'riladi. Prof. M. X. Hakimov ishlab chiqqan kengayuvchi kirish tili (KKT) formalizmiga asoslangan o'zbekcha–inglizcha tizim negizida integratsiyaning olti arxitektura varianti yagona tasnifga keltiriladi. Tasnif ikki savolga javob beradi: har konfiguratsiyada KKT nima qiladi va YTM nima qiladi. Sof qoidaviy va sof neyron chegara-holatlardan tortib, YTM realizator, tahlil-iste'molchi, verifikatsiya sikli ostidagi tarjimon va yozma qoida-matn ijrochisi vazifasini bajaradigan konfiguratsiyalar ko'rib chiqiladi; har biri uchun mehnat taqsimoti, kutiladigan afzallik va ishlash-sharti aniqlanadi. Tasnifga ortogonal ravishda tarjima vaqtidan tashqari yo'nalish — tasdiq-darvozali oflayn resurs to'ldirish — ajratiladi. Gibrid tizimlarni baholashda yagona avtomatik ball yetarli emasligi asoslanib, to'rt o'qli mezon ramkasi taklif etiladi: sifat, formalizm-sodiqlik, determinizm va terminologik izchillik. Miqdoriy taqqoslash natijalari oldindan ro'yxatga olingan protokol asosida alohida ishda e'lon qilinadi.*

Kalit so'zlar: *mashina tarjimasi, KKT formalizmi, yirik til modellari, gibrid arxitektura, tasdiq darvozasi, baholash mezonlari, o'zbek tili.*

Аннотация. *В тезисах рассматривается задача объединения правилового машинного перевода и больших языковых моделей (LLM) для малоресурсного узбекского языка в условиях цифровизации социально значимых сфер. На базе узбекско-английской системы, основанной на формализме расширяемого входного языка (KKT, разработан проф. М. Х. Хакимовым), предлагается единая классификация шести архитектурных вариантов интеграции: от чисто правилового и чисто нейронного предельных случаев до конфигураций, где LLM выступает реализатором, потребителем анализа, переводчиком под циклом верификации или исполнителем письменного свода правил. Отдельно выделяется ортогональное направление — офлайн-пополнение ресурсов через ворота подтверждения. Обосновывается недостаточность единого автоматического балла и предлагается четырёхосевая система критериев: качество, верность формализму, детерминизм и терминологическая согласованность. Количественные результаты будут опубликованы в отдельной работе по предварительно зарегистрированному протоколу.*

Ключевые слова: *машинный перевод, формализм KKT, большие языковые модели, гибридная архитектура, ворота подтверждения, критерии оценки, узбекский язык.*

Abstract. *The paper addresses the integration of rule-based machine translation with large language models (LLMs) for the low-resource Uzbek language under the digitalisation of socially significant domains. Based on an Uzbek–English system grounded in the Extensible Input Language (KKT) formalism developed by Prof. M. Kh. Khakimov, six architecture variants are brought into a single classification answering two questions for each configuration: what the rule-based component does and what the LLM does. The variants range from the purely rule-based and purely neural boundary cases to configurations in which the LLM acts as a realiser, an analysis consumer, a translator under a verification loop, or an executor of an automatically compiled written rule text. For each variant the division of labour, expected advantage and operating condition are identified. Orthogonally to the classification, an offline resource-augmentation direction with a confirmation gate is distinguished. A four-axis evaluation framework is substantiated: quality, formalism fidelity, determinism and terminological consistency. Quantitative results will be reported in a separate study under a pre-registered protocol.*

Keywords: *machine translation, KKT formalism, large language models, hybrid architecture, confirmation gate, evaluation criteria, Uzbek language.*

1. Kirish

Ijtimoiy tizimlarning raqamli transformatsiyasi — davlat hujjatlari aylanmasi, ta'lim resurslari va huquqiy matnlarning raqamlashtirilishi — tarjima texnologiyasiga alohida talab qo'yadi: bunday sohalarda natijaning nafaqat sifati, balki har qarorning asoslanishi, takrorlanishi va nazorat qilinishi ham muhim. Neyron mashina tarjimasi ravonlik bo'yicha yuqori natija beradi, biroq chiqishi nodeterministik, terminologik jihatdan notekis va ichki qarorlari izohlab bo'lmaydigan tarzda shakllanadi. Qoidaviy tizimlar esa determinizm, izohlanuvchanlik va terminologik nazoratni ta'minlaydi, ammo qamrovi qoidalar va lug'at hajmi bilan chegaralanadi.

Kam resursli o'zbek tili uchun ikkala yondashuvning kuchli tomonlarini birlashtirish tabiiy yo'l; savol qoidaviy tizimni yirik til modeli (YTM) bilan almashtirishda emas — ularni qanday birlashtirishda. Gibrid tizimlarning mavjud tasniflari odatda quvurdagi birlashtirish nuqtasiga yoki komponentlar ketma-ketligiga tayanadi. Ushbu ishda tasnifning asosi sifatida boshqa mezon olinadi: har konfiguratsiyada nazorat manbai qayerda — kirish tomonida, chiqish tomonida yoki umuman yo'q. Shu bilan birga baholash mezonlariga formalizm-sodiqlik o'qi kiritiladi, ya'ni chiqishning qoidaviy tahlilga mosligi mustaqil o'lchov sifatida qaraladi.

Tezisning maqsadi — integratsiya konfiguratsiyalarini yagona tasnifga keltirish va ularni baholashning ko'p-mezonli ramkasini taklif etish.

2. Qoidaviy asos

Tadqiqot kengayuvchi kirish tili (KKT) formalizmi doirasida olib boriladi. Formalizm prof. M. X. Hakimov tomonidan ishlab chiqilgan [Aripov M., 2022: 81; Hakimov M., 2026: 104]: so'z va gap tuzilishi maxsus belgilar qatori (matematik model) ko'rinishida yoziladi, belgilar to'plami yangi til hodisalari uchun kengaytirilishi mumkin. Ushbu tezisning hissasi formalizmning o'zi emas, balki uning YTM bilan integratsiya konfiguratsiyalarini tasniflash va baholash ramkasidir.

Shu asosda o'zbek–ingliz yo'nalishida ishlaydigan qoidaviy tarjima tizimi ishlab chiqilgan: o'zbek so'zi o'ngdan chapga morfologik segmentatsiya bilan tahlil qilinadi, gap fe'lli bo'laklarga ajratiladi va har bo'lak o'z modeli bo'yicha ingliz tiliga o'giriladi. Tizim uch xossaga ega: takrorlanuvchanlik (bir kirishga doim ayni bir chiqish), izohlanuvchanlik (har qaror

aniq qoidaga bog‘lanadi) va boshqariluvchanlik (lug‘at va qoidalarni maqsadli, nazorat ostida kengaytirish). Quyidagi tasnifda ayni shu xossalar turli nuqtalarda nazorat manbai vazifasini o‘taydi.

3. Integratsiya arxitekturalarining tasnifi

Olti konfiguratsiya ikki savol bo‘yicha farqlanadi: har birida KKT nima qiladi va YTM nima qiladi. 1-jadvalda mehnat taqsimoti umumlashtirilgan, quyida har variant afzalligi va ishlash-sharti bilan tavsiflanadi. Jadvalning oxirgi ustuni variantning hozirgi holatini ko‘rsatadi — tajriba tizimida amalga oshirilgan, qisman amalga oshirilgan yoki hozircha konseptual.

1-jadval

	KKT nima qiladi	**YTM nima qiladi**	**Holat**
T1	tahlil, transfer va sintez — tarjimani to‘liq bajaradi	ishtirok etmaydi	amalga oshirilgan
T2	ishtirok etmaydi	tarjimani to‘liq bajaradi	amalga oshirilgan
T3	tahlil qiladi, gap bo‘laklarini va so‘z tarjimalarini beradi	so‘zlarni joylashtirib, jumla yig‘adi	konseptual
T4	tahlil natijasini tavsiya sifatida beradi	tavsiyani qabul qilish yoki qilmaslikni o‘zi hal qiladi, tarjimani bajaradi	qisman
T5	tarjimani tekshiradi; xatolik topilsa jarayonni T4 holatiga o‘tkazadi	tarjima qiladi, signal kelsa qayta tarjima qiladi	konseptual
T6	qoidalarini yozma matn ko‘rinishida beradi	shu qoidalarga tayanib tarjima qiladi	qisman

T1 (nazorat-etalon). Tahlil, transfer va sintez — hammasi qoidaviy quvurda bajariladi; neyron model umuman ishtirok etmaydi. Determinizm, izohlanuvchanlik va terminologik nazorat bu variantda maksimal, shuning uchun u qolgan beshtasi uchun taqqos-nuqta bo‘lib xizmat qiladi. Cheklovi: ochiq matnda ravonlik va qamrov qoidalar hamda lug‘at hajmiga qat’iy bog‘liq.

T2 (ikkinchi chegara-holat). Tarjimani butunlay neyron model bajaradi, KKT ishtirok etmaydi. Ravonlik va leksik qamrov yuqori. Cheklovi: chiqish nodeterministik, terminologik jihatdan notekis va uni ichki qoidalarga tayanib tekshirib bo‘lmaydi.

T3 (realizator). KKT gapni tahlil qiladi — gap bo‘laklarini ajratadi, so‘zlarning tarjimasini va grammatik belgilarini (zamon, son, sintaktik rol) beradi; YTM esa shu tayyor birliklarni joylashtiradi va ravon jumla yig‘adi. Mazmun qarori qoidaviy tomonda, so‘z tartibi va sirt-realizatsiya neyron tomonda qoladi. Ishlash-sharti: tahlil paketi majburiy shartnoma sifatida qaraladi — undagi xato bevosita chiqishga ko‘chadi, shuning uchun YTM ning paketga sodiqligi alohida o‘lchanishi kerak. Bu variant o‘lchangan ehtiyoj ustiga quriladi: qoidaviy chiqishdagi qoldiq nuqsonlar bir necha qatlamga taqsimlanadi va ularning bir qismi (son–ot kelishuvi, determiner joylashuvi, fe’l zanjiri) aynan sirt-realizatsiya qatlamiga tegishli.

T4 (tahlil-iste'molchi). KKT morfologik tahlilni — o'zak, affiks tarkibi, turkum kodlari — tavsiya sifatida beradi; tarjimani YTM bajaradi va tavsiyani qabul qilish yoki qilmaslikni o'zi hal qiladi. T3 dan farqi shunda: u yerda tahlil paketi majburiy, bu yerda ixtiyoriy kontekst. Ishlash-sharti: tahlil aniqligi yuqori bo'lishi va model uni haqiqatan e'tiborga olishi; aks holda tahlil foydasiz yoki chalg'ituvchi bo'lib qoladi.

T5 (verifikatsiya sikli). Tarjimani YTM bajaradi, KKT esa tarjima qilmaydi — u manba gapni va chiqishni tahlil qilib, yuqori aniqlikdagi nomuvofiqlik signallarini (inkor, son, zamon, majburiy termin, mazmun tushishi) qidiradi. Signal topilmasa chiqish o'zgarishsiz qoladi; topilsa jarayon T4 holatiga o'tadi — KKT tahlili izoh sifatida qaytariladi va YTM qayta tarjima qiladi. Ikki shart muhim. Birinchisi: tarjima bosqichida YTM KKT ni ko'rmasligi kerak, aks holda variant boshidanoq T4 ga aylanadi va tekshiruv mustaqil bo'lmaydi. Ikkinchisi: soxta signal ulushi past bo'lishi kerak — bu farazi xavf emas, dastlabki tajribalarimizda verifikatsiya sikli to'g'ri chiqishlarni ham o'zgartirgani kuzatilgan. Bunday tekshiruvlarning urug'lari qoidaviy tizimning ichki nazorat darvozalari (majburiy token saqlanishi, inkor polaritesi, kesim chegarasi) ko'rinishida allaqachon ishlab turibdi.

T6 (qoida-matn ijrochisi). KKT qoidalari — affiks inventari, so'z va gap modellari, tartib cheklovlari — ma'lumotlar bazasidan avtomatik ravishda yaxlit yozma matnga aylantiriladi va YTM ko'rsatmasi tarkibida beriladi; model tarjimani ana shu yozma formalizmga tayanib bajaradi. T4 dan farqi: u yerda modelga shu gapning tahlili beriladi (dinamik, har gapga boshqacha), bu yerda esa butun formalizm matni (statik, hamma gap uchun bir xil). Afzalligi: qoidaviy bilim modelga kod-integratsiyasiz uzatiladi, baza yangilanganda matn avtomatik qayta yig'iladi. Ishlash-sharti: qoida matnining to'liqligi, bir ma'noliligi va model kontekst oynasiga sig'ishi.

Tasnifga ortogonal yo'nalish: tasdiq-darvozali resurs to'ldirish. Yuqoridagi olti variant tarjima vaqtidagi mehnat taqsimotini qamraydi. Amaliyotimizda bulardan tashqari tarjima vaqtidan tashqari (oflayn) yo'nalish markaziy o'rin tutdi: YTM lug'at va qoidalar uchun nomzodlar taklif qiladi, har nomzod inson ko'rigi va avtomatik qoida nazoratidan (tasdiq darvozasi) o'tgachgina resurs bazasiga kiradi. Bu yo'nalish T1–T6 ga ortogonal: tizim qaysi konfiguratsiyada ishlashidan qat'i nazar, uning resurs bazasini boyitadi va qoidaviy yadroning uch xossasini buzmaydi. Amaliy tajriba ikki sifat kuzatuvini berdi: mustaqil modellar o'rtasidagi konsensus to'g'rilik kafolati emas — ikkala model bir xil kontekst tuzoqqa birga tushishi mumkin; va konsensus ilgari o'lchab qabul qilingan qarorni bekor qila olmaydi — yakuniy so'z tasdiq darvozasida qoladi. Shu bois darvoza protokoli (muzlatilgan nomzod ro'yxati, mustaqil ko'r annotatsiya, arbitraj, avtomatik qo'riqchilar, o'lchov) integratsiyaning majburiy tarkibi sifatida qaraladi.

4. Baholash mezonlari

Yagona etalonli avtomatik ballar gibrid konfiguratsiyalarni solishtirish uchun yetarli emas: ular sirt mosligini o'lchaydi, qonuniy tarjima variativligini jazolaydi [Papineni K., 2002: 311; Freitag M., 2020: 61] va nazorat xossalarini umuman aks ettirmaydi. Kam resursli tillar uchun qo'shimcha qiyinchilik — etalon to'plamlarning cheklangan hajmi va domen qamrovi [Goyal N., 2022: 522]. YTM ni baholovchi sifatida qo'llash ham o'z muammolariga ega [Kocmi T., 2023: 193].

Shu bois taqqoslashni to'rt mustaqil o'q bo'yicha olib borish taklif etiladi: (1) sifat — mazmun muvofiqligi va ravonlik, tizim nomi yashirilgan (ko'r) inson baholash orqali; (2) formalizm-sodiqlik — chiqishning KKT tahliliga va qoidalariga mosligi: morfologik moslik,

tartib cheklovlariga rioya, berilgan leksik ekvivalentlardan chetlashish darajasi; (3) determinizm — ayni kirishga bayt darajasida takrorlanadigan chiqish; (4) izchillik — hujjat bo‘ylab bir terminning barqaror tarjimasi. Qo‘shimcha amaliy mezon sifatida verifikatsiya qulayligi — xatoni komponentgacha lokalizatsiya qilish imkoni — qayd etiladi.

Ikkinchi o‘q — formalizm-sodiqlik — mavjud baholash amaliyotida odatda uchramaydi va ushbu ramkaning asosiy farqi shundadir: u chiqishning etalon bilan emas, tizimning o‘z tahlili bilan mosligini o‘lchaydi. Har konfiguratsiya bu o‘qlarda o‘ziga xos profil hosil qiladi; eng yaxshi variant emas, vazifa talabiga mos variant tanlanadi.

5. Muhokama: nazorat–erkinlik o‘qi

Olti variant nazorat–erkinlik o‘qiga tabiiy joylashadi: T1 va T2 — chegara nuqtalari; T3, T4 va T6 nazoratni kirish tomonidan (mazmun paketi, tahlil yoki yozma qoida orqali), T5 esa chiqish tomonidan (verifikatsiya orqali) kiritadi. Umumiy qonuniyat: nazorat qanchalik qattiq bo‘lsa, tizim xulqi shunchalik bashoratli, biroq sirt erkinligi shunchalik cheklangan; va aksincha.

Amaliy tanlov qoidasi sifatida quyidagini taklif etamiz: terminologiya va tuzilma kafolatlari hal qiluvchi bo‘lgan, izohlanuvchanlik boshqaruv talabi darajasiga ko‘tarilgan domenlarda (davlat hujjatlari aylanmasi, huquqiy, sug‘urta, tibbiy matnlar) nazorat-og‘ir variantlar (T1, T3, T6) asos qilib olinadi; umumiy ma’lumot matnlarida erkinlik-og‘ir variantlar (T2, T5) maqsadga muvofiq; T4 tahlil aniqligi yuqori tor domenlar uchun oraliq yechim. Kam resursli til uchun yana bir jihat muhim: qoidaviy yadro to‘liq oflayn ishlaydi, ya’ni nazorat-og‘ir konfiguratsiyalar ma’lumot suverenitetini ham ta’minlaydi.

6. Cheklovlar

Ushbu ish tasnif va ramka taklif etadi, miqdoriy taqqoslash emas. Uch cheklov ochiq qayd etiladi. Birinchidan, tasnif bitta til juftidan (o‘zbek–ingliz) va bitta formalizmdan chiqarilgan; uning boshqa juftlar va formalizmlarga ko‘chishi tekshirilmagan. Ikkinchidan, oltita variant nazariy jihatdan to‘liq to‘plam emas — ular kuzatilgan va amalda sinalgan konfiguratsiyalarni qamraydi. Uchinchidan, variantlarning ikkitasi (T3, T5) hozircha konseptual bosqichda, ikkitasi (T4, T6) qisman amalga oshirilgan; ularning ishlash shartlari to‘liq empirik tekshiruvdan o‘tmagan. Olib borilayotgan tajribalar har variantning foydasi o‘z ishlash sharti bilan qattiq bog‘liqligini ko‘rsatmoqda; miqdoriy natijalar — ehtimoliy salbiy holatlar bilan birga — oldindan ro‘yxatga olingan protokol asosida alohida ishda e’lon qilinadi.

7. Xulosa

Tezisda KKT formalizmi va yirik til modellari integratsiyasining olti arxitektura varianti yagona tasnifga keltirildi; har biri uchun KKT va YTM o‘rtasidagi mehnat taqsimoti, kutiladigan afzallik va ishlash sharti aniqlandi. Tasnifga ortogonal ravishda tasdiq-darvozali oflayn resurs to‘ldirish yo‘nalishi ajratildi hamda to‘rt o‘qli baholash ramkasi taklif etildi. Ramka ikki vazifani bajaradi: tadqiqot uchun — taqqoslashni bitta avtomatik ballga emas, ko‘p-mezonli o‘lchovga qurish; amaliyot uchun — vazifa profiliga qarab konfiguratsiya tanlashga asos berish. Kelgusi ish barcha variantlarni yagona muzlatilgan sinov to‘plamlarida, oldindan e’lon qilingan mezonlar bo‘yicha empirik taqqoslashdan iborat.

FOYDALANILGAN ADABIYOTLAR

1. Papineni K., Roukos S., Ward T., Zhu W.-J. BLEU: a method for automatic evaluation of machine translation // Proceedings of the 40th Annual Meeting of the ACL. — 2002. — pp. 311–318.

2. Freitag M., Grangier D., Caswell I. BLEU might be guilty but references are not innocent // Proceedings of EMNLP 2020. — 2020. — pp. 61–71.
3. Goyal N., Gao C., Chaudhary V. et al. The FLORES-101 evaluation benchmark for low-resource and multilingual machine translation // Transactions of the ACL. — 2022. — Vol. 10. — pp. 522–538.
4. Aripov M., Khakimov M., Matlatipov S., Sirojiddinov Z. Analysis and processing of the Uzbek language on the multi-language modelled computer translator technology // Lecture Notes in Computer Science. — Vol. 13212. — Cham: Springer, 2022. — pp. 81–95.
5. Hakimov M., Sirojiddinov Z. Towards the technology of a multi-language models computer translator // O'zbekiston Respublikasi Fanlar akademiyasining ma'ruzalari. — 2026. — №4. — B. 104–109.
6. Kocmi T., Federmann C. Large language models are state-of-the-art evaluators of translation quality // Proceedings of the 24th Annual Conference of the EAMT. — 2023. — pp. 193–203.

XITOY VA O'ZBEK TILLARIDAGI EKOTURIZM TERMINLARINING STRUKTUR XUSUSIYATLARI VA RAQAMLI TERMINOLOGIK BAZANI SHAKLLANTIRISH MASALALARI

Sobirova Madinaxon Raxim qizi
Toshkent davlat Sharqshunoslik universiteti Toshkent, O'zbekiston
ORCID ID: 0009-0008-0123-7495
https://doi.org/10.5281/zenodo.23253819

Annotatsiya. *Mazkur tezis xitoy va o'zbek tillaridagi ekoturizm terminlarining struktur xususiyatlari hamda ularni raqamli terminologik bazaga tizimlashtirish masalalariga bag'ishlangan. Tadqiqotda ekoturizm terminlarining struktur modellari chog'ishtirma jihatdan tahlil qilindi. Natijalar xitoy tilida qo'shma terminlar, o'zbek tilida esa termin birikmalarining ustunligini ko'rsatdi. Shuningdek, ekoturizm terminlarini tizimlashtirish va raqamli terminologik bazani shakllantirishning ayrim jihatlari yoritildi. Olingan natijalar ekoturizm terminologiyasini chog'ishtirma o'rganish hamda elektron terminologik resurslarni yaratishda muhim ahamiyatga ega.*

Kalit so'zlar: *ekoturizm terminlari, struktur tahlil, terminologiya, xitoy tili, o'zbek tili, raqamli terminologik baza.*

Аннотация. *Данный тезис посвящён изучению структурных особенностей экотуристических терминов в китайском и узбекском языках, а также вопросам их систематизации в цифровой терминологической базе. В исследовании проведён сопоставительный анализ структурных моделей экотуристических терминов. Результаты показали преобладание сложных терминов в китайском языке и терминологических словосочетаний в узбекском языке. Кроме того, рассмотрены отдельные аспекты систематизации экотуристической терминологии и формирования цифровой терминологической базы. Полученные результаты имеют важное значение для сопоставительного изучения экотуристической терминологии и создания электронных терминологических ресурсов.*

Ключевые слова*: экотуристические термины, структурный анализ, терминология, китайский язык, узбекский язык, цифровая терминологическая база.*

Abstract. *This thesis is devoted to the study of the structural features of ecotourism terms in Chinese and Uzbek, as well as issues related to their systematization in a digital terminological database. The research provides a comparative analysis of the structural models of ecotourism terms. The results reveal the predominance of compound terms in Chinese and terminological word combinations in Uzbek. In addition, certain aspects of systematizing ecotourism terminology and developing a digital terminological database are discussed. The findings are significant for the comparative study of ecotourism terminology and the development of electronic terminological resources.*

Keywords*: ecotourism terms, structural analysis, terminology, Chinese language, Uzbek language, digital terminological database.*

So'nggi yillarda ekologik muammolarning kuchayishi va barqaror rivojlanish tamoyillarining keng qo'llanilishi natijasida ekoturizm sohasi jadal rivojlanmoqda (Weber, A. S.,

2024, 465–471). Ushbu jarayon ekoturizmga oid terminologik birliklar sonining ortishiga ham sabab bo‘lmoqda (Buckley, R., & Ollenburg, C., 2011, 307-309). Terminologik tizimning izchil shakllanishi esa terminlarning struktur va semantik xususiyatlarini o‘rganishni talab etadi (Hasanova G., 2019,54 – 56)

Xitoy va o‘zbek tillari turli tipologik tizimlarga mansub bo‘lgani sababli ulardagi ekoturizm terminlarining tuzilishida ham muayyan farqlar kuzatiladi (Turg‘unov R., 2017, 68). Tadqiqot davomida ekoturizm sohasiga oid terminlar struktur jihatdan tahlil qilinib, sodda terminlar, qo‘shma terminlar hamda termin birikmalari ajratildi. Tahlil natijalari terminlarning shakllanish modellari va ularni raqamli terminologik bazaga tizimlashtirish imkoniyatlarini aniqlashga xizmat qiladi.

Tahlil qilingan ekoturizm terminlari tarkibida xitoy tilida qo‘shma terminlar ustunlik qiladi. Masalan, 生态旅游 (shēngtài lǚyóu) termini 生态 (shēngtài – ekologiya, ekologik) va 旅游 (lǚyóu – turizm) komponentlaridan tashkil topgan bo‘lib, “ekoturizm” ma’nosini ifodalaydi (李艳., 2003, 177页). Xuddi shu model 生态环境 (shēngtài huánjìng) – “ekologik muhit”, 生态资源 (shēngtài zīyuán) – “ekologik resurslar”, 生态平衡 (shēngtài pínghéng) – “ekologik muvozanat” terminlarida ham kuzatiladi. Ushbu terminlarda 生态 (shēngtài) komponenti termin yasovchi asosiy unsur vazifasini bajaradi.

Ekologik obyektlarni ifodalovchi terminlar tarkibida ham ko‘p komponentli modellar faol qo‘llanadi (Choi, H. C., 2016, 916–917). Masalan, 自然保护区 (zìrán bǎohùqū) termini 自然 (zìrán – tabiat), 保护 (bǎohù – muhofaza qilish) va 区 (qū – hudud) komponentlaridan tashkil topib, “qo‘riqxona” tushunchasini ifodalaydi. 生物多样性 (shēngwù duōyàngxìng) termini esa 生物 (shēngwù – biologik organizmlar) va 多样性 (duōyàngxìng – xilma-xillik) komponentlari asosida hosil bo‘lgan.

O‘zbek tilida ekoturizm terminlari tarkibida termin birikmalari faol qo‘llanadi. Jumladan, “biologik xilma-xillik”, “ekologik muvozanat”, “tabiiy resurslar”, “milliy bog‘”, “ekologik ta’lim” terminlari sifat + ot hamda ot + ot modellari asosida shakllangan. Xitoy tilidan farqli ravishda ushbu birliklar mustaqil so‘zlarning sintaktik birikishi natijasida hosil bo‘ladi.

Tahlil qilingan materiallar orasida xitoy tilida ot + ot modeli asosida shakllangan terminlar ham salmoqli o‘rin egallaydi. Jumladan, 森林旅游 (sēnlín lǚyóu) – “o‘rmon turizmi”, 湿地保护 (shīdì bǎohù) – “botqoqliklarni muhofaza qilish”, 环境教育 (huánjìng jiàoyù) – “ekologik ta’lim”, 绿色发展 (lǜsè fāzhǎn) – “yashil rivojlanish” kabi birliklar ekoturizm sohasining turli tushunchalarini ifodalaydi. Ushbu terminlarda komponentlar o‘rtasidagi munosabat aniqlovchi va aniqlanmish asosida qurilgan.

O‘zbek tilida esa “yashil turizm”, “ekologik nazorat”, “tabiatni muhofaza qilish”, “barqaror rivojlanish”, “ekologik madaniyat” kabi birliklar keng qo‘llanadi. Mazkur terminlar ekoturizmning ekologik, ijtimoiy va ta’limiy yo‘nalishlariga oid tushunchalarni ifodalaydi. Tahlil natijasida o‘zbek tilida ko‘p so‘zli terminlarning qo‘llanish chastotasi xitoy tiliga nisbatan yuqoriroq ekanligi kuzatildi.

Tahlil jarayonida xitoy tilidagi ekoturizm terminlarida 旅游 (lǚyóu – turizm), 环境 (huánjìng – muhit), 保护 (bǎohù – muhofaza) va 资源 (zīyuán – resurs) komponentlarining yuqori faolligi kuzatildi. Masalan, 旅游资源 (lǚyóu zīyuán) – “turistik resurslar”, 环境保护 (huánjìng bǎohù) – “atrof-muhitni muhofaza qilish”, 生态保护 (shēngtài bǎohù) – “ekologik muhofaza”, 绿色旅游 (lǜsè lǚyóu) – “yashil turizm” terminlari ekoturizm terminologiyasining

turli semantik yo‘nalishlarini ifodalaydi. Mazkur birliklar tarkibidagi asosiy komponentlarning takrorlanishi xitoy tilida termin hosil qilishning tizimli xarakterga ega ekanligini ko‘rsatadi.

O‘zbek tilidagi materiallarda esa “ekologik turizm”, “tabiiy meros”, “ekologik xavfsizlik”, “yashil iqtisodiyot”, “rekreatsion hudud” kabi birliklar uchraydi. Ushbu terminlarning bir qismi xalqaro terminologiya asosida shakllangan bo‘lsa, ayrimlari milliy til vositalari yordamida hosil qilingan. Bu holat o‘zbek tilidagi ekoturizm terminologiyasining shakllanishida ichki va tashqi manbalarning birgalikda ishtirok etganligini ko‘rsatadi.

Chog‘ishtirma tahlil natijalari xitoy tilida terminlarning morfemalar kombinatsiyasi asosida ixcham shaklda ifodalanishini, o‘zbek tilida esa termin birikmalarining qo‘llanish darajasi yuqoriligini ko‘rsatdi. Bu holat har ikki tilning tipologik xususiyatlari bilan izohlanadi hamda ekoturizm terminlarini raqamli terminologik bazaga kiritishda struktur modellarni alohida hisobga olish zarurligini ko‘rsatadi.

Tahlil qilingan terminlarni raqamli terminologik bazaga joylashtirishda ularning tili, ekvivalenti va semantik guruhi asosiy tasnif mezonlari sifatida xizmat qiladi. Masalan, 生态旅游 (shēngtài lǚyóu) — ekoturizm, 国家公园 (guójiā gōngyuán) — milliy bog‘, 生物多样性 (shēngwù duōyàngxìng) — biologik xilma-xillik kabi birliklar o‘zaro ekvivalent terminlar sifatida bir bazada jamlanishi mumkin.

Bunday bazada har bir terminning xitoycha shakli, pinyin yozuvi, o‘zbekcha muqobili va semantik guruhi alohida qayd etiladi. Natijada terminlarni qidirish, saralash hamda xitoy va o‘zbek tillaridagi ekoturizm terminlarini chog‘ishtirma o‘rganish imkoniyati kengayadi.

Xulosa. Tadqiqot natijalari xitoy va o‘zbek tillaridagi ekoturizm terminlarining struktur jihatdan muayyan farqlarga ega ekanligini ko‘rsatdi. Xitoy tilida terminlar asosan ikki va undan ortiq morfemaning birikishi orqali hosil bo‘lsa, o‘zbek tilida termin birikmalari faol qo‘llanadi. Chog‘ishtirma tahlil natijasida har ikki tilda ekoturizm sohasining asosiy tushunchalarini ifodalovchi terminlar mavjudligi, biroq ularning struktur ifodalanish usullari turlicha ekanligi aniqlandi. Shuningdek, terminlarni xitoycha shakli, pinyin transkripsiyasi va o‘zbekcha ekvivalentlari asosida tizimlashtirish raqamli terminologik bazani shakllantirish uchun muhim asos bo‘lib xizmat qiladi. Olingan natijalar ekoturizm terminologiyasini chog‘ishtirma o‘rganish hamda elektron terminologik resurslarni yaratishda qo‘llanishi mumkin.

FOYDALANILGAN ADABIYOTLAR

1. Buckley, R., & Ollenburg, C. (2011). Trends and lags in the use of ecotourism terminology in the scholarly and popular press. Tourism Recreation Research, 36(3), 307–309. https://doi.org/10.1080/02508281.2011.11081677
2. Choi, H. C. (2016). Sustainable Tourism. In Sustainable Tourism (pp. 916–917). https://doi.org/10.1007/978-3-319-01384-8_193
3. Weber, A. S. (2024). Disambiguation of the terminology of sustainable tourism and ecotourism. In Lecture notes in networks and systems (pp. 465–471). https://doi.org/10.1007/978-3-031-74828-8_40
4. Hasanova G. Turizm terminologiyasining tarjimasi: O‘zbek va rus tillari misolida. – Toshkent: O‘zbekiston Milliy Universiteti, 2019. – 54 – 56 b.
5. Turg‘unov R. Turizm va ekologiya: Ekoturizm atamalari va alarming rivojlanishi. – Samarqand: Samarqand davlat universiteti. – 2017. – 68 b.
6. 李艳. 中国生态旅游术语的发展：语言学视角. – 北京: 北京语言文化大学出版，2003. – 177页.

O‘ZBEK TILIDA SENTIMENT TAHLILNING NAZARIY VA LINGVISTIK MASALALARI

Jumagul Sulaymonova
Kompyuter lingvistikasi mutaxassisligi 2-kurs magistranti
O ‘zbekiston Milliy universiteti
https://doi.org/10.5281/zenodo.23253834

Annotatsiya. *Ushbu maqolada tabiiy tilni qayta ishlashning muhim yo‘nalishlaridan biri bo‘lgan sentiment tahlilni o‘zbek tilida amalga oshirishning nazariy, lingvistik va amaliy muammolari tadqiq etiladi. O‘zbek tilining agglyutinativ tabiati, morfologik shakllarning serqirraligi, inkor konstruksiyalari mos keladi: ular qoidaga asoslangan yondashuvda keng leksik bazaning zarurligini ko‘rsatadi, barqaror emotsional birikmalar hamda milliy-madaniy nutq xususiyatlarining sentiment aniqlash tizimlariga ta’siri ko‘rib chiqilgan. Shuningdek, o‘zbek tilida sentiment tahlil uchun sifatli dataset yaratish, annotatsiya mezonlarini belgilash va leksik hamda morfologik resurslardan foydalanishning ahamiyati yoritilgan.*

Kalit so‘zlar: *sentiment tahlil, tabiiy tilni qayta ishlash, o‘zbek tili, agglyutinativ til, morfologik tahlil, dataset, kompyuter lingvistikasi.*

Abstract. *In this article, the theoretical, linguistic and practical problems of sentiment analysis in Uzbek, which is one of the important areas of natural language processing, are investigated. The effect of the agglutinative nature of the Uzbek language, the variety of morphological forms, negative constructions, stable emotional compounds, and national-cultural speech features on sentiment detection systems is considered. Also, the importance of creating a quality dataset for sentiment analysis in Uzbek, defining annotation criteria, and using lexical and morphological resources is highlighted.*

Keywords: *sentiment analysis, natural language processing, Uzbek language, agglutinative language, morphological analysis, dataset, computational linguistics.*

Аннотация. *В статье рассматриваются теоретические, лингвистические и практические аспекты реализации анализа тональности (sentiment analysis) на узбекском языке как одного из ключевых направлений обработки естественного языка. Особое внимание уделяется влиянию агглютинативной природы узбекского языка, многообразия морфологических форм, отрицательных конструкций, устойчивых эмоционально-оценочных сочетаний, а также национально-культурных особенностей речевой коммуникации на эффективность систем автоматического определения тональности. Кроме того, обосновывается значимость формирования качественных датасетов для узбекского языка, разработки критериев аннотирования и использования лексических и морфологических ресурсов в задачах анализа тональности.*

Ключевые слова: *анализ тональности, обработка естественного языка, узбекский язык, агглютинативный язык, морфологический анализ, датасет, компьютерная лингвистика.*

Kirish. Sentiment tahlil matnda ifodalangan hissiy munosabat, baho va kayfiyatni avtomatik aniqlashga qaratilgan yo‘nalish bo‘lib, bugungi kunda tabiiy tilni qayta ishlashning eng amaliy sohalaridan biri hisoblanadi. U ijtimoiy tarmoqlar monitoringi, mijozlar fikrini

o‘rganish, siyosiy kommunikatsiya, ta’lim, media tahlili va tavsiya tizimlarida keng qo‘llanmoqda. Yuqori resursli tillar, ayniqsa ingliz tilida, bu yo‘nalish uchun katta hajmli korpuslar, annotatsiyalangan datasetlar va tayyor modellar mavjud. Biroq o‘zbek tili kabi kam resursli tillarda sentiment tahlilning samaradorligi ko‘p jihatdan lingvistik resurslarning cheklanganligi bilan belgilanadi[1].

O‘zbek tilida sentiment tahlilni yaratishning murakkabligi faqat ma’lumot yetishmasligi bilan cheklanmaydi. Muammo tilning agglyutinativ tuzilishi, so‘z shakllarining ko‘pligi, inkorning turli ko‘rinishlari, kuchaytirgich va yumshatgich birliklar, idiomatik ifodalar hamda milliy nutq etiketiga xos emotsional konstruksiyalarning avtomatik qayta ishlanishida namoyon bo‘ladi[6]. Shu bois sentiment tahlilni o‘zbek tilida rivojlantirish faqat algoritmik emas, balki chuqur lingvistik yondashuvni ham talab qiladi.

Adabiyotlar tahlili. Sentiment tahlilning asosiy vazifasi matndagi ijobiy, salbiy yoki neytral munosabatni aniqlashdan iborat bo‘lsa-da, zamonaviy tadqiqotlar ushbu jarayonni oddiy “musbat–manfiy” oppozitsiyasi bilan chegaralab bo‘lmasligini ko‘rsatadi. Xususan, Bo Pang va Lillian Lee sentiment tahlilni kontekstga bog‘liq semantik va pragmatik hodisa sifatida talqin qilib, hissiy baholash ko‘pincha so‘z darajasidan yuqoriroq — birikma, gap hamda diskurs darajasida shakllanishini ta’kidlaydi. Shu sababli hissiy mazmunni aniqlash jarayonida kontekstual omillar, ironiyaviy ifodalar (gapning yuzaki (to‘g‘ridan-to‘g‘ri) ma’nosi bilan asl ma’nosi bir-biriga zid bo‘lgan holat) va yashirin semantik signallarni hisobga olish muhim hisoblanadi.

Bundan tashqari, B.Liu sentiment tahlilni “fikrlarni qazib olish” (opinion mining) doirasida ko‘rib, unda subyektivlikni aniqlash, baholovchi birliklarni ajratish va ularning yo‘nalishini belgilash asosiy bosqichlar ekanini qayd etadi. Bu yondashuv sentimentni faqatgina leksik birliklarga tayanib emas, balki kengroq semantik tizim doirasida o‘rganish zarurligini ko‘rsatadi.

Kam resursli tillar kontekstida esa muammo yanada murakkablashadi. Emily M.Bender va B.Friedman ma’lumotlar to‘plamining tavsifi, annotatsiya mezonlarining aniqligi va tushunarli ifodalanganligi model natijalarining ishonchliligiga bevosita ta’sir qilishini alohida ta’kidlaydi. Ularning fikricha, noto‘g‘ri yoki noaniq annotatsiya nafaqat model aniqligini pasaytiradi, balki natijalarning noto‘g‘ri talqin qilinishiga ham olib keladi.

O‘zbek tili misolida ham ushbu muammolar yaqqol kuzatiladi. E.Kuriyozov va hammualliflar tomonidan olib borilgan tadqiqotlarda kam resursli tillarda asosiy muammo sifatida belgilangan va muvozanatli korpuslarning yetishmasligi ko‘rsatiladi. Ular tomonidan o‘zbek tili uchun qo‘lda annotatsiyalangan o‘rta hajmli dataset hamda avtomatik tarjima asosida kengaytirilgan ma’lumotlar to‘plami taklif etilgan bo‘lib, bu ish milliy til uchun sentiment tahlilning dastlabki ilmiy poydevorini shakllantirishda muhim ahamiyat kasb etadi.

Shuningdek, zamonaviy yondashuvlarda neyron tarmoqlarga asoslangan modellar, xususan, J.Devlin tomonidan taklif etilgan BERT kabi transformator modellar kontekstni chuqurroq anglash imkonini berib, sentiment tahlil sifatini sezilarli darajada oshirgan. Biroq bunday modellarni kam resursli tillarga moslashtirish uchun yetarli hajmdagi sifatli annotatsiyalangan ma’lumotlar zarurligi dolzarb masala bo‘lib qolmoqda.

Shunday qilib, o‘zbek tilida sentiment tahlilning nazariy va amaliy asoslari uch muhim tayanchga tayanadi: birinchidan, til birliklarining hissiy yukini lingvistik va pragmatik jihatdan to‘g‘ri talqin qilish; ikkinchidan, sifatli annotatsiyalangan ma’lumotlar bazasini yaratish; uchinchidan esa, ushbu lingvistik bilimlarni zamonaviy hisoblash modellariga mos shaklda

integratsiya qilish zaruratidir. Bu uch omilning uyg‘unligi milliy til uchun samarali sentiment tahlil tizimlarini ishlab chiqishda hal qiluvchi ahamiyatga ega.

Shu bilan birga, so‘nggi yillarda o‘zbek tilidagi matnlarni avtomatik qayta ishlash, jumladan, emotsional va baholovchi leksikani tahlil qilishda sentiment xususiyatlarini tortish hamda matn elementlarini sinfga ajratish modellari ham takomillashtirilmoqda. [11]

Tahlil va natijalar. O‘zbek tilida sentiment tahlilning o‘ziga xosligi *agglyutinativ* tabiatidir. O‘zakka ketma-ket qo‘shiladigan affikslar natijasida bitta leksemaning ko‘plab morfologik shakllari yuzaga keladi. Bu esa so‘z shakllarining sonini keskin oshiradi va model uchun ma’lumot fregmentatsiyasi kuchaytiradi. Shu sababli stemmizatsiya, lemmatizatsiya va morfologik teglash sentiment tahlilning ajralmas bosqichiga aylanadi. U.Salaevning 2024-yildagi UzMorphAnalyser modeli aynan o‘zbek tilidagi affiksal tizimni, lemmatizatsiya va morfologik belgilashni hisobga olgan holda ishlab chiqilgan va 5,3 ming so‘zli sinov to‘plamida 91 foizdan yuqori so‘z-darajali aniqlik ko‘rsatgan[6].

N.Abduraxmonova va boshqalar muallifligida MorphUz ishlanmasi ham o‘zbek tilida morfologik analizator yaratish yo‘nalishidagi muhim qadamlardan biri sifatida tilga olinadi[2]. Bu kabi vositalar sentiment tahlilda bir xil o‘zakka mansub, lekin turli grammatik shakllarda kelgan birliklarni umumlashtirish imkonini beradi. Aks holda model “yaxshi”, “yaxshimas”, “yaxshiroq”, “yaxshilandi” kabi shakllarni alohida birlik sifatida ko‘rib, emotsional yaqinlikni to‘liq anglay olmasligi mumkin.

Masalan, “tartib”, “tartibli”, “tartibsiz”, “tartibsizlik” birliklari formal jihatdan bir ildizga bog‘liq bo‘lsa-da, ularning bahoviy yo‘nalishi turlicha. Demak, sentiment tahlilda morfologik tahlil faqat so‘zni qisqartirish emas, balki hosila shaklning yangi semantik-bahoviy funksiyasini aniqlash vositasi hamdir. Shu jihatdan o‘zbek tilida sentiment modellarini qurishda morfologik analizatorni integratsiya qilish ixtiyoriy emas, balki zarur bosqich hisoblanadi. [12]

Sentiment tahlilda *emotsional ma’no* doimo alohida so‘z darajasida yuzaga chiqmaydi. Ko‘p hollarda u butun birikma, frazeologik tuzilma yoki nutq etiketi orqali namoyon bo‘ladi. S.Allanazarova va D.Elovaning o‘zbek matnlari sentiment bazasini yaratish haqidagi ishida hissiy rang beruvchi so‘zlar, sharhlar va ijtimoiy tarmoq materiallari bilan ishlashda kontekstning ahamiyati qayd etiladi[4].

O‘zbek tilida duo, olqish, tabrik, maqtov, koyish, kinoya va yumshoq rad etish kabi pragmatik shakllar juda faol. Masalan, “baraka toping”, “rahmat, zo‘r bo‘libdi”, “hech ham yomon emas”, “gap yo‘q”, “shunaqasi ham bo‘ladimi” kabi birliklar yuzaki shaklda turlicha ko‘rinsa-da, ularning haqiqiy ma’nosi kontekst va madaniy kod orqali belgilanadi. Ayniqsa “hech ham yomon emas” kabi konstruktsiyalar formal jihatdan inkor elementini o‘z ichiga olgan bo‘lsa-da, real kommunikatsiyada ijobiy baho ifodalashi mumkin. Shuning uchun o‘zbek tilidagi sentiment tahlil sof leksik yondashuv bilan cheklanib qolsa, ko‘plab holatlarda noto‘g‘ri xulosa chiqaradi. Bu xulosa B.Mengliev va hammualliflarning leksikografik resurslarga tayangan sentiment tasnifi ishiga ham

Bundan tashqari, inkor o‘zbek tilida nafaqat “emas” va “yo‘q” yordamida, balki “-ma/-me” affiksi, ba’zan esa kontekstual cheklash vositalari orqali ifodalanadi. Sentiment tizimi bu holatlarni ajrata olmasa, masalan, “yoqmadi”, “yoqtirmadim”, “omon qolmadi”, “zo‘r emas”, “unchalik yaxshi emas” kabi tuzilmalarni noto‘g‘ri tasniflashi mumkin. Shu sababli inkor va modallikni sentiment tahlil modelining alohida lingvistik moduli sifatida ko‘rish maqsadga muvofiq.

1-jadval. O'zbek tilida sentiment tahlilga ta'sir qiluvchi lingvistik omillar

Lingvistik omil	Tavsifi	Muammo	Sentimentga ta'siri
Agglyutinativlik	So'zga ketma-ket qo'shimchalar qo'shilishi	So'z shakllari soni juda ko'p	Model umumlashtira olmaydi
Affiksatsiya	Qo'shimchalar orqali ma'no o'zgarishi	Bir o'zak turli sentimentga ega bo'ladi	Noto'g'ri klassifikatsiya
Inkor konstruksiyasi	"emas", "yo'q", "-ma" orqali inkor	Yashirin yoki murakkab inkor	Tonallik teskarilanadi
Kuchaytirgichlar	"juda", "nihoyatda", "rosa"	Intensivlik hisobga olinmaydi	Sentiment kuchi noto'g'ri baholanadi
Pragmatik birliklar	Duo, olqish, kinoya	Formal va real ma'no farqi	Model chalkashadi
Frazeologizmlar	Barqaror iboralar	Literal tarjima ishlamaydi	Noto'g'ri sentiment chiqadi
Kontekst	Gap darajasidagi ma'no	So'z darajasida yetarli emas	Global sentiment noto'g'ri aniqlanadi

O'zbek tilida sentiment tahlilni rivojlantirishning asosiy amaliy sharti – sifatli dataset (ma'lumotlar bazasi) yaratishdir. Dataset tarkibida qo'lda annotatsiyalangan korpusdan foydalanish, zarur hollarda esa avtomatik tarjima asosida ma'lumotlarni kengaytirish mumkin. Bunday yondashuv kam resursli tillarda dastlab kichik, lekin ishonchli dataset yaratib, keyingi bosqichlarda kattaroq modellarni o'qitish uchun mustahkam poydevor yaratishga xizmat qiladi.

Biroq annotatsiya jarayonida bir nechta metodologik talablar saqlanishi kerak. Birinchidan, ijobiy, salbiy va neytral sinflar muvozanatli bo'lishi lozim. Ikkinchidan, inkor, kinoya, kuchaytirgich, frazeologizm va emotsional nutq etiketi alohida qayd etilishi kerak. Uchinchidan, matn manbasi, yozilish uslubi, mavzu va auditoriya ko'rsatilishi kerak. Aynan mana shu prinsiplarga E.M.Bender va B.Friedmanning data statements yondashuvi ham urg'u beradi[3].

Shu nuqtai nazardan, dastlabki bosqichda 500–1000 gapdan iborat, qo'lda tekshirilgan, muvozanatli va izohli mikro-dataset yaratish eng samarali yo'l bo'lishi mumkin. [13] Unda kamida quyidagi qatlamlar belgilansa foydali bo'ladi: umumiy tonallik, inkor mavjudligi, kuchaytirgich yoki yumshatgich birlik, emotsional leksema turi, pragmatik formula mavjudligi. Ana shundan keyingina kattaroq korpuslar va neyron modellar bilan ishlash ilmiy jihatdan barqaror natija beradi. Bu yondashuv S.Allanazarova, D.Elova va B.Menglievlar ishlaridagi leksik va ma'lumotlar bazasiga tayangan amaliy yo'nalishlar bilan hamohangdir[4].

Bugungi kunda o'zbek tilida sentiment tahlilni rivojlantirish uchun bir nechta parallel yo'nalish zarur. Birinchi yo'nalish – ochiq va standartlashtirilgan datasetlar yaratish. Ikkinchisi – morfologik analizator, lemmatizator va tokenizatorlarni barqarorlashtirish. Uchinchi yo'nalish – o'zbek tiliga xos emotsional lug'at va pragmatik formulalar bazasini ishlab chiqish. To'rtinchisi esa qoidaga asoslangan hamda neyron yondashuvlarni birlashtirgan gibrid modellarni yaratishdir. B.Mengliev va hammualliflarning qoidaviy-leksik yondashuvi hamda Salaevning morfologik modeli bunday integratsiya uchun real tayanch bo'la oladi[5].

2-jadval. O'zbek tilida sentiment dataset yaratish talablari

Mezoni	Tavsifi	Tavsiya etilgan yondashuv
Dataset hajmi	Boshlang'ich o'lcham	500–1000 gap (manual)
Klass balansi	Ijobiy/salbiy/neytral	1:1:1 nisbat
Annotatsiya turi	Label berish	Qo'lda (lingvist tomonidan)
Morfologik belgilash	Lemma, POS	MorphUz yoki rule-based
Inkor belgisi	Negation flag	Alohida ustun
Intensivlik	Kuch darajasi	weak / medium / strong
Pragmatik belgilar	Duo, kinoya, yumshatish	Binary (0/1)
Ma'lumot manbasi	Matn kelib chiqishi	Real (ijtimoiy tarmoq, sharhlar)
Dataset formati	Tuzilishi	CSV / Excel

Kelgusida sentiment tahlilni faqat uch sinfli klassifikatsiya bilan cheklamasdan, emotsiya turi, kuchi, nishoni va aspektiga ko'ra ham rivojlantirish mumkin. Biroq bunday murakkab bosqichga o'tishdan oldin o'zbek tilida lingvistik jihatdan puxta tayyorlangan bazaviy korpus va annotatsiya standarti shakllanishi shart.

Xulosa. Xulosa qilib aytganda, o'zbek tilida sentiment tahlilni yo'lga qo'yish sof texnik masala emas, balki lingvistika, ma'lumotlar muhandisligi va sun'iy intellekt kesishgan murakkab fanlararo vazifadir. O'zbek tilining agglyutinativ xususiyati, inkor tizimi, hosila shakllarning bahoviy yuklamasi, pragmatik formulalar va milliy-madaniy ifodalar sentiment tahlilni yuqori resursli tillarga qaraganda murakkablashtiradi. Shu bois bu yo'nalishda muvaffaqiyatga erishish uchun morfologik tahlil vositalari, leksikografik resurslar, sifatli annotatsiyalangan datasetlar va aniq metodologiya birgalikda rivojlantirilishi kerak.

O'zbek tilida sentiment tahlilning istiqbollari mavjud, biroq barqaror va ishonchli natijalarga erishish uchun, avvalo, kichik hajmli, sifatli hamda lingvistik jihatdan puxta tayyorlangan korpuslarni shakllantirish muhimdir. Ushbu tadqiqot doirasida sentiment tahlil uchun tanlangan ma'lumotlar bazasi qo'lda tahlil qilinib, matnlar sentiment yo'nalishiga ko'ra ijobiy, salbiy va neytral toifalarga ajratildi. Mazkur annotatsiyalangan ma'lumotlar bazasi kelgusida sentiment tahlil modelini yaratish va uning samaradorligini baholash uchun asos bo'lib xizmat qiladi. Tadqiqotning keyingi bosqichida muayyan domenni belgilash hamda ushbu domen doirasidagi matnlarni avtomatik ravishda ijobiy, salbiy va neytral sentimentlarga ajratib tahlil qila oladigan lingvistik platformani yaratish rejalashtirilgan. Bunday bosqichma-bosqich yondashuv o'zbek tilining lingvistik xususiyatlarini hisobga olgan holda samarali sentiment tahlil tizimlarini shakllantirish, shuningdek, kelgusida yirik modellar va amaliy dasturiy vositalarni ishlab chiqish uchun mustahkam ilmiy-amaliy asos yaratadi.

FOYDALANILGAN ADABIYOTLAR

1. E. Kuriyozov, S. Matlatipov, M. A. Alonso, and C. Gómez-Rodríguez, "Construction and Evaluation of Sentiment Datasets for Low-Resource Languages: The Case of Uzbek," in Human Language Technology. Challenges for Computer Science and Linguistics, LTC 2019, LNCS 13212, Springer, 2022.
2. N. Abdurakhmonova, A. Ismailov, and R. Sayfullayeva, "MorphUz: Morphological Analyzer for the Uzbek Language," 2022. Bibliographic details should be finalized from the original publication copy.

3. E. M. Bender and B. Friedman, "Data Statements for Natural Language Processing: Toward Mitigating System Bias and Better Science," Transactions of the Association for Computational Linguistics, vol. 6, pp. 587–604, 2018.
4. S. Allanazarova and D. Elova, "Uzbek Texts Sentiment Analysis: Database Development," in Proceedings of PAMIR-2 2023, pp. 744–747, 2024.
5. D. B. Mengliev, E. Y. Akhmedov, V. B. Barakhnin, Z. A. Hakimov, and O. M. Alloyorov, "Utilizing Lexicographic Resources for Sentiment Classification in Uzbek Language," in Proc. APEIE 2023, pp. 1720–1724, 2023.
6. U. Salaev, "UzMorphAnalyser: A Morphological Analysis Model for the Uzbek Language Using Inflectional Endings," arXiv:2405.14179, 2024.
7. S. G. Matlatipov, J. Rajabov, E. Kuriyozov, and M. Aripov, "UzABSA: Aspect-Based Sentiment Analysis for the Uzbek Language," in Proceedings of the 3rd Annual Meeting of the ELRA/ISCA SIG on Under-Resourced Languages (SIGUL 2024), 2024.
8. N. Abdurakhmonova, R. Shirinova, R. Sayfullayeva, D. Mengliev, et al., "An Annotated Morphological Dataset for Uzbek Word Forms: Towards Rule-Based and Machine Learning Approaches," Data in Brief, vol. 61, Art. 111702, 2025.
9. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT, 2019, pp. 4171–4186.
10. Liu, B. Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, 2012.
11. Mengliev, D., Abdurakhmonova, N., Barakhnin, V., Vasliddinova, K., Rahimov, H., & Djalolova, K. (2024, June). Enhancing sentiment analysis in Uzbek language texts through weighted lexical features. In *2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM)* (pp. 2450-2453). IEEE.
12. Liu, S., Maimaiti, M., Yolwas, N., Abdurakhmonova, N., Le, W., Xie, Z., ... & Silamu, W. (2026). Dependency-Aware Graph Transformer with Hierarchical Syntax Encoding for Aspect-Level Sentiment Classification. *Expert Systems with Applications*, 132574.
13. Mengliev, D., Abdurakhmonova, N., Shirinova, R., Ibragimov, B., Khudayberganova, D., & Suyunova, M. (2025, October). A comparative study of annotation schemes for named entity recognition in Uzbek legal texts. In *AIP Conference Proceedings* (Vol. 3377, No. 1, p. 070005). AIP Publishing LLC.

ДИАГНОСТИКА КАК ИНСТРУМЕНТ ФОРМИРУЮЩЕГО ЭКСПЕРИМЕНТА: АВТОРСКАЯ МЕТОДИКА И ПЛАТФОРМА «ДИАГКОМП-РУС»

Султонова Сайёра Холмирзаевна
старший преподаватель Ташкентский институт инженеров ирригации и механизации сельского хозяйства Национальный исследовательский университет кафедра языков
г. Ташкент, Республика Узбекистан
ORCID: 0009-0001-0530-7437
https://doi.org/10.5281/zenodo.23253853

Аннотация. *Статья посвящена описанию и обоснованию основного формирующего эксперимента диссертационного исследования, направленного на совершенствование методики формирования цифровой компетентности будущих учителей русского языка в педагогических вузах Узбекистана. Представлена авторская четырёхкомпонентная модель методики, реализуемая через курс «Цифровые технологии в обучении русскому языку» (72 часа). Описан диагностический инструментарий эксперимента, включающий анкету самооценки и тест цифровой компетентности. Особое внимание уделяется авторской веб-платформе «ДиагКомп-Рус», разработанной специально для онлайн-проведения диагностики на базе четырёх вузов Узбекистана. Обосновывается методологическая логика сопоставления предтеста и посттеста как основного механизма верификации эффективности разработанной методики.*

Ключевые слова: *цифровая компетентность, учитель русского языка, формирующий эксперимент, LMS-платформы, искусственный интеллект, диагностическая веб-платформа, методика обучения, Узбекистан.*

Abstract*: The article describes and substantiates the main formative experiment within a doctoral research project aimed at improving the methodology for developing digital competence among future Russian language teachers at pedagogical universities in Uzbekistan. The author's four-component instructional model, implemented through a 72-hour course "Digital Technologies in Russian Language Teaching", is presented. The diagnostic toolkit of the experiment — comprising a self-assessment questionnaire and a digital competence test — is described in detail. Special attention is given to the original web platform "DiagComp-Rus", developed specifically for online diagnostics across four Uzbek universities. The methodological rationale for comparing pre-test and post-test results as the primary mechanism for verifying the effectiveness of the proposed methodology is substantiated.*

Keywords: *digital competence, Russian language teacher, formative experiment, LMS platforms, artificial intelligence, diagnostic web platform, teaching methodology, Uzbekistan*

Annotatsiya. *Ushbu maqolada O‘zbekistondagi pedagogika universitetlarida bo‘lajak rus tili o‘qituvchilari orasida raqamli kompetentsiyani rivojlantirish metodologiyasini takomillashtirishga qaratilgan dissertatsiya tadqiqot loyihasining asosiy shakllantiruvchi eksperimenti tasvirlangan va asoslangan. 72 soatlik "Rus tilini o‘qitishda raqamli texnologiyalar" kursi orqali amalga oshirilgan metodologiyaning to‘rt komponentli modeli taqdim etilgan. Eksperiment uchun diagnostika vositalari, jumladan, o‘zini baholash so‘rovnomasi va raqamli kompetentsiya testi tasvirlangan. O‘zbekistondagi to‘rtta universitetda*

onlayn diagnostika uchun maxsus ishlab chiqilgan "DiagComp-Rus" xususiy veb-platformasiga alohida e'tibor qaratilgan. Ishlab chiqilgan metodologiyaning samaradorligini tekshirishning asosiy mexanizmi sifatida testdan oldingi va testdan keyingi natijalarni taqqoslashning metodologik mantig'i asoslanadi.

Kalit so'zlar: *raqamli kompetentsiya, rus tili o'qituvchisi, shakllantiruvchi eksperiment, LMS platformalari, sun'iy intellekt, diagnostika veb-platformasi, o'qitish usullari, O'zbekiston.*

1. ВВЕДЕНИЕ

Цифровизация педагогического образования в Республике Узбекистан, закреплённая в Стратегии «Цифровой Узбекистан — 2030» и Указе Президента РУз № УП-60 (2022), выдвигает принципиально новые требования к профессиональной компетентности будущих учителей. Особую значимость данная тенденция приобретает применительно к подготовке учителей русского языка: специалисты данного профиля должны не только владеть предметным знанием и традиционными методическими подходами, но и системно интегрировать цифровые технологии — LMS-платформы и инструменты искусственного интеллекта — в свою профессиональную деятельность.

Вместе с тем констатирующий эксперимент автора (n = 285, октябрь 2025 — февраль 2026, 4 вуза Узбекистана) зафиксировал тревожную картину: лишь 18% будущих учителей русского языка продемонстрировали высокий уровень цифровой компетентности, тогда как 47% показали средний уровень и 35% — низкий (Султонова, 2026). При этом 68% студентов испытывают затруднения при создании интерактивных цифровых заданий, лишь 24% регулярно применяют ИИ-инструменты в учебных целях, а 71% указывают на нехватку практических занятий по цифровой методике.

Перечисленные факты определяют как практическую актуальность, так и методологическую логику основного формирующего эксперимента, описываемого в настоящей статье.

2. АВТОРСКАЯ МОДЕЛЬ МЕТОДИКИ: ТЕОРЕТИЧЕСКОЕ ОСНОВАНИЕ

Разработанная модель методики формирования цифровой компетентности будущих учителей русского языка опирается на взаимодополняющую теоретическую рамку, включающую три концепции.

Модель TPACK (Koehler & Mishra, 2009) постулирует, что профессиональная компетентность учителя определяется не суммой, а интеграцией предметного (CK), педагогического (PK) и технологического (TK) знания. Применительно к учителю русского языка это означает: недостаточно уметь пользоваться LMS — необходимо понимать, каким образом конкретный инструмент трансформирует методику преподавания конкретного языкового явления.

Концепция смешанного обучения (Garrison & Vaughan, 2008) обеспечивает организационный принцип курса: 36 часов аудиторного взаимодействия и 36 часов онлайн-работы в LMS. Модель трёх присутствий — когнитивного, социального и учебного — определяет требования к педагогическому дизайну каждого модуля.

Конструктивистская парадигма (Vygotsky, 1978; Jonassen, 1994) обосновывает ключевой принцип методики: компетентность не передаётся в готовом виде, а строится студентом через активную деятельность по созданию цифровых образовательных продуктов.

На пересечении данных концепций обоснован авторский принцип критической ИИ-грамотности — обязательное формирование у студентов способности верифицировать ИИ-генерированный контент, распознавать фактологические ошибки нейросетей и сохранять авторскую ответственность за итоговый текст. Данный принцип не представлен в существующих зарубежных и региональных моделях цифровой педагогики и составляет методологическую новизну настоящего исследования.

3. СОДЕРЖАНИЕ ОСНОВНОГО ФОРМИРУЮЩЕГО ЭКСПЕРИМЕНТА

3.1. Дизайн и выборка

Основной формирующий эксперимент проводится на базе четырёх педагогических вузов Узбекистана: Узбекского государственного университета мировых языков (УзГУМЯ), Ташкентского государственного педагогического университета имени Низами, Бухарского государственного университета (БухГУ), Ферганского государственного университета (ФерГУ), Национального университета Узбекистана им. М.Улугбека. Период проведения: сентябрь 2026 — февраль 2027 учебного года. Выборка: 285 студентов 3–4 курсов и магистрантов направления «Русская филология» / «Русский язык и литература». Дизайн эксперимента — квазиэкспериментальный, схема «предтест → формирующее воздействие → посттест» — позволяет верифицировать изменения компетентности, непосредственно обусловленные авторской методикой, контролируя фоновые эффекты.

3.2. Структура авторского курса

Экспериментальное воздействие реализуется через авторский курс «Цифровые технологии в обучении русскому языку» объёмом 72 академических часа, структурированный в четыре последовательных модуля с нарастающей сложностью:

Модуль 1 (14 ч.) — «Введение в цифровую экосистему обучения»: концептуальное освоение трёхуровневой модели цифровой экосистемы (инфраструктурный, интеллектуальный, коммуникативный уровни), нормативно-правовые и этические основания применения ИИ.

Модуль 2 (24 ч.) — «Проектирование цифрового курса по русскому языку»: практическое создание учебного контента в Moodle и Google Classroom, разработка аналитических рубрик оценивания.

Модуль 3 (20 ч.) — «Инструменты искусственного интеллекта в методической и письменной деятельности»: поэтапная работа с ChatGPT, Grammarly AI, DeepL Write с обязательным протоколом рефлексивного документирования каждого взаимодействия с нейросетью.

Модуль 4 (14 ч.) — «Управление учебным процессом и профессиональная рефлексия»: образовательная аналитика в LMS, организация смешанного обучения, рефлексия профессионального опыта.

Ключевым методическим решением Модуля 3 выступает протокол рефлексивного документирования: студент фиксирует в LMS-дневнике каждое обращение к нейросети — введённый промпт, полученный ответ и развёрнутое обоснование того, как ответ был принят, изменён или отклонён. Данная практика, эмпирически апробированная на пилотном этапе (Sarbon University, n = 78; η^2 = 0,38), продемонстрировала высокую эффективность в профилактике пассивной технологической зависимости студентов.

4. ДИАГНОСТИЧЕСКИЙ ИНСТРУМЕНТАРИЙ И ПЛАТФОРМА «ДИАГКОМП-РУС»

4.1. Диагностический инструментарий

Измерение цифровой компетентности осуществляется с помощью авторского двухинструментального комплекса, разработанного в соответствии с четырёхкомпонентной структурой целевой модели (когнитивный, деятельностный, коммуникативный, рефлексивно-критический компоненты).

Инструмент 1 — **Анкета самооценки** (пять блоков):

Блок I — общие сведения (вуз, курс, направление);

Блок II — матрица самооценки по шкале Лайкерта (6 утверждений × 5 градаций);

Блок III — частота использования LMS, интерактивных платформ и ИИ-инструментов;

Блок IV — затруднения (множественный выбор из пяти вариантов);

Блок V — открытый вопрос.

Инструмент 2 — **Тест цифровой компетентности** (максимум 100 баллов, три части):

Часть А — 15 вопросов закрытого типа (20 баллов, автоматическая проверка);

Часть Б — кейс-задание с развёрнутым ответом (30 баллов, проверка экспертом);

Часть В — практическое задание по созданию фрагмента цифрового образовательного продукта (50 баллов, проверка экспертом).

Итоговая шкала: высокий уровень — 80–100 баллов; средний — 50–79 баллов; низкий — 0–49 баллов.

4.2. Авторская диагностическая платформа «ДиагКомп-Рус»

Для проведения онлайн-диагностики в масштабах четырёх вузов автором совместно с разработчиком создана специализированная веб-платформа «ДиагКомп-Рус» (URL: https://diagnostika-sayti.vercel.app/). Платформа реализует полный диагностический цикл без необходимости установки какого-либо программного обеспечения.

С точки зрения архитектуры платформа включает две функциональные зоны. Клиентская зона (доступна студентам по прямой ссылке без регистрации) реализует пятишаговый интерфейс анкеты и трёхчастный интерфейс теста с прогресс-индикатором, автосохранением ответов (на случай разрыва соединения) и финальным экраном благодарности, намеренно не показывающим баллы во избежание соревновательного эффекта, способного исказить честность ответов на предтесте.

Административная зона (защищённая паролем) предоставляет исследователю: фильтрацию ответов по вузу, курсу и волне исследования; форму ручного выставления баллов за Части Б и В с автоматическим пересчётом итогового балла и уровня компетентности; панель автоматической статистики с интерактивными диаграммами (распределение по уровням, средние баллы по утверждениям Блока II, процентное распределение затруднений, сравнение баллов по вузам); выгрузку данных в формат Excel/CSV.

Принципиально важной методологической функцией платформы является встроенная поддержка волн исследования: предтест (сентябрь 2026) и посттест (февраль 2027) регистрируются как отдельные волны с независимыми датами, что обеспечивает автоматическую попарную сопоставимость данных без риска их смешения и позволяет отслеживать динамику компетентности конкретного студента от первого к второму срезу.

5. МЕТОДОЛОГИЯ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ

Количественный анализ данных будет осуществляться с применением следующих методов: t-критерий Стьюдента для зависимых выборок (сравнение предтеста и посттеста в экспериментальных группах каждого вуза); критерий Манна–Уитни (сравнение баллов между вузами с проверкой нормальности распределения); размер эффекта η^2 (Cohen, 1988) для оценки практической значимости зафиксированных изменений; тематический анализ ответов на открытый вопрос Блока V и развёрнутых ответов Части Б теста. Минимально значимым размером эффекта принят $\eta^2 \geq 0{,}06$ (средний эффект по Cohen), что согласуется со стандартами, применявшимися на пилотном этапе исследования и позволяет обеспечить сопоставимость результатов двух этапов.

6. ЗАКЛЮЧЕНИЕ

Представленный основной формирующий эксперимент представляет собой методологически строгую верификацию авторской четырёхкомпонентной модели методики формирования цифровой компетентности будущих учителей русского языка. Принципиальная новизна описанного подхода состоит в трёх взаимосвязанных элементах: системном объединении LMS-платформ и инструментов искусственного интеллекта в единую педагогическую систему; обязательном протоколе рефлексивного документирования взаимодействия студента с нейросетью как инструменте формирования критической ИИ-грамотности; и авторской веб-платформе «ДиагКомп-Рус», обеспечивающей надёжный, масштабируемый и методологически корректный сбор и анализ диагностических данных в условиях многовузовского эксперимента.

СПИСОК ЛИТЕРАТУРЫ

1. Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.
2. Garrison, D. R., & Vaughan, N. D. (2008). Blended learning in higher education. Jossey-Bass.

   Jonassen, D. H. (1994). Technology as cognitive tools: Learners as designers. IT Forum Paper, (1).
3. Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? Contemporary Issues in Technology and Teacher Education, 9(1), 60–70.
4. Султонова, С. Х. (2026). Современное состояние цифровой компетентности будущих учителей русского языка в педагогических вузах Узбекистана. Безопасность. Управление. Искусственный интеллект, (S3), 141–148.
5. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
6. Указ Президента Республики Узбекистан № УП-60 от 28 января 2022 года «О стратегии развития Нового Узбекистана на 2022–2026 годы».
7. Гафурова С., Султонова С. Х. Языковая ситуация в Узбекистане и ее влияние на процесс освоения русского произношения // Американский журнал филологических наук. — 2025. — Т. 5, № 5. — С. 104–109. — DOI: 10.37547/ajps/Volume05Issue05-28
8. Султонова С. Х. Коммуникативные аспекты в методике преподавания РКИ в школах Узбекистана // Международный журнал конференций серии по образованию и общественным наукам (онлайн). — 2022. — Том. 2, № 3.

9. Султонова С. Х. Русский мир Азии : материалы конференции «Русский мир Азии», Ханой, Социалистическая Республика Вьетнам, 25–27 ноября 2024 г. Северо-Восточный федеральный университет им. М. К. Аммосова. — Якутск : Изд. дом СВФУ, 2024.
10. Султонова С. Х. Русский язык в Узбекистане: вчера и сегодня // Гуманитарный договорт. — 2018. — № 25. — С. 8–10.
11. Султонова С. Х., Насириддинова М. М. Использование электронных технологий в практике преподавания русского языка как иностранного (на примере онлайн-доски) // Новости образования: исследование в XXI веке. — 2023. — Т. 1, № 9. — С. 1346–1352.
12. Саримсакова Н. Сравнительный анализ лингвистических особенностей русского и иностранных языков // Журнал «БГ Пульс». — 2026. — Т. 2, № 5. — С. 812–816.
13. Султанова С. Х. Особенности коммуникативного подхода к работе с текстом на уроках русского языка как иностранного в узбекских группах // Актуальный научный журнал педагогики. — 2023. — Т. 4, № 9. — С. 30–40. — DOI: 10.37547/pedagogics-crjp-04-09-06.
14. Султонова С.Х. Цифровая экосистема обучения: Роль Lms-платформ и искусственного интеллекта в профессиональной подготовке преподавателей русского языка. Материалы конференции "International scientific and practical conference “innovative approaches and pedagogical theory in practice: prospects for improving the quality of education”, may 29-30, 2026. стр 268-273. https://doi.org/10.5281/zenodo.20645233

КУЛЬТУРНАЯ ОБУСЛОВЛЕННОСТЬ РЕЧЕВЫХ АКТОВ В УЗБЕКСКОМ ЯЗЫКЕ: КОРПУСНАЯ ВЕРИФИКАЦИЯ

Суюнова Мохинур Илхом кизи
Аспирант 3-го курса Национального университета Узбекистана
Ташкент, Узбекистан
https://doi.org/10.5281/zenodo.23253888

Аннотация. *Статья посвящена культурной обусловленности речевых актов и проверке на корпусном материале положения о преимущественно косвенном характере директивов в узбекском языке. Материалом послужил корпус узбекской письменной цифровой коммуникации (1 797 текстовых сообщений, 11 912 словоупотреблений); сбалансированная по типу отношений выборка объёмом 271 речевой акт размечена по иллокутивному классу, грамматической форме и девятиуровневой шкале косвенности CCSARP. Установлено, что степень косвенности не является постоянным свойством языка, а представляет собой функцию социальной дистанции: индекс косвенности монотонно возрастает от 0,10 при разовом обращении до 1,00 при обращении к непосредственному руководителю, составляя в среднем 0,39. Доля директивов при этом от типа отношений не зависит ($\chi^2(3) = 2{,}10$; $p = 0{,}552$), то есть дистанция определяет форму побуждения, а не его частотность. Бенефактивная конструкция -(i)b ber- употребляется втрое чаще лексического смягчителя iltimos (27,9 против 9,7 на 10 000 словоупотреблений), что количественно подтверждает морфологический характер смягчения в узбекском языке. Ряд узбекских речевых актов, прежде всего благопожелание (duo), не имеет полных эквивалентов в английском языке.*

Ключевые слова: *лингвокультурология; речевой акт; корпусный анализ; вежливость; косвенность; социальная дистанция; CCSARP; узбекский язык.*

Abstract. *The paper addresses the cultural conditioning of speech acts and tests, against corpus data, the claim that directives in Uzbek are predominantly indirect. The material is a corpus of Uzbek written digital communication (1,797 text messages; 11,912 tokens); a relationship-balanced sample of 271 speech acts was annotated for illocutionary class, grammatical form and degree of indirectness on the nine-point CCSARP scale. Indirectness proves not to be a constant property of the language but a function of social distance: the index rises monotonically from 0.10 in one-off requests to 1.00 in exchanges with an immediate superior, averaging 0.39. The proportion of directives, by contrast, does not vary by relationship type ($\chi^2(3) = 2.10$; $p = 0.552$): distance governs the form of a directive, not its frequency. The benefactive -(i)b ber- is three times as frequent as the lexical mitigator iltimos (27.9 vs 9.7 per 10,000 tokens), which quantitatively confirms the morphological locus of mitigation in Uzbek. Several Uzbek speech acts, above all the benediction (duo), have no full English equivalents.*

Keywords: *linguoculturology; speech act; corpus analysis; politeness; indirectness; social distance; CCSARP; Uzbek language.*

1. Введение

Лингвокультурология, сформировавшаяся на стыке языкознания, этнолингвистики, культурологии и прагмалингвистики, исследует способы выражения культурных ценностей и национального сознания посредством языка. Теория речевых актов, восходящая к работам Дж. Сёрла [15], описывает язык как форму действия, однако рассматривает иллокутивные типы как универсальные, тогда как кросс-культурные исследования [3; 4; 10] показали существенные различия в способах их реализации.

В литературе о речевом этикете узбекского языка устойчиво повторяется положение о том, что директивные акты выражаются здесь преимущественно косвенно, а прямой отказ почти не употребляется. Это положение формулировалось на основании отдельных иллюстративных примеров и количественно не проверялось. Цель настоящей работы — проверить его на размеченном корпусе и установить, является ли косвенность признаком языка или переменной, зависящей от параметров коммуникативной ситуации. Новизна работы состоит в том, что косвенность узбекских директивов впервые описывается как измеряемая функция социальной дистанции.

2. Материал и методы

Эмпирическим материалом послужил корпус узбекской письменной цифровой коммуникации [19, 20], сформированный на основе личной переписки в мессенджере (таблица 1). Медиа-, голосовые и пересланные сообщения из разметки исключены как не являющиеся собственной речевой продукцией участника.

Таблица 1. Характеристика корпуса

Параметр	Значение	Параметр	Значение
Диалогов	8	Узбекоязычная часть, сл.	8 229
Текстовых сообщений	1 797	Участников	9
Общий объём, словоупотреблений	11 912	Период сбора	2024–2026 гг.
Размеченная выборка	271 акт / 1 147 сл.	Страт по типу отношений	4

Этические основания. Участники дали согласие на использование материала в научных целях; имена, номера телефонов и иные идентифицирующие сведения удалены; переписка, содержащая сведения о здоровье и частной жизни третьих лиц, из корпуса исключена.

Единицей анализа принята реплика или её часть, выражающая одну иллокутивную цель. Разметка велась по трём параметрам: иллокутивный класс по Дж. Сёрлу [15] с выделением фатического обращения в отдельный класс; грамматическая форма; степень косвенности директивов действия по девятиуровневой шкале CCSARP [3]. Индекс косвенности — доля актов, реализованных конвенционально косвенными стратегиями (уровни 6–7) и намёками (8–9). Применялись критерий χ^2, коэффициент V Крамера, скорректированные стандартизованные остатки, перестановочная оценка Монте-Карло (20 000 повторений) и точный критерий Фишера; частоты нормализованы на 10 000 словоупотреблений [12].

3. Теоретические основания

Проблема взаимосвязи языка и культуры была поставлена В. фон Гумбольдтом [7], развита Э. Сепиром [14] и Б. Уорфом [17] и оформилась в теорию лингвистической

относительности. А. Вежбицкая [18] предложила изучать эту связь через универсальные семантические примитивы и культурно-специфические концепты. В российской традиции Н.Д. Арутюнова [1] исследовала проявление семантических единиц в культурном мышлении, В.И. Карасик [8] рассматривал язык как концептуальное пространство, Ю.С. Степанов [16] описывал «коды культуры», а Дж. Лакофф и М. Джонсон [9] показали метафорическую природу мышления.

В узбекском языкознании Ш. Сафаров [13] показал, что при реализации языковых единиц решающую роль играют социальный статус говорящего, ситуация общения и культурная среда; Н. Махмудов [11] рассмотрел речевые акты в рамках речевого этикета; Г. Хакимова [5] установила зависимость прямых и косвенных актов от социальной дистанции и статуса участников. Исследования М. Хошимова и Х. Каримова показывают, что такие лингвопрагматические единицы, как «уважение», «благословение» и «гостеприимство», тесно связаны с культурой.

Согласно концепции Э.Т. Холла [6], узбекское общество относится к высококонтекстным культурам, где широко используются косвенная речь и иносказания. Т.В. Ларина [10] связывает неудачи межкультурной коммуникации с различным пониманием социальной дистанции и речевых конвенций, а П. Браун и С. Левинсон [4] описывают смягчение директивов как набор стратегий негативной вежливости, ориентированных на защиту «негативного лица» адресата.

4. Национально-культурная специфика узбекских речевых актов

Многие речевые акты узбекского языка формируются под влиянием ценностей уважения, почитания старших, социальной солидарности и религиозных представлений. Благословение *(duo qilish)*, добрые напутствия и выражение почтения выполняют не только коммуникативную, но и культуросохраняющую функцию (таблица 2).

Таблица 2. Узбекские речевые формулы: оригинал, перевод и иллокутивный тип

Формула	Перевод	Иллокутивный тип	Культурная основа
Xudo yo‘lingni ochsin	*‘Пусть Всевышний откроет тебе дороги’*	экспрессив: благопожелание	религиозная
Umringiz uzoq bo‘lsin	*‘Пусть ваша жизнь будет долгой’*	экспрессив: благопожелание	возрастная иерархия
Sizga qarab o‘rgandik	*‘Мы учились, глядя на Вас’*	экспрессив: выражение уважения	почитание старших
Shuni qilsangiz yaxshi bo‘lardi	*‘Было бы хорошо, если бы Вы это сделали’*	директив: косвенная просьба	минимизация давления
Xudo xohlasa, qilamiz	*‘Ин шаа Аллах, сделаем’*	комиссив с условной модальностью	религиозная
Boshqa safar, Xudo xohlasa, boramiz	*‘В другой раз, если будет угодно Аллаху, сходим’*	косвенный отказ	избегание прямого отказа
Bezovta qilmadikmi?	*‘Надеюсь, мы Вас не беспокоим?’*	экспрессив: превентивное извинение	защита «лица» адресата

Большинство этих актов не имеет прямых эквивалентов в других языках: английские *Bless you* и *May God help you* близки к узбекскому *duo* по религиозному содержанию, но не передают его социальной и духовной нагрузки, что свидетельствует о лакуне на данном участке речевого этикета.

5. Результаты корпусного анализа

Распределение иллокутивных классов по четырём стратам оказалось устойчивым: $\chi^2(15) = 24{,}51$; $p = 0{,}057$; $V = 0{,}174$; перестановочная оценка Монте-Карло даёт $p = 0{,}053$. Анализ скорректированных остатков локализует единственное значимое расхождение — фатическое обращение: у равностатусных партнёров оно встречается существенно чаще ожидаемого ($z = 3{,}85$), а в институциональной иерархии не фиксируется вовсе ($z = -2{,}56$). Речь идёт не о различии репертуара, а об изменении единицы разметки: при равном статусе обращение отправляется отдельным сообщением и образует самостоятельный ход, тогда как в иерархическом общении оно растворяется внутри формулы «Ассалому алайкум + [обращение] + яхшимисиз». Таким образом, статусная дистанция изменяет не только выбор стратегии, но и секвенциальную организацию эпизода.

Доля директивных актов по стратам не различается ($\chi^2(3) = 2{,}10$; $p = 0{,}552$), тогда как их распределение по шкале косвенности различается систематически (таблица 3).

Таблица 3. Степень косвенности директивов действия по шкале CCSARP

Тип отношений	n	Прямые (1–5)	Конв. косвенные (6–7)	Намёк (8–9)	Индекс косвенности
Разовое обращение	10	9 (90 %)	1 (10 %)	0	0,10
Равный коллега	7	5 (71 %)	2 (29 %)	0	0,29
Иерархия (институт)	16	8 (50 %)	7 (44 %)	1 (6 %)	0,50
Иерархия (руководитель)	3	0 (0 %)	3 (100 %)	0	1,00
Всего	**36**	**22 (61 %)**	**13 (36 %)**	**1 (3 %)**	**0,39**

Полученные данные уточняют исходное положение. Средний индекс косвенности 0,39 означает, что большинство директивов действия реализуется прямыми стратегиями. Вместе с тем значение 1,00 при обращении к непосредственному руководителю показывает, что исходное положение верно применительно к страте с наибольшей статусной дистанцией. Индекс монотонно возрастает по мере её увеличения: 0,10 → 0,29 → 0,50 → 1,00. Следовательно, косвенность в узбекском языке следует описывать не как признак языка, а как функцию социальной дистанции.

Внутри одного диалога направление статусного различия сохраняется: доля прямых директивов у вышестоящего участника составляет 0,73 (16 из 22), у нижестоящего — 0,43 (6 из 14); точный критерий Фишера даёт $p = 0{,}092$ при отношении шансов 3,56. Различие не достигает порога значимости, однако его направление теоретически ожидаемо и требует проверки на расширенной выборке.

Таблица 4. Частотность средств смягчения (узбекоязычная часть корпуса, 8 229 сл.)

Средство	Абс.	На 10 000 сл.	Функция
Модальность возможности -a ol- / -a olma-	39	47,4	косвенность и отказ
Бенефактив -(i)b ber-	23	27,9	смягчение просьбы
Xudo xohlasa / InshaAllah	21	25,5	условное согласие

-(i)b qo‘y-	18	21,9	снижение веса действия
mayli / maylimi	13	15,8	запрос разрешения
bo‘ladimi	10	12,2	запрос разрешения
mumkin / mumkinmi	9	10,9	запрос разрешения (офиц.)
iltimos	8	9,7	лексический смягчитель
bir (дискурсивное употребление)	7	8,5	минимизация запроса

Примечание: частоты получены с помощью регулярных выражений с последующей ручной верификацией; bir отделено от числительного вручную, лексикализованные сочетания типа javob ber- исключены из подсчёта бенефактива.

Частотные данные подтверждают морфологический характер смягчения: бенефактив -(i)b ber- употребляется втрое чаще лексического iltimos, а формула Xudo xohlasa вдвое частотнее iltimos, что указывает на большую активность средств смягчения обязательства по сравнению со средствами смягчения просьбы. Дискурсивное bir встречается редко (8,5) и представляет собой факультативный слой смягчения. Материал подтверждает и косвенный характер отказа: прямое *yo‘q* ‘нет’ фиксируется преимущественно в институциональных ответах, тогда как на личную просьбу отвечают конструкцией невозможности — *boshqatdan yozolmiyman-da* ‘я ведь не смогу переписать заново’. Показательно, что конструкция -a ol-, смягчающая просьбу в утвердительной форме, в отрицательной (-a olma-) становится основным средством отказа, представляя его не как личное нежелание, а как объективную невозможность; в типологии Л. Биби и её соавторов [2] это непрямой отказ.

6. Обсуждение и выводы

1. Тезис о преимущественно косвенном характере узбекских директивов подтверждается частично. Средний индекс 0,39 показывает преобладание прямых стратегий; значение 1,00 достигается только при наибольшей статусной дистанции. Корректная формулировка: степень косвенности закономерно изменяется в зависимости от социальной дистанции — это превращает исходный тезис в проверяемое утверждение.

2. Социальная дистанция определяет форму, а не количество побуждений. Доля директивов по стратам не различается ($p = 0,552$), тогда как распределение по шкале косвенности различается систематически. Это уточняет модель П. Брауна и С. Левинсона [4]: набор стратегий универсален, а средства их реализации ситуативно-специфичны.

3. Статусная дистанция изменяет организацию диалогического эпизода. Фатическое обращение у равностатусных партнёров образует самостоятельный ход ($z = 3,85$), а в институциональной иерархии становится лексическим компонентом приветственной формулы ($z = -2,56$).

4. Смягчение в узбекском языке сосредоточено на морфологическом уровне. Бенефактив -(i)b ber- втрое частотнее лексического iltimos (27,9 против 9,7), что количественно подтверждает типологический тезис и уточняет выводы Т.В. Лариной [10] о синтаксическом кодировании дистанции в английской лингвокультуре.

5. Одно морфологическое средство обслуживает две противоположные иллокутивные функции. Показатель -a ol- смягчает просьбу, а -a olma- становится основным средством отказа.

6. Ряд узбекских речевых актов не имеет полных эквивалентов в европейских языках. Прежде всего это относится к благопожеланию (duo), что указывает на лакуну в данной области речевого этикета и имеет непосредственное значение для перевода.

Ограничения. При наблюдаемой величине эффекта (V = 0,174) для надёжного обнаружения слабых различий требуется около 731 акта, поэтому статистически незначимые результаты не следует интерпретировать как доказательство отсутствия различий. Во всех диалогах участвует один постоянный коммуникант, вследствие чего результаты не полностью свободны от идиолектных особенностей. Коэффициент согласия кодировщиков не вычислялся: для этого необходима независимая повторная разметка. Сопоставительная количественная проверка на английском и русском материале составляет задачу дальнейшей работы.

Практическая значимость результатов состоит в возможности их применения в преподавании узбекского языка как иностранного, в переводоведении при работе с лакунарными единицами речевого этикета и в подготовке специалистов по межкультурной коммуникации.

СПИСОК ЛИТЕРАТУРЫ

1. Арутюнова Н.Д. Язык и мир человека. – М.: Языки русской культуры, 1999. – 896 с.
2. Биби Л.М., Такахаси Т., Улисс-Уэлц Р. Pragmatic Transfer in ESL Refusals // Developing Communicative Competence in a Second Language. – New York: Newbury House, 1990. – P. 55–73.
3. Блум-Кулка Ш., Хаус Дж., Каспер Г. Cross-Cultural Pragmatics: Requests and Apologies. – Norwood, NJ: Ablex, 1989. – 300 p.
4. Браун П., Левинсон С.К. Politeness: Some Universals in Language Usage. – Cambridge: Cambridge University Press, 1987. – 345 p.
5. Хакимова Г. Ўзбек тилида нутқий мулойимлик категорияси. – Тошкент: Фан, 2019. – 156 б.
6. Холл Э.Т. Beyond Culture. – New York: Anchor Press / Doubleday, 1976. – 298 p.
7. Гумбольдт В. фон. On Language. – Cambridge: Cambridge University Press, 1999. – 300 p.
8. Карасик В.И. Языковой круг: личность, концепты, дискурс. – Волгоград: Перемена, 2002. – 477 с.
9. Лакофф Дж., Джонсон М. Metaphors We Live By. – Chicago: University of Chicago Press, 1980. – 242 p.
10. Ларина Т.В. Категория вежливости и стиль коммуникации: сопоставление английских и русских лингвокультурных традиций. – М.: Языки славянских культур, 2009. – 512 с.
11. Маҳмудов Н. Нутқ маданияти асослари. – Тошкент: Ўқитувчи, 2007. – 192 б.
12. Рэйсон П., Гарсайд Р. Comparing Corpora using Frequency Profiling // Proceedings of the Workshop on Comparing Corpora, 38th Annual Meeting of the ACL. – Hong Kong, 2000. – P. 1–6.
13. Сафаров Ш. Прагмалингвистика. – Тошкент: Ўзбекистон Миллий Энциклопедияси, 2008. – 318 б.
14. Сепир Э. Language: An Introduction to the Study of Speech. – New York: Harcourt, Brace and Company, 1921. – 258 p.
15. Сёрл Дж.Р. Speech Acts. – Cambridge: Cambridge University Press, 1969. – 203 p.

16. Степанов Ю.С. Константы: словарь русской культуры. – М.: Академический проект, 2004. – 992 с.
17. Уорф Б.Л. Language, Thought and Reality. – Cambridge, MA: MIT Press, 1956. – 278 p.
18. Вежбицкая А. Understanding Cultures through Their Key Words. – New York: Oxford University Press, 1997. – 317 p.
19. Abdurakhmonova, N. (2021). Formal-Functional Models of The Uzbek Electron Corpus. ANGLISTICUM. Journal of the Association-Institute for English Language and American Studies, 10(8), 59-66.
20. Abduraxmonova, N. (2021). O ‘zbek tili elektron korpusining kompyuter modellari. Filologiya fanlari doktori (DSc) dissertatsiyasi avtoreferati.

USLUBIYAT VA NUTQ MADANIYATI

Z. Taxirov
O‘zMU professori,
filologiya fanlari doktori
https://doi.org/10.5281/zenodo.23253903

Annotatsiya. *Mazkur maqolada nutqiy muloqotda til birliklari va kategoriyalaridan foydalanish haqida so‘z boradi. Nutq madaniyati, ortologiya (to‘g‘ri nutq), til me’yorlari, ijtimoiy lingvistika, adabiy nutq nazariyasi haqidagi ma’lumotlar berilgan.*

Аннотация. *В данной статье рассматриваются вопросы использования языковых единиц и категорий в речевой коммуникации. Представлены сведения о культуре речи, ортологии (правильной речи), языковых нормах, социолингвистике, а также теории литературной речи.*

Abstract. *This article examines the use of linguistic units and categories in speech communication. It provides an overview of speech culture, orthology (correct and normative speech), language norms, sociolinguistics, and the theoretical foundations of literary language.*

Tayanch terminlar: *Uslubiyat, muloqot, nutq madaniyati, til birliklari, kategoriyalar, funksiya, kommunikativlik, og‘zaki nutq, yozma nutq, konstruksiya, nominativ ma’no, adabiy nutq.*

Uslubiyat nutqiy (og‘zaki va yozma) muloqot vaqtida til birliklari va kategoriyalaridan foydalanish haqidagi fandir.

Uslubiyatda, asosan, *“til birliklaridan foydalanish”* yoki *“til birliklarining muayyan vazifani bajarishi”* haqida so‘z boradi. Ana shu vazifa bajarishligiga ko‘ra u vazifaviy (funksiyaviy) uslubiyat deb ataladi. Til vositalaridan foydalanishga tildan foydalanuvchi, ya’ni nutq egasi nuqtayi nazaridan qarash lozim [4].

Tildan, xususan, adabiy nutqdan foydalanishning barcha usullari uslubiyatning predmeti hisoblanadi [1].

Uslubiyat tilshunoslik fanlari orasida alohida o‘rinni egallaydi. Boshqa tilshunoslik fanlari tilda *nima* borligini o‘rgansa, uslubiyat til birliklari va kategoriyalaridan *qanday* foydalanishni tadqiq etadi. Bir so‘z bilan aytganda, *harakatdagi* tilni o‘rganadi.

Tilda shakllangan birliklar va kategoriyalar nutqiy muloqot jarayonida quyidagilarga bog‘liq ravishda muayyan vazifani bajaradi:

– xabarlar mazmuniga;
– matn (umumiy nutqiy muloqot)ga;
– nutqiy muloqot amalga oshirilayotgan sharoitlarga.

Tilning qanday vositalari birikmaga kirishishi, bir-birlari bilan bog‘lanishi, bir-biriga moslashishi, muloqot sohasi, nutq shakli (og‘zaki, yozma), matn mavzui, muhokamaning mohiyati, “gaplashish” va h. ga ko‘ra uslubiyat quyidagilarni aniqlashtiradi:

– kishilarning birgalikda faoliyat olib borishlari, ish yuritishlari, umuman, turmush kechirishlari kabi jamiyatning muhim ehtiyojini qondirishda tildan qanday foydalanilishini;
– til har bir birligi qanday vazifani bajarishligini;

– muayyan matn tarkibida (yoki muayyan matn turida), u yoki bu muloqiy sharoitda til birliklari mutanosiblashishi yoki bir-biriga ta'sir etishligini.

Uslubiyat tilshunoslikning funksiyaviy-kommunikativ (vazifaviy-muloqiy) turi hisoblanadi va nutq madaniyatini, ortologiya (to'g'ri nutq haqidagi fan)ni, til me'yorlari to'g'risidagi ma'lumotni, ijtimoiy lingvistikani, adabiy nutq nazariyasini o'z ichiga oladi.

Nutq egasi quyidagilarni a'lo darajada bilishi nazarda tutiladi.

1. Til "material"ini, kategoriyalarini, birliklarini — uning tizimini tashkil etuvchilarni;
2. Til tovush tizimi, leksikasi, grammatikasi (morfologiya, sintaksis), frazeologiyasi, so'z yasalishi va b. ni;
3. Uslubga ko'ra ega va kesim moslashuvini;
4. Sodda va qo'shma gaplar tuzilishini;
5. Nutqdagi obyektiv va subyektiv munosabatni;
6. O'timli va o'timsiz fe'llarni;
7. Qaratqich va tushum kelishiklarining farqini;
8. Mavhum otlar, shakllar omonimiyasi va b. ni.

Uslubiyatda, avvalo, nutqiy matn, shuningdek, nutq egasining umumiy fikri, kontekst ta'siriga ko'ra bayon etilayotgan nutqda til birliklaridan foydalanishdagi me'yordan chetlanish — sintaktik konstruksiyalardagi kamchiliklar, frazeologizmlarning xato qo'llanganligi, so'zlarning birikmaga noto'g'ri kiritilganligi, so'zning o'z ma'nosida qo'llanilmaganligi, ma'no bo'yog'idan noto'g'ri foydalanilganligi e'tiborga olinadi.

Havo bulut edi. Oq bulutlar bo'lganlikdan quyoshning nuri uncha kesilmagan, quyosh to'r ro'mol ichida o'tirgan kelinlar kabi yer yuziga suzilib qarar edi. (A. Qodiriy).

Tabiat manzarasini chizar ekan, adib *quyosh kelinlar kabi suzilib qarar edi* o'xshatish konstruksiyasini tuzadi.

Uslubiy nuqtayi nazardan bu konstruksiya e'tiborni o'ziga tortadi. Grammatika bo'yicha mutaxassisning fikricha, bu konstruksiya grammatik jihatdan to'g'ri hisoblanadi. Leksikolog esa, *suzilmoq* so'zini qo'llashda me'yordan chekinish bor, deydi.

Uslubiyat bo'yicha mutaxassis esa mazkur konstruksiyadagi me'yordan chekinishlikni aniqlay boshlaydi: nima uchun, nega? Matn uchun bu nima beradi, muayyan aniq ifodaga yana qanday yangi ifoda (ottenka) bag'ishlaydi.

Mazkur konstruksiya aslida, *quyosh kelinlar kabi qaraydi*. Konstruksiya to'rt komponentdan tashkil topadi: *quyosh, kelin, kabi, qaraydi.* Bu konstruksiyada *quyosh, kelin* — ot, *kabi* — ko'makchi, *qaraydi* — fe'l (o'tgan zamon). Ulug' adib quyoshning nur taratishini inson qarashiga o'xshatar ekan, qarashga xos ko'plab belgilar orasidan faqat bittasini tanlaydi: *suzilib*. Ana endi tasviriy tilshunoslikdan uslubiyatga o'tiladi: nega? Inson qarashiga turli belgilar xos: *kulib qarash, xo'mrayib qarash, bezrayib qarash, tirjayib qarash, jilmayib qarash, tumshayib qarash, tersayib qarash, ola qarash, irjayib qarash, sirli qarash, shubhali qarash, ishonqiramay qarash, ishonib qarash, chaqchayib qarash, baqrayib qarash, yovqarash, do'stona qarash, bo'zrayib qarash, iljayib qarash, yotsirab qarash, yig'lamsirab qarash* va h.

Nega yozuvchi qarashga xos ko'pgina belgilar ichidan faqat suzilibni tanladi. Ana shu tanlash ayni uslubni anglatadi.

Agar *suzilib* so'zi — ravishdoshning nominativ (atash) ma'nosiga e'tibor bersak, u insonga xos qarash belgisini anglatadi, ma'no bo'yog'i, tabiiyki, ijobiy. Bunday qarashga insonlar yoshi, jinsidan qat'i nazar, barchasi ham ega bo'lishi mumkin. Lekin ular barchasi birdek ijobiy emas. Adib yuqorida ayol zotiga mansub tur — *kelin qarashiga* quyoshning *nur*

*sochishi*ni o‘xshatadi. Nega kelinning, nega qizlarning, nega katta yoshdagilar (kampirlar)ning? Vaholanki, ular ham ayol! Chunki qizlarning *suzilib qarashi — odobsizlik*, yoshi kattalarniki — *yarashmagan qiliq*. Kelinning bu xil qarashlari ham ma’lum davrda (yangi kelinlik davrida), shunda ham o‘z jufti haloliga nisbatan bo‘lsa, *ijobiy* sanaladi.

Yuqoridagi parcha badiiy matn ekanligini hisobga olsak, unda muallif quyosh nurlarini oq bulut to‘sib qo‘yganligi va uning ham yorug‘ligi xira tortgan, hovuri pasayganligi va bu kishiga yoqishlik darajasida ekanligini ta’kid etmoqchi bo‘ladi va buni kelinlarning suzilib boqishligi yoqimliligiga o‘xshatadi.

Uslubiyat nutqdagi bu hodisani tadqiq etarkan, tasviriy tilshunoslik bilan o‘zaro jips bog‘liqdir, chunki u mohiyatan tilni tadqiq etishni nihoyasiga yetkazadi. Uslubiyat tilni tizimli-tuzilmaviy tahlil qilish vazifasi (tasviriy tahlil)dan o‘tib, lingvistika qismlariga asoslangan holda quyidagilarni tadqiq etadi:

1. Fikrni ifoda etishning yondosh (parallel) usullaridan, til vositalari yordamida yuzaga chiqadigan his-hayajonlarni;
2. Turli maqsadlarda va nutqiy muloqot sharoitlarida til vositalarining turli yo‘sinda birikuvi, qo‘shiluvi usullari, tartibini;
3. Nutqiy muloqotda, yozma va og‘zaki nutqda til unsurlari, kategoriyalarining amalda qo‘llanilish qonuniyatlari va o‘ziga xosliklarini.

U darvozaga yaqinlashayotgan edi, nimadandir cho‘chib chetga burildi. Chunki darvozadan chiqib qolgan ikki kishini ko‘rib, bularning birini qayin otasi bo‘lib chiqmagidan hurkkan edi. (Abdulla Qodiriy).

Mazkur parchadagi jumlalar grammatika nuqtayi nazaridan to‘g‘ri. Leksikolog so‘z qo‘llash xususida me’yordan ma’lum darajada chetlashish borligini qayd etadi va shu bilan cheklanadi. Faqat uslubiyatchi mutaxassisgina birinchi jumlada *cho‘chimoq,* ikkinchi jumlada *hurkmoq* so‘zining qo‘llanilganiga *nega* deya e’tibor beradi va uning tahliliga kirishadi, bunday ikki xil qo‘llash mohiyatini ochib beradi.

Birinchi jumladagi *cho‘chimoq* fe’lining genetik ma’nosi *kutilmagan holatdan xavotirga tushmoq* ifodasini beradi, ya’ni holat, harakat xavotir bilan bog‘liq va bu holat ayni insonlarga xos.

Ikkinchi jumladagi *hurkmoq* so‘zining genetik ma’nosi jonivorlarga, ularning ba’zi turlariga — ot, qora mol, tuya va b. xos holat, harakat ifodasini beradi va bu holat, harakat *qo‘rquv* bilan bog‘liq bo‘ladi.

Yuqoridagi har ikki jumla grammatika qoidalariga muvofiq tuzilgani, ya’ni ega, kesim, ikkinchi darajali bo‘laklar o‘z o‘rnida, so‘zlarning birikmaga kirishuvi ham me’yorida ekanligi tilshunoslik nuqtayi nazaridan to‘g‘ri, deya tan olinadi va bular uning o‘rganish obyekti hisoblanadi.

Leksikologiyada *cho‘chimoq, hurkmoq* so‘zlarining bir-biriga sinonim ekanligi, faqat *cho‘chimoq* so‘zining ma’no ifodasi betaraf, *hurkmoq* so‘zi ma’nosi (insonga nisbatan qo‘llanganda) bo‘yoqqa egaligini ta’kidlash bilan kifoyalanadi.

Uslubiyatda esa, boshqacha aytganda, semik tahlilda har ikki so‘z, garchi sinonim bo‘lsa-da, nega qo‘llanildi, degan savol qo‘yiladi va leksik birliklar, grammatik kategoriyalar semik tahlil qilinadi.

Birinchi jumlada *cho‘chib* ravishdoshining qo‘llanishining boisi, cho‘chitgan narsa, hodisa hali ma’lum emas. Albatta, cho‘chishlikka xavotir sabab bo‘ladi. Keyingi jumlada esa

hurkib ravishdoshi — *cho'chib*ning sinonimi qo'llanadi. Nega? Uslubiyat ana shu negaligini tadqiq etadi va quyidagi xulosaga keladi:

– *cho'chib* so'zi ma'no ifodasida hali *qo'rqish* hissi yo'q, u *xavotirlanish*, *sergaklanish* ma'no ifodasini beradi;
– *hurkib* so'zi ma'no ifodasida esa *qo'rqish* hissi ma'no ifodasi ayon seziladi;
– birinchi holatda xavotirga solgan, sergaklantirgan obyekt noma'lum;
– ikkinchi holatda qo'rquvga sabab bo'lgan obyekt — ikki kishi ma'lum;
– ikki kishidan biri "qayin otasi bo'lib chiqmag'i".

Biz yuqorida uslubiyat nutq madaniyati bilan jips aloqada degan edik. Zero, uslubiyat funksiyaviy kommunikativ fanlar sirasiga kiradi.

Demak, uslubiyat, nutq madaniyati, ortologiya, ritorika til unsurlari va kategoriyalaridan adabiy nutq doirasida oqilona, me'yorida foydalanishni tahlil etadi.

Nutq madaniyati haqidagi ta'limda til birliklaridan kundalik nutqiy muloqotda yozma va og'zaki tarzda, adabiy nutq doirasida foydalanish ko'rib chiqiladi.

Nutq madaniyati deyilganda alohida olingan bir kishi va aholi nutq madaniyatining umumiy darajasi ham nazarda tutiladi.

Nutq madaniyati bo'yicha mutaxassislar amaldagi adabiy nutq me'yorlari tizimi nuqtayi nazaridangina emas, balki, individual hamda umumiy nutqiy madaniyatga uning qanchalik mantiqiy bayonidagi aniq-tiniqlik va ifodalilik (estetik) ekanligini ham nazarda tutadilar.[5]

Adabiy me'yorlar nuqtayi nazaridan nutq madaniyatiga quyidagi talablar qo'yiladi:

– til vositalaridan foydalanishning asoslangan bo'lishi;
– matn tuzilishi mantiqan izchillikka ega bo'lishi;
– bayon aniq-tiniq, tushunarli bo'lishi.

Shunday qilib, nutq madaniyati o'z ichiga faqat adabiy nutqnigina emas, balki xalq so'zlashuv nutqini ham to'liq qamrab oladi, demak, oddiy og'zaki so'zlashuv unsurlari, xalq shevalari, jargonlar, jinoyatchilar argosidan tortib, to kasb-hunarga oid nutqdagi bilgan-bilmagan holda og'zaki, yozma nutqqa, shu jumladan adabiy og'zaki nutqqa o'tadigan barcha-barcha unsurlar tahlil etiladi.

Individual nutq ham, umumiy nutqiy madaniyat ham nutq egasining saviyasiga, ijtimoiy ongga, adabiy nutq rivoji va uning hozirgi darajasiga bog'liq.

Nutq madaniyati, aytish mumkinki, baholovchi fan hisoblanadi. U nutq asoslarini baholaydi, me'yoriylashganlik darajasini aniqlaydi. Muayyan nutqiy vaziyatda mazmunan o'rinli, o'rinsiz ekanligini tahlil etadi.

Uslubiyat nutq madaniyati sohasidagi tadqiqotlarga nazariy asos bo'lib hisoblanadi. O'rganilayotgan nutq unsurlari uslubiyat nuqtayi nazaridan aniq tavsifga ega bo'ladi.

"Til madaniyati yoki nutq madaniyati haqidagi fan ham nazariy, ham amaliy fan (yoki tadqiqot sohasi), til uslubiyati va nutq uslubiyati bilan turdosh hisoblanadi" [2].

FOYDALANILGAN ADABIYOTLAR

1. Винокур Г. О. Филологические исследования. — Москва, 1989. — 221 с.
2. Виноградов В. В. Проблемы русского языка. — Москва, 1981. — 175 с.
3. Taxirov Z. T. O'zbek tilining amaliy uslubiyati: darslik. — T.: Tafakkur avlodi, 2021. — 240 b.

4. Abdurahmonova, N. (2025). Leksikografiyada Terminologik Lug'atlarning o'rni Va Ahamiyati. *Yangi Davr Ilm-Fani: Inson Uchun Innovatsion G 'Oya Va Yechimlar*, *1*(9), 47-49.
5. Mengliev, D., Abdurakhmonova, N., Barakhnin, V., Vasliddinova, K., Rahimov, H., & Djalolova, K. (2024, June). Enhancing sentiment analysis in Uzbek language texts through weighted lexical features. In *2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM)* (pp. 2450-2453). IEEE.

RAQAMLI TA'LIM MUHITIDA BO'LAJAK PEDAGOGLARNING KASBIY RIVOJLANISHINI BAHOLASHNING INTELLEKTUAL TEXNOLOGIYALAR

Go'zal S, Tojiyeva
Mustaqil izlanuvchi
Termiz davlat pedagogika instituti, Surxondaryo, O'zbekiston.
E mail: guzaltojieva65@gmail.com
ORCID: 0009-0001-5423-2304
https://doi.org/10.5281/zenodo.23254125

Annotatsiya. *Mazkur tezisda bo'lajak pedagoglarning kasbiy rivojlanishini baholashda intellektual texnologiyalardan foydalanish masalasi ko'rib chiqiladi. Unda ta'lim tahlili va sun'iy intellektga asoslangan baholash texnologiyalarining imkoniyatlari hamda ularni kasbiy rivojlanishni aniqlashda qo'llash yo'llari yoritiladi.*

Kalit so'zlar: *raqamli ta'lim, bo'lajak pedagog, kasbiy rivojlanish, ta'lim tahlili, sun'iy intellekt, intellektual baholash texnologiyalari.*

Аннотация. *В данной статье рассматривается вопрос использования интеллектуальных технологий для оценки профессионального развития будущих педагогов. Раскрываются возможности аналитики обучения и технологий оценивания на основе искусственного интеллекта, а также пути их применения для определения профессионального развития будущих педагогов.*

Ключевые слова: *цифровое образование, будущие педагоги, профессиональное развитие, аналитика обучения, искусственный интеллект, интеллектуальные технологии оценивания.*

Annotation. *This thesis examines the use of intelligent technologies for assessing the professional development of future teachers. It highlights the potential of learning analytics and artificial intelligence-based assessment technologies, as well as ways of applying them to assess the professional development of future teachers.*

Keywords: *digital education, future teachers, professional development, learning analytics, artificial intelligence, intelligent assessment technologies.*

Kirish. Raqamli texnologiyalarning ta'lim jarayoniga kirib kelishi bilan talabalarning o'quv faoliyatini tashkil etish va baholash imkoniyatlari sezilarli darajada kengaydi. Elektron ta'lim platformalari, raqamli o'quv resurslari, elektron portfoliolar va turli baholash vositalari talabalarning o'quv faoliyati to'g'risidagi ma'lumotlarni yig'ish hamda tahlil qilish uchun yangi imkoniyatlar yaratdi. Shu bilan birga, ta'lim jarayonining raqamli shaklga o'tishi bo'lajak pedagoglarning kasbiy rivojlanishini baholash mazmuni va usullarini ham takomillashtirishni taqozo etmoqda.

O'zbekistonda raqamlashtirish va sun'iy intellekt texnologiyalarini rivojlantirishga qaratilgan davlat siyosati ham mazkur jarayonning dolzarbligini kuchaytirmoqda. Xususan, O'zbekiston Respublikasi Prezidentining 2024-yil 14-oktabrdagi PQ-358-son qarori bilan Sun'iy intellekt texnologiyalarini 2030-yilgacha rivojlantirish strategiyasi tasdiqlanib, sun'iy intellektni

ijtimoiy sohalarga joriy etish, ma'lumotlarni qayta ishlash imkoniyatlarini kengaytirish va kadrlar salohiyatini rivojlantirish ustuvor yo'nalish sifatida belgilangan [1].

Biroq raqamli muhitda ma'lumotlarning ko'pligi oddiy qayd etish orqali talabaning kasbiy rivojlanishini to'liq aniqlash imkonini bermaydi. O'quv faoliyati, topshiriqlarni bajarish natijalari, test ko'rsatkichlari, elektron portfoliodagi ishlar va boshqa raqamli ma'lumotlarni tizimli tahlil qilish zarurati yuzaga keladi. Ayniqsa, bo'lajak pedagoglarning kasbiy bilim, ko'nikma va kompetensiyalarining shakllanish dinamikasini aniqlashda bunday ma'lumotlardan samarali foydalanish muhim ahamiyat kasb etadi.

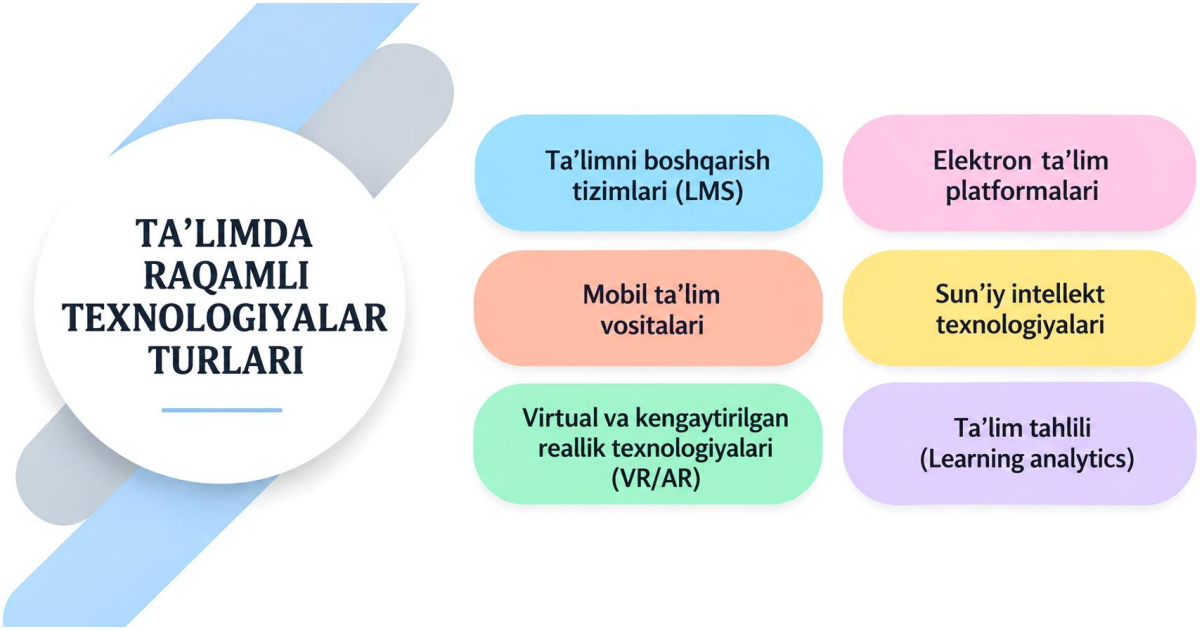
Shu sababli raqamli ta'lim muhitida baholashning asosiy e'tibori faqat yakuniy natijani qayd etishga emas, talabaning kasbiy rivojlanish jarayonini tahlil qilish va uning individual xususiyatlarini aniqlashga qaratilishi lozim. Bunda ta'lim tahlili raqamli ma'lumotlarni tizimlashtirish va rivojlanish dinamikasini aniqlashga, sun'iy intellektga asoslangan baholash esa ushbu ma'lumotlar asosida baholash jarayonini takomillashtirishga xizmat qilishi mumkin.

Asosiy qism. Bo'lajak pedagoglarning kasbiy rivojlanishini baholashda ilmiy ma'lumotlardan foydalanishga asoslangan yondashuvlar ta'lim jarayonidagi baholash imkoniyatlarini kengaytirmoqda. Xususan, tadqiqotchi R. Ferguson ta'lim tahlili orqali ta'lim oluvchilarning faoliyati va rivojlanishiga oid ma'lumotlarni yig'ish, tahlil qilish hamda ulardan ta'lim jarayonini takomillashtirishda foydalanish mumkinligini ta'kidlaydi [Ferguson R., 2012:2]. Bu esa bo'lajak pedagoglarning kasbiy rivojlanishini baholashda faqat yakuniy natija emasligini bildiradi. Jumladan, rivojlanish jarayonida kuzatiladigan o'zgarishlarga ham e'tibor qaratish imkonini beradi.

Shuningdek, V. González-Calatayud, P. Prendes-Espinosa va R. Roig-Vila qarashlarida sun'iy intellektning ta'lim oluvchilarni baholashdagi imkoniyatlarini tizimli ravishda o'rgangan. Tadqiqotchilar sun'iy intellektdan baholash jarayonlarini avtomatlashtirish, qayta ishlash va ta'lim oluvchiga fikr-mulohaza taqdim etishda foydalanish mumkinligini ko'rsatgan (González-Calatayud, Prendes-Espinosa & Roig-Vila, 2021:3).

R. Luckin nazariyasiga ko'ra, sun'iy intellektga asoslangan baholash tizimlari ta'lim oluvchining rivojlanishini muntazam kuzatish va natijalarga mos fikr-mulohaza shakllantirish imkoniyatlarini kengaytirishini qayd etadi (Luckin, 2017 4). M. I. Jordan va T. M. Mitchell esa mashinaviy o'rganish texnologiyalarining katta hajmdagi ma'lumotlardan qonuniyatlarni aniqlash va bashoratli tahlil qilish imkoniyatlarini asoslagan (Jordan & Mitchell, 2015:255-260: 5).

Ko'rsatilib o'tilgan bu qarashlar intellektual texnologiyalarning bo'lajak pedagoglarning kasbiy rivojlanishini baholashdagi imkoniyatlarini umumlashtirish uchun asos bo'lib xizmat qiladi. Bunda ma'lumotlarni yig'ish va tahlil qilish, baholash natijalarini aniqlash hamda ular asosida keyingi rivojlanish yo'nalishlarini belgilash o'zaro bog'liq jarayon sifatida qaraladi (1-rasm).



1-rasm. Bo‘lajak pedagoglarning kasbiy rivojlanishini baholashda intellektual texnologiyalardan foydalanish jarayoni

Bo‘lajak pedagoglarning kasbiy rivojlanishini baholashda raqamli ma’lumotlardan foydalanishning asosiy bosqichlari aks ettirilgan. Unda o‘quv va kasbiy faoliyatga oid ma’lumotlarni yig‘ishdan boshlab, tahlil qilish, rivojlanish holatini aniqlash va baholash natijalariga asoslangan individual tavsiyalarni shakllantirishgacha bo‘lgan jarayon izchil ifodalangan. Bunday ilmiy qarashlar bo‘lajak pedagoglarning kasbiy rivojlanishini baholashda ta’lim tahlili va sun’iy intellektga asoslangan baholash texnologiyalaridan foydalanish uchun nazariy asos yaratishi mumkin. Bunda ta’lim tahlili bo‘lajak pedagoglarning o‘quv va kasbiy faoliyatiga oid ma’lumotlarni tizimlashtirish hamda rivojlanish dinamikasini aniqlashga, sun’iy intellekt esa ushbu ma’lumotlarni qayta ishlash, baholash va individual tavsiyalar shakllantirishga xizmat qiladi.

Tahlil natijalaridan samarali foydalanish baholash va talabaning keyingi rivojlanish yo‘nalishlarini aniqlash mexanizmlarini talab etadi. Bu jarayonda sun’iy intellektga asoslangan baholash texnologiyalari muhim imkoniyat yaratadi. To‘plangan ma’lumotlarni qayta ishlash, turli ko‘rsatkichlar o‘rtasidagi bog‘liqliklarni aniqlash va baholash natijalariga asoslangan individual tavsiyalar shakllantirishga xizmat qilishi mumkin.

Ta’lim tahlili orqali olingan ma’lumotlarning sun’iy intellekt asosida qayta ishlanishi baholash jarayonining izchilligini ta’minlash imkonini beradi. Bunda birinchi texnologiya bo‘lajak pedagoglarning faoliyatiga oid ma’lumotlarni aniqlash va tahlil qilishga xizmat qilsa, ikkinchisi ushbu ma’lumotlar asosida baholash natijalarini qayta ishlash va individual rivojlanish ehtiyojlarini belgilash imkonini beradi. Mazkur texnologiyalarning vazifalari va o‘zaro bog‘liqligi quyidagicha aks ettirilgan (2-rasm).



2-rasm. Bo‘lajak pedagoglarning kasbiy rivojlanishini baholashda ta’lim tahlili va sun’iy intellekt texnologiyalarining o‘zaro bog‘liqligi

Tahlil ma’lumotlarni yig‘ish va tahlil qilishni ta’minlasa, sun’iy intellekt ushbu ma’lumotlardan baholash natijalarini shakllantirish va individual tavsiyalar ishlab chiqishda foydalanadi. Natijada bo‘lajak pedagogning kasbiy rivojlanishini baholash jarayoni muntazam ma’lumotlar tahliliga asoslanadi.

Xulosa. Tahlillar asosida bo‘lajak pedagoglarning kasbiy rivojlanishini baholashda raqamli ma’lumotlardan foydalanish baholash mazmunini kengaytirishi aniqlandi. Bunda kasbiy bilimlarning qo‘llanishi, amaliy topshiriqlarni bajarish, o‘quv faolligi va rivojlanishdagi o‘zgarishlarni ham hisobga olish imkoniyati yuzaga keladi. Intellektual baholash texnologiyalari mazkur ko‘rsatkichlarni aniqlash, o‘zgarishini kuzatish va olingan natijalarni tahlil qilish imkonini beradi. Ta’lim tahlili ma’lumotlarni tizimlashtirishga, sun’iy intellekt esa ularni qayta ishlab, bo‘lajak pedagogning rivojlanish holati va individual ehtiyojlarini aniqlashga yordam beradi. Shu jihatdan, intellektual baholashning asosiy ahamiyati baholash natijasini qayd etish bilan cheklanmay, bo‘lajak pedagogning kasbiy rivojlanishidagi mavjud holatni aniqlash va keyingi rivojlanish yo‘nalishlarini belgilash imkoniyatida namoyon bo‘ladi.

FOYDALANILGAN ADABIYOTLAR

1. O‘zbekiston Respublikasi Prezidentining "Sun'iy intellekt texnologiyalarini 2030-yilgacha rivojlantirish strategiyasini tasdiqlash to‘g‘risida" 2024-yil 14-oktabrdagi PQ-358-son Qarori.
2. Ferguson, R. The State of Learning Analytics in 2012: A Review and Future Challenges. Technical Report KMI-12-01, Knowledge Media Institute, The Open University, 2012. – 22 p.
3. González-Calatayud, V., Prendes-Espinosa, P., Roig-Vila, R. Artificial Intelligence for Student Assessment: A Systematic Review. Applied Sciences, 2021. – Vol. 11, No. 12. – P. 5467.
4. Jordan, M. I., Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. Science, 2015. – Vol. 349, No. 6245. – P. 255–260.
5. Luckin, R. Towards artificial intelligence-based assessment systems. Nature Human Behaviour, 2017. – Vol. 1. – Article 0028.

GENERATIONAL DIGITAL KNOWLEDGE GAP IN UZBEK ORGANIZATIONS

Kamola Toshtemirova
"Silk Road" International University of Tourism and Cultural Heritage, Samarkand, Uzbekistan
Email: [email placeholder]
https://doi.org/10.5281/zenodo.23254150

Abstract. *This paper documents a generational gradient in digital knowledge among employees of Uzbek enterprises. Using cross-sectional survey data from 288 respondents, the Composite Digital Knowledge Index (CDKI) is compared across four age bands. The 18–25 group records the highest mean (3.04), the 46+ group the lowest (2.66). The gap of 0.38 index points is statistically significant. Since older cohorts occupy the majority of senior decision-making positions, the pattern implies a strategic knowledge gap at the top of the organisational hierarchy that constrains data-driven decision-making.*

Keywords: *generational gap, digital knowledge, workforce, human capital, Uzbekistan, age cohorts.*

Аннотация. *В работе фиксируется поколенческий градиент цифровых знаний среди сотрудников узбекских организаций. На данных кросс-секционного опроса 288 респондентов Композитный индекс цифровых знаний (CDKI) сопоставлен по четырём возрастным группам. Группа 18–25 показывает максимальный средний CDKI (3,04), группа 46+ - минимальный (2,66). Разница в 0,38 балла статистически значима. Поскольку старшие когорты занимают большинство руководящих позиций, обнаруженный паттерн указывает на стратегический разрыв цифровых знаний на верхнем уровне управленческой иерархии, ограничивающий принятие решений на основе данных.*

Ключевые слова: *поколенческий разрыв, цифровые знания, персонал, человеческий капитал, Узбекистан, возрастные когорты.*

Annotatsiya. *Ushbu maqolada O'zbekiston korxonalari xodimlari o'rtasidagi raqamli bilimlarning avlodlararo gradienti aniqlanadi. 288 respondent ishtirok etgan kesma tarzdagi so'rovnoma ma'lumotlari asosida Raqamli bilimlar kompozit indeksi (CDKI) to'rt yosh guruhi bo'yicha taqqoslanadi. 18–25 yoshdagilar eng yuqori o'rtacha ko'rsatkichni (3.04), 46+ yoshdagilar esa eng pastini (2.66) qayd etadi. 0.38 indeks farqi statistik jihatdan ahamiyatli. Katta yoshdagi guruhlar rahbarlik lavozimlarining aksariyatini egallagani sababli, natija boshqaruv iyerarxiyasining yuqori qismidagi strategik raqamli bilim uzilishini va ma'lumotga asoslangan qaror qabul qilishga cheklovni ko'rsatadi.*

Kalit so'zlar: *avlodlararo farq, raqamli bilim, xodimlar, inson kapitali, O'zbekiston, yosh guruhlari.*

Main text

The generational dimension of digital transformation is an under-analysed feature of workforce dynamics in emerging economies. While the assumption that younger workers hold more digital knowledge is intuitive, quantitative evidence for Central Asia is scarce. This thesis provides empirical measurement for Uzbekistan.

Using data from a cross-sectional online survey of 288 respondents, the Composite Digital Knowledge Index (CDKI) is compared across four age bands: 18–25 (n = 126; 43.8%), 26–35 (n = 82; 28.5%), 36–45 (n = 62; 21.5%) and 46+ (n = 18; 6.2%). The CDKI aggregates five Likert-scale items covering perceived operational efficiency, paperwork reduction, use of digital data in decision-making, decision speed and decision quality; Cronbach's $\alpha = 0.797$.

Mean CDKI values decline monotonically with age: 18–25 = 3.04; 26–35 = 2.98; 36–45 = 2.79; 46+ = 2.66 (Figure 1). The 0.38-point difference between the youngest and oldest groups is small in absolute terms but analytically meaningful: it corresponds to roughly one third of the standard deviation of the CDKI. A Kruskal–Wallis test confirms that the age effect is statistically significant ($H = 8.27$; $df = 3$; $p = 0.041$).

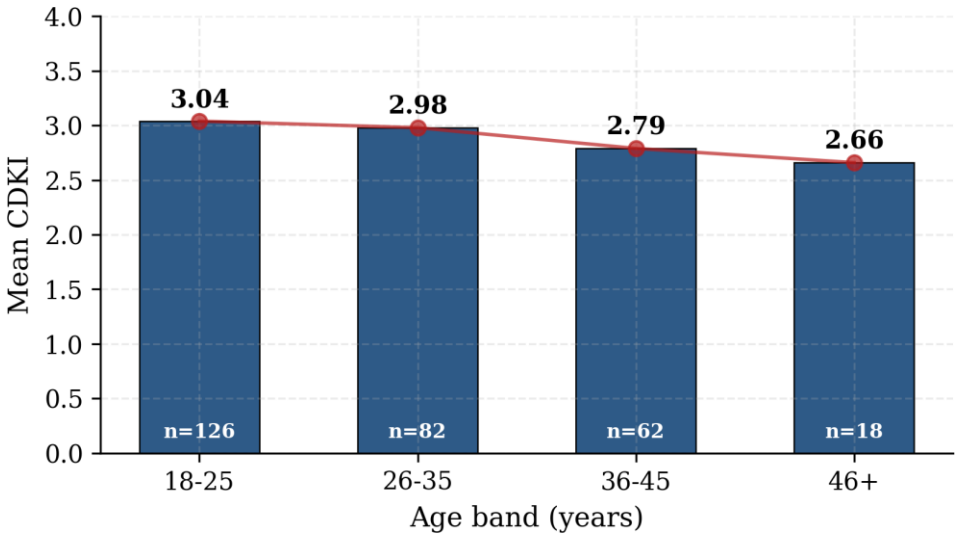


Figure 1. Mean CDKI across four age bands. The monotonic decline is statistically significant (Kruskal–Wallis p = 0.041).

The pattern has a specific organisational consequence. In the sample, executives and business owners are drawn disproportionately from the 36+ age brackets, in line with typical career trajectories. This means that the workforce cohort with the lowest CDKI is also the cohort making the majority of strategic decisions. Younger employees, better equipped with digital knowledge, occupy positions with less strategic authority.

The implication for enterprise practice is direct. Reverse mentoring programmes, in which junior digital-native employees coach senior decision-makers on the use of digital tools, are a low-cost lever with a plausible return. National policy should complement this with targeted digital reskilling programmes for the 36+ workforce, particularly in sectors with high organisational lag. The generational gap is not a problem that will resolve by attrition alone within the relevant policy horizon; active intervention is required.

REFERENCES

1. Brynjolfsson, E., & McAfee, A. (2014). The second machine age. Norton.
2. Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2019). Accelerating digital innovation inside and out. MIT Sloan Management Review.
3. Legner, C., et al. (2017). Digitalization: Opportunity and challenge for the business and information systems engineering community. Business & Information Systems Engineering, 59(4), 301–308.
4. OECD. (2021). The digital transformation of SMEs. OECD Publishing.
5. World Economic Forum. (2023). The Future of Jobs Report 2023. WEF.

IJTIMOIY-MADANIY TIZIMLARDA O‘ZBEK MILLIY LIBOSLARI DIZAYNINI RAQAMLI TRANSFORMATSIYA QILISH VA VIZUAL KODLARNI PARAMETRIK MODELLASHTIRISH METODOLOGIYASI

Toshtemirova Mohichehra G‘olibjon qizi
tadqiqotchi
ORCID: 0009-0001-9834-659X
https://doi.org/10.5281/zenodo.23255120

Annotatsiya. *Mazkur izlanishda ijtimoiy-madaniy tizimlar kontekstida o‘zbek milliy liboslarining etnik vizual kodlarini parametrik modellashtirish va mahalliy matolarning raqamli egizagini "Digital Twin" shakllantirish masalalari konseptual jihatdan tadqiq etiladi. Taklif etilayotgan uslubiy yondashuv dizayn jarayonlarida material tejamkorligini ta’minlash barobarida, milliy madaniy meros elementlarini zamonaviy raqamli muhitda ilmiy asosda interpretatsiya qilishga xizmat qiladi.*
Kalit so‘zlar: *raqamli transformatsiya, parametrik modellashtirish, raqamli egizak, etnik vizual kod.*

Аннотация.*В данном исследовании концептуально рассматриваются вопросы параметрического моделирования этнических визуальных кодов узбекского национального костюма и формирования цифрового двойника «Digital Twin» локальных тканей в контексте социокультурных систем. Предложенный методологический подход позволяет обеспечить ресурсосбережение в процессе проектирования, а также способствует научно обоснованной интерпретации элементов культурного наследия в современной цифровой среде.*
Ключевые слова: *цифровая трансформация, параметрическое моделирование, цифровой двойник, этнический визуальный код.*

Abstract.*This study conceptually examines the parametric modeling of ethnic visual codes in Uzbek national costumes and the development of a "Digital Twin" for local fabrics within socio-cultural systems. The proposed methodological approach ensures material efficiency in the design process and contributes to a scientifically grounded interpretation of cultural heritage elements within the modern digital environment.*
Keywords: *digital transformation, parametric modeling, digital twin, ethnic visual code.*

Kirish. Jamiyat hayotining jadallik bilan raqamlashtirilishi an’anaviy san’at turlari va yengil sanoat tarmoqlari o‘rtasidagi chegaralarni yo‘qotib, ma’lumotlar integratsiyasiga asoslangan yangi ijodiy-muhandislik muhitini shakllantirmoqda. Ushbu sharoitda O‘zbekiston Respublikasi Prezidentining 2020-yil 5-oktabrdagi “Raqamli O‘zbekiston — 2030” strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to‘g‘risida”gi PF-6079-son Farmoni [1] hamda 2024-yil 1-maydagi “To‘qimachilik va tikuv-trikotaj sanoatini rivojlantirishni yangi bosqichga olib chiqish chora-tadbirlari to‘g‘risida”gi PF-71-son Farmonida[2] belgilangan ustuvor vazifalar mamlakatimiz iqtisodiyotining barcha sohalarini, jumladan, milliy yengil sanoat tarmoqlarini raqamlashtirishning konseptual va huquqiy asosini belgilab beradi.

Mazkur normativ-huquqiy hujjatlar talablaridan kelib chiqqan holda, global To‘rtinchi sanoat inqilobi (*Industry 4.0*) texnologiyalaridan foydalanib, milliy madaniy meros obektlarining o‘ziga xos vizual va konstruktiv xususiyatlarini saqlagan holda, ularni zamonaviy 3D parametrik loyihalash muhitiga ilmiy asoslangan holda adaptatsiya qilish metodologiyasini shakllantirish strategik ahamiyat kasb etadi.Ushbu sharoitda yengil sanoat tizimlari an’anaviy ishlab chiqarishdan intellektual va avtomatlashtirilgan kiber-fizik tizimlar (*Cyber-Physical Systems*) modeliga o‘tmoqda. Kiyim-kechak sanoatida Industry 4.0 tamoyillari loyihalash jarayonlariga sun'iy intellekt, bulutli ma’lumotlar ombori va uch o‘lchamli virtual simulyatsiyalarning integratsiyalashuvida namoyon bo‘ladi[3]. Zamonaviy dizayn metodologiyasida ushbu raqamli transformatsiya jarayonlari shunchaki avtomatlashtirilgan chizmachilik dasturlaridan foydalanish bilan cheklanmaydi. U g‘oyaning tug‘ilishidan tortib, uning virtual simulyatsiyasi va maqsadli ishlab chiqarilishigacha bo‘lgan barcha bosqichlarni yagona ma'lumotlar bazasiga (*data-driven*) asoslangan mukammal ekotizimga integratsiyalashni ko‘zda tutadi[4].

Muammoning qo‘yilishi. An’anaviy dizayn algoritmlarining chiziqli strukturasi yuqori resurs sarfi (overproduction) va maketlash zanjirining invariantliligi bilan xarakterlanadi. Shu bilan birga, generativ neyrotarmoqlarning ekstensiv rivojlanishi milliy ornamentlar (anor, qalampir, bodom guli, iroqi kashta) geometriyasini chuqur tahlil qilmasdan, raqamli makonda chuqur yondashilmagan anomal naqsh konseptlarini ishlab chiqilishiga sabab bo‘lmoqda[5].

Mavjud CAD platformalarining standart matematik apparati mahalliy matolarning (ayniqsa, qo‘lda to‘qilgan shoyi va adras kabi materiallarning) teksturaviy, plastik va anizotrop xususiyatlarini to‘liq qamrab olmaganligi sababli, milliy vizual kodlarni parametrik modellashtirish metodikasini tizimlashtirish zaruriyati yuzaga kelmoqda. Bu esa o‘z navbatida, matoning fizik-mexanik xususiyatlarini aniq raqamli konstantalarga o‘tkazish va virtual model ustida uning geometrik deformatsiyalarini algoritmlash vazifasini qo‘yadi.



1-rasm. Muallif tomonidan *CLO 3D* muhitida loyihalangan milliy-etnik ornamentli parametrik liboslar kolleksiyasi

Asosiy qism. Tadqiqotda CAD/3D muhitida milliy liboslarning vizual va jismoniy konstantalarini sintezlovchi uch bosqichli metodologiya mualliflik kolleksiyasi (1-rasm) obyektlari asosida qiyosiy-nazariy tahlil qilish mumkin:

1. **Vizual semantikaning parametrik invariantlari (1 va 2-modellar):** Birinchi va ikkinchi obyektlarda etnik ornamentlarning kompozitsion o‘qlari hamda silliq invariant maydonlar (oq va sariq rang zonalari) monoxrom va koloristik (*CIE L*a*b**) koordinatalar tizimida modellashtirilgan. Vertikal konstruktiv bo‘linishlar inson antropometrik korrelyatsiyasini vizuallashtirib, vizual kodlarning semantik barqarorligini ta’minlaydi. Kolleksiyadagi invariantlilikni va dizayn tizimining moslanuvchanligini isbotlash maqsadida, yagona baza karkasi asosida ikki xil mato va etnik ornament simulyatsiyasi amalga oshirildi. Bu raqamli muhitda andoza o‘lchamlarini o‘zgartirmasdan, material va naqsh teksturasini bir necha soniya ichida variativ almashtirish imkoniyatini ko‘rsatadi.
2. **Dinamik draperiyaning Raqamli Egizagi (3-model):** Uchinchi modelda milliy matolarga xos elastiklik, bukilish qattiqligi (bending stiffness) va sirt zichligi koeffitsiyentlari raqamli konstantalar (N/m va g/m² birliklarida) ko‘rinishida simulyatsiya qilingan. Virtual avatar harakatlanganda hosil bo‘luvchi klosh burmalarining nelineyar deformatsiyasi matoning real hayotdagi fizik tabiati bilan adekvatlikni *Digital Twin* fenomenini[4] namoyon etadi. Bu model kompyuter muhitidagi kiyim dinamikasining haqiqiy kiyim ergonomikasiga yuqori ko‘rsatkichda mos kelishini ta’minlashga xizmat qiladi;
3. **Teksturaning konstruktiv integratsiyasi (4-model):** To‘rtinchi modelda palak va doira nusxa xarakteridagi murakkab kompozitsiyalarni andoza chiziqlari bilan parametrik jihatdan sinxronlash algoritmi qo‘llanilgan. Bu algoritm jismoniy ishlab chiqarishning lazerli kesish (CNC) bosqichida naqshlar simmetriyasining buzilishini avtomatik tarzda bartaraf etadi. Natijada, andozalarni mato yuzasiga joylashtirish optimal sxemaga keltirilib, sanoatda ekologik va iqtisodiy jihatdan yuqori samara beruvchi “Zero-waste” (nol chiqindi) formulasini reallashga zamin yaratadi.

Xulosa. O‘zbek milliy liboslari dizaynini raqamli transformatsiya qilish muammolariga bag‘ishlangan mazkur tadqiqot, an’anaviy madaniy merosni zamonaviy kiber-fizik tizimlar muhitiga adaptatsiya qilishning istiqbolli ilmiy-amaliy yo‘nalishini belgilab beradi. Muallif tomonidan ilgari surilgan vizual kodlarni parametrik modellashtirish hamda matoning plastik xususiyatlariga tayangan holda Raqamli Egizak (Digital Twin) modelini shakllantirish metodologiyasi yengil sanoatda resurs tejovchi "Zero-waste" texnologiyalarini joriy etishda muhim konseptual asos bo‘lib xizmat qilishi mumkin. CLO 3D platformasida amalga oshirilgan mualliflik eksperimentlari etnik ornamentlar semantikasi va kiyim ergonomikasi o‘rtasidagi mutanosiblikni virtual muhitda muvaffaqiyatli sintezlash imkoniyatini namoyon etishga harakat qilindi. Ushbu yo‘nalishdagi izlanishlarni kelgusida mahalliy matolarning o‘ziga xos jismoniy konstantalarini raqamli algoritmlar bilan yanada chuqurroq integratsiyalash orqali boyitish, milliy brendlarimizning global raqamli bozordagi o‘rnini mustahkamlashga xizmat qilishi kutiladi.

FOYDALANILGAN ADABIYOTLAR

1. **“O‘zbekiston Respublikasi Prezidentining Farmoni, 2020”** — O‘zbekiston Respublikasi Prezidentining 2020-yil 5-oktabrdagi “«Raqamli O‘zbekiston — 2030» strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to‘g‘risida”gi PF-6079-son Farmoni. *Qonunchilik ma'lumotlari milliy bazasi*, www.lex.uz.

2. **"O'zbekiston Respublikasi Prezidentining Farmoni, 2024"** — O'zbekiston Respublikasi Prezidentining 2024-yil 1-maydagi "To'qimachilik va tikuv-trikotaj sanoatini rivojlantirishni yangi bosqichga olib chiqish chora-tadbirlari to'g'risida"gi PF-71-son Farmoni. *Qonunchilik ma'lumotlari milliy bazasi*, www.lex.uz.
3. Papachristou, E., & Bilalis, N., 2019 — Papachristou, E., & Bilalis, N. (2019). 3D virtual prototyping in apparel industry: Challenges and future trends. International Journal of Clothing Science and Technology, 31(5), 610-625.
4. Chun, J., & Kim, S., 2021— Chun, J., & Kim, S. (2021). Digital Transformation in Fashion Design: 3D Virtual Garment Simulation and Fabric Physics. Journal of Fashion Technology & Textile Engineering, 9(3), 145-152.
5. Sayfullayev, N. S., 2020" — Sayfullayev, N. S. (2020). O'zbek an'anaviy liboslari va ornamentikasi semantikasi. Toshkent: San'at, 42-56 b.

BO‘LAJAK BOSHLANG‘ICH SINF O‘QITUVCHILARINI ILMIY-IJODIY FAOLIYATGA YO‘NALTIRISH PEDAGOGIK MUAMMO SIFATIDA

Umarov Xusniddin Egamboyevich
Guliston davlat pedagogika instituti
Boshlang‘ich ta’lim metodikasi kafedrasi katta o‘qituvchisi
https://doi.org/10.5281/zenodo.23255228

Annotatsiya. *Ushbu maqolada bo‘lajak boshlang‘ich sinf o‘qituvchilarini ilmiy-ijodiy faoliyatga yo‘naltirishning pedagogik asoslari, dolzarbligi va tizimli muammolari tahlil qilingan. Zamonaviy ta’lim talablaridan kelib chiqqan holda, talabalarda tadqiqotchilik ko‘nikmalarini va ijodiy fikrlashni shakllantirishning asosiy bosqichlari yoritilgan. Maqolada bo‘lajak pedagoglarning kasbiy kompetensiyasini oshirishda ilmiy-ijodiy faoliyatning o‘rni va uni tashkil etishning samarali metodika hamda shakllari taklif etilgan.*

Kalit so‘zlar: *ilmiy, ijodiy, kompetentlik, pedagogik muammo, ta’lim, tarbiya, metodika, dolzarb mavzu, tadqiqot, shaxsiy sifatlar.*

Аннотация. *в статье анализируются педагогические основы, актуальность и системные вопросы, связанные с вовлечением будущих учителей начальных классов в научно-творческую деятельность. С учетом современных образовательных требований определены ключевые этапы формирования у студентов исследовательских навыков и творческого мышления. Рассматривается роль научно-творческой деятельности в повышении профессиональной компетентности будущих педагогов, а также предлагаются эффективные методики и формы организации такой деятельности.*

Ключевые слова: *научно-творческая деятельность, компетентность, педагогическая проблема, образование, воспитание, методика, актуальный вопрос, исследования, личностные качества.*

Abstract. *this article analyzes the pedagogical foundations, relevance, and systemic issues associated with guiding prospective primary school teachers toward scientific and creative activities. Based on modern educational requirements, the key stages of fostering research skills and creative thinking in students are outlined. The article highlights the role of scientific and creative activities in enhancing the professional competence of future educators and proposes effective methodologies and forms for organizing such activities.*

Keywords: *scientific, creative, competence, pedagogical issue, education, upbringing, methodology, topical issue, research, personal qualities.*

Jahon ta’lim va ilmiy tadqiqot muassasalarida bo‘lajak boshlang‘ich sinf o‘qituvchilarini tayyorlash jarayonining asosiy bosqichi sifatida ko‘rilayotgan ilmiy-ijodiy kompetensiyani rivojlantirish bo‘yicha Bayjing universiteti (Xitoy), Singapur Milliy universiteti hamda ―Ta‘lim sifatini baholash va pedagog kadrlar malakasini oshirish‖ (RF, Janubiy Federal universitet) va ilmiy tadqiqot akademiyalarida bo‘lajak mutaxassislarning pedagogik mahoratini oshirish va ularning intellektual ilmiy-ijodiy kompetentligini takomillashtirish bo‘yicha ilmiy tadqiqotlar amalga oshirilmoqda. Bu borada bo‘lajak o‘qituvchilarning ijodiy kompetentligini rivojlantirishda ta’lim muassasalarida ijodiy yondashuvdan foydalanish, ijodkorlik

kompetensiyasini rivojlantirishning didaktik ta'minotini yaratish, bo'lajak mutaxassislarni tayyorlashda ish beruvchining talablarini inobatga olgan holda kasbiy va shaxsiy sifatlarni innovatsion texnologiyalar asosida rivojlantirish bo'yicha ilmiy tadqiqotlar olib borilmoqda.

O'zbekistonda yosh avlodni mustaqil, ijodiy fikrlashga o'rgatishning davlat darajasidagi masalaga aylanishi bejiz emas. Chunki mustaqil fikr fuqarolarining erkinligi ta'minlangan jamiyatda rivojlanadi, fikrning erkinligi esa jamiyatning kuchli bo'lishi garovidir. O'zbekiston Respublikasi birinchi Prezidenti I.A.Karimov shunday degan edilar: "Demokratik jamiyatda bolalar, umuman, har bir inson erkin fikrlaydigan etib tarbiyalanadi. Agar bolalar erkin fikrlashni o'rganmasa, berilgan ta'lim samarasi past bo'lishi muqarrar. Albatta bilim kerak. Ammo… ijodiy fikrlash katta boylikdir",- deya millat rivoji va jamiyat taraqqiyotida ijodiy fikrlashning ahamiyati nechog'lik zarurligini ko'rsatib berdi[7]. "Ta'lim to'g'risida"gi Qonuni va Milliy dastur hamda uzluksiz ta'lim tizimini modernizatsiya qilish bo'yicha keyingi yillarda amalga joriy etilgan direktiv hujjatlar talablari asosida bo'lajak o'qituvchilarni ijodiy fikrlashga o'rgatish masalasi pedagogika fani va amaliyoti zimmasiga yuklatildi[2].

Ilmiy adabiyotlarda "kompetensiya" va "kompetentlik" tushunchalarini ta'riflashda turli qarashlar mavjud. F.Delamare va J.Winterton kompetensiyani ma'lum bir faoliyat talab qiladigan standart xulq-atvor, o'zini tutish, kompetentlikni ushbu talabga moslik darajasi, ya'ni kompetensiyani namoyish qilishning pirovard natijasi sifatida tavsiflaydi.

Ijodiy qobiliyat pedagogikada keng o'rganilgan bo'lib, ushbu tadqiqot ishida, shaxs, qobiliyat, ijod, ijodiy qobiliyat, o'qituvchining ijodiy qobiliyati tushunchalari o'rganildi. Adabiyotlarda "ijod", "ilmiy ijod" tushunchalari ilmiy kategoriya sifatida o'rganilayotgan bo'lsa-da, hali bu masalaning tadqiqiga umumiy qarash shakllangan emas. Hozirda ko'plab xorijiy mutaxassislar fikriga ko'ra butun dunyoda kompetentlik yondashuv asosida kompetensiyalarni baholashga qiziqtirib borish orqali innovatsion jarayon oliy ta'limning sifatini ta'minlash muhimdir[4].

N.Bayto'rayevaning fikricha, kompetentlikni tushunchalar tizimi asosida talab etiladigan amaliy faoliyatni amalga oshiruvchi qobiliyat orqali yuzaga keladigan muammo va vazifalarni tezlikda hal etish uchun imkoniyat yaratish asosida fikrlashning mos ravishdagi bosqichidir[3]. A.V.Xutorskoy kompetentlik ma'lum bir qolipdagi predmetlar va jarayonlarga doir zarur bo'lgan sifatli mahsullik faoliyati asosida shaxsning o'zaro bog'liq jihatlari jamlanmasi (bilim, ko'nikma va malakalari, faoliyat usullari) deb izoh berilgan [7].

Kompetentlik mutaxassislarning kasbiy tayyorgarligi orqali kasbiy faoliyatga, bilim, ko'nikma va malakalarini o'zlashtirish jarayoni asosida amaliyotga tatbiq etilib, baholanadi. Bu faoliyatning qisman bosqichi har bir shaxs xususiyatlaridan kelib chiqqan holda kasbiy, ijtimoiy, psixologik, pedagogik tayyorgarlik bosqichlariga qo'yiladigan malakaviy talablar bilan mutanosib hisoblanadi[5].

Ijodiy kompetentlilik muammosini tushunish uchun «ijod» va "kompetensiya" kategoriyalarining tarkibiy qismlarini tahlil qilish lozim. Bunda shaxs va ijodning o'zaro bog'liklik faoliyatining yo'llari, ta'lim va pedagogikani turli xil gumanistik falsafa g'oyalarini jamlashtiruvchi ijod muhitiga aylantirish muhim ahamiyat kasb etadi[3].

N.A.Muslimovning fikricha, kompetentlik bu – talabaning shaxsiy va ijtimoiy ahamiyatga ega didaktik faoliyatni amalga oshirilishi uchun zarur bo'lgan bilim, ko'nikma va malakalarning egallanishi hamda ularni didaktik faoliyatda qo'llay olish qobiyati natijasidir. Kompetentlik alohida bilim va malakalarning egallanishi emas, balki har bir mustaqil yo'nalish bo'yicha integrativ bilimlar va harakatlarning o'zlashtirilishidir[6].

Ijod deganda keng ma'noda obyektga (butun atrof muhit yoki uning ozmi, ko'pmi ajratib olingan, alohida bo'lagiga) subyektning (alohida shaxs, ijtimoiy tabaqa, jamiyatning) faol, izchil ta'siri tushuniladi. Mazkur ta'sir jarayonida subyekt o'zini qurshagan muhitni o'zgartiradi, shu paytgacha ko'rilmagan, bilinmagan, o'rganilmagan, g'aroyib va jozibali yangilikni yaratadi yoki kashf etadi. Ijodiy jarayon davomida shaxs ham dunyoni o'zgartiradi, ham o'zini yaratuvchi, ijodkor sifatida kashf etadi. Ijod shaxsning bunyodkorlik, yaratuvchilik salohiyatini ro'yobga chiqaradi. Shaxs ijod usuli bilan yangilik yaratadi, xilma-xil muammolarni qo'yadi va hal qiladi, ularning o'ziga xos yechimlarini, ba'zan bunday yechimlarga nisbatan yondashuvlar, metodlarni topadi. Ijodiy faollikda tadqiqotchi, yaratuvchi, olim shaxsining o'ziga xosligi va betakrorligi namoyon bo'ladi.

Ijod – yaratish, kashf etish; vujudga keltirish. 1) Badiiy, ilmiy, san'atga oid asar yoki moddiy boylik yaratish; 2) yaratuvchilik faoliyati mahsuli, samarasi.

Abu Nasr Farobiyning ta'rifiga ko'ra Ijod – bilish jarayonida shunday ulug' fazilatki, inson uni egallashi uchun boshqa hamma fazilatlarini ishga solishi kerak.

Yuqori sifatli ta'lim natijasiga erishish uchun pedagogik jamoalar tajribasini ta'lim muassasasining ta'lim va tarbiya muhiti orqali ta'lim jarayonidagi ishtirokchilarining ijodiy qobiliyatlarini rivojlantirish uchun yordam berish lozimdir. Bugungi kunda bilimlarni nostandart qo'llash qobiliyatini o'z ichiga oluvchi ijodiy shaxsni rivojlantirishga ko'maklashuvchi bunday muhitni yaratish uchun maqsadga yo'naltirilgan ishlarni tashkil etish zarurati paydo bo'lmoqda. Ijodkorlik ikkita o'zaro aloqadorlikni o'z ichiga oladi: birinchidan, ijodkorlik insonning ijtimoiy ahamiyatida ma'naviy va moddiy qadriyatlar yaratish faoliyati bo'lsa, ikkinchi tomondan natijaga erishish jarayoni, individual, psixologik, intellektual kuchlar va ijodiy qobiliyatlarni amalga oshirish va tasdiqlash jarayoni orqali uning rivojlanishi demakdir.

Ijodiy faoliyatning falsafiy, psixologik, pedagogik jarayonlarini tushuntirishimizda boshqa jihatlari ajralib turiladi. Falsafiy va psixologik tadqiqotlarda ijod faoliyatning o'zini rivojlantirish uchun faollashtirish bilan bog'liqligi sifatida shaxsga xos terminlar, ijodiy jarayonlar va yangilik yaratishda faoliyat natijasi hisoblanadi. Falsafiy adabiyotlarda, ayniqsa, ijodiy faoliyatga oid tadqiqotlar hisobotlarida, ijodkorlikning sub'ektiv asoslarini o'rganishga va eng avvalo, insonning mohiyat kuchlari masalasini ishlab chiqishga katta e'tibor qaratiladi. Ushbu kuchlarning rivojlanganligi va amalga oshirilishi tufayli ijodiy faoliyatni amalga oshirish mumkin bo'ladi.

Ijodiy jarayonning g'oya "g'unchasining" paydo bo'lishidan boshlab, uning bashorat qilib bo'lmaydigan yetilib, tug'ilishi lahzasigacha ko'plab modellari paydo bo'lib, ularni o'rganish bizning tadqiqotlarimizni boyitdi. Grem Uollec ijodkorlikning tayyorgarlik, yetilish, yashnab, ravshanlashish va tekshirish bosqichlarini ajratdi. Bu fazalar ijodiy faoliyatning vazifani belgilash; doimiy ravishda ongsiz kechadigan ijodiy ish; topilgan ongsiz natijani ongga uzatish payti; ongli sinchkovlik bilan tekshirishdan iborat asosiy ketma-ketligini ifoda etadi. Psixologiyada tadqiqotchilar ijodning quyidagi bosqichlarini ajratib ko'rsatadilar: muammoning paydo bo'lgandagi lahzasini anglash; kuchlarni bir joyga yig'ish, ma'lumot qidirish; muammoning yechimi; yechimni tekshirish, o'ylangan rejani batafsil ishlab chiqish; amalga oshirish, yechimni amaliyot vositalari orqali tekshirish[5].

Ijodiy jarayonning tuzilishi, tarkibiy qismlari va bosqichlarining tahlili B.Ziyamuxammedov, Sh.Pozilova, Sh.S.Sharipov, A.Dj.Rou, G.T.Abdullin, C.L.Rubinshteyn, V.A. Slastenin, Y.S. Shelestova va boshqalarning ishlarida berilgan. Ijodiy fikrlash fenomeni

G.A.Sultonova, N.Saidaxmedov, B.S.Abdullayeva, E.G'oziyev O.S.Anisimov, Y.A.Ponomarev, S.L.Rubinshteyn, A.F.Esaulov ning ishlarida bayon qilingan.

Demak, ijodiy kompetentlilik kasbiy ta'lim jarayonida shakllanadi; ta'lim berishning yakuniy natijasini tavsiflash uchun qo'llaniladi; bu rag'batlantiruvchi, kognitiv, qadriyat, shaxsiy-faoliyat sohalarida, jumladan, berilgan sohasida kasbiy faoliyatida minimal tajribada ham shakllangan shaxsiy sifatni anglatadi.

O'qituvchilik kasbi doimo ijodiy va mas'uliyatli hisoblanib kelinganligini ta'kidlab o'tishimiz zarur. O'tmish mutafakkirlarining ilmiy asarlarini o'rganimizda o'qituvchilik kasbini tanlagan inson uchun kasbiy tayyorgarligi va pedagogik faoliyatiga asoslanganligi kuzatiladi. Pedagogik faoliyatga tayyorlanganlik darajasiga erishishning mazmuni va yo'llari ko'plab amaliyotchi o'qituvchilar va nazariy pedagoglar avlodlari tomonidan ishlab chiqilib kelingan. O'qituvchi ishining mazmun tarkibiy qismi (komponenti), pedagogik faoliyat subyektining ajralmas sifat xususiyati sifatida uning kasbiy ko'nikma-malaka xususiyatlarida yetarli darajada to'la namoyon bo'ladi[3].

B.R.Jo'rayeva o'z ilmiy tadqiqotida asosiy e'tiborni pedagogika o'quv yurtlari faoliyatida pedagogika turkum fanlarining tutgan o'rni va ularning imkoniyatlarini o'rganishga qaratib, pedagogik turkum fanlarni o'qitish asosida bo'lajak o'qituvchilarning pedagogik madaniyatini shakllantirish masalasining o'z ijobiy yechimiga ega bo'lishi, o'qituvchilarning pedagogik madaniyatini shakllantirishda uning ijodiy faoliyatiga yo'naltirilganliklari muhim ahamiyatga ega ekanligini ta'kidlaydi. Bizningcha, tadqiqotning nazariy va amaliy ahamiyati muallif tomonidan o'qituvchining pedagogik madaniyati modelining yaratilganligi hamda tadqiqot muammosi mohiyatining yoritilishida texnologik g'oyalar asosida ish ko'rilganligi bilan belgilanadi. Biroq mazkur ilmiy tadqiqot ishida bo'lajak o'qituvchining ilmiy-ijodiy faoliyati shakllanishining pedagogik asoslari alohida tadqiqot muammosi sifatida qaralmagan. Bizning tadqiqotimiz ayni shu jihati bilan xarakterlanadi.

H.Abdukarimov o'z tadqiqot ishida uzluksiz pedagogik ta'lim jarayonida o'qituvchi shaxsining kasbiy tarbiyalash muammosini tadqiq etadi. Muallif bo'lajak mutaxassisga xos bo'lgan kasbiy-pedagogik sifatlarning shakllanishi uning maxsus ta'lim olgunga qadar kechgan kasbiy-pedagogik tayyorgarligiga bevosita bog'liq ekanligini qayd etib o'tadi hamda ushbu tayyorgarlik negizida quyidagi holatlar in'ikosi samarali bo'lishini ko'rsatadi: a) kasb tanlash mayl (motiv)lari; b)tanlangan kasbning asoslanganligi; v)kasbiy intilish, qiziqish va hayotiy rejalarning qat'iyligi; g)bo'lajak o'qituvchining o'z mehnati natijalarini oldindan ko'ra olish qobiliyati; d)shaxsiy-kasbiy takomillashuvni tahlil eta olish qobiliyati; ye)yangi va kasbiy ishchanlik layoqatiga egalik; j)o'qituvchilik kasbining ijtimoiy ahamiyati va o'zining pedagogik faoliyatga munosibligini anglash.

Ta'lim-tarbiya jarayonini samaradorlik darajasini belgilovchi asosiy omillardan biri ta'limni ilmiy-ijodiy asosda tashkil etishdan iboratdir.O'zbekiston Respublikasi mustaqillikka erishgan dastlabki kunlardan boshlab ta'lim sohasida ijtimoiy, ilmiy –ijodiy faol jamiyat a'zolarini shakllantirishni asosiy ta'lim taraqqiyot yo'li sifatida e'tirof etadi, ijtimoiy hayotning barcha sohalarida chuqur va keng ko'lamli islohotlarni amalga oshirish masalasi kun tartibiga qo'yiladi. Ayni vaqtda bu borada samarali faoliyat tashkil etilib, dastlabki yutuqlar qo'lga kiritildi. Mazkur yutuqlardan muhimlari: ta'lim mazmunini belgilashda jahonning ilg'or mamlakatlarining ilm-fan taraqqiyoti tajribasi asosida ta'lim asoslarini talaba va talabalarga o'rgatish kabilar. Chunki ta'limni tashqil etishda ilmiylik prinsiplariga rioya qilish jamiyat va davlat taraqqiyotining asosiy omiliga aylanmoqda. Ta'limni jarayonida ilmiylik prinsiplari

ustuvor o‘rin egallashi, avvalo, talaba va talaba-yoshlar ongida ta’lim mohiyatining to‘g‘ri talqin etilishi, ularga nisbatan hurmat va e’tibor bilan munosabatda bo‘lishdan iboratdir.

Shuning uchun ham hozirda butun e’tibor talaba va talaba-yoshlar orasida ta’limni tashkil etishda talaba faoliyatini ilmiy-ijodiy yondashuv asosida tashkil etishiga oid pedagogik chora-tadbirlarni kuchaytirish, yoshlarning ilmiy-ijodiy faoliyatini shakllantirish va izchil rivojlantirib borish, ularda mustaqil ilmiy-ijodiy faoliyat ko‘rsatish ko‘nikmalarini tarkib toptirishga butun e’tiborni qaratishdan iboratdir. Bo‘lajak o‘qituvchida ilmiy ijodiy faoliyatini shakllantirishning izchil tizimini yaratish, uni nazariy-pedagogik jihatdan asoslash hamda oliy o‘quv o‘quv yurtlari talabalariga ta’lim berish samaradorligini oshirish pedagogika fanida o‘z yechimini kutayotgan dolzarb muammolardan biridir.

Bu sohadagi mavjud manbalar mazmunida talaba hamda talaba-yoshlar orasida ta’limni tashkil etish, ilmiy bilimlarni targ‘ib etish yo‘llari, ilmiy bilimning tashkiliy shakllari, metod va vositalari chuqur tahlil etilishi zarur.

Ilmiy tashkil etilgan ta’lim natijasida shaxsda ijodiy faoliyat shakllanadi. “ilmiy-ijodiy faoliyat” tushunchasining mohiyati, jamiyatda ijtimoiy faol shaxsni shakllantirish jarayoni hamda bu borada namoyon bo‘lgan qarashlar so‘nggi yillarda olib borilgan ushbu muammoga oid tadqiqotlarning asosini tashkil etadi.

FOYDALANILGAN ADABIYOTLAR

1. “Ta’lim to‘g‘risida”gi O‘zbekiston Respublikasi qonuni. 2020-yil 23- sentabr: https://lex.uz/docs/50130071.
2. Karimov I.A.Barkamol avlod orzusi.–T.: “SHarq” nashriyot-matbaa konserni, 1998. 33-b.
3. Bayto‘rayeva N. I. Mustaqil ta’lim faoliyatida bo‘lajak boshlang‘ich sinf o‘qituvchilarining ijodiy kompetentligini rivojlantirish. Ped. fan. dok. (PhD) diss. – Toshkent: 2023, 165b.
4. Berdibekov P., Rashidova G. Ta’lim-tarbiya jarayonini rivojlantirishning ilmiy-pedagogik jihatlari/ Xalq ta’limi. 2016 №2.29-b.
5. Казанцева В.В. Формирование творческой компетентности учителя начальной школы в самообразовательной деятельности. Диссертация на соискание ученой степени кандидата педагогических наук. Воронеж – 2013. 205 с.
6. Muslimov N.A. Kasb ta’limi o‘qituvchilarining didaktik kompetentligini shakllantirish texnologiyasi. –T.: Fan va texnologiya, 2013.– 128 b.
7. Хуторской А.В. Ключевые компетенции как компонент личностно-ориентированной парадигмы образования / Доклады 4-й Всероссийской дистанционной августовской педагогической конференции «Обновление российской школы» (26 августа-10сентября 2002г.). - http://www.eidos.ru/conf/

V INOVATSION TA'LIM SHAROITIDA MAKTABGACHA TA'LIM TAYYORLOV GURUHI BOLALARIDA ATROF MUHITGA MUNOSABATNI SHAKILLANTIRISH

Umarova Dildora Nizom qizi
+998991003498
umarovadildora@icloud.com
https://doi.org/10.5281/zenodo.23255288

Annotatsiya. *Maktabgacha ta'lim muassasasi tarbiyalanuvchilari borliq va unga tutash hodisalar va tabiatni sevishga qaratilgan holatlarni ko'rish mumkin. Inavatsion ta'lim sharoitida tabiatdagi barcha yer osti va yer usti boyliklarini asrab avaylash, bola tarbiyasida tabiatning o'z o'rni borligi va undan oqilona foydalanishni tushuntirish zarurligi alohida to'xtalib o'tilgan bo'lib katta yoshdagi bolalarda inavatsion ta'lim sharoitida tabiat haqidagi tushunchalarini shakllantirish orqali ularda axloqiy sifatlarni tarbiyalashning ahamiyati, tabiat bilan tanishtirish ish shakllari va metodlari haqida ma`lumotlar keltirilgan.*

Kalit so'zlar: *bolalar, maktabgacha ta`lim tashkiloti, inavatsion muhit tabiat, ekologik qadriyat, ekologik bilim, tabiatga muhabbat, inson, rol, muhofaza, fikr, tarix, tanishtirish, metod, fan, qizil kitob, nafosat.*

Аннотация. *В статье рассматриваются способы формирования у детей в дошкольных образовательных учреждениях ценностного отношения к окружающему миру и природным явлениям. В контексте инновационного образования обосновывается необходимость сохранения природных ресурсов (как подземных, так и поверхностных), подчеркивается значимость природы для развития ребенка и необходимость обучения ее рациональному использованию. Кроме того, в статье обсуждается важность воспитания нравственных качеств у детей старшего дошкольного возраста через формирование у них представлений о природе, а также описываются различные формы и методы приобщения детей к миру природы.*

Ключевые слова: *дети, дошкольная образовательная организация, инновационная среда, природа, экологические ценности, экологические знания, любовь к природе, человек, роль, охрана природы, мышление, история, приобщение, метод, наука, Красная книга, эстетика.*

Abstract. *This article examines how children in preschool educational institutions can be encouraged to appreciate the natural world and the phenomena surrounding them. Within the context of innovative education, it highlights the necessity of preserving natural resources—both subterranean and surface—and emphasizes the importance of nature in child development and the need to teach its rational use. Furthermore, the article discusses the significance of fostering moral qualities in older children by shaping their understanding of nature, and outlines various forms and methods for introducing them to the natural world.*

Keywords: *children, preschool educational organization, innovative environment, nature, ecological values, ecological knowledge, love for nature, human, role, conservation, thought, history, introduction, method, science, Red Book, aesthetics.*

Tabiatning jamiyat bilan yaqindan aloqa bog‘laydigan, uning taraqqiyotiga ma‘lum darajada ta‘sir ko‘rsatadigan qismi atrof-muhit deb ataladi. Maktabgacha yoshdagi tayyorlov guruhi bolarini inavatsion muhitda tabiat bilan tanishtirish orqali estetik, aqliy, axloqiy jihatdan tarbiyalash, jismoniy rivojlantirish barcha zamonlarda ham o‘ta dolzarb muammolardan hisoblangan. Tabiat bilan tanishtirish metodika fanida bolalarga ekologik ta`lim-tarbiya berish usullari, mavjud inovatsiyalardan foydalanib «Qizil kitob» haqida qisqacha ma`lumot va «Qizil kitob»ga kiritilgan, noyob o‘simlik va hayvonot olami haqida videoroliklar orqali ham bilim berishga katta o‘rin berilgan.

Tabiat bilan tanishtirish metodikasining asosiy vazifalari quyidagilardan iborat bo‘lishi lozim:

- Bolalarni o‘z o‘lkasini sevishga o‘rgatish.
- O‘zining, tengdoshlari va kattalarning mehnatini qadirlashga o‘rgatish.
- Tabiat orqali psixik jarayonlarni rivojlantirish (sezgilar,idrok, xotira, xayol, tasavvur, nutq diqqat).
- Bolalarni xissiyotini irodasini rivojlantirish.
- Tabiat in`om etgan ne`matlarni avaylab asrashga o‘rgatish.

Yuqoridagi vazifalardan kelib chiqqan xolda, bolalarni har tomonlama inavatsion muhitda tarbiyalashda tabiat orqali – Aqliy tarbiya; - Axloqiy tarbiya; - Nafosat tarbiyasi; - Jismoniy tarbiya; - Ekologik tarbiya; Iqtisodiy tarbiya berish lozim.

Innovatsion tabiat burchagi uchun quyiladigan talablar:

1. O‘simlik yoki hayvon biron bir anik, tur uchun xos bo‘lishi kerak va har biri nomlangan ozi haqida ma’lumot jamlagan qog‘ozchasi bolishi zarur.

2. Tabiat burchagida yashovchilar "beor" ko‘p mehnat talab qilmaydigan, ovqat tanlamaydigan, bolalar yoshiga mos bulishi kerak.

3. Tabiat burchagida yashovchilar yorqin, jozibador, bolalar diqqatini uziga jalb qilib tura oladigan bulishi zarur.

4. Bir turdagi o‘simlik va hayvonlarni bir necha xili bulishi kerak.

5.Tabiat burchagida yashovchilar tamoman xavfsiz va bolalar sog‘ligiga hech qanday zarar etkazmasligi shart.

6.Tabiat burchagida yashovchi o‘simliklar va hayvonlar soni cheklangan va biologik xususiyatlari xisobga olingan holda joylashtirilishi kerak.

O‘rta asrlarda O‘rta Osiyoda yashab ijod etgan olimlardan Muxammad Muso al - Xorazmiy, Abu Nasr Farobiy, Abu Rayhon Beruniy, Abu Ali Ibn Sino va boshqalar tabiat fanining rivojlanishiga ulkan hissa qo‘shganlar. Ular hali ekologiya fani dunyoga kelmagan davrda tabiat va undagi muvozanat, o‘simliklar va hayvonot dunyosi, tabiatni e‘zozlash haqida qimmatli fikrlar aytganlar. Mana shunday zabardast buyuk allomalarimizdan Muhammad Muso al-Xorazmiy, o‘zining "Kitob surat al - arz" asarini Abu Nasr Farobiy (870-810) "al - hajmi va al - mikdor", "Kitob almabodi al - inson" asarlarini Abu Rayhon Beruniy "Saydana", "Minerologiya" asarlarini Abu Ali ibn Sino, u "Kitob al - shifo", "Tadbir ul - manzil" kabi asarlarini yozganlar.

Hadisi sharifimizdagi tabiat, tabiatni avaylab asrash, uni ne’matlari tabiiy boyliklariga ehtiyotkorona munosabatda bo‘lish, parvarish qilishga oid bandlari bilan ham tanishib. uni chuqur tahlil qilib o‘rganib chiqiladi.

Ilm - fan, ma’rifat - madaniyat asrlar davomila insoniyat olamida so‘nmas mash’al bo‘lib, yoritib kelgan. Bizga ma‘lum bo‘lmagan tarix zarvaraqlari qatida pinhona yotgan

qadriyatlarimiz, noyob qo‘lyozma asarlarimiz, qadimiy yodgorliklarimiz istiqlol sharofati bilan tadqiqotchi olimlarimiz tomonidan teran o‘rganilmoqda.

Ushbu mavzuni o‘rganishda biz G‘arb pedagoglaridan: Y.A.Komenskiy, J.J.Russo, I.G.Pestolotsii kabi olimlarning bola tarbiyasida tabiatning tutgan o‘rni haqidagi ilmiy asarlari bilan tanishamiz.

Rus pedagoglaridan K.D.Ushinskiyning, "Inson tarbiya predmeti sifatida" asari bilan tanishamiz.

Ushinskiy "Meni pedagogikada var var hisoblasangiz ham, lekin men o‘z hayot tajribamda shunday xulosaga keldim. Go‘zal tabiat yosh qalbga shunday katta tarbiyaviy ta’sir ko‘rsatadiki, hatto pedagogikaning ta’siri ham u bilan raqobatlashishga ojizdir", -degan edi.

Bizga ma‘lumki, bola dunyoga kelar ekan ilk yoshligidan boshlab tabiat bilan muloqotda bo‘ladi. Biz hozirda o‘tmishdagi pedagoglarimizning dono fikrlari va ilmiy asarlariga suyanganmiz ozimizning pedagogik faoliyatimizda foydalanamiz.

Bolalarni tabiat bilan tanishtirishning ahamiyati. Tabiatga muhabbat - juda keng qamrovli va murakkab his-tuyg‘udir. U yuksak ruhiy va aqliy qatlamlarni o‘z ichiga oladigan murakkab bir butunlikni tashkil etadi. Maktabgacha yoshdagi bolalarda atrof-muxitga mas‘uliyatli munosabatni shakllantirishga doir tadqiqot ishlarini olib borgan M. Umarovaning fikricha, bu hissiyotni tarbiyalashni bolalikning erta yoshidan boshlash muhim pedagogik ahamiyat kasb etadi. Har bir bolada o‘zi yashab turgan mahalliy o‘lka va O‘zbekiston tabiatiga muhabbat bilan qarash, unga ko‘ngil qo‘yish ishtiyoqi o‘sib boradi. Bolalikda tug‘ilgan bu tuyg‘u maktaabgacha ta’lim yoshida rivojlanib, boyib boradi.

Ana shuning uchun ham P.Yusupova o‘zining «Formirovaniye nachal materialisticheskogo miroponimaniya u starshix doshkolnikov v protsesse oznakomleniya ix s rasteniyami v usloviyax Uzbekistana» deb nomlangan tadqiqot ishida metodist olim A.V.Zaporojetsning quyidagi fikrini o‘rinli keltiradi: "... maktabgacha yosh didaktikasini shunday yo‘lga qo‘yish kerakki, bolalar tevarak-atrofdagi buyumlarning tashqi jihatlari bilangina emas, balki ular orasidagi eng oddiy boglanishlar, kuzatilayotgan hodisalarning eng yaqin sabablari, o‘zgarish va rivojlanishlaridagi ayrim xususiyatlar bilan ham tanishsinlar, bunday material saviyaga mos bo‘libgina qolmay, maktabgacha yoshdagi bola uchun qiziqarli hamdir". Bola ana shu bog‘lanishlarni anglash orqali tabiat qonunlarini bilib boradi.

Yuqoridagi fikrlardan kelib chiqib aytish mumkinki, **maktabgacha ta’lim yoshidagi bolalarni inavatsion muhitda tabiat bilan tanishtirish** - bu ularda tabiatga to‘gri munosabatni tarbiyalash va hissiy tajribaga asoslangan, atrof-muhit haqidagi real bilimlarni anglashning ta‘limiy vositasi.

Kuzatishlarning tasdiqlashicha, bolalarning atrof-muhitdagi faoliyati asosan o‘z-o‘zidan, kutilmagan tarzda kechadi va buni ulardagi tabiatga bo‘lgan bunday munosabatning asosiy sababi deb hisoblash to‘gri bo‘ladi Oila, mahalla, bir so‘z bilan aytganda, tevarakdagilarning, ayniqsa, maktabgacha ta‘lim muassasasi tarbiyachilarining pedagogik ta‘siri bilan bunday faoliyat tuzatib boriladi. Mana shuning uchun ham maktabgacha yoshdayoq bolalarning tabiat haqida ishonarli ma‘lumotlarni egallashlari juda muhim. Bolalar tabiat hodisalarini to‘gri idrok etishlari uchun ularni **ertalabki sayr** hamda **tushdan keying sayrlar** to‘g‘ri yo‘lga qo‘yi’ishi zarur. Bolalarni tabiatga oshno etmasdan va maktabgacha ta‘lim muassasasidagi ta‘lim-tarbiya ishlarida undan keng foydalanmasdan, maktabgacha ta‘lim yoshidagi bolalarni har tomonlama - aqliy, estetik, axloqiy, jismoniy jihatdan tarbiyalash vazifasini muvafaqqiyatli hal etib bo‘lmaydi.

Bolalarning innovatsion muhitda tabiat haqidagi bilimlarni egallashlari ularning bilish qobiliyatlari - sensor apparati, mantiqiy fikrlashi, diqqat, nutq, kuzatuvchanlik**,** bilishga qiziqishi kabilar bilan mustahkam bog‘langan bo‘lishi kerak. Fikrlashni rivojlantirish va ilmiy dunyoqarashni shakllantirish uchun bolalarni sayr vaqtida tabiat narsa-hodisalari bilan tanishtirish, ularni kuzatilayotgan narsa-hodisalarni tushunishga o‘rgatish va ular orasidagi aloqa va munosabatlarni anglashlariga olib kelish kerak. Anglash jarayonidagi tabiat hodisalari orasidagi sababiy aloqalar va munosabatlar fikrlashni rivojlantiradi.

Bolalarda mehnatsevarlik, tabiatga ehtiyotkorona munosabatni tarbiyalash uchun ularni o‘simliklarni sugorish va hayvonlarni ovqatlantirishning oddiy usullarini o‘rgatish zarur. Eng muhimi, bolalarning o‘zlari bajarayotgan ish jarayoni va natijalaridan quvonish o‘ziga bo‘lgan ishonchni oshirishga erishish lozim. Mehnatsevarlik sifatini shakllantirish uchun bolalarni kattalarning tabiatdagi mehnati bilan tanishtirib ularga yordamlashish ularning faoliyatiga hurmatni tarbiyalash zarur.

Tabiat bolalarni estetik va aqliy tarbiyalashning asosiy vositalaridan biridir. Tabiatdagi go‘zalliklar hatto eng kichik yoshdagi bolalarni ham o‘ziga rom qilmasdan qolmaydi. Ularni tabiatni bilish jarayonida e‘tiborlarini bog‘u –rog‘larning shovqinlari, qushlarning sayrashi, barglarning shitirlashlari, shakli va rangi, gullarning hidlari, hayvonlarning harakatlariga qaratish kerak. Buning muhim jihati bola kelajakda o‘z hayotida dunyoning xilmaxil go‘zalliklarini idrok eta olishiga sharoit yaratishdir.

Maktabgacha ta‘lim muassasasi tarbiyachilaridan avvalo tabiatda bo‘ladigan voqea hodisalar fasillar almashinuvi to‘g‘risida kerakli elektron ma’lumotlarni to‘plab ularni turli multimediya vositalari bilan qiziqarli shakilda bolalarga taqdim eta olishdir. To‘plangan ma’lumotlardan foydalanib bolalarni innovatsion muhitda tabiat bilan tanishtirish uchun tabiat haqidagi turli fan sohalaridagi katta bilim zahiralariga ega bo‘lishlari, ularning rivojlanish qonuniyatlarini tushunishlari, aql sohasi (noosfera)ning imkoniyatlarini to‘g‘ri baholay olishlari, tabiatda yuz beradigan narsa-hodisalar orasidagi aloqadorliklarni izoxlay olishlari talab etiladi. Ayniqsa, o‘zi yashab turgan joy tabiati: fasl almashinuvlarini kuzata olishi, tabiat kalendarini tuza olishi, o‘simlik va gullarni ko‘paytirish ko‘nikmasiga o‘rganishi muhim.

Bolalarda innovatsion muhitda tabiatga faol va ehtiyotkorona munosabatni tarbiyalash maktabgacha ta‘lim muassasasida mavjud bo‘lgan hayvon va o‘simliklarni oziqlantirishga yordam beradi. Tarbiyachidan jonli tabiat burchagini tashkil eta olish, o‘simlik va hayvonlar uchun ularning tabiatda yashashlariga yaqin shart-sharoitlar yaratish talab etiladi. Bundan tashqari u innovatsion muhitda maktabgacha ta‘lim muassasasi maydonchasi va tabiat burchagida o‘simliklarni sug‘orish va hayvonlarni boqishning agrotexnik usullarini egallay olishi lozim.

FOYDALANILGAN ADABIYOTLAR

1. Maktabagcha ta’lim muassasalarining “Ilk qadam” davlat o‘quv dasturi. -T., -2018.

2. Ilk va maktabgacha yoshdagi bolalar rivojlanishiga qo‘yiladigan davlat talablari. -T., -2018.

3. Djanpeisova G.E. Maktabgacha yoshdagi bolalarning matematik ta’limida modellashtirish usullaridan foydalanish.
Zamonaviy ta’lim. 2017. №5. 94 b.

4. Mirzakarimova A.N. Maktabgacha yoshdagi bolalarda ijodiy xususiyatlarni rivojlantirish. Zamonaviy ta’lim. 2019. №
75-80 b.

AUDIOKORPUS MATERIALLARIGA ISHLOV BERISHNING TEXNOLOGIK BOSQICHLARI: O‘ZBEK TILI MISOLIDA

Urazaliyeva Mavluda Yangiboyevna
PhD, Toshkent davlat Sharqshunoslik universiteti
mavludaurazaliyeva@gmail.com
ORCID: 0009-0004-4673-6539
https://doi.org/10.5281/zenodo.23255353

ANNOTATSIYA. *Maqolada o‘zbek tili audiokorpusi materiallariga ishlov berishning to‘rt bosqichli texnologik modeli tavsiflanadi. Model xom audio materiallarni yig‘ish va ularning texnik formatini standartlashtirish, annotatsiya platformasini (Label Studio) tanlash va korpus interfeysini shakllantirish, ko‘p qatlamli annotatsiya (fonetik va orfografik transkripsiya, dialektal tasniflash, mavzuiy kategoriyalashtirish, lug‘aviy izohlash) hamda metama’lumotga asoslangan qidiruv tizimini yaratish bosqichlarini o‘z ichiga oladi. Ko‘p qatlamli belgilash jarayoni o‘g‘uz lahjasi materiali misolida ko‘rsatiladi. Annotatsiya qatlamlarining o‘zaro mustaqil, ammo bog‘liq holda saqlanishi korpusdan fonetik, dialektologik va sotsiolingvistik tadqiqotlarda qayta foydalanish imkonini kengaytiradi. Tavsiflangan model o‘zbek tili uchun audiokorpuslar yaratishda ochiq metodologiya sifatida qo‘llanilishi mumkin.*

Kalit so‘zlar: *audiokorpus, annotatsiya, ko‘p qatlamli belgilash, Label Studio, metama’lumot, qidiruv tizimi, dialektal tasniflash, fonetik transkripsiya, o‘zbek tili.*

АННОТАЦИЯ. *В статье описывается четырёхэтапная технологическая модель обработки материалов аудиокорпуса узбекского языка. Модель включает сбор исходных аудиоматериалов и стандартизацию их технического формата, выбор платформы аннотирования (Label Studio) и формирование интерфейса корпуса, многоуровневое аннотирование (фонетическая и орфографическая транскрипция, диалектная классификация, тематическая категоризация, лексическое комментирование), а также построение поисковой системы, основанной на метаданных. Процесс многоуровневой разметки демонстрируется на материале огузского диалекта. Независимое, но взаимосвязанное хранение уровней аннотации расширяет возможности повторного использования корпуса в фонетических, диалектологических и социолингвистических исследованиях. Предложенная модель может применяться в качестве открытой методологии при создании аудиокорпусов узбекского языка.*

Ключевые слова: *аудиокорпус, аннотирование, многоуровневая разметка, Label Studio, метаданные, поисковая система, диалектная классификация, фонетическая транскрипция, узбекский язык.*

ABSTRACT. *The paper describes a four-stage technological model for processing audio corpus materials of the Uzbek language. The model comprises the collection of raw audio materials and the standardisation of their technical format, the selection of an annotation platform (Label Studio) and the design of the corpus interface, multi-layer annotation (phonetic and orthographic transcription, dialectal classification, thematic categorisation, lexical commentary), and the construction of a metadata-based retrieval system. The multi-layer annotation procedure is illustrated with material from the Oghuz dialect. Storing annotation layers independently yet in an interrelated manner extends the reusability of the corpus in*

phonetic, dialectological and sociolinguistic research. The proposed model may serve as an open methodology for building Uzbek audio corpora.

Keywords: *audio corpus, annotation, multi-layer markup, Label Studio, metadata, retrieval system, dialectal classification, phonetic transcription, Uzbek language.*

1. Kirish

So‘nggi yillarda til texnologiyalarining jadal rivojlanishi korpus lingvistikasi oldiga yangi vazifalarni qo‘ymoqda: endi korpus nafaqat yozma matnlarni, balki tovush shaklidagi og‘zaki nutqni ham qamrab olishi, bunda esa xom audio material yig‘ilishidan tortib qidiruv tizimida foydalanuvchiga taqdim etilishigacha bo‘lgan barcha bosqichlar aniq texnologik tartibga solinishi talab etiladi. T. McEnery va A. Wilson korpusni maxsus mezonlar asosida tanlangan hamda mashina o‘qiy oladigan shaklda tashkil etilgan matnlar to‘plami sifatida ta’riflaydi [1, 29–32-b.]; ushbu ta’rifni og‘zaki materialga tatbiq etish esa qo‘shimcha texnologik shartlarni yuzaga keltiradi. M. Musaev va hammualliflarining o‘zbek tili uchun USC nutq korpusini yaratish tajribasi [3, 438-b.] ham buni tasdiqlaydi: audiokorpus shunchaki tovush yozuvlari yig‘indisi emas, balki har bir birlik tilshunoslik nuqtai nazaridan izohlangan, qidirilishi va tahlil qilinishi mumkin bo‘lgan tuzilgan resurs bo‘lishi lozim.

O‘zbek tilshunosligida audiomatnlarga asoslangan korpuslarni shakllantirish tajribasi hali dastlabki bosqichda turibdi. Amaldagi uzbekcorpus.uz elektron korpusi asosan yozma matnlarga tayanadi [2, 1–2-b.], og‘zaki matnlar subkorpusini yaratish masalasi esa nazariy jihatdan qo‘yilgan bo‘lsa-da [6, 645–647-b.], uning texnologik zanjiri izchil tavsiflanmagan. Shu bois bunday korpusni yaratish jarayonini boshqa tadqiqotchilar ham qo‘llay oladigan darajada tizimlashtirish alohida amaliy ahamiyat kasb etadi. Ushbu tezisning maqsadi – audiokorpus materiallariga ishlov berishning to‘rt bosqichli texnologik modelini (format standartlashtirish, annotatsiya platformasini tanlash, ko‘p qatlamli belgilash va qidiruv tizimini shakllantirish) yaxlit tarzda tavsiflashdan iborat.

2. Materiallar va metodlar

2.1. Material yig‘ish va formatni standartlashtirish

Texnologik zanjirning birinchi bosqichi xom audio materiallarni turli manbalardan (ommaviy axborot portallari, intervyu yozuvlari, nutq namunalari) yig‘ishni o‘z ichiga oladi. Bu bosqichda hal qilinishi zarur bo‘lgan asosiy masala — audio fayllarning texnik formatini (mono yoki stereo kanal, namunaviy chastota, bit chuqurligi) yagona standartga keltirishdir. USC korpusi mualliflari ham yozuvlarni yagona namunaviy chastota va bit chuqurligiga keltirishni mustaqil tayyorgarlik bosqichi sifatida ajratgan [3, 439–440-b.]. Format bir xilligi ta’minlanmasa, keyingi avtomatlashtirilgan qayta ishlov bosqichlarida (masalan, ovoz sifatini tahlil qilish yoki nutqni segmentlashda) texnik nomuvofiqliklar yuzaga kelishi, bu esa annotatsiya sifatiga bevosita ta’sir qilishi mumkin. Muallifning audio fayllarni saqlash mexanizmlariga bag‘ishlangan oldingi kuzatishlari [7, 41–43-b.] ham aynan shu nomuvofiqlik keyingi barcha bosqichlarga tarqalishini ko‘rsatgan edi. Shu sababli format standartlashtirish nafaqat texnik, balki metodologik ahamiyatga ega bosqich sifatida qaraladi.

2.2. Annotatsiya platformasini tanlash va korpus interfeysini shakllantirish

Standartlashtirilgan audio materiallar korpus interfeysiga – annotatsiya platformasiga – yuklanadi. Ushbu tadqiqotda Label Studio muhiti tanlab olindi, chunki u bitta audio birlik ustida bir vaqtning o‘zida bir nechta bog‘liq bo‘lmagan belgilash turini parallel saqlash imkonini beradi. Platformaning moslashuvchan tuzilishi tufayli annotatsiya sxemasi loyihaning

ehtiyojlariga qarab kengaytirilishi – masalan, yangi dialektal kategoriya yoki mavzuiy yorliq qo‘shilishi – mumkin, bu esa uni katta hajmdagi, uzoq muddatli korpus loyihalari uchun qulay vositaga aylantiradi.

3. Natijalar

3.1. Ko‘p qatlamli annotatsiya jarayoni

Uchinchi, eng ko‘p mehnat talab qiladigan bosqichda har bir audiomatn bir nechta mustaqil qatlam bo‘yicha belgilanadi: (a) fonetik transkripsiya – tovush darajasidagi talaffuz xususiyatlarini, jumladan unlilarning cho‘zilishi va sifat o‘zgarishini qayd etish; (b) orfografik transkripsiya – adabiy me’yordagi yozma ekvivalentni fiksatsiya qilish; (v) dialektal tasniflash – nutqning qaysi lahja guruhiga (qipchoq, qarluq-chigil yoki o‘g‘uz) mansubligini ko‘rsatish; (g) mavzuiy kategoriyalashtirish; (d) lug‘aviy izohlash. Bunday ko‘p qatlamli yondashuv turkiy tillar korpuslarida allaqachon sinovdan o‘tgan: D. Suleymanov va hammualliflari tatar tilining “Tugan Tel” milliy korpusida annotatsiyani bir necha darajada tashkil etishni asoslab bergan [5, 69–71-b.], o‘zbek tili uchun esa shunga yaqin sxema ilgari taklif qilingan edi [6, 646–648-b.]. Quyidagi jadvalda ushbu ko‘p qatlamli annotatsiya jarayonining amaliy namunasi – o‘g‘uz lahjasiga xos uchta so‘z misolida – keltirilgan (1-jadval).

Adabiy shakl	Shevadagi shakl	Transkripsiya	Lahja	Annotatsiya qatlami
bayram	bayräm	bäyräm	o‘g‘uz	fonetik + dialektal
zo‘r	zo‘r	zör	o‘g‘uz	fonetik + dialektal
bo‘ldi	bo‘ldi	boldi	o‘g‘uz	fonetik + dialektal

1-jadval. Ko‘p qatlamli annotatsiya jarayonining namunasi (o‘g‘uz lahjasi materiali asosida)

Jadvaldan ko‘rinib turibdiki, bitta so‘z birligi bir vaqtning o‘zida bir nechta annotatsiya qatlamida – fonetik va dialektal – mustaqil ravishda qayd etiladi. Bunga parallel ravishda so‘zlovchining ijtimoiy-demografik xususiyatlari (jinsi, yoshi, ta’lim darajasi, hududiy kelib chiqishi) ham metama’lumot sifatida tizimga kiritiladi; bu jihatdan tavsiflanayotgan sxema “Tugan Tel” korpusida qo‘llanilgan metama’lumot modeliga [5, 70–72-b.] yaqin turadi. Aynan shu ko‘p qatlamlilik tufayli korpus nafaqat fonetik yoki dialektologik, balki sotsiolingvistik tadqiqotlar uchun ham qimmatli manbaga aylanadi.

3.2. Metama’lumotga asoslangan qidiruv tizimini shakllantirish

Texnologik zanjirning yakuniy bosqichida yig‘ilgan annotatsiya va metama’lumotlar asosida qidiruv tizimi shakllantiriladi. Tatar tili korpusi tajribasida ta’kidlanganidek, metama’lumotga asoslangan filtrlash korpusdan foydalanish samaradorligini sezilarli darajada oshiradi [5, 72–73-b.]. Shunga muvofiq, bu bosqichda foydalanuvchiga audiomatnlarni bir nechta parametr – jins, yosh guruhi, hududiy mansublik yoki mavzu – bo‘yicha filtrlab qidirish imkoni taqdim etiladi; tovush fayllarini saqlash va ularni qayta chaqirib olishning texnik mexanizmlari muallifning oldingi ishida batafsil ko‘rib chiqilgan [7, 44–47-b.]. Natijada xom tovush yozuvi bosqichma-bosqich tuzilgan, izlanuvchan va lingvistik jihatdan izohlangan korpus birligiga aylanadi. Qidiruv tizimining bunday tuzilishi, xususan, dialektologik tadqiqotlarda ma’lum bir hududga yoki demografik guruhga oid materiallarni tezkor ajratib olish imkonini beradi.

4. Muhokama

Tavsiflangan to‘rt bosqichli model bir qator afzalliklarga ega. Birinchidan, annotatsiyaning turli qatlamlari (fonetik, dialektal, mavzuiy, ijtimoiy-demografik) bir-biridan

mustaqil, ammo o‘zaro bog‘liq holda saqlanadi, bu esa korpusdan kelgusida turli maqsadlarda – lingvostatistik tahlil, dialektologik tadqiqot yoki nutqni avtomatik tanish tizimlarini sinovdan o‘tkazish uchun – qayta foydalanish imkonini kengaytiradi; shunga o‘xshash ko‘p maqsadli foydalanish “Tugan Tel” korpusi tajribasida ham qayd etilgan [5, 71–73-b.]. Ikkinchidan, format standartlashtirish bosqichining alohida ajratilishi korpusning texnik barqarorligini ta’minlaydi [7, 46–48-b.] va uni kelgusida kengaytirishni osonlashtiradi. Shu bilan birga, ko‘p qatlamli qo‘lda annotatsiyalash jarayoni sezilarli vaqt va inson resursini talab qiladi; USC korpusini yaratishda ham ona tili sohiblari tomonidan qo‘lda tekshirish eng mehnat talab bosqich bo‘lgani qayd etilgan [3, 441–443-b.]. Bu holat katta hajmdagi korpuslar uchun jarayonni qisman avtomatlashtirish zaruratini keltirib chiqaradi. A. Baevski va hammualliflari taklif qilgan wav2vec 2.0 kabi o‘z-o‘zini nazorat qiluvchi modellar [4, 12450–12452-b.] dastlabki segmentlash va transkripsiya uchun istiqbolli vosita bo‘lishi mumkin – bu masala kelgusi tadqiqotlarning alohida yo‘nalishi sifatida qaralishi lozim.

5. Xulosa

O‘tkazilgan tahlil shuni ko‘rsatadiki, o‘zbek tili audiokorpusini yaratishning texnologik zanjiri – format standartlashtirish, annotatsiya platformasini tanlash, ko‘p qatlamli belgilash va metama’lumotga asoslangan qidiruv tizimi – korpusning ham ilmiy-tadqiqot, ham amaliy-texnologik qiymatini belgilovchi asosiy omil hisoblanadi. Hozirgacha o‘zbek tili uchun yaratilgan korpus resurslari [2, 2–3-b.; 3, 437–438-b.] asosan alohida vazifalarga yo‘naltirilgan bo‘lib, ular o‘rtasidagi metodologik uzilishni bartaraf etishda aynan shunday yaxlit zanjir zarur bo‘ladi. Muallifning oldingi izlanishlari davomi sifatida [6, 648–650-b.; 7, 47–49-b.] taklif etilayotgan to‘rt bosqichli modelni tizimlashtirish va uni ochiq metodologiya sifatida taqdim etish o‘zbek tili uchun audiokorpuslar yaratish sohasini yanada rivojlantirishga, shuningdek kelgusida nutq texnologiyalari (avtomatik nutqni tanish, dialekt identifikatsiyasi) uchun sifatli ma’lumotlar bazasi shakllantirishga xizmat qiladi.

ADABIYOTLAR RO‘YXATI

[1] McEnery T., Wilson A. Corpus Linguistics: An Introduction. – 2nd ed. – Edinburgh: Edinburgh University Press, 2001. – 256 p.

[2] Абдурахмонова Н., Тулиев У., Gatiatullin A.R. Linguistic functionality of Uzbek Electron Corpus: uzbekcorpus.uz // 2021 International Conference on Information Science and Communications Technologies (ICISCT). – 2021. – P. 1–4.

[3] Musaev M., Mussakhojayeva S., Khujayorov I., Khassanov Y., Ochilov M., Atakan Varol H. USC: An Open-Source Uzbek Speech Corpus and Initial Speech Recognition Experiments // Speech and Computer. SPECOM 2021. Lecture Notes in Computer Science, vol. 12997. – Springer, Cham, 2021. – P. 437–447.

[4] Baevski A., Zhou H., Mohamed A., Auli M. Wav2Vec 2.0: Self-supervised learning for speech recognition // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 12449–12460.

[5] Suleymanov D., Nevzorova O., Gatiatullin A., Gilmullin R., Khakimov B. National corpus of the Tatar language “Tugan Tel”: grammatical annotation and implementation // Procedia – Social and Behavioral Sciences. – 2013. – Vol. 95. – P. 68–74.

[6] Abduraxmonova N.Z., Urazaliyeva M.Y. O‘zbek tili elektron korpusida (http://uzbekcorpus.uz/) og‘zaki matnlar korpusini yaratishning nazariy va amaliy masalalari // Academic Research in Educational Sciences. – 2022. – Vol. 3, № 3. – P. 644–650.

SUN'IY INTELLEKT TEXNOLOGIYALARINING ZAMONAVIY MUAMMOLARI: IJTIMOIY, AXLOQIY VA TEXNOLOGIK TAHLIL

Xanxodjayeva Gavxar
O'zbekiston Milliy universiteti magistranti
https://doi.org/10.5281/zenodo.23255403

Annotatsiya: *Ushbu maqolada sun'iy intellekt texnologiyalarining jadal rivojlanishi va zamonaviy jamiyatga integratsiyalashuvi natijasida yuzaga kelayotgan ijtimoiy, axloqiy va texnologik muammolar kompleks tarzda tahlil qilinadi. Tizimli yondashuv va qiyosiy tahlil usullari asosida algoritmik shaffoflik, algoritmik diskriminatsiya va adolat, shaxsiy ma'lumotlar maxfiyligi, kiberxavfsizlik, dezinformatsiya hamda mehnat bozorining transformatsiyasi masalalari o'rganildi. Shuningdek, ta'lim tizimida sun'iy intellektdan foydalanishning imkoniyat va xavflari, noto'g'ri qarorlar uchun javobgarlik masalasi ko'rib chiqildi. Tadqiqot natijasida sun'iy intellektni nafaqat texnik yangilik, balki ijtimoiy institutlarga ta'sir etuvchi murakkab raqamli fenomen sifatida baholash zarurligi asoslandi va mas'uliyatli sun'iy intellekt kontseptsiyasini rivojlantirishga doir takliflar ilgari surildi.*

Kalit so'zlar: *sun'iy intellekt, mashinali o'qitish, chuqur o'qitish, algoritmik shaffoflik, algoritmik adolat, ma'lumotlar maxfiyligi, raqamli etika.*

Аннотация. *В статье проводится комплексный анализ социальных, этических и технологических проблем, возникающих в результате стремительного развития и интеграции технологий искусственного интеллекта (ИИ) в современное общество. На основе системного и сравнительного анализа научной литературы рассматриваются вопросы алгоритмической прозрачности, алгоритмической дискриминации и справедливости, конфиденциальности персональных данных, кибербезопасности, дезинформации, порождаемой генеративным ИИ, а также трансформации рынка труда. Особое внимание уделено возможностям и рискам применения ИИ в образовании и вопросам этической и правовой ответственности за решения, принимаемые системами ИИ. Обосновывается необходимость рассмотрения ИИ не только как технологической инновации, но и как сложного цифрового феномена, влияющего на социальные институты и человеческие отношения. По результатам исследования предложены практические рекомендации по развитию концепции ответственного искусственного интеллекта.*

Ключевые слова: *искусственный интеллект, машинное обучение, глубокое обучение, нейронные сети, генеративный ИИ, алгоритмическая прозрачность, цифровая этика.*

Abstract. *This article provides a comprehensive analysis of the social, ethical, and technological problems arising from the rapid development and integration of artificial intelligence (AI) technologies into modern society. Based on a systematic and comparative analysis of scientific literature, the study examines issues of algorithmic transparency, algorithmic discrimination and fairness, personal data privacy, cybersecurity, disinformation generated by generative AI, and the transformation of the labor market. Particular attention is paid to the opportunities and risks of using AI in education and to the ethical and legal*

responsibility for decisions made by AI systems. The article substantiates the need to view AI not merely as a technological innovation, but as a complex digital phenomenon that affects social institutions and human relations. Based on the findings, practical recommendations are proposed for developing a responsible AI concept.

Keywords: *artificial intelligence, machine learning, deep learning, neural networks, generative AI, algorithmic transparency, digital ethics.*

Ming yillar davomida insoniyat aqliy faoliyatining to‘planishi va rivojlanishi insoniyatga cheksiz imkoniyatlar yaratib bermoqda. Biz yashayotgan davr insoniyatning sun’iy intellekt tizimlari bilan faol hamkorlik qiladigan davr bilan ustma-ust tushmoqda. Ushbu maqolada sun’iy intellekt texnologiyalarining jadal rivojlanishi va ularning zamonaviy jamiyat hayotiga integratsiyalashuvi bilan bog‘liq ijtimoiy, iqtisodiy, axloqiy, huquqiy va texnologik muammolar kompleks tarzda tadqiq etiladi. Sun’iy intellektning mashinali o‘qitish, chuqur o‘qitish, neyron tarmoqlar, tabiiy tilni qayta ishlash va generativ modellar kabi texnologik yo‘nalishlari inson faoliyatining turli sohalarida yangi imkoniyatlarni shakllantirmoqda [Mitchell, 1997]. Biroq mazkur texnologiyalarning keng joriy etilishi algoritmik shaffoflik, ma’lumotlar maxfiyligi, mehnat bozorining transformatsiyasi hamda insonning texnologik tizimlarga bog‘liqligi kabi muammolarni ham yuzaga keltirmoqda. Tadqiqotda sun’iy intellektni texnologik innovatsiya sifatidagina emas, balki ijtimoiy institutlar va inson munosabatlariga ta’sir etuvchi murakkab raqamli fenomen sifatida baholash zarurligi asoslanadi.

XXI asrning texnologik taraqqiyot tendentsiyalari raqamli transformatsiya jarayonlarining jadallashuvi bilan tavsiflanmoqda. Raqamlashtirishning navbatdagi bosqichida sun’iy intellekt texnologiyalari iqtisodiyot, ta’lim, tibbiyot, jurnalistika, davlat boshqaruvi, sanoat, moliya va ijtimoiy kommunikatsiya tizimlariga tobora chuqur kirib bormoqda. Sun’iy intellekt insonning kognitiv faoliyatiga xos bo‘lgan ayrim jarayonlarni — ma’lumotlarni tahlil qilish, tasniflash, prognozlash, tilni qayta ishlash, obrazlarni tanish va qaror qabul qilishni — kompyuter algoritmlari asosida amalga oshirish imkonini beruvchi texnologik tizim sifatida shakllanmoqda. Xalqimizda «Ilm — taraqqiyotning poydevori» degan hikmat mavjud. Bugungi raqamli tsivilizatsiya sharoitida mazkur fikrning mazmuni yanada kengayib, ilm-fan va innovatsion texnologiyalarni samarali o‘zlashtirish davlat va jamiyat taraqqiyotining muhim omiliga aylanmoqda. Sun’iy intellektning rivojlanishi inson mehnatini avtomatlashtirish, qaror qabul qilish samaradorligini oshirish va katta hajmdagi ma’lumotlarni qisqa vaqt ichida qayta ishlash imkonini bermoqda. Biroq texnologiyaning imkoniyatlari qanchalik kengaysa, uning ijtimoiy va axloqiy oqibatlarini ilmiy jihatdan baholash zarurati ham shunchalik ortadi. Shu nuqtai nazardan, sun’iy intellekt fenomenini faqat texnik taraqqiyot natijasi sifatida emas, balki ijtimoiy-texnologik transformatsiyaning muhim determinanti sifatida tadqiq etish dolzarb ilmiy masala hisoblanadi.

Tadqiqotning asosiy maqsadi — sun’iy intellekt texnologiyalarining zamonaviy jamiyatda qo‘llanilishi natijasida shakllanayotgan asosiy muammolarni tizimli va kompleks tahlil qilishdan iborat. Ushbu maqsadga erishish uchun quyidagi vazifalar belgilandi: sun’iy intellekt texnologiyalarining zamonaviy rivojlanish tendentsiyalarini tahlil qilish; shaffoflik va adolat muammolarini o‘rganish; shaxsiy ma’lumotlar maxfiyligi va axborot xavfsizligi masalalarini tahlil qilish; dezinformatsiya muammosini tadqiq etish; ta’lim jarayonida sun’iy intellektdan foydalanishning imkoniyat va xavflarini aniqlash; mas’uliyatli sun’iy intellekt kontseptsiyasining ahamiyatini asoslash.

Mazkur tadqiqotda tizimli yondashuv, qiyosiy tahlil, ilmiy adabiyotlarni kontent-tahlil qilish, mantiqiy umumlashtirish va kontseptual tahlil usullaridan foydalanildi. Sun'iy intellekt bilan bog'liq muammolar texnologik, ijtimoiy, iqtisodiy, huquqiy va axloqiy omillarning o'zaro ta'siri nuqtai nazaridan o'rganildi. Metodologik jihatdan tadqiqotda «mas'uliyatli sun'iy intellekt», «inson-mashina hamkorligi», «algoritmik adolat», «raqamli etika» va «texnologik mas'uliyat» kontseptsiyalariga tayanildi.

1. Sun'iy intellektning texnologik transformatsiyadagi o'rni

Sun'iy intellektning zamonaviy rivojlanishi mashinali o'qitish va neyron tarmoq texnologiyalarining takomillashuvi bilan bevosita bog'liq. Mashinali o'qitish algoritmlari katta hajmdagi ma'lumotlardan muayyan qonuniyatlarni aniqlash va ular asosida prognoz yoki qaror shakllantirish imkonini beradi. Chuqur o'qitish texnologiyalari esa ko'p qatlamli neyron tarmoqlardan foydalangan holda murakkab ma'lumotlarni qayta ishlash imkoniyatlarini kengaytiradi. Generativ sun'iy intellektning paydo bo'lishi esa mazkur jarayonni yangi bosqichga olib chiqdi: generativ modellar matn, tasvir, audio, video va dasturiy kod kabi kontentlarni yaratish imkoniyatiga ega [Goodfellow, Bengio, Courville, 2016]. Shu bilan birga, texnologik imkoniyat va ijtimoiy mas'uliyat o'rtasidagi muvozanatni ta'minlash zamonaviy raqamli siyosatning muhim vazifasiga aylanmoqda [Russell, Norvig, 2020].

2. Algoritmik shaffoflik muammosi

Sun'iy intellekt tizimlarining muhim muammolaridan biri — algoritmlar tomonidan qabul qilinayotgan qarorlarning tushunarli va izohlanadigan bo'lishidir. Ilmiy adabiyotlarda bu holat ko'pincha «qora quti» (black box) muammosi sifatida talqin qilinadi. Mazkur muammo tibbiyot, moliya, ta'lim va ishga qabul qilish kabi inson hayotiga bevosita ta'sir ko'rsatuvchi sohalarda yanada muhim ahamiyat kasb etadi. Shu sababli Explainable AI (XAI), ya'ni tushuntiriladigan sun'iy intellekt kontseptsiyasini rivojlantirish muhim ilmiy yo'nalishlardan biri hisoblanadi: algoritm natijasi nafaqat aniq, balki tushuntiriladigan, tekshiriladigan va asoslanadigan bo'lishi lozim [Floridi, Cowls, 2019].

3. Algoritmik diskriminatsiya va adolat muammosi

Sun'iy intellekt modellarining sifati ularni o'qitishda foydalanilgan ma'lumotlarning sifatiga bevosita bog'liq. Agar ma'lumotlar bazasida statistik nomutanosiblik yoki biryoqlama yondashuv mavjud bo'lsa, algoritm ushbu xususiyatlarni o'zlashtirishi va qaror qabul qilish jarayonida qayta ishlab chiqarishi mumkin. Bu holat algoritmik diskriminatsiya deb ataluvchi muammoni yuzaga keltiradi [Jobin, Ienca, Vayena, 2019]. Shuning uchun sun'iy intellekt tizimlarini ishlab chiqishda ma'lumotlarning reprezentativligi, modellarning adolatliligi va natijalarning muntazam auditi ta'minlanishi zarur. «Adolat — jamiyat barqarorligining asosiy mezoni» degan ilmiy-ijtimoiy tamoyil sun'iy intellekt algoritmlariga nisbatan ham tatbiq etilishi mumkin.

4. Shaxsiy ma'lumotlar maxfiyligi va kiberxavfsizlik

Sun'iy intellekt texnologiyalarining samaradorligi ko'p hollarda katta hajmdagi ma'lumotlarga tayanadi. Shu bilan birga, ma'lumotlarning haddan tashqari markazlashuvi shaxsiy hayot daxlsizligi va kiberxavfsizlik bilan bog'liq xavflarni oshiradi. Shaxsiy ma'lumotlarning ruxsatsiz yig'ilishi, qayta ishlanishi yoki tarqatilishi insonning raqamli huquqlariga salbiy ta'sir ko'rsatishi mumkin. Mazkur kontekstda data privacy, data governance, cybersecurity va information security kabi tushunchalar alohida ilmiy-amaliy ahamiyatga ega [NIST, 2023; European Union, 2024]. Shuning uchun sun'iy intellekt tizimlarida ma'lumotlarni

minimallashtirish, maqsadli qayta ishlash, kriptografik himoya va foydalanuvchining xabardorligini ta'minlash muhimdir.

5. Generativ sun'iy intellekt va dezinformatsiya

Generativ sun'iy intellekt axborot ishlab chiqarish jarayonini sezilarli darajada tezlashtirmoqda. Biroq uning muhim muammolaridan biri — faktik jihatdan noto'g'ri, ammo ishonchli ko'rinishda taqdim etilgan axborotni shakllantirish ehtimolidir. Ayrim generativ modellar mavjud bo'lmagan faktlar yoki manbalarni ishonchli uslubda taqdim qilishi mumkin; bu holat ilmiy adabiyotlarda AI hallucination tushunchasi bilan ifodalanadi [Kaplan, Haenlein, 2019]. Bundan tashqari, sintetik media va deepfake texnologiyalari orqali real insonlarning ovozi, tasviri yoki harakatlariga o'xshash kontent yaratish imkoniyati axborot xavfsizligi uchun yangi chaqiriqlarni yuzaga keltirmoqda. Shu sababli zamonaviy axborot muhitida mediaverifikatsiya, faktcheking va raqamli mediasavodxonlikning ahamiyati tobora ortib bormoqda. Xalqimizda «Yetti o'lchab, bir kes» degan maqol bor; raqamli axborot muhitida bu tamoyilni «axborotni tekshirib, keyin ishonish» printsipi sifatida talqin qilish mumkin.

6. Sun'iy intellektning mehnat bozoriga ta'siri

Sun'iy intellektning avtomatlashtirish imkoniyatlari mehnat bozorining institutsional tuzilishiga ta'sir ko'rsatmoqda. Takroriy, standartlashtirilgan va ma'lum darajada algoritmlashtirilishi mumkin bo'lgan vazifalar avtomatlashtirilishi ehtimoli yuqori. Biroq bu jarayonni faqat ish o'rinlarining qisqarishi bilan izohlash biryoqlama yondashuv bo'ladi: sun'iy intellekt yangi kasblar, yangi mehnat funktsiyalari va yangi kompetentsiyalarga bo'lgan talabni ham shakllantiradi. Shu munosabat bilan mehnat bozorida reskilling va apskilling jarayonlarini rivojlantirish muhim ahamiyat kasb etadi. Demak, kelajak mehnat bozorida inson va texnologiya raqobati emas, balki inson-mashina sinergiyasi ustuvor modelga aylanishi mumkin.

7. Ta'lim tizimida sun'iy intellektdan foydalanish

Sun'iy intellekt ta'lim tizimida o'quv materiallarini tahlil qilish, ta'lim natijalarini baholash va o'qituvchi faoliyatini qo'llab-quvvatlashda katta imkoniyatlarga ega. Shu bilan birga, undan ortiqcha foydalanish o'quvchi va talabalarda kognitiv faollikning pasayishiga olib kelishi ehtimoli mavjud. Shuning uchun ta'limda sun'iy intellektning asosiy vazifasi inson intellektual faoliyatini almashtirish emas, balki uni qo'llab-quvvatlashdan iborat bo'lishi kerak. «Bilim — kuch, ammo bilimdan to'g'ri foydalanish — mas'uliyat» degan tamoyil raqamli ta'lim muhitida alohida ahamiyat kasb etadi.

8. Sun'iy intellektning axloqiy va huquqiy muammolari

Sun'iy intellektning axloqiy va huquqiy jihatdan eng murakkab masalalaridan biri — sun'iy intellekt tomonidan qabul qilingan noto'g'ri qaror uchun kim javobgar bo'lishi masalasidir. Bunda algoritmik javobgarlik, texnologik mas'uliyat va inson nazorati muhim ahamiyatga ega. Sun'iy intellekt tizimlari inson huquqlari va qadr-qimmatini himoya qilish tamoyillariga zid kelmasligi kerak [UNESCO, 2021]. Ayniqsa, yuqori xavfli sohalarda inson nazoratini saqlab qolish printsipi muhimdir [OECD, 2019].

9. Sun'iy intellekt va inson-mashina hamkorligi

Inson-mashina hamkorligining asosiy maqsadi insonni mashina bilan almashtirish emas, balki ularning kuchli jihatlarini birlashtirishdan iborat. Demak, sun'iy intellektning istiqboli insonni to'liq almashtirishdan ko'ra, inson salohiyatini kengaytirish texnologiyasi sifatida namoyon bo'lishi ehtimoldan xoli emas. Generativ sun'iy intellektning rivojlanishi axborot ishlab chiqarish samaradorligini oshirgani holda, dezinformatsiya va sintetik kontent xavfini ham

kuchaytirmoqda. Muammoning optimal yechimi texnologik, huquqiy, institutsional, axloqiy va ta'limiy mexanizmlarning integratsiyalashgan tizimini shakllantirishdan iborat.

Ilmiy-amaliy takliflar

Tadqiqot natijalari asosida quyidagi takliflarni ilgari surish mumkin:

– sun'iy intellekt tizimlarini xavf darajasi bo'yicha tasniflash va yuqori xavfli tizimlar uchun majburiy algoritmik audit tizimini joriy etish;

– algoritmlarning shaffofligi, tushuntirilishi va tekshirilishiga qaratilgan Explainable AI mexanizmlarini rivojlantirish;

– ma'lumotlar maxfiyligi va kiberxavfsizlikni ta'minlaydigan texnik va huquqiy vositalarni takomillashtirish;

– ta'lim tizimida sun'iy intellektdan foydalanish bo'yicha uslubiy ko'rsatmalar va etik standartlarni ishlab chiqish;

– mas'uliyatli sun'iy intellekt kontseptsiyasini amaliyotga joriy etish orqali inson nazoratini saqlab qolish.

Xulosa

«Har bir yangilik imkoniyat bilan birga mas'uliyat ham olib keladi» degan tamoyil sun'iy intellekt rivojiga nisbatan ham dolzarbdir. Mazkur texnologiyalar katta hajmdagi ma'lumotlarni qayta ishlash, prognozlash, avtomatlashtirish va qaror qabul qilish samaradorligini oshirish orqali inson faoliyati uchun yangi imkoniyatlar yaratmoqda. Shu bilan birga, sun'iy intellektni samarali boshqarish uchun texnologik innovatsiyalar bilan bir qatorda raqamli etika, algoritmik adolat, shaffoflik, inson nazorati, ma'lumotlar xavfsizligi va huquqiy tartibga solish mexanizmlari ham rivojlantirilishi zarur.

FOYDALANILGAN ADABIYOTLAR

1. Russell S., Norvig P. Artificial Intelligence: A Modern Approach. 4th ed. – Pearson, 2020.
2. UNESCO. Recommendation on the Ethics of Artificial Intelligence. – Paris: UNESCO, 2021.
3. OECD. Recommendation of the Council on Artificial Intelligence. – OECD, 2019.
4. NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). – National Institute of Standards and Technology, 2023.
5. European Union. Artificial Intelligence Act. – European Union, 2024.
6. Floridi L., Cowls J. A Unified Framework of Five Principles for AI in Society // Harvard Data Science Review, 2019.
7. Jobin A., Ienca M., Vayena E. The Global Landscape of AI Ethics Guidelines // Nature Machine Intelligence, 2019.
8. Mitchell T. M. Machine Learning. – McGraw-Hill, 1997.
9. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press, 2016.
10. Kaplan A., Haenlein M. Siri, Siri, in my hand: Who's the fairest of them all? // Business Horizons, 2019.

TALABALAR YOZMA ISHLARIDAGI GRAMMATIK XATOLARNI O'ZBEK TILIDA AVTOMATIK ANIQLASH VA TUZATISH TIZIMINING YETISHMASLIGI

Xojimurodova Ominaxon Bobomurod qizi
e-mail: ominaxonxojimurodova270@gmail.com
+998999874478
https://doi.org/10.5281/zenodo.23255567

Annotatsiya. *Ushbu maqolada talabalarning o'zbek tilidagi yozma ishlarida uchraydigan grammatik xatolarni avtomatik aniqlash va tuzatish muammosi ko'rib chiqiladi. Hozirgi kunda ingliz va rus tillari uchun bunday tizimlar yetarlicha rivojlangan bo'lsa-da, o'zbek tili uchun bunday vosita amalda deyarli mavjud emas. Tadqiqotda mavjud xalqaro tajribalar tahlil qilinib, o'zbek tilining morfologik xususiyatlari sababli yuzaga keladigan qiyinchiliklar aniqlandi. Natijada, qoidaga asoslangan va statistik usullarni birlashtirgan gibrid yondashuv taklif etildi. Ushbu yechim ta'lim muassasalarida amaliy jihatdan joriy etilishi mumkin. Xulosa qilib aytganda, milliy tilga moslashtirilgan grammatika tekshiruvchisi ta'lim sifatini oshirishda muhim vosita bo'la oladi.*

Kalit so'zlar: *tabiiy tilni qayta ishlash, grammatik xatolarni aniqlash, o'zbek tili morfologiyasi, avtomatik tuzatish tizimi, ta'lim texnologiyalari.*

Аннотация. *В данной статье рассматривается проблема автоматического обнаружения и исправления грамматических ошибок в письменных работах студентов на узбекском языке. Хотя для английского и русского языков подобные системы уже достаточно развиты, для узбекского языка такой инструмент практически отсутствует. В исследовании проанализирован имеющийся международный опыт и выявлены трудности, связанные с морфологическими особенностями узбекского языка. По результатам предложен гибридный подход, объединяющий методы, основанные на правилах, и статистические методы. Данное решение может быть внедрено в образовательных учреждениях на практике. В заключение отмечается, что система проверки грамматики, адаптированная к национальному языку, способна повысить качество образования.*

Ключевые слова: обработка естественного языка, обнаружение грамматических ошибок, морфология узбекского языка, система автоматической коррекции, образовательные технологии.

Abstract. *This article examines the problem of automatically detecting and correcting grammatical errors in students' written works in the Uzbek language. While such systems are already well developed for English and Russian, a comparable tool is nearly absent for Uzbek. The study analyzes existing international experience and identifies difficulties arising from the morphological features of the Uzbek language. Based on the findings, a hybrid approach combining rule-based and statistical methods is proposed. This solution can be practically implemented in educational institutions. In conclusion, a grammar-checking system adapted to the national language can serve as an important tool for improving the quality of education.*

Keywords: natural language processing, grammatical error detection, Uzbek language morphology, automatic correction system, educational technology.

Kirish

Bugungi kunda raqamli transformatsiya ta'lim tizimining barcha bosqichlariga kirib bormoqda. Talabalarning yozma nutq malakasini rivojlantirish oliy ta'lim sifatini belgilovchi muhim omillardan biri hisoblanadi. Biroq, talabalar tomonidan yoziladigan insho, referat va kurs ishlarida grammatik xatolar tez-tez uchraydi. Bunday xatolarni qo‘lda tekshirish o‘qituvchidan ko‘p vaqt va kuch talab qiladi.

Xorijiy tillar, xususan ingliz va rus tillari uchun grammatik xatolarni avtomatik aniqlaydigan tizimlar allaqachon keng qo‘llanilmoqda. Bu tizimlar tabiiy tilni qayta ishlash (Natural Language Processing, NLP) sohasining yutuqlariga asoslanadi. NLP sohasining nazariy asoslari N. Chomskiyning generativ grammatika nazariyasidan boshlangan bo‘lib, u tilning ichki grammatik qoidalarini formal tarzda tavsiflash mumkinligini ko‘rsatib bergan edi [1]. Keyinchalik bu g‘oyalar hisoblash lingvistikasi va zamonaviy NLP tizimlari rivojlanishiga asos bo‘ldi [2].

Biroq, o‘zbek tili uchun bunday tizim hozircha amalda mavjud emas. O‘zbek tilining boy morfologik tuzilishi, ya'ni bitta so‘z ildizidan o‘nlab shakllar yasalishi, mavjud xorijiy vositalarni to‘g‘ridan-to‘g‘ri qo‘llash imkonini bermaydi [3]. Shu sababli, ushbu maqolada talabalar yozma ishlaridagi grammatik xatolarni o‘zbek tilida avtomatik aniqlash va tuzatish tizimining yetishmasligi muammosi ko‘rib chiqiladi hamda uni bartaraf etishning amaliy yo‘llari taklif etiladi.

Tadqiqot uslubi

Tadqiqotda qiyosiy tahlil va adabiyotlarni tizimli ko‘rib chiqish usullaridan foydalanildi. Dastlab, ingliz va boshqa tillar uchun yaratilgan mavjud grammatika tekshiruvchi tizimlarning ishlash tamoyillari o‘rganildi [4]. Bu tizimlar odatda ikki asosiy yondashuvga tayanadi: qoidaga asoslangan usul va statistik, ya'ni neyron tarmoqqa asoslangan usul.

Qoidaga asoslangan usulda til mutaxassislari tomonidan tuzilgan grammatik qoidalar ro‘yxatidan foydalaniladi. Statistik usulda esa katta hajmdagi matn korpusi asosida model o‘qitiladi, u to‘g‘ri va noto‘g‘ri qurilgan gaplar orasidagi farqni ehtimollik hisob-kitoblari orqali aniqlaydi [5]. So‘nggi yillarda bu ikki yondashuvni birlashtirgan gibrid modellar ham keng tarqalmoqda [6].

Shuningdek, o‘zbek tilining morfologik xususiyatlarini hisobga olgan holda, mavjud morfologik tahlilchilar haqidagi ilmiy manbalar ko‘rib chiqildi. Turkiy tillar guruhiga oid, resurslari kam bo‘lgan tillar uchun yaratilgan tizimlar tajribasi alohida tahlil qilindi, chunki ular o‘zbek tiliga tuzilishi jihatidan yaqin turadi [7].

Natijalar va tahlil

O‘tkazilgan tahlil natijasida bir qancha muhim omillar aniqlandi.

Birinchidan, o‘zbek tili qo‘shimchalovchi (agglyutinativ) til turiga kiradi. Bitta fe'l yoki ot ildizidan grammatik qoidalarga ko‘ra o‘nlab, hatto yuzlab so‘z shakllari hosil bo‘lishi mumkin. Bu holat so‘zlarni avtomatik tahlil qilishda katta miqdordagi variantlarni hisobga olishni talab qiladi, aks holda tizim ko‘plab noto‘g‘ri signal berishi mumkin.

Ikkinchidan, o‘zbek tilida grammatik xatolar belgilab chiqilgan, katta hajmdagi maxsus matn korpusi mavjud emas. Zamonaviy statistik yoki neyron tarmoqqa asoslangan tizimlarni yaratish uchun aynan shunday korpus zarur bo‘ladi [8]. Korpusning yo‘qligi sababli, mavjud xalqaro modellarni o‘zbek tiliga to‘g‘ridan-to‘g‘ri moslashtirib bo‘lmaydi.

Uchinchidan, talabalar yozma ishlarida ko‘pincha o‘zbek, rus va ingliz tillari elementlari aralashib keladi. Bu holat, ayniqsa ilmiy-texnik matnlarda, grammatik tahlilni yanada murakkablashtiradi.

To‘rtinchidan, mavjud oddiy imlo tekshiruvchilar, masalan matn muharrirlaridagi standart vositalar, faqat so‘z darajasidagi xatolarni aniqlaydi. Ular gap tuzilishi yoki so‘z birikmalaridagi grammatik nomuvofiqliklarni aniqlay olmaydi.

Amaliy joriy etish yechimlari

Yuqoridagi tahlil asosida quyidagi amaliy yechimlar taklif etiladi.

Birinchi qadam sifatida, talabalarning haqiqiy yozma ishlaridan iborat, grammatik xatolar belgilab chiqilgan maxsus o‘quv korpusini yaratish lozim. Bu korpus kelajakda istalgan avtomatik tizimni o‘qitish uchun asos bo‘lib xizmat qiladi. Korpusni yaratishda til mutaxassislari va IT mutaxassislarining hamkorligi zarur.

Ikkinchi qadam — gibrid tizim ishlab chiqish. Boshlang‘ich bosqichda mavjud o‘zbek tili morfologik qoidalari asosida qoidaga asoslangan modul yaratiladi. Bu modul kelishik qo‘shimchalari, fe’l shakllari, ko‘plik-birlik mosligi kabi tez-tez uchraydigan xatolarni aniqlaydi. Korpus hajmi yetarli darajaga yetgach, statistik yoki neyron tarmoqqa asoslangan modul qo‘shib boriladi, bu esa tizim aniqligini oshiradi.

Uchinchi qadam — tizimni amaliyotga tadbiq etish. Yaratilgan vosita dastlab matn muharriri uchun qo‘shimcha dastur yoki oddiy veb-sahifa shaklida ishlab chiqilishi mumkin. Bu talabalarga o‘z ishlarini topshirishdan oldin mustaqil tekshirib olish imkonini beradi.

To‘rtinchi qadam — ta’lim jarayoniga integratsiya. Oliy ta’lim muassasalari ushbu vositani masofaviy ta’lim platformalariga joriy etishi mumkin. Bu o‘qituvchining yuklamasini kamaytiradi va talabaga o‘z xatolari ustida mustaqil ishlash imkonini yaratadi.

Xulosa

Shaxsiy fikrimcha, o‘zbek tili uchun grammatik xatolarni avtomatik aniqlash tizimining yaratilishi endi ixtiyoriy emas, balki zaruriy vazifaga aylanmoqda. Raqamli ta’lim tizimlari tez sur’atlar bilan rivojlanayotgan bir paytda, davlat tilida ishlaydigan bunday vositaning yo‘qligi jiddiy bo‘shliq hisoblanadi. Ushbu muammoni bosqichma-bosqich, avval maxsus korpus yaratish, so‘ngra gibrid model ishlab chiqish orqali hal qilish mumkin, deb hisoblayman. Bu yo‘nalishdagi tadqiqotlar nafaqat ta’lim sifatini oshiradi, balki o‘zbek tilini raqamli makonda mustahkamlashga ham xizmat qiladi. Kelgusi tadqiqotlarda ushbu tizimning dastlabki namunasini (prototipini) yaratish va uni real ta’lim muhitida sinovdan o‘tkazish muhim vazifa bo‘lib qoladi.

FOYDALANILGAN ADABIYOTLAR

1. Chomsky N. Aspects of the Theory of Syntax. Cambridge, MA: MIT Press, 1965.
2. Jurafsky D., Martin J.H. Speech and Language Processing. 3rd ed. (draft). Stanford University, 2023.
3. Manning C.D., Schütze H. Foundations of Statistical Natural Language Processing. Cambridge, MA: MIT Press, 1999.
4. Leacock C., Chodorow M., Gamon M., Tetreault J. Automated Grammatical Error Detection for Language Learners. San Rafael: Morgan & Claypool Publishers, 2010.
5. Church K.W. A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text // Proceedings of the 2nd Conference on Applied Natural Language Processing. 1988. P. 136–143.

6. Sennrich R., Haddow B., Birch A. Neural Machine Translation of Rare Words with Subword Units // Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. 2016. P. 1715–1725.
7. Tyers F.M., Washington J.N. Towards a Free/Open-Source Universal-Dependency Treebank for Kazakh // Proceedings of the 3rd International Workshop on Computational Linguistics for Uralic Languages. 2015.
8. Bryant C., Felice M., Briscoe T. Automatic Annotation and Evaluation of Error Types for Grammatical Error Correction // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017. P. 793–805.

IJTIMOIY TIZIMLARNI MODELLASHTIRISH VA RAQAMLI TRANSFORMATSIYA: ILMIY TAJRIBALAR, TEXNOLOGIYALAR VA RIVOJLANISH ISTIQBOLLARI

Xolmurotova Muxlisa Botirovna
Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti Filologiya fakulteti 1-kurs talabasi
E-mail:XolmurotovaMuxlisa050@gmail.com
https://doi.org/10.5281/zenodo.23255539

Annotatsiya. *Ushbu maqolada zamonaviy jamiyatda ijtimoiy tizimlarni modellashtirish va raqamli transformatsiyalash jarayonlarining ilmiy-nazariy hamda amaliy asoslari tadqiq etiladi. Axborot-kommunikatsiya texnologiyalari, katta ma'lumotlar (Big Data) va sun'iy intellekt vositalarining ijtimoiy jarayonlarni boshqarish hamda bashorat qilishdagi o'rni tahlil qilingan. Tadqiqot davomida raqamli texnologiyalarni ijtimoiy sohalarga tatbiq etishning mavjud tajribalari o'rganilib, raqamlashuv sharoitida ijtimoiy tizimlar barqarorligini ta'minlash va rivojlantirish istiqbollari bo'yicha ilmiy xulosalar ilgari surilgan.*

Kalit so'zlar: *ijtimoiy tizimlar, modellashtirish, raqamli transformatsiya, sun'iy intellekt, katta ma'lumotlar (Big Data), raqamli texnologiyalar, jamiyat boshqaruvi, rivojlanish istiqbollari.*

Abstract. *This article investigates the scientific-theoretical and practical foundations of modeling and digitally transforming social systems in modern society. The role of information and communication technologies, Big Data, and artificial intelligence tools in managing and forecasting social processes is analyzed. During the study, existing practices of implementing digital technologies in social spheres were examined, and scientific conclusions were proposed regarding the prospects for ensuring the sustainability and development of social systems in the context of digitalization.*

Keywords: *social systems, modeling, digital transformation, artificial intelligence, Big Data, digital technologies, societal management, development prospects.*

1. KIRISH

XXI asrda ro'y berayotgan globallashuv va jadal texnologik taraqqiyot jamiyatning barcha jabhalarida tub o'zgarishlarni keltirib chiqarmoqda. Bugungi kunda ijtimoiy tizimlarni samarali boshqarish, ularning rivojlanish tendensiyalarini oldindan ko'ra bilish va yuzaga kelishi mumkin bo'lgan xatarlarning oldini olish eng dolzarb masalalardan biriga aylandi. Yangi O'zbekistonni barpo etish sharoitida davlat va jamiyat boshqaruvida raqamli texnologiyalarni keng joriy etish, ya'ni raqamli transformatsiyani amalga oshirish ustuvor vazifa qilib belgilangan.

Ijtimoiy tizimlarni modellashtirish — bu murakkab ijtimoiy-iqtisodiy, madaniy va ma'rifiy jarayonlarni matematik, mantiqiy hamda kompyuter modellari yordamida aks ettirish usulidir. Raqamli transformatsiya esa ushbu modellarni amaliyotga tatbiq etib, qarorlar qabul qilish jarayonini avtomatlashtirish va aniqlik darajasini oshirishga xizmat qiladi. Ushbu tadqiqotning maqsadi ijtimoiy tizimlarni modellashtirishning ilmiy tajribalarini umumlashtirish, zamonaviy raqamli texnologiyalarning imkoniyatlarini baholash hamda sohaning kelgusi rivojlanish istiqbollarini ko'rsatib berishdan iborat.

2. TADQIQOT METODLARI

Tadqiqot jarayonida tizimli tahlil, struktural-funktsional modellashtirish, qiyosiy-analitik hamda prognozlash usullaridan foydalanildi.

Ijtimoiy tizimlarning dinamikasini o‘rganishda xalqaro va mahalliy tajribalar, raqamli platformalar hamda axborot-kommunikatsiya texnologiyalarining jamiyat hayotiga ta’siri tahlil qilindi. Tadqiqot metodologiyasi ijtimoiy jarayonlarni murakkab moslashuvchan tizim (complex adaptive system) sifatida qarashga va ularni raqamli modellar orqali tahlil qilishga asoslanadi.

3. NATIJALAR

Ijtimoiy tizimlarni modellashtirish va raqamli transformatsiyalash tajribalari o‘rganilganda quyidagi asosiy natijalar hamda yo‘nalishlar shakllanganligi aniqlandi:

1. Bashorat qilish modellarining shakllanishi: Katta ma’lumotlar (Big Data) tahlili orqali jamiyatdagi ehtiyojlar, aholi kayfiyati va ijtimoiy talablarni oldindan bashorat qilish aniqligi keskin oshgani kuzatildi.
2. Boshqaruv samaradorligining ortishi: Davlat va ijtimoiy xizmatlarni raqamlashtirish (masalan, "E-hukumat" va raqamli ta'lim platformalari) natijasida ijtimoiy tizimlardagi byurokratiya darajasi kamayib, shaffoflik ta'minlanmoqda.
3. Sun'iy intellekt texnologiyalarining integratsiyasi: Ijtimoiy sohalarda sun'iy intellekt vositalaridan foydalanish qaror qabul qilish vaqtlarini qisqartirib, jarayonlarni optimallashtirish imkonini bermoqda.

4. MUHOKAMA

Ijtimoiy tizimlarni raqamli modellashtirish tajribasi shuni ko‘rsatadiki, texnologik taraqqiyot faqatgina texnik vositalarni qo‘llash bilan cheklanmaydi. U, eng avvalo, inson va jamiyat manfaatlariga xizmat qilishi lozim.

Raqamli transformatsiya sharoitida axborot xavfsizligi, kiberxavfsizlik hamda raqamli tengsizlikni bartaraf etish kabi yangi ijtimoiy masalalar yuzaga kelmoqda. Raqamli modellashtirish jarayonida milliy va ma'naviy qadriyatlarni hisobga oluvchi ijtimoiy algoritmlarni ishlab chiqish hamda axloqiy mezonlarga rioya qilish texnologiyalardan to‘g‘ri foydalanishning garovidir.

5. XULOSA

O‘tkazilgan tadqiqot shuni ko‘rsatadiki, ijtimoiy tizimlarni modellashtirish va raqamli transformatsiyalash — zamonaviy jamiyat barqarorligi va taraqqiyotining muhim omilidir. Zamonaviy texnologiyalarni ijtimoiy sohalarga bosqichma-bosqich va tizimli tatbiq etish boshqaruv qarorlarining sifatini oshiradi.

Kelajakda ijtimoiy tizimlarni modellashtirishda sun’iy intellekt imkoniyatlarini yanada kengaytirish, milliy raqamli ekotizimlarni rivojlantirish hamda yosh avlodning raqamli savodxonligini oshirish bo‘yicha amaliy chora-tadbirlarni davom ettirish maqsadga muvofiqdir.

FOYDALANILGAN ADABIYOTLAR :

1. Karimov I.A. Yuksak ma’naviyat – yengilmas kuch. – Toshkent: Ma’naviyat, 2008.
2. Qosimov B. Milliy uyg‘onish: jasorat, ma’rifat, fidoyilik. – Toshkent: Ma’naviyat, 2002.
3. Abduroziqov A. "Ijtimoiy jarayonlarni modellashtirishda raqamli texnologiyalarning o‘rni". // *O‘zbekiston ilmiy-amaliy jurnali*, 2023.
4. "Science and Innovation" xalqaro ilmiy jurnali materiallari to‘plami. – Toshkent, 2024.

SEMANTIK KORPUS VA UNING NAZARIY ASOSLARI

Munisa Xudoyberganova
TDTU assistent o'qituvchisi
https://doi.org/10.5281/zenodo.23255606

Annotatsiya. *Mazkur maqolada semantik korpus tushunchasi, uning nazariy asoslari hamda zamonaviy kompyuter lingvistikasidagi o'rni tahlil qilinadi. Korpus lingvistikasining rivojlanish bosqichlari, semantik annotatsiyaning mazmuni va vazifalari, semantik korpuslarning tuzilishi hamda ularning tabiiy tilni qayta ishlash texnologiyalaridagi ahamiyati yoritiladi. Shuningdek, dunyoda yaratilgan semantik korpuslar va ularning ilmiy-amaliy qo'llanilishi tahlil qilinib, o'zbek tili uchun semantik korpus yaratish zarurati asoslab beriladi.*

Kalit so'zlar: *semantik korpus, korpus lingvistikasi, semantik annotatsiya, tabiiy tilni qayta ishlash, NLP, sun'iy intellekt, leksik semantika, Word Sense Disambiguation.*

Аннотация. *В данной статье рассматривается понятие семантического корпуса, его теоретические основы и роль в современной компьютерной лингвистике. Анализируются этапы развития корпусной лингвистики, содержание и функции семантической аннотации, структура семантических корпусов, а также их значение в технологиях обработки естественного языка. Кроме того, исследуются наиболее известные семантические корпуса, созданные в мировой практике, их научное и прикладное применение, а также обосновывается необходимость создания семантического корпуса для узбекского языка.*

Ключевые слова: *семантический корпус, семантическая аннотация, обработка естественного языка (NLP), искусственный интеллект, лексическая семантика, разрешение лексической многозначности (Word Sense Disambiguation, WSD).*

Abstract. *This article examines the concept of a semantic corpus, its theoretical foundations, and its role in modern computational linguistics. It discusses the development of corpus linguistics, the concept and functions of semantic annotation, the structure of semantic corpora, and their significance in natural language processing technologies. In addition, the paper reviews major semantic corpora developed worldwide and their scientific and practical applications, while highlighting the necessity of developing a semantic corpus for the Uzbek language.*

Keywords: *semantic corpus, semantic annotation, natural language processing (NLP), artificial intelligence, lexical semantics, Word Sense Disambiguation (WSD).*

Kirish. Tabiiy tilda ko'plab so'zlar bir xil yozilishiga qaramay, turli ma'nolarni ifodalashi mumkin. Bunday holatlarni kompyuter to'g'ri anglashi uchun so'zning faqat shaklini emas, balki uning matndagi ma'nosini ham aniqlash zarur. Aynan shu vazifani semantik korpus bajaradi.

Masalan, "olma" so'zini olaylik.

1) *Men bozordan olma sotib oldim.*
2) *Kitobni olma.*

Har ikki gapda ham "olma" so'zi bir xil shaklda yozilgan. Biroq ularning ma'nosi mutlaqo boshqacha. Birinchi gapda "olma" meva nomini bildiruvchi ot bo'lsa, ikkinchi gapda

"olma" "olmoq" fe'lining inkor shakli hisoblanadi. Inson bu farqni gap mazmunidan osongina anglay oladi, ammo kompyuter uchun ikkala so‘z ham dastlab bir xil ko‘rinishga ega.

Oddiy korpuslarda ushbu ikki qo‘llanish faqat so‘z shakli sifatida qayd etiladi. Natijada kompyuter ularning ma'nolarini farqlay olmaydi. Semantik korpusda esa har bir so‘z kontekstiga qarab alohida semantik belgi (annotatsiya) bilan ta'minlanadi. Ya'ni birinchi gapdagi **"olma" "meva"** konsepti bilan, ikkinchi gapdagi **"olma" esa "olmoq"** fe'lining inkor shakli sifatida belgilab qo‘yiladi.

Demak, semantik korpus — bu matndagi so‘z va iboralarning nafaqat grammatik xususiyatlarini, balki ularning kontekstga mos ma'nolarini ham aks ettiruvchi maxsus annotatsiyalangan lingvistik resursdir. Bunday korpus yordamida kompyuter so‘zlarning haqiqiy ma'nosini aniqlash, o‘xshash shakldagi birliklarni farqlash hamda matn mazmunini chuqurroq tahlil qilish imkoniyatiga ega bo‘ladi.

Mazkur maqolaning maqsadi semantik korpus tushunchasining nazariy asoslarini yoritish, uning tarkibiy tuzilishi, asosiy xususiyatlari hamda zamonaviy tadqiqotlardagi o‘rnini tahlil qilishdan iborat.

Ingliz tili semantik annotatsiyalangan korpuslarni yaratish yo‘nalishida yetakchi tillardan biri hisoblanadi. Ushbu yo‘nalishdagi dastlabki yirik loyihalar qatoriga **SemCor** (Miller va boshq., 1994) hamda **FrameNet** (Fillmore va boshq., 2003) kiradi. Keyinchalik ishlab chiqilgan **PropBank** (Palmer va boshq., 2003) esa semantik korpuslarni rivojlantirishda muhim bosqich bo‘ldi. Mazkur loyiha fe'l predikatlari va ularning argumentlarini annotatsiyalashga asoslanishi bilan ajralib turadi. PropBank sintaktik jihatdan belgilangan **Penn Treebank II** korpusiga semantik qatlamni qo‘shish orqali yaratilgan bo‘lib, uning asosiy vazifasi predikat–argument munosabatlarini tizimli ravishda ifodalovchi hamda semantik tadqiqotlar uchun ishonchli empirik ma'lumotlar bazasini shakllantirishdan iborat.

SemCor asosan so‘z ma'nolarini annotatsiyalashga qaratilgan bo‘lsa, PropBank predikat–argument tuzilmasini, FrameNet freym semantikasini, OntoNotes esa ko‘p qatlamli (named entity, word sense, semantic role va coreference) annotatsiyani o‘z ichiga oladi.

MultiSemCor (Bentivogli va Pianta, 2005) SemCor korpusi asosida yaratilgan ingliz–italyan parallel semantik korpus hisoblanadi. Ushbu resurs SemCor tarkibidagi semantik annotatsiyalarni saqlab qolish bilan birga, ularni italyan tilidagi **WordNet** ma'nolari bilan bog‘laydigan tillararo aloqalarni ham o‘z ichiga oladi. MultiSemCor tajribasi SemCor matnlarini boshqa tillarga tarjima qilish orqali yangi semantik korpuslarni yaratish imkoniyatini ko‘rsatdi. Bunday yondashuv, ayniqsa, **Word Sense Disambiguation (WSD)** masalalarini hal etishda samarali hisoblanadi. Shu bilan birga, fe'llarning semantik ma'nolari, argument tuzilmalari va seleksion cheklovlarini tahlil qilishda ayrim cheklovlar mavjud. Buning asosiy sababi korpusdagi fe'llar uchun misollar sonining bir xil taqsimlanmaganligidir.

Ispan tilida yaratilgan muhim semantik resurslar qatoriga **Adesse** (García-Miguel va Albertuz, 2005) hamda **AnCora** (Taulé, Martí va Recasens, 2008) kiradi. Har ikkala loyiha semantik annotatsiyalangan korpus va unga asoslangan fe'llar bazasini o‘zida mujassam etgan bo‘lib, fe'l semantikasi va argument tuzilishini tadqiq etishga xizmat qiladi.

Adesse korpusi **Arthus** korpusining semantik jihatdan kengaytirilgan shakli bo‘lib, Santyago-de-Kompostela universitetida ishlab chiqilgan. Ushbu resursda fe'lli gaplarning sintaktik va semantik xususiyatlari batafsil annotatsiya qilingan. Jumladan, har bir gap uchun sintaktik funksiya, grammatik kategoriya hamda ishtirokchilarning asosiy semantik belgilari (jonli yoki jonsiz) ko‘rsatilgan. Bundan tashqari, semantik rollar, fe'l ma'nolari va ularning

semantik sinflari ham qayd etilgan. Shu ma'lumotlar asosida fe'llarning turli sintaktik konstruksiyalarda qo'llanishi va ularning chastotasini aks ettiruvchi ma'lumotlar bazasi shakllantirilgan. Adesse fe'l ma'nolarini tasniflasa-da, bu tasnif nisbatan umumiy xarakterga ega bo'lib, WordNet ma'nolari bilan bevosita bog'lanmagan.

AnCora esa ispan va katalan tillari uchun yaratilgan ikki parallel korpusdan iborat bo'lib, unda ko'p qatlamli sintaktik va semantik annotatsiyalar amalga oshirilgan. Korpus tarkibida morfologik belgilar, sintaktik tuzilma, sintaktik funksiyalar, argument strukturasi, tematik rollar, fe'llarning semantik sinflari, WordNet ma'nolari hamda referensial obyektlar haqidagi ma'lumotlar mavjud. Ushbu annotatsiyalar asosida har ikki til uchun fe'llarning semantik tavsifi, argument tuzilishi va sintaktik subkategorizatsiyasini o'z ichiga olgan maxsus leksikonlar yaratilgan.

	SemCor	BBN	NomBank	VerbNet/SemLink	PropBank	OntoNotes 4 (5)	FrameNet Full Text
values (times, quantities)	–	✓	–	–	–	✓	✓
named entities	CNE: The Fold/ORG	BNE: The Fold/ORG:OTHER	–	–	–	ONNE: The Fold/ORG	CNE: The Fold/ORG
nouns	WNS: fold.n.01	BED	NBF: folding.01	–	–	~ONS: fold-n.01, (ONF)	FNF: Reshaping, Endeavor_failure
verbs	WNS: fold_up.v.01	–	–	VNC: bend-45.2	PBF: fold-v.03	ONS: fold-v.01, ONF: fold-v.03	
anaphoric coreference	–	✓	–	–	–	~	–
noun coreference	–	–	–	–	–	~	–

1-**rasm. Ingliz tilidagi mashhur semantik korpuslarda qo'llanilgan annotatsiya turlarining taqqoslanishi.**

Yuqorida ko'rib chiqilgan semantik korpuslar shuni ko'rsatadiki, dunyoning ko'plab tillarida semantik annotatsiya ishlari bir necha o'n yillardan buyon izchil rivojlanib kelmoqda. SemCor, FrameNet, PropBank, MultiSemCor, Adesse va AnCora kabi loyihalar so'z ma'nolarini belgilash, predikat–argument munosabatlarini aniqlash, semantik rollarni tavsiflash hamda ko'p qatlamli annotatsiya tizimlarini yaratish borasida muhim tajriba to'plagan. Mazkur korpuslar bugungi kunda mashina tarjimasi, avtomatik matn tahlili, savol-javob tizimlari va sun'iy intellekt modellarini o'qitishda asosiy lingvistik resurslardan biri sifatida keng qo'llanilmoqda.

Shunga qaramay, bunday yondashuvlar turkiy tillar, xususan o'zbek tilida hali to'liq shakllanmagan. So'nggi yillarda o'zbek tili uchun elektron korpuslar va ayrim lingvistik resurslar yaratilayotgan bo'lsa-da, ular asosan matnlarni yig'ish, morfologik tahlil yoki sintaktik belgilash bilan cheklanadi. So'z ma'nolarini kontekst asosida ajratish, semantik rollarni belgilash hamda leksik birliklar o'rtasidagi semantik munosabatlarni aks ettiruvchi maxsus semantik korpuslar esa hozircha yetarli darajada ishlab chiqilmagan.

Shu bois o'zbek tili uchun semantik annotatsiyalangan korpusni yaratish bugungi kompyuter lingvistikasining dolzarb yo'nalishlaridan biri hisoblanadi. Bunday korpus nafaqat o'zbek tilining semantik xususiyatlarini tizimli ravishda tavsiflash, balki mashina tarjimasi, tabiiy tilni qayta ishlash, katta til modellari (LLM), semantik qidiruv tizimlari va boshqa sun'iy intellekt texnologiyalarini rivojlantirish uchun ham muhim lingvistik baza vazifasini bajaradi.

FOYDALANILGAN ADABIYOTLAR

1. Abdurakhmonova, Nilufar. (2023). SEMANTIK ANNOTATSIYALANGAN KORPUS YARATISH TAJRIBASIDAN.
2. Bentivogli, L., & Pianta, E. (2005). Exploiting parallel texts in the creation of multilingual semantically annotated resources: The MultiSemCor corpus. *Natural Language Engineering*, *11*(3), 247–261.
3. Fillmore, C. J., Johnson, C. R., & Petruck, M. R. L. (2003). Background to FrameNet. García-Miguel, J. M., & Albertuz, F. J. (2005). Verbs, semantic classes and semantic roles in the ADESSE project. *Proceedings of the Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes.*
4. Miller, G. A., Leacock, C., Tengi, R., & Bunker, R. T. (1993). A semantic concordance. *Proceedings of the ARPA Human Language Technology Workshop*, 303–308.
5. Palmer, M., Gildea, D., & Kingsbury, P. (2005). The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, *31*(1), 71–106. https://doi.org/10.1162/0891201053630264
6. Taulé, M., Martí, M. A., & Recasens, M. (2008). AnCora: Multilevel annotated corpora for Catalan and Spanish. *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008).*

O‘ZBEK TILI TA’LIMIY LUG‘ATLARINI RAQAMLI MUHITDA MODELLASHTIRISH

Yusupova Iroda Ravshanbek qizi
Kompyuter lingvistikasi magistratura mutaxassisligi 2 - kurs magistranti
O‘zbekiston Milliy universiteti, Toshkent, O‘zbekiston.
https://doi.org/10.5281/zenodo.23255645

Annotatsiya. *Ushbu maqolada o‘zbek tili ta’limiy lug‘atlarini raqamli muhitda modellashtirish masalasi zamonaviy raqamli leksikografiya, pedagogik lug‘atchilik hamda til texnologiyalari kesishmasida tadqiq etiladi. Tadqiqotning maqsadi mavjud ilmiy izlanishlarni tizimli tahlil qilish asosida raqamli ta’limiy lug‘atning asosiy funksional komponentlarini aniqlash va ular o‘rtasidagi munosabatni ifodalovchi konseptual modelni taklif etishdan iborat. Tadqiqotda scoping review, qiyosiy tahlil va konseptual modellashtirish usullaridan foydalanildi.*

Kalit so‘zlar: *ta’limiy leksikografiya, raqamli leksikografiya, o‘zbek tili, o‘quv lug‘ati, NLP, elektron korpus, multimodallik, raqamli ta’lim.*

Аннотация. *В статье рассматривается проблема моделирования учебных словарей узбекского языка в цифровой среде на пересечении современной цифровой лексикографии, педагогической лексикографии и языковых технологий. Цель исследования заключается в выявлении основных функциональных компонентов цифрового учебного словаря и разработке концептуальной модели их взаимосвязи на основе систематического анализа научной литературы. Использованы методы scoping review, сравнительного анализа и концептуального моделирования.*

Ключевые слова: учебная лексикография, цифровая лексикография, узбекский язык, учебный словарь, NLP, корпус, мультимодальность, цифровое обучение.

Annotation. *This article examines the issue of modeling Uzbek language educational dictionaries in a digital environment at the intersection of modern digital lexicography, pedagogical lexicography, and language technologies. The aim of the study is to identify the main functional components of a digital educational dictionary and propose a conceptual model representing the relationships among them, based on a systematic analysis of existing scientific research. The study utilized scoping review, comparative analysis, and conceptual modeling methods.*

Keywords: *educational lexicography, digital lexicography, Uzbek language, learner's dictionary, NLP, electronic corpus, multimodality, digital education.*

KIRISH

Raqamli texnologiyalarning til ta’limiga kirib kelishi lug‘atning an’anaviy vazifasi va tuzilishiga ta’sir ko‘rsatmoqda. Bugungi kunda lug‘at mustaqil ma’lumotnomagina emas, matn muharriri, ta’limiy platforma yoki mobil ilova ichida faoliyat yurituvchi lingvistik vosita sifatida ham qaralmoqda va bu o‘zgarish ma’lumotning qanday tanlanishi, taqdim etilishi, ta’limiy faoliyat bilan bog‘lanishini alohida metodologik masalaga aylantiradi. Huang va Tarp buni ingliz tili o‘quv ilovalaridagi lug‘atlar misolida ma’lumot ortiqchaligi va noqulay interfeys muammolari orqali ko‘rsatib bergan [Huang F., Tarp S., 2021: 68–92]. O‘zbek tili uchun bu

masala alohida dolzarblik kasb etadi. So‘nggi yillarda elektron korpus, morfologik tahlil vositalari va leksik-semantik resurslar shakllanayotgan bo‘lsa-da, mavjud ta’limiy lug‘atlarning aksariyati, asosan, matnni raqamli formatga o‘tkazish va qidiruv funksiyasi bilan cheklanadi. Ularda foydalanuvchi modeli, moslashuvchan taqdimot yoki pedagogik qatlam amalda mavjud emas.

Shundan kelib chiqib, maqolaning maqsadi o‘zbek tili o‘quv lug‘atlarini raqamli muhitda modellashtirishning nazariy-texnologik asoslarini aniqlash va mavjud tadqiqotlar sintezi asosida uning konseptual modelini taklif etishdan iborat bo‘lib, bu maqsadga erishish uchun quyidagi vazifalar belgilandi: raqamli ta’limiy lug‘atchilikning rivojlanish tendensiyalarini aniqlash, foydalanuvchiga yo‘naltirilganlik, korpus, pedagogik integratsiya, multimodallik va intellektual texnologiyalar bo‘yicha tadqiqotlarni tahlil qilish, o‘zbek tili uchun mavjud raqamli resurslarning modeldagi o‘rnini belgilash hamda aniqlangan komponentlarni yagona konseptual modelga birlashtirish.

ADABIYOTLAR SHARHI

Zamonaviy tadqiqotlarni tahlil qilish raqamli ta’limiy lug‘atchilikning bir necha o‘zaro bog‘liq tendensiyalar asosida rivojlanayotganligini ko‘rsatadi. Ular orasida foydalanuvchiga yo‘naltirilganlik, korpusga va NLP texnologiyalariga asoslanganlik, o‘quv muhitiga integratsiya qilish, multimodal taqdimot hamda sun’iy intellektdan foydalanish alohida ajralib turadi.

Raqamli lug‘atning samaradorligi undagi ma’lumot hajmiga emas, foydalanuvchi uchun zarur ma’lumotning kerakli vaqtda, tegishli shaklda taqdim etilishiga bog‘liq. Gouws va Tarp raqamli leksikografiyada information overload va data overload muammolariga e’tibor qaratib, ma’lumot hajmining keskin oshishi foydalanuvchiga barcha mavjud ma’lumotni bir vaqtning o‘zida ko‘rsatish zaruratini tug‘dirmasligini ta’kidlaydi [Gouws R.H., Tarp S., 2017: 389–415]. Leksikografik ma’lumotlarni alohida lug‘at ko‘rinishida emas, bevosita yozish jarayonining o‘ziga integratsiya qilish konsepsiyasi Frankenberg-Garcia tomonidan ColloCaid loyihasida ishlab chiqilgan. Muallif “feed forward” yondashuvi bilan bog‘lagan bu tizim foydalanuvchi hali qidirmagan, ayni faoliyat uchun foydali bo‘lishi mumkin bo‘lgan ma’lumotni ham taqdim etadi [Frankenberg-Garcia A., 2020: 29–43].

O‘zbek tilida shakllanayotgan elektron korpus va NLP vositalari bu yo‘nalish uchun zarur infratuzilmani yaratmoqda: Abduraxmonova va hamkorlari uzbekcorpus.uz platformasining qidiruv va tahlil imkoniyatlarini, shuningdek, turkiy tillar uchun ishlab chiqilgan NLP vositalari orqali o‘zbek tilining morfologik xususiyatlarini avtomatik qayta ishlash imkoniyatini ko‘rsatgan [Abdurakhmonova N. va boshq., 2021: 66–69; Abdurakhmonova N.Z. va boshq., 2022: 1790–1793].

Leksik-semantik darajada esa Mengliev, Baraxnin va Abduraxmonova 200 ta so‘z asosida sinovdan o‘tkazilgan morfologik tahlil tizimini, Agostini va hamkorlari esa 28 140 synset va 75,98% aniqlikka ega UzWordNet resursini taqdim etadi — bular o‘zbek tili uchun avtomatik qayta ishlanadigan lingvistik infratuzilmaning mavjudligini tasdiqlaydi [Mengliev D. va boshq., 2021: 9117; Agostini A. va boshq., 2021: 8–19].

Ba’zi olimlar raqamli lug‘atni ta’limiy ilovaga joylashtirishning o‘zi yetarli emasligini ta’kidlashadi. Huang va Tarp xitoylik ingliz tili o‘rganuvchilari foydalanadigan o‘quv ilovasi tarkibidagi lug‘atlarni tahlil qilib, unda ma’lumotlarning ortiqchaligi, ma’lumotga kirishning murakkabligi va interfeys bilan bog‘liq muammolarni aniqlaydi va kontekstga moslashtirilgan, minimal va o‘rganuvchining o‘qish jarayonini to‘xtatmaydigan lug‘aviy taqdimotni taklif qiladi [Huang F., Tarp S., 2021: 68–92]. Tarp sun’iy intellektga asoslangan yozish yordamchilari

loyihasida leksikograflar vazifasi til modelini tayyorlashdan tizim-foydalanuvchi aloqasini loyihalashga siljiyotganini ko‘rsatadi [Tarp S., 2023: 107–131].

Texnologiya yordamida lug‘at o‘rganish bo‘yicha Yu va Traininning meta-tahlili 34 ta tadqiqot, 2 511 ishtirokchi va 49 ta effect sizeni qamrab oladi. Umumiy effekt hajmi d = 0,64, mobile-assisted vocabulary learning uchun d = 0,85 deb aniqlangan [Yu A., Trainin G., 2022: 235–252]. Zhou tomonidan amalga oshirilgan 2026-yilgi tahlil esa mobil lug‘atlarni o‘rganishga oid 65 ta tadqiqotni qamrab olib, unda mualliflar uzoq muddatli jarayonda effekt hajmi d = 1,28 ga yetganini ko‘rsatadi [Zhou Y., Zhou M., 2026: 75–93]. Natijalar raqamli lug‘atni multimodal muhitda tashkil etishning pedagogik salohiyati mavjudligini ko‘rsatsa-da, bunday imkondan foydalanishda multimedia elementlarining ko‘paytirilishi emas, o‘quv maqsadiga mosligi muhimdir.

Sun’iy intellektning rivojlanishi o‘quv lug‘atlarida izoh va boshqa ma’lumotlarni avtomatik generatsiya qilish imkoniyatlarini ochdi. Shu sohada Rees va Lew tadqiqotida GPT yaratgan ta’riflarni Macmillan English Dictionary izohlari bilan taqqoslaganda ta’rifsiz guruhdan ustun bo‘lgan, lekin lug‘at izohlaridan sezilarli farq qilmagan [Rees G.P., Lew R., 2024: 50–74].

TADQIQOT METODOLOGIYASI

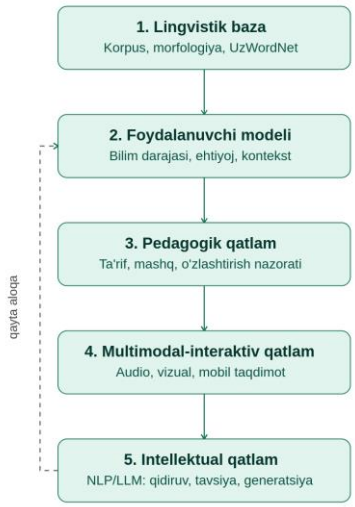
Tadqiqot sifat jihatidan scoping review, qiyosiy tahlil va konseptual modellashtirish metodlariga asoslanadi. Qidiruv 2025-yil noyabr oyida Google Scholar, Scopus va Lexikos jurnalining elektron arxivi bo‘yicha o‘tkazildi. Qidiruv so‘z birikmalari sifatida “digital lexicography”, “learner’s dictionary”, “e-dictionary user needs”, “NLP for lexicography”, “mobile-assisted vocabulary learning”, “dictionary integration language learning apps” hamda ularning o‘zbek va rus tilidagi muqobillari ishlatildi. Dastlabki qidiruv natijasida 52 ta potentsial manba aniqlandi. Ularni saralashda uchta mezon qo‘llanildi:

1) tadqiqotning ta’limiy lug‘at yoki raqamli leksikografiyaga bevosita aloqadorligi;
2) tadqiqotda aniq metod, empirik natija yoki nazariy konsepsiyaning mavjudligi;
3) natijaning o‘zbek tili ta’limiy lug‘atini raqamli modellashtirish uchun metodologik jihatdan qo‘llanish imkoniyati.

Ikkinchi bosqichda tanlangan manbalar beshta tendensiya bo‘yicha guruhlandi: foydalanuvchiga yo‘naltirilganlik, korpus/NLP, pedagogik integratsiya, multimodallik va intellektual texnologiyalar. Keyingi bosqichda esa har bir tendensiyaning raqamli ta’limiy lug‘atdagi funksiyasi aniqlanib, ular o‘rtasidagi aloqa konseptual modellashtirish orqali ifodalandi.

NATIJALAR

Adabiyotlar sintezi natijasida raqamli ta’limiy lug‘atni shakllantiruvchi besh komponentli konseptual model ishlab chiqildi. Model quyidagi rasmda umumlashtirilgan (1-rasm):



1-rasm. Raqamli ta'limiy lug'atning besh komponentli konseptual modeli

Dastlabki komponent lug‘atning lingvistik ma’lumotlar qatlamini tashkil etadi. U korpus ma’lumotlari, morfoanaliz, leksik-semantik munosabatlar va boshqa lingvistik resurslarni qamrab oladi. O‘zbek tili uchun bu qatlamni shakllantirishda elektron korpus, morfologik analizator va UzWordNet kabi resurslardan foydalanish mumkin [Abdurakhmonova N. va boshq., 2021; Mengliev D. va boshq., 2021; Agostini A. va boshq., 2021]. Ularning mavjudligi raqamli ta’limiy lug‘atni qo‘lda tuzilgan ma’lumotlar majmuasi bilan birgalikda qayta ishlanuvchi lingvistik ma’lumotlar bazasiga aylantirishga yordam beradi.

Foydalanuvchining ta’limiy ehtiyojlarini hisobga olish lozim bo‘lganligi uchun lug‘at bir xil ma’lumotni barcha foydalanuvchilarga bir xil shaklda berish o‘rniga uning bilim darajasi, topshiriq turi va foydalanish kontekstiga mos ma’lumotni tanlashi kerak bo‘ladi. Yuqoridagi natijalarda ma’lumotning ortiqchaligi va noqulay interfeys o‘quv jarayoniga xalaqit berishi mumkinligini ko‘rsatilgani sababli foydalanuvchi modeli raqamli lug‘at arxitekturasining markaziy qismlaridan biri sifatida belgilanadi [Huang F., Tarp S., 2021: 68–92].

Uchinchi komponentda ta’rif, misol, kontekst, mashq, takrorlash va o‘zlashtirishni tekshirish kabi funksiyalar joylashadi. Oldingi qatlamlar individual empirik tadqiqotga tayansa, ushbu qatlam foydalanuvchi modeli hamda multimodal qatlam bo‘yicha ko‘rib chiqilgan o‘rganishlarning umumlashmasidan iborat mualliflik taklifidir. Lug‘at foydalanuvchi ehtiyojini aniqlab, uni multimodal vositalar orqali taqdim etganda pedagogik qatlam ushbu ikki jarayonni o‘zaro bog‘lovchi vositachi vazifasini bajaradi. Shu bois uning tarkibiy elementlari (mashq, takrorlash, o‘zlashtirishni tekshirish) hozircha nazariy jihatdan asoslangan, ammo mustaqil empirik sinovdan o‘tkazilmagan. Lug‘at maqolasi “so‘z+ma’no+kontekst+ o‘quv harakati” ko‘rinishida modellashtiriladi.

Keyingi komponent lug‘aviy ma’lumotni turli axborot shakllari orqali namoyish etadi. Audio talaffuz, vizual material, interaktiv topshiriq va mobil foydalanish imkoniyatlari matn bilan birgalikda ushbu qatlamga kiradi. Yu va Trainin meta-tahlilida texnologiya yordamidagi L2 vocabulary learning uchun umumiy d = 0,64 effekt aniqlangani, mobile-assisted learning uchun esa d = 0,85 ko‘rsatkichi qayd etilgani ushbu qatlamning pedagogik salohiyatini empirik jihatdan qo‘llab-quvvatlaydi [Yu A., Trainin G., 2022: 235–252].

So‘nggi komponent NLP va LLM texnologiyalaridan foydalanish imkoniyatlarini qamrab olgan holda avtomatik qidiruv, kontekstni aniqlash, morfologik qayta ishlash, mos ma’lumotni

tavsiya qilish va kelajakda ta'rif yoki misollarni avtomatik generatsiya qilish kabi funksiyalarni o'z ichiga olishi mumkin. Biroq Rees va Lewning tajribasi LLM yaratgan ta'riflar pedagogik tayyor lug'at ta'riflarining avtomatik o'rnini bosuvchi vosita bo'lishiga asos yetarli emasligini aytib o'tishganini ham hisobga olish zarur [Rees G.P., Lew R., 2024: 50–74]. Natijada taklif etilayotgan model quyidagi ketma-ketlik asosida ifodalanadi: Lingvistik ma'lumot → leksikografik qayta ishlash → foydalanuvchiga moslashtirish → pedagogik taqdimot → multimodal/interaktiv foydalanish → foydalanuvchi faoliyati haqidagi ma'lumot → tizimni takomillashtirish.

MUHOKAMA

Olingan natijalar elektron lug'at va raqamli o'quv lug'at o'rtasidagi farqni yana ham aniqroq ko'rsatadi. Elektron shaklga o'tkazilgan lug'at, asosan, mavjud ma'lumotni raqamli formatda qidirishga ko'mak beradi.

O'zbek tili uchun taklif etilayotgan model amaliyotdagi bo'shliqni to'ldirishga qaratilgan bo'lib, u mavjud lingvistik resurslarni alohida saqlash emas, balki pedagogik qatlam orqali yagona ta'limiy tizimga bog'lashni nazarda tutadi. Amaldagi ko'p o'zbekcha elektron lug'atlar real vaqtda foydalanuvchiga moslashmaydi, lingvistik baza darajasida qolib ketmoqda. Tizimda avtomatlashtirish darajasi ehtiyotkorlikni talab etadi: Rees va Lewning 43 nafar ishtirokchi bilan o'tkazgan tajribasi GPT-generated ta'riflar learner's dictionary ta'riflaridan avtomatik ustun emasligini ko'rsatdi [Rees G.P., Lew R., 2024: 50–74]. Demak, LLM leksikografni almashtirmasdan, faqat yordamchi tamoyilda qo'llanishi lozim. Undan tashqari, multimodal imkoniyatlar ham o'z-o'zidan samaradorlik bermaydi. Mayerning multimedia o'qitish konsepsiyasiga ko'ra, kognitiv yuklamani hisobga olib, audio, rasm va animatsiyalar o'quv maqsadi bilan funksional bog'lanishi shart. Modelning asosiy farqi texnologiyalarni sanashda emas, balki korpus, NLP, multimedia va LLMni yagona ta'limiy tizim qatlamlari sifatida funksional ketma-ketlikda birlashtirishidadir [Liu, S., 2026]. Mazkur yondashuv o'zbek tili ta'limiy lug'atlari uchun uchta metodologik prinsipni ilgari suradi: pedagogik maqsadning ustuvorligi, kontekstga moslashtirish, inson va sun'iy intellekt hamkorligi.

Ushbu tadqiqot bir qator cheklovlarga ham ega bo'lib, ular keyingi ishlarda hisobga olinishi lozim. Masalan, taklif etilgan besh komponentli model hozircha faqat nazariy-sintetik xarakterga ega bo'lib, u haqiqiy foydalanuvchi yoki texnik prototip ustida sinovdan o'tkazilmagan, shu sababli uning pedagogik samaradorligi haqida yakuniy xulosa chiqarishga erta. Pedagogik qatlam komponenti mustaqil empirik tadqiqotga emas, muallifning nazariy umumlashmasiga asoslangan. Bu uning zaif nuqtasi hisoblanadi. Manbalarning katta qismi ingliz tili konteksti bo'yicha olib borilganligi uchun o'zbek tilining o'ziga xos morfologik murakkabligi bu natijalarni tanqidiy moslashtirishni talab qiladi.

XULOSA

Tadqiqot natijalari raqamli ta'limiy lug'atni oddiy elektron lug'atning texnik jihatdan takomillashtirilgan shakli emas, balki lingvistik, pedagogik, foydalanuvchiga oid va texnologik jihatlar integratsiyalashgan tizim sifatida qarash zarurligini ko'rsatdi. Taklif etilgan besh komponentli model shu integratsiyani ifodalaydi, ammo LLM asosidagi avtomatik generatsiya inson leksikografik ekspertizasining o'rnini to'liq bosa olmasligi sababli, kelajakdagi tizimlarda inson–AI hamkorligi tamoyili ustuvor bo'lishi lozim. O'zbek tili o'quv lug'atlarini raqamli muhitda modellashtirishning asosiy vazifasi mavjud materiallarni elektron shaklga ko'chirishgina emas, lingvistik bilimni foydalanuvchining real ta'limiy faoliyatida qo'llash mumkin bo'lgan intellektual, kontekstual tizimga aylantirishdir.

FOYDALANILGAN ADABIYOTLAR:

1. Abdurakhmonova N. va boshq. (2021). Linguistic functionality of the Uzbek electronic corpus. — Toshkent, 66–69-b.
2. Abdurakhmonova N.Z., Ismailov B., Mengliev D. (2022). Developing NLP tools for Turkic languages: morphological processing of Uzbek. 1790–1793-b.
3. Agostini A. va boshq. (2021). UzWordNet: a lexical-semantic database for the Uzbek language. 8–19-b.
4. Frankenberg-Garcia A. (2020). Providing Ready-Made and Do-It-Yourself Corpus Resources for Language Learners: ColloCaid. ReCALL, 29–43-b.
5. Ge X., Zhang S., Xu H., Zhang X. (2024). A Scoping Review of Studies into Dictionary Use and Language Learning. Lexikos, 34(1), 331–354.
6. Gouws R.H., Tarp S. (2017). Information overload and data overload in lexicography. International Journal of Lexicography, 389–415-b.
7. Huang F., Tarp S. (2021). Dictionaries Integrated into English Learning Apps: Critical Comments and Suggestions for Improvement. Lexikos, 31, 68–92.
8. Mayer R.E. (2014). Cognitive Theory of Multimedia Learning. Cambridge University Press, 43–71-b.
9. Mengliev D., Barakhnin V., Abdurakhmonova N. (2021). Intelligent web-system for morphological analysis of Uzbek words. 9117-b.
10. Rees G.P., Lew R. (2024). The Effectiveness of OpenAI GPT-Generated Definitions Versus Definitions from an English Learners' Dictionary in a Lexically Orientated Reading Task. International Journal of Lexicography, 50–74-b.
11. Tarp S. (2023). Eppur si muove: Lexicography is Becoming Intelligent! Lexikos, 33(2), 107–131.
12. Yu A., Trainin G. (2022). Technology-assisted second language vocabulary learning: A meta-analysis. ReCALL, 235–252-b.
13. Zhou Y., Zhou M. (2026). Mobile-assisted vocabulary learning: An updated meta-analysis. ReCALL, 75–93-b.
14. Liu, S., Maimaiti, M., Yolwas, N., Abdurakhmonova, N., Le, W., Xie, Z., ... & Silamu, W. (2026). Dependency-Aware Graph Transformer with Hierarchical Syntax Encoding for Aspect-Level Sentiment Classification. Expert Systems with Applications, 132574.

MECHANISMS FOR DEVELOPING MEDICAL STUDENTS' PROFESSIONAL COMPETENCE THROUGH DIGITAL TECHNOLOGIES

Davronov Shakhbos Erkin ugli[1,*], Kakhkhorov Siddik K.[2]
[1] Doctoral student, Bukhara State University, Bukhara, Uzbekistan
[2] Department of Heliophysics, Renewable Energy Sources and Electronics, Bukhara State University, Bukhara, Uzbekistan
* Corresponding author: shaxbos.davronov.91@gmail.com
https://doi.org/10.5281/zenodo.23256036

Abstract. *Digital technologies are widely used in medical universities, but their effect on students' professional competence depends on how they are used. This paper describes the mechanisms through which digital technologies contribute to the development of medical students' professional competence. Based on constructivist, experiential and self-regulated learning theories and an analysis of research in medical education, five mechanisms are identified: contextualisation (digital tasks based on real clinical situations), simulation (safe practice with virtual patients and data), collaboration (joint work in shared digital spaces), feedback (rapid, data-based feedback through learning platforms) and reflection (self-assessment and planning through e-portfolios). For each mechanism, the paper shows its theoretical basis, the digital tools that support it, the competence components it develops and the typical mistakes in its use. The mechanisms are presented as a cycle, and an example of its application in the course "Information Technologies in Medicine" is given.*

Keywords: *medical education, professional competence, digital technologies, virtual patients, feedback, self-regulated learning.*

Introduction

Over the last decade, medical universities have introduced learning management systems, electronic libraries, simulation software and online assessment. These tools are now a normal part of student life. However, research shows that digital health competencies are still taught unevenly and are often separated from clinical subjects [1]. In many cases technology is used to deliver the same information in a new format: lectures become video recordings, and paper tests become online tests. The tools change, but the way students learn remains the same.

The DigComp 2.2 framework defines digital competence as the confident, critical and responsible use of digital technologies for learning, work and participation in society [2]. For a future doctor, this means being able to find reliable medical information, work with patient data, communicate in digital health systems and protect confidential information. These abilities cannot be developed only by listening or reading; they require practice, feedback and reflection.

Therefore, the question is not whether to use digital technologies, but through which mechanisms they actually change students' competence. By "mechanism" we mean a repeated process through which the use of a digital tool leads to a change in the student's knowledge, skills or attitudes. Understanding these mechanisms helps teachers choose tools for a clear purpose instead of using technology for its own sake. The aim of this paper is to identify these mechanisms, describe how each of them works and show how they can be combined in teaching.

Theoretical background

Four groups of theories explain how learning with digital technologies leads to competence. Constructivist theory states that knowledge is actively built by the learner through work on meaningful problems [3]. Social constructivism adds that many concepts are first mastered in interaction with others before they become the learner's own [4]. Experiential learning theory describes learning as a cycle of concrete experience, reflection, conceptualisation and active experimentation [5]. Finally, theories of self-regulated learning show that successful learners set goals, monitor their progress and adjust their strategies [6].

In medical education, these ideas are reflected in the use of simulation, which allows students to gain experience safely [7, 8], and in the use of portfolios, which support reflective practice [9]. Research on feedback shows that it is one of the strongest influences on learning, but only when it is specific, timely and focused on the task [10, 11]. The concept of the reflective practitioner explains why professionals need to think about their actions both during and after practice [12]. Taken together, these theories suggest that digital technologies develop competence when they support active problem solving, safe practice, interaction, feedback and reflection.

Methods

The study is theoretical. It is based on an analysis of research in medical education and educational technology, a comparison of theoretical approaches (constructivism, experiential learning, self-regulated learning, feedback theory) and the authors' experience of teaching the course "Information Technologies in Medicine". Each mechanism was identified when three conditions were met: it had a clear theoretical basis, its effect was reported in research, and it could be supported by digital tools available in a typical medical university.

Results

Five mechanisms were identified (Table 1).

Table 1. Mechanisms for developing professional competence through digital technologies

Mechanism	What happens	Digital tools	Main competence component
Contextualisation	Digital tasks are built on real clinical situations	Medical databases, case collections, LMS assignments	Cognitive, activity-based
Simulation	Students practise decisions safely, without risk to patients	Virtual patients, data sets, simulators	Activity-based
Collaboration	Students solve problems together and share knowledge	Cloud documents, forums, video conferencing	Activity-based, communicative
Feedback	Students quickly see their mistakes and progress	Online tests, rubrics, learning analytics	Cognitive, reflective
Reflection	Students evaluate their own work and plan next steps	E-portfolio, reflective journal, self-assessment forms	Reflective, motivational

Contextualisation. According to constructivist theory, knowledge is built through active work on meaningful problems [3]. When a digital task is connected to a clinical situation, for example searching for evidence on a disease or analysing a table of laboratory results, the

student sees why the skill matters. This mechanism also increases motivation, because the link to the future profession is clear. A typical mistake is to use general tasks ("create a table", "make a presentation") without a medical context. Such tasks develop technical skills, but they do not develop professional competence.

Simulation. Virtual patients and simulated data sets allow students to practise decisions that would be unsafe or impossible with real patients. A meta-analysis has shown that computerised virtual patients are associated with large positive effects on knowledge and skills compared with no intervention [7]. A systematic review of high-fidelity simulation found that feedback, repetitive practice and integration into the curriculum are among the features that make simulation effective [8]. In the context of digital competence, simulation is also useful for practising data protection, for example anonymising patient records. A typical mistake is to use simulation once, as a demonstration, instead of as repeated practice.

Collaboration. Learning is a social process, and many concepts are first mastered in interaction with others [4]. Shared digital spaces make this interaction visible and lasting: the history of a cloud document shows who contributed what. Collaborative learning develops critical thinking, communication and responsibility [13], and project-based work develops both professional and interpersonal skills [3]. For future doctors, who will work in multidisciplinary teams, this mechanism is especially important. A typical mistake is to give a group task without clear roles, so that one student does all the work.

Feedback. Feedback is one of the strongest influences on learning, but only when it is specific and timely [10]. Formative feedback is most effective when it is focused on the task and tells the student what to do next [11]. Digital platforms make it possible to give feedback immediately after a test, to use rubrics that explain criteria, and to track progress with learning analytics [14]. As a result, students and teachers can respond to difficulties before the final assessment. A typical mistake is to give only a score without explaining what was wrong and how to improve.

Reflection. Self-regulated learners set goals, monitor their progress and adjust their strategies [6]. An e-portfolio and a reflective journal support this process: the student collects evidence of their work, compares it with criteria and plans further development. Reviews of portfolios in medical education show that they are most successful when their goals are clear, when they are supported by mentoring and when they are not overloaded with paperwork [9]. This mechanism is especially important for doctors, who must continue learning throughout their careers [12].

Discussion

The five mechanisms work as a cycle rather than separately (Figure 1). A contextualised task leads to simulation or collaborative work; the results receive feedback; reflection on the feedback leads to a new, more complex task. This cycle is similar to the experiential learning cycle [5], but each stage is supported by specific digital tools. When one link is missing, for example when there is no feedback, the cycle weakens and digital activity becomes routine.

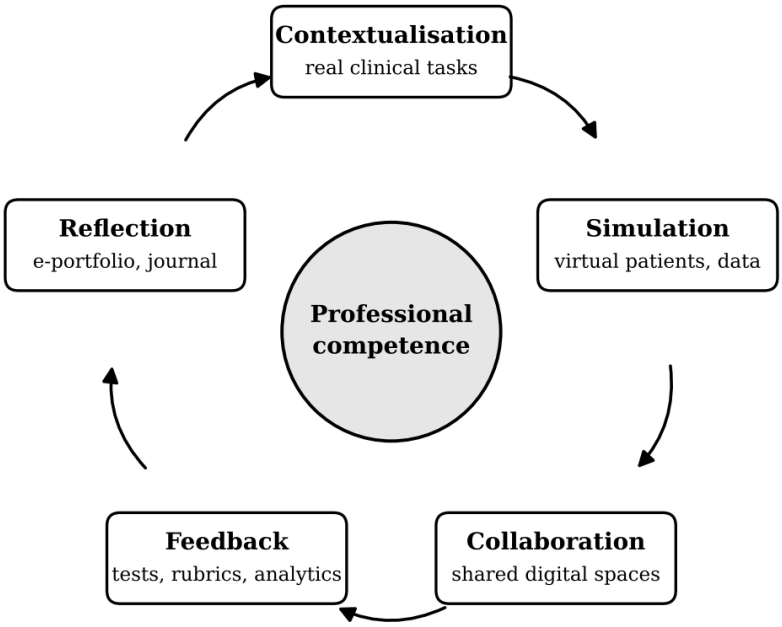


Figure 1. Cycle of mechanisms for developing professional competence through digital technologies

The cycle can be illustrated with a practical lesson in the course "Information Technologies in Medicine". First, students receive a clinical case: a table of blood test results for a group of patients (contextualisation). Second, they work with an anonymised data set in a spreadsheet, calculate indicators and build charts (simulation). Third, they prepare a short report together in a shared cloud document, dividing the roles of analyst, designer and presenter (collaboration). Fourth, the teacher evaluates the report with a rubric, and an online quiz checks understanding of the key concepts (feedback). Finally, each student writes a short reflection in the e-portfolio: what they learned, what was difficult and what they will do differently next time (reflection). In the next lesson, the task becomes more complex, for example including data protection requirements.

The mechanisms also show why the same technology can have very different effects. An LMS used only to upload lecture files supports none of the mechanisms. The same LMS, used for contextual tasks, discussion forums, quizzes with feedback and e-portfolios, supports all five. In this sense, the effect depends not on the tool itself but on the mechanism it is used for.

Based on this analysis, the following recommendations can be made for teachers:

1. Start lesson planning from the mechanism, and only then choose the digital tool.
2. Build every digital task on a real or realistic clinical situation.
3. Use simulation and data sets for repeated practice, not only for demonstration.
4. Give clear roles in group work and use shared documents to make contributions visible.
5. Combine quick automatic feedback with criteria-based feedback from the teacher.
6. Make reflection a regular, short activity rather than a large final report.

Conclusion

Digital technologies develop medical students' professional competence through five interrelated mechanisms: contextualisation, simulation, collaboration, feedback and reflection. Each mechanism has a theoretical basis, is supported by specific digital tools and develops particular components of competence. The mechanisms form a cycle, and the effect of digital learning weakens when any link is missing.

For practical teaching, this means that the choice of a digital tool should start from the mechanism it is meant to support. The proposed cycle can be used to design practical lessons in the course "Information Technologies in Medicine" and similar courses. In future research, the contribution of each mechanism can be tested empirically, for example by comparing groups that use the full cycle with groups that use only some of its elements.

REFERENCES

1. Khurana, M. P., Raaschou-Pedersen, D. E., Kurtzhals, J., Bardram, J. E., Ostrowski, S. R., & Bundgaard, J. S. (2022). Digital health competencies in medical school education: a scoping review and Delphi method study. *BMC Medical Education, 22*(1), 129. https://doi.org/10.1186/s12909-022-03163-7
2. Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The Digital Competence Framework for Citizens*. Publications Office of the European Union. https://doi.org/10.2760/115376
3. Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: a review of the literature. *Improving Schools, 19*(3), 267–277. https://doi.org/10.1177/1365480216659733
4. Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
5. Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Prentice Hall.
6. Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2
7. Cook, D. A., Erwin, P. J., & Triola, M. M. (2010). Computerized virtual patients in health professions education: a systematic review and meta-analysis. *Academic Medicine, 85*(10), 1589–1602. https://doi.org/10.1097/ACM.0b013e3181edfe13
8. Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., Gordon, D. L., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher, 27*(1), 10–28. https://doi.org/10.1080/01421590500046924
9. Driessen, E., van Tartwijk, J., van der Vleuten, C., & Wass, V. (2007). Portfolios in medical education: why do they meet with mixed success? A systematic review. *Medical Education, 41*(12), 1224–1233. https://doi.org/10.1111/j.1365-2923.2007.02944.x
10. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487
11. Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795
12. Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.

FACTORS INFLUENCING THE DEVELOPMENT OF STUDENTS' PROFESSIONAL COMPETENCE IN A DIGITAL LEARNING ENVIRONMENT

Davronov Shakhbos Erkin ugli[1,*], Kakhkhorov Siddik K.[2]
[1] Doctoral student, Bukhara State University, Bukhara, Uzbekistan
[2] Department of Heliophysics, Renewable Energy Sources and Electronics, Bukhara State University, Bukhara, Uzbekistan
* Corresponding author: shaxbos.davronov.91@gmail.com
https://doi.org/10.5281/zenodo.23256057

Abstract. *The digital transformation of higher education has changed what professional competence means for future specialists. This paper identifies and classifies the factors that influence the development of students' professional competence in a digital learning environment, with particular attention to medical universities. Based on an analysis of the literature and international frameworks (DigComp 2.2, DigCompEdu, TPACK and the Technology Acceptance Model), the factors are grouped into three levels: personal (motivation, digital self-efficacy, prior digital experience), pedagogical (content integration, teaching methods, assessment, teacher digital competence) and institutional (digital infrastructure, regulatory framework, organisational support). The paper describes how each factor works, shows the hierarchical relationship between the three levels and proposes practical recommendations for universities. It argues that the development of professional competence depends less on the availability of technology than on its pedagogical use.*

Keywords: *professional competence, digital learning environment, influencing factors, DigComp 2.2, medical education, higher education.*

Introduction

Digitalisation has become one of the main drivers of change in higher education. Employers now expect graduates not only to know their subject but also to use digital tools confidently, critically and responsibly in their professional work. The European DigComp 2.2 framework describes this ability through five areas: information and data literacy, communication and collaboration, digital content creation, safety, and problem solving [1]. Together with subject knowledge and practical skills, digital competence has become an integral part of professional competence.

In medicine, this demand is especially strong. Electronic health records, telemedicine, laboratory information systems and digital diagnostic tools are now part of daily clinical practice. A doctor who cannot find reliable information, interpret digital data or protect patient confidentiality is less prepared for modern healthcare. Nevertheless, a review of medical curricula shows that digital health competencies are often taught as optional topics or in separate informatics courses, rather than being integrated into professional training [2].

In Uzbekistan, the "Digital Uzbekistan – 2030" Strategy has made the development of digital education and students' digital competence a national priority [3]. Universities have invested in computer laboratories, learning management systems (LMS) and online resources. However, teaching experience shows that investment in technology does not automatically lead

to higher professional competence. In some groups, the same equipment and the same LMS produce very different results. This raises a practical question: which factors actually determine whether a digital learning environment develops students' professional competence?

The aim of this paper is to identify and classify these factors, to show how they interact and to formulate recommendations for their use in teaching. The study focuses on medical universities, but most of the conclusions can be applied to other professional fields.

Theoretical background

The concept of digital competence has developed over the last decade. The first DigComp framework described digital competence as a set of knowledge, skills and attitudes needed to use technology in work, learning and everyday life [4]. Later versions added proficiency levels and examples of use, and DigComp 2.2 included detailed examples of knowledge, skills and attitudes for each competence [1]. For teachers, the DigCompEdu framework describes how educators use technology for professional engagement, teaching, assessment and empowering learners [5].

Several theories explain why people use or do not use technology. The Technology Acceptance Model (TAM) states that the use of a technology depends mainly on its perceived usefulness and perceived ease of use [6]. The Unified Theory of Acceptance and Use of Technology (UTAUT) adds social influence and facilitating conditions [7]. Bandura's concept of self-efficacy explains why a person's belief in their own ability affects whether they attempt difficult tasks [8]. For teachers, the TPACK model shows that effective use of technology requires the integration of technological, pedagogical and content knowledge [9].

Research on technology integration also distinguishes between first-order barriers (lack of equipment, time, support) and second-order barriers (teachers' beliefs and attitudes) [10]. Later studies confirm that teachers' pedagogical beliefs strongly shape how they use technology in class [11]. These theories show that the development of competence in a digital environment is influenced by factors at different levels: the individual student, the teaching process and the institution.

Methods

The study is theoretical. It is based on an analysis of scientific and pedagogical literature, a comparative analysis of international digital competence frameworks and the generalisation of the authors' teaching experience in the course "Information Technologies in Medicine". The literature search focused on three topics: digital competence frameworks, technology acceptance and integration in education, and digital competence in medical education. The factors found in the literature were grouped by the level at which they operate and by their role in the learning process.

Results

The analysis made it possible to group the factors into three levels: personal, pedagogical and institutional (Table 1).

Table 1.

Factors influencing the development of students' professional competence in a digital learning environment

Level	Factor	How it influences competence development
Personal	Motivation and perceived usefulness of digital tools	Students who see a tool as useful for their future profession use it more actively and more deeply [6]

Level	Factor	How it influences competence development
	Digital self-efficacy	Belief in one's own ability determines whether a student attempts complex digital tasks [8]
	Prior digital experience	Everyday digital skills (social media, messaging) do not automatically transfer to professional tasks [12]
Pedagogical	Integration of digital tasks into professional content	Competence develops when digital tools are used to solve profession-specific problems, not in isolation
	Teaching methods	Active, project-based and collaborative methods develop more competence areas than lectures [13]
	Assessment	Criteria-based assessment shows students and teachers which areas need improvement
	Teacher digital competence	The teacher's own readiness (DigCompEdu, TPACK) limits what students can achieve [5, 9, 14]
Institutional	Digital infrastructure	Access to devices, internet and an LMS is a necessary but not a sufficient factor
	Regulatory framework	National strategies and curricula set the priority and the time given to digital competence [3]
	Organisational support	Teacher training, technical support and a culture of innovation sustain long-term change [7]

Personal factors. The first personal factor is motivation, which is closely linked to the perceived usefulness of digital tools. When a medical student understands that a skill, for example searching a medical database, will be needed in clinical work, they are more likely to practise it seriously. The second factor is digital self-efficacy. Students with low confidence often avoid complex tasks, such as working with data or creating digital content, even when they have enough basic skills. The third factor is prior digital experience. It is often assumed that today's students are "digital natives" who already know how to use technology. However, research shows that this assumption is not supported by evidence: everyday use of smartphones and social media does not guarantee the ability to use technology for learning or professional tasks [12].

Pedagogical factors. The most important pedagogical factor is the integration of digital tasks into professional content. When digital skills are taught in isolation, for example as general computer exercises, students rarely transfer them to their profession. When the same skills are used to solve medical problems, such as analysing a table of laboratory results or preparing health information for patients, competence develops together with professional understanding. The second factor is the choice of teaching methods. Project-based and collaborative methods require students to plan, search, analyse, create and present, which covers several areas of digital competence at once [13]. The third factor is assessment. If only knowledge is assessed, students focus on memorising; if practical tasks are assessed with clear criteria, they focus on doing. The fourth factor is the teacher's own digital competence, which is discussed below.

Institutional factors. Digital infrastructure includes computers, internet access, an LMS and access to electronic resources. Without it, digital learning is impossible, but its presence alone does not guarantee results. The regulatory framework, including national strategies, state

educational standards and curricula, determines how much time and attention is given to digital competence. Organisational support includes teacher training, technical help and a culture that encourages innovation. According to UTAUT, such "facilitating conditions" directly affect whether technology is actually used [7].

Discussion

Three observations follow from this classification (Figure 1).

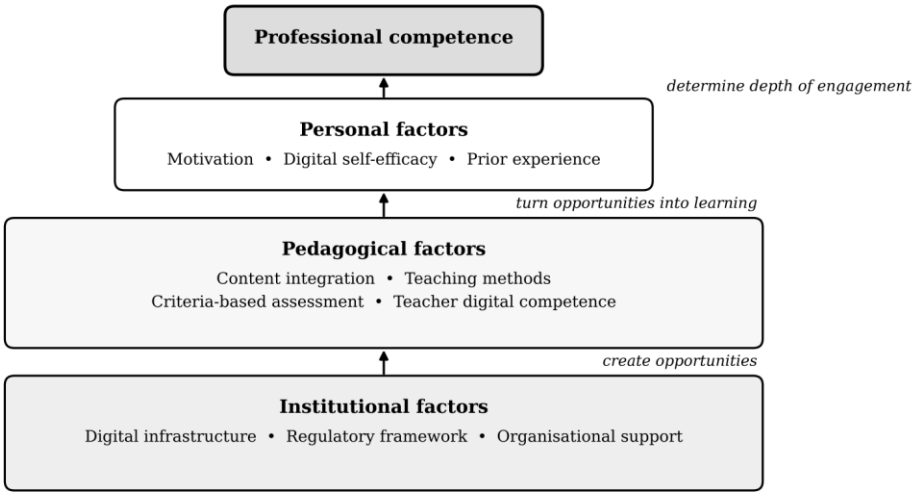


Figure 1. Hierarchical relationship between the levels of factors

First, the factors are hierarchical. Institutional factors create opportunities, pedagogical factors turn these opportunities into learning activities, and personal factors determine how deeply each student engages. A university with excellent infrastructure but weak pedagogical integration will see limited results. This is consistent with research showing that teachers' acceptance and use of technology is shaped by perceived usefulness and ease of use, not by access alone [15]. It also matches the distinction between first-order and second-order barriers: removing equipment barriers is easier than changing teaching practice [10].

Second, the teacher is the key link between the levels. Studies of academics in higher education show that their digital competence is uneven [14]. If a teacher lacks the knowledge to combine technology, pedagogy and subject content, digital tools are used only for presentation and information transfer. Teachers' beliefs about teaching also matter: a teacher who sees learning as information transfer will use technology differently from a teacher who sees learning as active construction [11]. For this reason, teacher professional development should be seen as a factor of student competence, not as a separate issue.

Third, in medical education the most important pedagogical factor is the professional context of digital tasks. A medical student develops professional digital competence when searching for reliable clinical information, analysing patient data or protecting medical confidentiality, not when performing general computer exercises. Health professionals' digital competence is increasingly measured with tools adapted to the healthcare context [16], and teaching should follow the same logic. The gap between clinical needs and current curricula noted in the literature [2] can be reduced mainly through this factor.

The factors also influence each other. For example, well-designed professional tasks (a pedagogical factor) increase perceived usefulness and motivation (a personal factor). Successful completion of such tasks increases self-efficacy, which leads students to attempt more complex

tasks. In the opposite direction, poorly organised digital activities can reduce motivation and confidence. Therefore, the pedagogical level should be the main point of intervention.

Based on this analysis, the following practical recommendations can be made for universities:

1. Plan investment in digital infrastructure together with teacher training, not separately from it.
2. Integrate digital tasks into professional subjects instead of teaching them only in separate informatics courses.
3. Use active teaching methods, such as project-based and collaborative learning, in practical lessons.
4. Assess digital competence with clear criteria and practical tasks, not only with knowledge tests.
5. Show students the professional value of each digital skill to strengthen motivation and self-efficacy.

Conclusion

The development of students' professional competence in a digital learning environment is influenced by three groups of interrelated factors: personal, pedagogical and institutional. These groups form a hierarchy: institutional factors create opportunities, pedagogical factors turn them into learning, and personal factors determine the depth of engagement. Technology is a necessary condition, but the decisive factors are pedagogical: integration of digital tasks into professional content, active teaching methods, criteria-based assessment and the teacher's own digital competence.

The proposed classification can be used to design a methodology for developing professional competence in the course "Information Technologies in Medicine" and similar courses. In future research, the relative weight of each factor can be measured empirically, for example through surveys of students and teachers and structural equation modelling.

REFERENCES

6. Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The Digital Competence Framework for Citizens*. Publications Office of the European Union. https://doi.org/10.2760/115376
7. Khurana, M. P., Raaschou-Pedersen, D. E., Kurtzhals, J., Bardram, J. E., Ostrowski, S. R., & Bundgaard, J. S. (2022). Digital health competencies in medical school education: a scoping review and Delphi method study. *BMC Medical Education, 22*(1), 129. https://doi.org/10.1186/s12909-022-03163-7
8. Decree of the President of the Republic of Uzbekistan No. PF-6079 of 5 October 2020 "On approval of the 'Digital Uzbekistan – 2030' Strategy and measures for its effective implementation".
9. Ferrari, A. (2013). *DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe*. Publications Office of the European Union. https://doi.org/10.2788/52966
10. Redecker, C. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. Publications Office of the European Union. https://doi.org/10.2760/159770
11. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

12. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540
13. Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191
14. Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: a framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
15. Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: strategies for technology integration. *Educational Technology Research and Development, 47*(4), 47–61. https://doi.org/10.1007/BF02299597
16. Tondeur, J., van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: a systematic review of qualitative evidence. *Educational Technology Research and Development, 65*(3), 555–575. https://doi.org/10.1007/s11423-016-9481-2
17. Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education, 67*, 135–142. https://doi.org/10.1016/j.tate.2017.06.001
18. Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: a review of the literature. *Improving Schools, 19*(3), 267–277. https://doi.org/10.1177/1365480216659733
19. Inamorato dos Santos, A., Chinkes, E., Carvalho, M. A. G., Solórzano, C. M. V., & Marroni, L. S. (2023). The digital competence of academics in higher education: is the glass half empty or half full? *International Journal of Educational Technology in Higher Education, 20*, 9. https://doi.org/10.1186/s41239-022-00376-0
20. Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): a meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education, 128*, 13–35. https://doi.org/10.1016/j.compedu.2018.09.009
21. Mainz, A., Nitsche, J., Weirauch, V., & Meister, S. (2024). Measuring the digital competence of health professionals: scoping review. *JMIR Medical Education, 10*, e55737. https://doi.org/10.2196/55737

MAKTABGACHA TA'LIM TASHKILOTI RAHBARINING SHAXSIY SIFATLARI VA EMOTSIONAL INTELLEKTINING KONFLIKTLARNI BOSHQARISHDAGI O'RNI

D.R. Raxmonova
NamDU tayanch doktoranti
https://doi.org/10.5281/zenodo.23256078

Annotatsiya. *Maqolada maktabgacha ta'lim tashkiloti rahbarining shaxsiy sifatlari va emotsional intellekti nizolarni boshqarish samaradorligiga qanday ta'sir etishi tahlil qilinadi. Emotsional intellektning tarkibiy qismlari nizoli vaziyatlar bilan bog'lab ko'rsatilib, uni rivojlantirishning amaliy yo'llari taklif etiladi.*

Kalit so'zlar: *emotsional intellekt, shaxsiy sifatlar, rahbar, maktabgacha ta'lim tashkiloti, nizolarni boshqarish, empatiya, o'zini boshqarish.*

Bir xil nizoli vaziyatga duch kelgan ikki rahbar butunlay boshqacha natijaga erishishi mumkin. Biri xotirjam suhbat orqali ikki tarbiyachini yarashtiradi, ikkinchisi xuddi shu holatda ovozini balandlatib, muammoni yanada chigallashtiradi. Farq bilim yoki tajriba miqdorida emas, ko'pincha rahbarning o'zini va boshqalarni his qila olish darajasida bo'ladi. So'nggi o'ttiz yilda psixologiyada bu qobiliyat emotsional intellekt nomi bilan o'rganilmoqda. Maktabgacha ta'lim tashkilotida (MTT) bu masala alohida ahamiyatga ega. Jamoa asosan tarbiyachilardan iborat, ular bolalar bilan ishlaganda hissiy jihatdan doimiy zo'riqishda bo'ladi, ota-onalar esa farzandi haqida gap ketganda his-tuyg'ularini jilovlay olmasligi mumkin.

Bunday muhitda rahbarning hissiy holati butun jamoaning kayfiyatiga ta'sir qiladi. Maqolaning maqsadi rahbarning shaxsiy sifatlari va emotsional intellekt nizolarni boshqarishda qanday rol o'ynashini ko'rsatishdir. Rahbarning nizoga munosabatini belgilaydigan shaxsiy sifatlar orasida bir necha xususiyat ayniqsa ko'zga tashlanadi. Birinchisi: hissiy barqarorlik; tanqid yoki ochiq e'tirozga duch kelganda o'zini yo'qotmaydigan rahbar muloqotni saqlab qola oladi. Ikkinchisi: adolatlilik: xodimlar rahbarning qaroridan ko'ra uning qaror qabul qilish tartibi adolatli bo'lganini muhimroq deb biladi. Uchinchisi: tinglash qobiliyati va ochiqlik, ya'ni boshqa fikrni darhol rad etmasdan eshitish odati. To'rtinchisi: mas'uliyatni o'z zimmasiga olish va xatosini tan olishga tayyorlik. Haddan tashqari avtoritarlik, tez jahl qilish, shubhalanuvchanlik yoki nizodan qochish odati boshqaruvda muammo tug'diradi. Avtoritar rahbar nizoni bostiradi, biroq uning sabablarini hal qilmaydi; nizodan qochuvchi rahbar esa muammoning o'zi yo'qolib ketishini kutadi. Ikkala holatda ham kelishmovchilik yashirin shaklga o'tadi. Shaxsiy sifatlar tug'ma xususiyat sifatida qaralmasligi kerak: ularning ko'pi ongli mashq va refleksiya orqali o'zgaradi. Emotsional intellekt tushunchasini ilmiy muomalaga Salovey va Mayer kiritgan. Ularning fikricha, u o'z va boshqalarning hissiyotlarini aniqlash, tushunish va ulardan fikrlash hamda xulq-atvorni yo'naltirishda foydalanish qobiliyatidir. Goleman bu g'oyani boshqaruv sohasiga kengaytirib, EIni o'zini anglash, o'zini boshqarish, motivatsiya, empatiya va ijtimoiy ko'nikmalar kabi tarkibiy qismlarga ajratgan.

Bar-On esa uni shaxsning stress bilan kurashish va muhitga moslashish qobiliyati bilan ham bog'lagan. Bu tarkibiy qismlarning har biri nizoli vaziyatda aniq vazifa bajaradi. O'zini anglash rahbarga nizo paytida qanday his qilayotganini ko'rishga yordam beradi: g'azabmi,

xafagarchilikmi yoki qo‘rquvmi. Bu bilan u hissiyot ta’sirida qaror qabul qilishdan saqlanadi. O‘zini boshqarish javob berishdan oldin to‘xtab olish imkonini beradi. Empatiya nizodagi ikkinchi tomonning nima uchun shunday harakat qilayotganini tushunishga xizmat qiladi, bu esa ko‘rinib turgan pozitsiya ortidagi haqiqiy manfaatni topishning asosiy sharti. Ijtimoiy ko‘nikmalar esa tomonlarni muzokara stoliga o‘tirg‘izish va kelishuvga erishishda kerak bo‘ladi.

Tadqiqotlar emotsional intellekt yuqori bo‘lgan a’zolarga ega jamoalarda nizolar odatda vazifaga yo‘naltirilgan holda, shaxsiy hujumlarsiz hal etilishini ko‘rsatgan. Shu bilan birga, emotsional intellekt hamma narsani hal qiladigan universal kalit emas: u huquqiy bilim, tashkiliy tuzilmaning aniqligi va muzokara texnikasi o‘rnini bosmaydi, ularni to‘ldiradi.

Shaxsiy sifatlar bilan emotsional intellekt o‘rtasida uzviy bog‘liqlik bor: hissiy barqarorlik o‘zini boshqarishning, tinglash qobiliyati esa empatiyaning amaldagi ko‘rinishidir. Shuning uchun rahbarni baholashda bu ikkalasini alohida-alohida emas, birgalikda ko‘rgan ma’qul. O‘zbek jamoaviy madaniyatida bu masalaning yana bir jihati bor: xodimlar ko‘pincha rahbar bilan ochiq bahslashishga jur’at etmaydi, e’tirozini yashirin saqlaydi. Bunday sharoitda rahbarning sezgirligi, ya’ni aytilmagan noroziliklarni ilg‘ay olishi alohida qimmatga ega bo‘ladi. Masalan, ota-ona ertalab guruh oldida tarbiyachiga e’tiroz bildirdi. Emotisonal intellekti past rahbar odatda yo himoyaga o‘tadi va tarbiyachini oqlashga urinadi, yo ota-onaning talabiga tezda rozi bo‘lib, xodimini xor qiladi. Emotsional intellekt rivojlangan rahbar avval o‘zining hissiy javobini nazorat qiladi, suhbatni jamoatchilik oldidan olib chiqadi, ota-onaning xavotirini eshitadi, so‘ng tarbiyachi bilan alohida gaplashib, vaziyatning ikkinchi tomonini aniqlaydi. Natijada har ikkala tomon o‘zini hurmat qilingan his qiladi va nizo uzoqqa cho‘zilmaydi. Emotsional intellektni rivojlantirish uchun bir necha amaliy yo‘l mavjud.

Rahbar har bir og‘ir suhbatdan keyin qisqa yozma refleksiya qilib borishi, ya’ni nima his qilgani va qanday harakat qilgani haqida yozishi mumkin. Rol o‘yinlari va keys tahlillari yordamida qiyin suhbatlarni oldindan mashq qilish foydali. Hamkasblar bilan o‘zaro fikr almashish guruhlari va tajribali mentor bilan ishlash ham yaxshi natija beradi. Bundan tashqari, rahbar o‘zining emotsional intellekt darajasini avval standartlashtirilgan asbob (masalan, EQ-i so‘rovnomasi) va xodimlarning yashirin fikri orqali aniqlab olsa, rivojlanish yo‘nalishi aniqroq bo‘ladi. Rahbarning hissiy holati jamoa iqlimiga bevosita ta’sir etishini ham unutmaslik kerak. Tarbiyachilar rahbarning kayfiyatiga qarab o‘z xulq-atvorini moslashtiradi va bu, o‘z navbatida, bolalar bilan muloqotda ham aks etadi. Demak, rahbar emotsional intellekt ni rivojlantirar ekan, faqat o‘z ish samaradorligiga emas, butun tashkilotning ijtimoiy-psixologik muhitiga, oxir-oqibat bolalarning muhitiga ham ta’sir ko‘rsatadi.

MTT rahbarining shaxsiy sifatlari va emotsional intellekti nizolarni boshqarishning asosiy shartlaridan biri hisoblanadi. Hissiy barqarorlik, adolatlilik, tinglash va empatiya rahbarga nizoni tahdid sifatida emas, jamoani mustahkamlash imkoniyati sifatida ko‘rish imkonini beradi. Bu sifatlar mashq va refleksiya orqali rivojlanadi, shu sababli rahbarlarni qayta tayyorlash va malaka oshirish dasturlariga emotsional intellektni shakllantirish bo‘yicha amaliy mashg‘ulotlarni kiritish maqsadga muvofiq.

FOYDALANILGAN ADABIYOTLAR

1. Goleman D. Emotional Intelligence: Why It Can Matter More Than IQ. –New York: Bantam Books, 1995.
2. Salovey P., Mayer J.D. Emotional intelligence // Imagination, Cognition and Personality, 1990, Vol. 9, No. 3. P. 185–211.

3. Bar-On R. The Bar-On Model of Emotional-Social Intelligence (ESI) // Psicothema, 2006, Vol. 18, Suppl. – P. 13–25.
4. Jordan P.J., Troth A.C. Managing emotions during team problem solving: Emotional intelligence and conflict resolution // Human Performance, 2004, Vol. 17, No. 2. – P. 195–218.
5. Гришина Н.В. Психология конфликта. –СПб.: Питер, 2008.
6. Mayer J.D., Salovey P., Caruso D.R. Emotional intelligence: Theory, findings, and implications // Psychological Inquiry, 2004, Vol. 15, No. 3. - P. 197–215.
7. Goleman D., Boyatzis R., McKee A. Primal Leadership: Realizing the Power of Emotional Intelligence. – Boston: Harvard Business School Press, 2002.
8. Lopes P.N., Salovey P., Straus R. Emotional intelligence, personality, and the perceived quality of social relationships // Personality and Individual Differences, 2003, Vol. 35, No. 3. – P. 641–658.
9. Rahim M.A. Managing Conflict in Organizations. 4th ed. – New Brunswick: Transaction Publishers, 2011.

PEDAGOGLARDA RAQAMLI KOMPETENSIYANI SHAKLLANTIRISH SAMARADORLIGI

Solihova Maftuna Sharoq qizi
Xalqaro Nordik Universiteti
ORCID: 0009-0009-1138-1914
https://doi.org/10.5281/zenodo.23256100

Annotatsiya. *Ushbu tezis* ***"Raqamli O‘zbekiston — 2030" strategiyasi*** *doirasida pedagoglarning texnologik va metodologik tayyorgarligini tubdan qayta ko‘rib chiqish muammolariga bag‘ishlangan. Tadqiqotda an’anaviy savodxonlikdan farqli o‘laroq, o‘qituvchilarning generativ sun’iy intellekt va data-tahlil (Data Literacy) vositalari bilan ishlash ko‘nikmalari ta’lim samaradorligini oshirishning* ***bosh strategik omili*** *sifatida tahlil qilingan. Maqolada pedagogik va texnologik bilimlarning integratsiyalashuv qonuniyatlari hamda raqamli kompetensiyaning dars sifatiga ko‘rsatadigan bevosita funksional ta’siri ilmiy jihatdan asoslangan.*

Kalit so‘zlar: *raqamli kompetensiya, "Raqamli O‘zbekiston — 2030", prompt-injenering, generativ intellekt, data-savodxonlik, TPACK modeli, ta’lim transformatsiyasi.*

Kirish. Bugungi kunda ta’lim makonini raqamlashtirish bosqichi oddiy kompyuter texnikasini joriy etish darajasidan o‘tib, generativ sun’iy intellekt, neyrotarmoqlar va katta ma’lumotlar (Big Data) bilan ishlash davriga qadam qo‘ydi. Zamonaviy pedagog nafaqat tayyor raqamli kontentdan foydalanuvchi, balki intellektual ta’lim muhitini ssenariylashtiruvchi va o‘quvchilarning kognitiv yuklamasini optimallashtiruvchi muhandis-pedagogga aylanishi zarur. Shu bois, o‘qituvchilar tarkibining raqamli kompetentligini yangi sifat bosqichida shakllantirish va uning samaradorligini baholash pedagogika fani oldidagi eng dolzarb vazifadir.

O‘zbekiston Respublikasining uzluksiz ta’limni rivojlantirishga oid davlat dasturlarida o‘qituvchilarning texnologik moslashuvchanligini oshirish masalasiga alohida e’tibor qaratilmoqda. Bugungi kun raqamli kompetensiyasi o‘z ichiga algoritmik fikrlash, kiberxavfsizlik madaniyati hamda dars jarayonida adaptiv ta’lim platformalarini qo‘llay olish qobiliyatini qamrab oladi. Ushbu ko‘nikmalarni tizimli shakllantirmasdan turib, zamonaviy maktab infratuzilmasidan samarali foydalanish imkonsizdir.

Asosiy qism: Tadqiqotning bosh maqsadi — pedagoglarning raqamli kompetensiyasini TPACK (Technological Pedagogical Content Knowledge) xalqaro modeli asosida rivojlantirish mexanizmini va uning dars unumdorligiga ta’sirini ilmiy jihatdan baholashdan iborat. Tadqiqot obyekti sifatida o‘qituvchilarning kasbiy faoliyatini tahlil qilish jarayonida quyidagi aniq vazifalar bajarildi:

1. Pedagogik neyromodellashtirish va raqamli pedagogika tamoyillari asosida "zamonaviy o‘qituvchining raqamli profili" mezonlarini ishlab chiqish.
2. O‘qituvchilarning dars samaradorligini oshirishda generativ sun’iy intellekt (AI) resurslaridan va prompt-injenering usullaridan foydalanish darajasini aniqlash.
3. Raqamli mikrokvalifikatsiyalar (Micro-credentials) tizimi orqali pedagoglarning malakasini uzluksiz oshirish modelini eksperimentdan o‘tkazish.

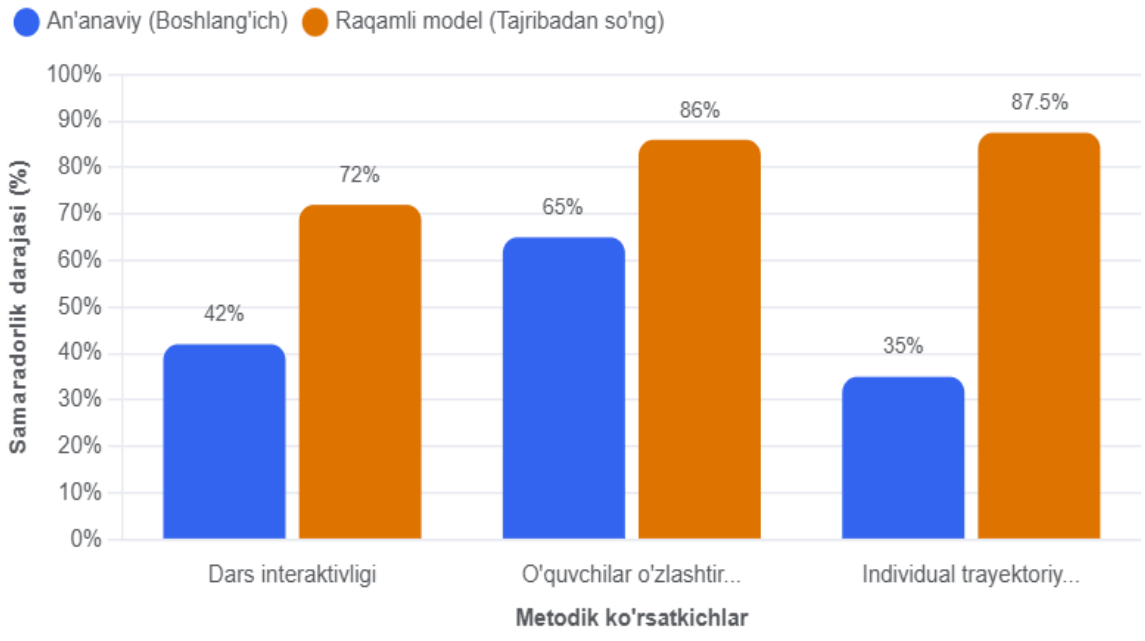
4. Olingan natijalarning dars sifatiga va o‘quvchilar motivatsiyasiga ko‘rsatgan ta’sirini tahlil qilish.

Tadqiqot doirasida tizimli-faoliyatli yondashuv, pedagogik ssenariylashtirish, TPACK diagnostik vositalari hamda o‘quv natijalarini korrelyatsion-statistik tahlil qilish metodlaridan foydalanildi.

Olib borilgan amaliy tajriba-sinov ishlari natijasida quyidagi yangi ilmiy xulosalar olindi:

Tadqiqot doirasida pedagoglarning raqamli kompetensiyasini rivojlantirish samaradorligi korrelyatsion-statistik hamda qiyosiy tahlil metodlari yordamida baholandi. Tajriba-sinov guruhlarida TPACK xalqaro modeli joriy etilgandan so‘ng olingan miqdoriy ko‘rsatkichlar quyidagi matematik-statistik natijalarni qayd etdi:

1. **Metodik transformatsiya unumdorligi:** O‘qituvchilarning generativ sun’iy intellekt (AI) va prompt-injenering vositalari bilan ishlash ko‘nikmalari tizimli oshirilishi natijasida, dars ssenariylarining interaktivlik koeffitsiyenti **an’anaviy 42% dan 72% gacha** ko‘tarildi (o‘sish ko‘rsatkichi **30% ni** tashkil etdi).
2. **Kognitiv va akademik o‘zlashtirish samaradorligi:** Raqamli pedagogik texnologiyalarni amaliyotda to‘g‘ri integratsiya qilgan pedagoglar dars bergan sinflarda o‘quvchilarning fundamental mavzularni o‘zlashtirish chuqurligi hamda mantiqiy xulosalar chiqarish ko‘nikmasi nazorat guruhlariga nisbatan **21% ga yuqoriligi** aniqlandi.
3. **O‘quv jarayonini shaxsiylashtirish va data-tahlil aniqligi:** Pedagoglarning raqamli platformalar tahliliy ma’lumotlari (Learning Analytics) bilan ishlash va data-savodxonlik (Data Literacy) darajasi oshishi natijasida, o‘quvchilarga individual ta’lim trayektoriyasini to‘g‘ri prognoz qilish va belgilash aniqligi **2.5 barobarga** (yoki **150% ga**) yaxshilandi.
4. **Vaqt resurslarini optimallashtirish:** Raqamli vositalar orqali darsni rejalashtirish, kontent yaratish va baholash jarayonlarini avtomatlashtirgan o‘qituvchilarning shaxsiy-metodik ustida ishlash va darsga tayyorlanish vaqti haftasiga o‘rtacha **4,5 soatdan 6 soatgacha** tejalgandir.



Xulosa: Pedagoglarda raqamli kompetensiyani shakllantirish an’anaviy malaka oshirish kurslari doirasidan chiqib, uzluksiz raqamli ekotizim shakliga o‘tishi zarur. Tadqiqot natijalari shuni ko‘rsatadiki, o‘qituvchining zamonaviy texnologik tayyorgarligi ta’lim sifatini oshirish, o‘quv motivatsiyasini barqarorlashtirish va raqamli avlod bilan samarali muloqot o‘rnatishning asosiy kafolatidir. Pedagoglarning raqamli kompetentligini baholash mezonlarini muntazam yangilab borish va maktab ichki pedagogik kengashlarida o‘zaro raqamli mentoring tizimini yo‘lga qo‘yish maqsadga muvofiqdir.

FOYDALANILGAN ADABIYOTLAR

1. To‘raqulov, R. (2022). Raqamli pedagogika va ta’lim texnologiyalari transformatsiyasi. Toshkent.
2. Qodirov, A. (2024). Sun’iy intellekt asrida pedagogik kompetensiyaning yangi qirralari. Ta’lim va innovatsiya, 5(2), 12–18.
3. O‘zbekiston Respublikasi Prezidenti. (2020). "Raqamli O‘zbekiston — 2030" strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to‘g‘risidagi Farmoni. Toshkent.
4. Ismoilov, B. (2025). Ta’limda data-savodxonlik va uning pedagogik ahamiyati. Pedagogik mahorat, 1(3), 88-95.

TABIIY FANLARNI O‘QITISHDA RAQAMLI TA’LIM MUHITINI TAKOMILLASHTIRISHNING PEDAGOGIK MEXANIZMLARI

Solihova Maftuna Sharof qizi
Xalqaro Nordik Universiteti
E-mail: m.solihova@nordicuniversity.org
ORCID: 0009-0009-1138-1914
Telefon raqami: +998970406611
https://doi.org/10.5281/zenodo.23256116

Annotatsiya. *Ushbu tezis zamonaviy axborot jamiyati sharoitida tabiiy fanlarni (fizika, kimyo, biologiya, geografiya) o‘qitishda raqamli ta’lim muhitini shakllantirish, uni tizimli takomillashtirish va dars jarayonlariga integratsiya qilishning pedagogik mexanizmlari tadqiqiga bag‘ishlangan. Maqolada an’anaviy ta’lim usullarini raqamli platformalar, elektron darsliklar, interaktiv virtual laboratoriyalar hamda masofaviy ta’lim elementlari bilan uyg‘unlashtirishning tashkiliy-pedagogik shartlari yoritilgan. Olib borilgan eksperimental tadqiqot natijalari asosida raqamli ta’lim muhitining o‘quvchilar bilish faolligini, mustaqil ta’lim olish hamda kreativ fikrlash kompetensiyalarini rivojlantirishdagi samaradorligi va ustuvor ahamiyati ilmiy jihatdan dalillab berilgan.*

Kalit so‘zlar: *raqamli ta’lim muhiti, tabiiy fanlar metodikasi, elektron darslik, masofaviy ta’lim, innovatsion pedagogika, pedagogik mexanizmlar, virtual laboratoriya, bilish faolligi.*

Kirish. Bugungi kunda jahon miqyosida axborot-kommunikatsiya texnologiyalarining shiddatli va jadal sur’atlar bilan rivojlanishi ijtimoiy hayotning barcha jabhalari qatori ta’lim makonining ham arxitektonikasini tubdan o‘zgartirmoqda. Globallashuv va raqamli transformatsiya davrida, ayniqsa, tajriba, kuzatish va amaliy tadqiqotlarga asoslangan tabiiy fanlarni o‘qitish tizimida raqamli ta’lim muhitini yaratish o‘quvchilarning intellektual salohiyatini hamda bilish faoliyatini faollashtiruvchi strategik omil bo‘lib xizmat qilmoqda. Zamonaviy raqamli resurslar yordamida o‘quvchilar makro va mikroolamdagi murakkab tabiiy jarayonlarni, kimyoviy reaksiyalarni yoki fizik hodisalarni vizual, dinamik hamda interaktiv tarzda, real vaqt rejimida modellashtirish hamda chuqur o‘zlashtirish imkoniyatiga ega bo‘lmoqdalar.

Postindustrial va yuqori texnologiyalarga asoslangan axborot jamiyatiga o‘tish zaruriyati jahon iqtisodiyoti va ilm-fanida mutlaqo yangi texnologik tartibni shakllantirdi. Axborot resurslari endilikda ijtimoiy-iqtisodiy va madaniy taraqqiyotning real harakatlantiruvchi kuchiga aylandi. Shu sababli, milliy ta’lim tizimini xalqaro andozalar va ilg‘or xorijiy tajribalar asosida modernizatsiya qilish, o‘quv jarayonlariga raqamli texnologiyalarni uzviy joriy etish davlat siyosatining ustuvor yo‘nalishlaridan biri hisoblanadi. So‘nggi yillarda olib borilgan pedagogik kuzatishlar shuni ko‘rsatadiki, raqamli va elektron ta’limning an’anaviy o‘qitish shakllariga qaraganda afzallik hamda qulaylik jihatlari ko‘proq namoyon bo‘lmoqda. Multimedia vositalari va intellektual tizimlar yordamida darslarni tashkil etishda uzatiladigan bilimlarning har tomonlama ilmiy, amaliy, muammoli bo‘lishi, o‘quv materiallarining yetarlicha ko‘rgazmali hamda tushunarli ilovalar bilan boyitilishi ta’lim sifatini kafolatlovchi bosh mezon hisoblanadi.

Asosiy qism: Mazkur tadqiqotning bosh maqsadi — umumiy o‘rta va oliy ta’lim tizimida tabiiy fanlarni o‘qitish jarayonida raqamli ta’lim muhitini takomillashtirishning dars samaradorligiga ta’sir etuvchi pedagogik mexanizmlarini aniqlash, modellashtirish va ilmiy-metodik jihatdan asoslab berishdan iborat. Belgilangan maqsadga erishish uchun tadqiqot doirasida quyidagi muhim pedagogik vazifalar bosqichma-bosqich hal etildi:

1. "Raqamli ta’lim muhiti" tushunchasining zamonaviy pedagogik va falsafiy mazmun-mohiyatini hamda uning tarkibiy qismlarini ilmiy adabiyotlar tahlili orqali ochib berish.

2. Tabiiy fanlar ta’limida mavjud elektron resurslar, platformalar va raqamli vositalardan foydalanishning bugungi holatini tahlil qilish, muammolarni aniqlash.

3. Raqamli ta’lim muhitining o‘quvchilarning mustaqil faoliyatiga, fanni o‘zlashtirish motivatsiyasiga va mantiqiy fikrlashiga ta’sirini pedagogik diagnostika qilish.

4. Tabiiy fanlarni o‘qitishda innovatsion va samarali raqamli muhit modelini (komponentlararo uzviylikni) ishlab chiqish va amaliyotga joriy etish.

5. O‘qituvchilar va metodistlar uchun raqamli darslarni loyihalash bo‘yicha amaliy va metodik tavsiyalar majmuasini tayyorlash.

Tadqiqot ishini olib borish va eksperimental ma’lumotlarni umumlashtirish jarayonida tizimli-strukturaviy va faoliyatga yo‘naltirilgan yondashuvlar, pedagogik diagnostika (anketa, so‘rovnoma, testlar), kontent-tahlil, o‘quv jarayonini kompyuterli loyihalash, pedagogik eksperiment (tajriba-sinov ishlari) hamda olingan natijalarni matematik-statistik tahlil qilish metodlaridan kompleks ravishda foydalanildi.

Xalqaro Nordik Universiteti va hamkor ta’lim muassasalarida o‘tkazilgan tajriba-sinov ishlari davomida quyidagi muhim ilmiy-amaliy natijalarga erishildi va ular tahlil qilindi:

Elektron darsliklar va multimedia qo‘llanmalari: Axborotni matnli, ovozli va video ko‘rinishda kompleks uzatish hisobiga o‘quvchilarning bilimlarni o‘zlashtirish tezligini oshirdi va mustaqil bilim olish iqtidorini tizimli ravishda kuchaytirdi.

Individual ta’lim traektoriyasi: Zamonaviy LMS (Learning Management Systems) va raqamli platformalar yordamida har bir o‘quvchining o‘zlashtirish tempi va individual qobiliyatiga mos keladigan shaxsga yo‘naltirilgan ta’lim muhiti muvaffaqiyatli yaratildi.

Virtual va masofaviy laboratoriyalar: Moddiy-texnik baza yetishmagan yoki real sharoitda o‘tkazish inson salomatligi uchun xavfli bo‘lgan kimyoviy, biologik hamda jismoniy eksperimentlarni xavfsiz, interaktiv shaklda bajarish orqali o‘quvchilarning amaliy-tajribaviy ko‘nikmalari doirasini keskin kengaytirdi.

Ko‘rsatkichlar o‘sishi: Matematik-statistik tahlillar shuni ko‘rsatdiki, raqamli ta’lim muhiti modelida tahsil olgan o‘quvchilarning mustaqil ishlash, muammoli vaziyatlardan chiqish va kreativ yechimlar topish ko‘nikmalari an’anaviy guruhlarga nisbatan o‘rtacha 18–22% ga oshganligi aniqlandi.

Raqamli o‘quv muhitining joriy etilishi o‘quvchilarni ta’lim jarayonida shunchaki tayyor axborotni qabul qiluvchi passiv obyektdan, muammolarni mustaqil tadqiq va tahlil qiluvchi faol hamda tashabbuskor o‘quv subyektiga aylantirish imkonini berdi.

Xulosa: Olib borilgan ilmiy-pedagogik tadqiqot yakunlari shuni ko‘rsatadiki, tabiiy fanlarni o‘qitishda raqamli ta’lim muhitini takomillashtirish va zamonaviy pedagogik mexanizmlarni ishga solish ta’lim samaradorligi hamda sifatini oshirishning eng muhim va kechiktirib bo‘lmaydigan pedagogik shartidir. Raqamli, vizual va interaktiv resurslarni ta’lim bosqichlarida maqsadli, tizimli va uzviy qo‘llash o‘quvchilarning fundamental bilimlarini mustahkamlab, ularda hayotiy ko‘nikma hamda XXI asr kompetensiyalarini shakllantirishga

xizmat qiladi. Tabiiy fanlarni o‘qitish metodikasi bugungi kunda faqatgina dars jarayonini ta’riflash yoki tushuntirish bilan cheklanmasligi lozim. U o‘qituvchining dars loyihasini yaratishidan boshlab, o‘quvchilarning darsdagi, uydagi, laboratoriyadagi, sinfdan tashqari hamda bevosita tabiat qo‘ynidagi amaliy va tadqiqotchilik ishlarini raqamli monitoring qilishgacha bo‘lgan barcha tizimli pedagogik jarayonlarni to‘liq qamrab olishi zarur.

FOYDALANILGAN ADABIYOTLAR

1. To‘raqulov, R. (2020). Raqamli pedagogika asoslari. Toshkent: "Fan va texnologiyalar" nashriyoti.
2. Qodirov, A. (2021). Ta’limda raqamli muhitni shakllantirish tendensiyalari. Ta’lim va innovatsiya, 2(1), 33–38.
3. O‘zbekiston Respublikasi Prezidenti. (2020). Ta’lim-tarbiya tizimini yanada takomillashtirish va raqamlashtirishga oid chora-tadbirlar to‘g‘risidagi qarorlari. Toshkent.
4. Ismoilov, B. (2022). Tabiiy fanlarni o‘qitishda innovatsion va virtual texnologiyalar. Pedagogik mahorat, 4(2), 112–118.

RAQAMLI TRANSFORMATSIYA SHAROITIDA TIJORAT BANKLARI FAOLIYATI SAMARADORLIGINI OSHIRISH MEXANIZMLARI

Zarnigor Komilova Xayriddin qizi
O‘zbekiston Respublikasi bank moliya akademiyasi magistranti
Bank ishi va auditi yo‘nalishi
https://doi.org/10.5281/zenodo.23340134

Annotatsiya. *Mazkur tezisda raqamli transformatsiya jarayonida tijorat banklari faoliyati samaradorligini oshirishning asosiy yo‘nalishlari tahlil qilinadi. Raqamli texnologiyalarni joriy etish, bank xizmatlarini avtomatlashtirish, sun’iy intellekt va katta ma’lumotlar (Big Data) texnologiyalaridan foydalanish banklarning operatsion samaradorligini oshirish, xarajatlarni kamaytirish hamda mijozlarga ko‘rsatilayotgan xizmatlar sifatini yaxshilashga xizmat qilishi yoritilgan. Shuningdek, raqamli transformatsiyani muvaffaqiyatli amalga oshirish uchun kiberxavfsizlikni ta’minlash, malakali kadrlarni tayyorlash va innovatsion boshqaruv tizimlarini rivojlantirish zarurligi asoslab berilgan.*

Kalit so‘zlar: *raqamli transformatsiya, tijorat banklari, bank samaradorligi, fintech, sun’iy intellekt, Big Data, mobil bank, internet bank, raqamli iqtisodiyot, kiberxavfsizlik.*

So‘nggi yillarda jahon moliya tizimida raqamli transformatsiya bank xizmatlarini ko‘rsatish usullarini tubdan o‘zgartirmoqda. Sun'iy intellekt (AI), Big Data, bulutli texnologiyalar (Cloud Computing), blokcheyn va Open Banking kabi innovatsion yechimlarning joriy etilishi tijorat banklari faoliyatining samaradorligini sezilarli darajada oshirmoqda. O‘zbekistonda ham bank tizimini raqamlashtirish davlat siyosatining ustuvor yo‘nalishlaridan biri sifatida belgilanib, ushbu yo‘nalishda qator normativ-huquqiy hujjatlar qabul qilindi.

2020-yil 5-oktabrda O‘zbekiston Respublikasi Prezidenti tomonidan "Raqamli O‘zbekiston – 2030" strategiyasi tasdiqlandi. Ushbu strategiyada bank va moliya sektorini raqamlashtirish, davlat xizmatlarini elektron shaklga o‘tkazish hamda raqamli iqtisodiyotni rivojlantirish ustuvor vazifa sifatida belgilandi. Strategiya doirasida tijorat banklarida masofaviy bank xizmatlari, mobil ilovalar, elektron identifikatsiya va raqamli to‘lov tizimlarini keng joriy etish ko‘zda tutilgan.

2025-yil 18-noyabrda O‘zbekiston Respublikasi Markaziy banki tomonidan 2025–2028-yillarga mo‘ljallangan bank nazorati va tartibga solishni takomillashtirish bo‘yicha "Yo‘l xaritasi" tasdiqlandi. Mazkur hujjatda Basel III standartlarini to‘liq joriy etish, riskka asoslangan bank nazoratini rivojlantirish, korporativ boshqaruvni takomillashtirish hamda banklarning moliyaviy barqarorligini mustahkamlash asosiy yo‘nalish sifatida belgilangan.

Raqamli transformatsiya tijorat banklari samaradorligini bir nechta yo‘nalishlarda oshiradi:

Birinchidan, bank operatsiyalarining avtomatlashtirilishi operatsion xarajatlarni kamaytiradi. Sun'iy intellekt asosidagi tizimlar kredit arizalarini tezkor baholaydi, firibgarlik holatlarini aniqlaydi hamda risklarni boshqarish sifatini oshiradi.

Ikkinchidan, mobil bank va internet-bank xizmatlari mijozlarga 24/7 rejimida xizmat ko‘rsatish imkonini beradi. Natijada filiallarga yuklama kamayadi, xizmat ko‘rsatish tezligi va sifati oshadi.

Uchinchidan, Big Data texnologiyasi yordamida mijozlarning moliyaviy xatti-harakatlari tahlil qilinadi va individual bank mahsulotlari taklif etiladi. Bu esa bank daromadlarining oshishiga xizmat qiladi.

To‘rtinchidan, Open Banking konsepsiyasi fintech kompaniyalari bilan hamkorlikni rivojlantirib, yangi raqamli xizmatlarni yaratishga imkon beradi.

Jahon bankining 2022-yil 18-may kuni e'lon qilingan "Digital Technology and the Future of Finance" hisobotida fintech texnologiyalarining joriy etilishi moliyaviy xizmatlarning qamrovini kengaytirishi, xarajatlarni kamaytirishi hamda banklar samaradorligini oshirishi qayd etilgan. Hisobotda pandemiyadan keyingi davrda raqamli moliyaviy xizmatlar rivojlanishi sezilarli tezlashgani alohida ta'kidlangan.

Jahon bankining 2024-yil 19-martda chop etilgan Digital Progress and Trends Report 2023 hisobotiga ko‘ra, raqamli texnologiyalar iqtisodiy o‘sishning asosiy omillaridan biriga aylangan. Hisobotda sun'iy intellekt, raqamli infratuzilma va raqamli davlat platformalarini rivojlantirish bank tizimining samaradorligini oshirishda muhim rol o‘ynashi ta'kidlangan.

Bugungi kunda tijorat banklarida quyidagi muammolar mavjud:

- kiberxavfsizlik tahdidlarining ortishi;
- ayrim hududlarda raqamli infratuzilmaning yetarli darajada rivojlanmaganligi;
- yuqori malakali IT mutaxassislariga ehtiyoj;
- raqamli texnologiyalarni joriy etish xarajatlarining yuqoriligi;
- ayrim mijozlarning raqamli savodxonligi pastligi.

Tijorat banklari faoliyati samaradorligini oshirish uchun quyidagi chora-tadbirlarni amalga oshirish maqsadga muvofiq:

- sun'iy intellekt asosidagi kredit skoring tizimlarini keng joriy etish;
- Open Banking va Open API platformalarini rivojlantirish;
- bulutli texnologiyalardan foydalanishni kengaytirish;
- kiberxavfsizlik bo‘yicha xalqaro standartlarni joriy etish;
- bank xodimlarining raqamli kompetensiyalarini muntazam oshirish;
- fintech kompaniyalari bilan hamkorlikni kengaytirish.

Raqamli transformatsiya natijasida Open Banking konsepsiyasi ham keng rivojlanmoqda. Mazkur yondashuv banklarning API texnologiyalari orqali fintech kompaniyalari bilan integratsiyalashuvini ta'minlaydi. Buning natijasida mijozlarga yangi innovatsion mahsulotlar, tezkor to‘lov xizmatlari, avtomatlashtirilgan moliyaviy boshqaruv tizimlari va investitsiya platformalari taklif etilmoqda.

Natijada banklar va fintech kompaniyalari o‘rtasidagi raqobat o‘rnini o‘zaro hamkorlik egallamoqda.

Shu bilan birga, raqamli transformatsiya banklar oldiga yangi xavflarni ham qo‘ymoqda. Kiberxavfsizlik masalasi bugungi kunda eng dolzarb muammolardan biri hisoblanadi. Bank axborot tizimlariga noqonuniy kirish, fishing hujumlari, zararli dasturlar va shaxsiy ma'lumotlarning sizib chiqishi moliyaviy barqarorlikka jiddiy xavf tug‘diradi. Shu sababli xalqaro amaliyotda ISO/IEC 27001 axborot xavfsizligi standarti, ko‘p bosqichli autentifikatsiya tizimlari va zamonaviy shifrlash algoritmlaridan foydalanishga alohida e'tibor qaratilmoqda.

Bugungi kunda sun'iy intellekt texnologiyalarining tijorat banklari faoliyatiga joriy etilishi bank boshqaruvining mutlaqo yangi bosqichini boshlab berdi. Sun'iy intellekt algoritmlari kredit oluvchilarning moliyaviy holatini qisqa vaqt ichida tahlil qilib, kredit riskini baholaydi, firibgarlik alomatlarini avtomatik aniqlaydi hamda mijozlarning moliyaviy xulq-atvorini prognoz qiladi. Natijada kredit portfeli sifati yaxshilanadi, muammoli kreditlar ulushi kamayadi va bank kapitalidan foydalanish samaradorligi ortadi. Bundan tashqari, sun'iy intellekt asosidagi virtual yordamchilar va chatbotlar mijozlarning murojaatlariga sutkaning istalgan vaqtida javob berish imkonini yaratmoqda. Bu esa xizmat ko‘rsatish sifatini oshirish bilan bir qatorda bank filiallariga tushadigan yuklamani ham kamaytiradi.

Raqamli transformatsiyaning yana bir muhim elementi Big Data texnologiyasidir. Zamonaviy tijorat banklari har kuni millionlab moliyaviy operatsiyalarni amalga oshiradi va ulardan katta hajmdagi ma'lumotlar shakllanadi. Ushbu ma'lumotlarni tahlil qilish orqali banklar mijozlarning ehtiyojlarini aniqlash, ularning xarid qobiliyatini baholash, individual moliyaviy mahsulotlarni ishlab chiqish hamda marketing strategiyasini takomillashtirish imkoniyatiga ega bo‘ladi. Big Data asosidagi tahlillar bank daromadlarini oshirish bilan birga kredit risklarini kamaytirishda ham muhim ahamiyat kasb etadi.

XULOSA

Raqamli transformatsiya tijorat banklari faoliyati samaradorligini oshirishning eng muhim omillaridan biri hisoblanadi. O‘zbekistonda amalga oshirilayotgan islohotlar bank tizimining barqarorligini, xizmatlar sifatini va xalqaro raqobatbardoshligini oshirishga xizmat qilmoqda. Markaziy bank tomonidan 2025-2028-yillarga mo‘ljallangan “Yo‘l xaritasi”, shuningdek Jahon banki tavsiyalari bank tizimida zamonaviy boshqaruv, risklarni samarali nazorat qilish va innovatsion texnologiyalarni joriy etishning muhim asosini yaratmoqda.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. O‘zbekiston Respublikasi Prezidenti. PF–6079-son Farmon. “Raqamli O‘zbekiston – 2030” strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to‘g‘risida. 2020-yil 5-oktabr.
2. O‘zbekiston Respublikasi Markaziy banki – Moliyaviy texnologiyalar strategiyasi (2026–2030). O‘zbekiston Respublikasi Markaziy banki. Yangilangan: 2026-yil 27-aprel.
3. World Bank – Digital Uzbekistan 2030 Strategy and Digital Inclusion Project. World Bank, 2023-yil 30-noyabr.
4. World Bank – Uzbekistan Country Partnership Framework (FY2022–FY2026). World Bank, 2022.

ОРГАНИЗАЦИЯ СБОРА ДЕНЕЖНЫХ ПОСТУПЛЕНИЙ И ЦЕННОСТЕЙ В КОММЕРЧЕСКИХ БАНКАХ: СТРАТЕГИИ УЧЕТА И ОПТИМИЗАЦИИ ПРОЦЕССОВ

Комилова Зарнигор Хайриддин кизи
Банковско-финансовая Академия Республики Узбекистан
Направление «Банковкое дело и аудит»
https://doi.org/10.5281/zenodo.23340217

I. **Введение**. В динамичном ландшафте коммерческих банков тщательная организация наличной выручки и инкассации ценностей является основой операционной целостности, финансовой стабильности и соответствия нормативным требованиям. Умелое управление денежными потоками и ценными активами не только защищает интересы финансовых учреждений, но и вызывает доверие у заинтересованных сторон. Данная диссертация посвящена изучению организационной структуры, регулирующей инкассацию денежной выручки и других ценностей в коммерческих банках, с особым акцентом на перспективы бухгалтерского учета.

Значимость данного исследования подчеркивается той важнейшей ролью, которую играет наличная выручка и ценности в повседневной деятельности коммерческих банков. На банки как на проводников финансовых операций и хранителей ценных активов возложена ответственность за обеспечение безопасного, эффективного и ответственного обращения с денежной наличностью и ценностями на протяжении всего их жизненного цикла. Кроме того, точность и прозрачность бухгалтерского учета, регулирующего эти операции, имеет первостепенное значение для соблюдения нормативных требований, управления рисками и составления финансовой отчетности.

На этом фоне цели данной работы многообразны. Во-первых, она направлена на всесторонний обзор существующей литературы, касающейся организационных основ инкассации наличной выручки и ценностей в коммерческих банках. Это предполагает изучение процессов инкассации, принципов бухгалтерского учета, нормативных требований и передовой практики, распространенной в банковской отрасли. Во-вторых, цель работы – выявить основные проблемы, возникающие при организации инкассации и учета денежных средств в коммерческих банках, и предложить эффективные решения для их устранения. В-третьих, исследование направлено на оценку влияния организационной и учетной практики на операционную эффективность, подверженность рискам и финансовые показатели коммерческих банков.

Методология, использованная в данном исследовании, включает в себя сочетание качественных и количественных исследовательских подходов. Качественные методы, включая интервью и тематические исследования, позволяют получить представление о тонкостях процессов инкассации денежных средств и практики учета с точки зрения практиков и экспертов отрасли. Количественные методы, такие как опросы и анализ данных, позволяют получить эмпирические данные для подтверждения выводов и рекомендаций, полученных на основе качественных исследований. Используя данную методологическую базу, исследование стремится представить целостное понимание

организационных и бухгалтерских аспектов инкассации наличной выручки и ценностей в коммерческих банках.

В данной диссертации ставится задача изучить организационную структуру инкассации денежной выручки и ценностей в коммерческих банках с особым акцентом на бухгалтерские аспекты. Вникая в тонкости процессов инкассации, принципы бухгалтерского учета, проблемы и решения, она стремится предложить действенные идеи, которые могут стать основой для принятия стратегических решений и операционной практики в банковской отрасли. Благодаря тщательному анализу и научному обсуждению, данное исследование призвано внести вклад в развитие знаний в области банковских операций, бухгалтерского учета и финансового менеджмента.

II. **Литературный обзор.** Литература, посвященная организационным основам сбора денежных средств и ценностей в коммерческих банках, представляет собой богатый набор теоретических представлений, эмпирических исследований и практических соображений. В этом разделе представлен обзор ключевых тем и результатов существующей литературы:

1. Процессы инкассации в коммерческих банках:

В литературе освещаются различные методологии и протоколы, используемые коммерческими банками при сборе денежных поступлений из различных источников, включая филиалы, банкоматы (АТМ), бизнес-клиентов и службы инкассации. Это включает в себя разработку надежных политик и процедур для управления обработкой, транспортировкой и хранением наличных денег, а также внедрение передовых технологических решений для оптимизации процессов инкассации.

2. Принципы и стандарты бухгалтерского учета:

Принципы и стандарты бухгалтерского учета играют ключевую роль в учете, оценке и отчетности о денежных поступлениях и ценностях в коммерческих банках. В литературе исследуется применение общепринятых принципов бухгалтерского учета (GAAP) и международных стандартов финансовой отчетности (МСФО) к деятельности по управлению денежными средствами, подчеркивая важность точности, прозрачности и соблюдения нормативных требований.

3. Нормативные требования и соответствие:

Нормативно-правовая база, регулирующая практику инкассации и бухгалтерского учета в коммерческих банках, широко обсуждается в литературе. Сюда входят правила, издаваемые банковскими органами, такими как Базельский комитет по банковскому надзору, а также правила по борьбе с отмыванием денег (AML) и финансированием терроризма (CTF), применяемые государственными учреждениями. В литературе подчеркивается необходимость соблюдения банками строгих нормативных требований для снижения рисков и поддержания целостности финансовых систем.

4. Передовой опыт и отраслевые стандарты:

В литературе освещаются передовые методы и отраслевые стандарты, касающиеся сбора денежных средств и ценностей в коммерческих банках. Это включает в себя принятие стандартизированных процедур выверки денежных средств, внутреннего контроля и контрольных журналов, а также внедрение технологических решений, таких как системы управления денежными средствами и биометрическая аутентификация, для повышения безопасности и эффективности.

5. Стратегии управления рисками:

Стратегии управления рисками, связанные с инкассацией и бухгалтерским учетом денежных средств, находятся в центре внимания научных исследований. В литературе исследуются выявление, оценка и снижение рисков, связанных с обработкой наличных денег, включая кражу, мошенничество, операционные ошибки и технологические уязвимости. Также обсуждаются стратегии повышения устойчивости и планирования действий в чрезвычайных ситуациях перед лицом возникающих рисков.

6. Технологические инновации и цифровизация:

Появление технологических инноваций и цифровизации изменило методы инкассации и бухгалтерского учета в коммерческих банках. В литературе подробно рассматривается интеграция цифровых решений, таких как мобильные банковские приложения, устройства для переработки наличных и технология блокчейна, для оптимизации процессов управления денежными средствами, улучшения качества обслуживания клиентов и снижения эксплуатационных расходов.

Таким образом, обзор литературы освещает многогранный характер сбора денежных средств и ценностей в коммерческих банках, охватывающий операционные, бухгалтерские, нормативные и технологические аспекты. Обобщая теоретические идеи, эмпирические исследования и отраслевую практику, этот обзор служит основой для последующего анализа и обсуждения в данной диссертации.

III. **Методология**. В данном разделе изложен методический подход к исследованию организационной структуры инкассации денежных средств и ценностей в коммерческих банках:

Дизайн исследования: применяется смешанный подход, объединяющий как качественные, так и количественные методологии для достижения всестороннего понимания предмета исследования. Этот подход позволяет проводить триангуляцию источников данных, повышая глубину и широту анализа.

Методы сбора данных: данные собираются с помощью различных средств, включая структурированные интервью с персоналом банка, ответственным за инкассацию денежных средств и учет, опросы, распространяемые среди выборки коммерческих банков для сбора количественных данных об организационной практике, а также тематические исследования, изучающие реальные примеры. Успешных стратегий управления денежными средствами.

Методы формирования выборки. Стратегия выборки предполагает целенаправленную выборку коммерческих банков, представляющих самые разные размеры, географическое расположение и операционные модели. Это обеспечивает учет точек зрения широкого спектра заинтересованных сторон отрасли.

Подход к анализу данных: Качественные данные, полученные в результате интервью и тематических исследований, подвергаются тематическому анализу для выявления повторяющихся закономерностей, тем и идей. Количественные данные опросов анализируются с использованием статистических методов, таких как описательная статистика и логический анализ, для получения значимых выводов.

IV. **Организационная структура сбора денежных поступлений**

В данном разделе подробно рассматривается организационная основа инкассации денежных средств в коммерческих банках:

Политика и процедуры: В исследовании изучается разработка и внедрение политик и процедур, регулирующих процессы инкассации наличных, включая протоколы обработки наличных, транспортную логистику и меры безопасности.

Интеграция технологий: исследуется роль технологий в обеспечении эффективных процессов инкассации денежных средств, включая развертывание систем управления денежными средствами, автоматизированных устройств обработки наличных денег и решений для цифровых платежей.

Стратегии управления рисками: анализируются стратегии выявления, оценки и снижения рисков, связанных со сбором денежных средств, таких как кражи, мошенничество и операционные ошибки. Это включает в себя создание надежных механизмов внутреннего контроля, протоколов безопасности и планов действий в чрезвычайных ситуациях.

V. **Методы бухгалтерского учета при работе с денежными средствами и ценностями.** В этом разделе рассматриваются методы бухгалтерского учета, связанные с обработкой и оценкой денежных поступлений и других ценностей в коммерческих банках:

Учет и сверка: в исследовании изучаются процедуры точного учета кассовых операций, сверки остатков денежных средств и выявления расхождений посредством регулярных проверок и внутреннего контроля.

Методы оценки: Различные методы оценки денежных средств и ценностей, такие как розничный метод, метод наименьшей из затрат или рыночный метод (LCM) и оценка справедливой стоимости, изучаются в контексте стандартов бухгалтерского учета и нормативных требований.

Финансовая отчетность: В исследовании оцениваются требования к отчетности по денежным средствам и ценностям в финансовой отчетности, включая раскрытие информации, связанной с эквивалентами денежных средств, отчетами о движении денежных средств и условными обязательствами.

VI. **Проблемы и решения.** В этом разделе определяются общие проблемы, возникающие в организационной структуре сбора денежных средств и ценностей в коммерческих банках, и предлагаются эффективные решения:

Риски безопасности: в исследовании рассматриваются уязвимости безопасности, связанные с процессами инкассации наличных, и рекомендуются стратегии по усилению физической безопасности, обучению сотрудников и системам наблюдения.

Операционная неэффективность: анализируются проблемы, связанные с операционной неэффективностью, такие как ошибки ручного ввода данных, задержки обработки и проблемы с распределением ресурсов. Предлагаются решения для оптимизации рабочих процессов, автоматизации повторяющихся задач и оптимизации использования ресурсов.

Соблюдение нормативных требований. В исследовании рассматриваются проблемы, связанные с соблюдением нормативных требований, включая правила AML/KYC, законы о конфиденциальности данных и стандарты финансовой отчетности. Обсуждаются стратегии обеспечения соблюдения нормативных требований при минимизации затрат на соблюдение требований и операционной нагрузки.

VII. **Реализация и оценка**

В данном разделе излагаются этапы реализации предлагаемых организационно-бухгалтерских практик и оценивается их эффективность:

План реализации: Очерчена дорожная карта для реализации рекомендуемых решений, включая сроки, распределение ресурсов и взаимодействие с заинтересованными сторонами.

Критерии оценки: определены критерии оценки эффективности внедренных решений, такие как экономия затрат, снижение рисков и повышение операционной эффективности.

Мониторинг и постоянное улучшение. Механизмы постоянного мониторинга, сбора обратной связи и постоянного улучшения созданы для обеспечения устойчивости и адаптируемости внедряемых практик.

VIII. **Обсуждение**

В этом разделе интерпретируются результаты, обсуждаются их последствия и дается представление о будущих направлениях исследований:

Интерпретация результатов: Результаты интерпретируются в свете целей исследования, проливая свет на эффективность предлагаемых решений в решении выявленных проблем.

Значение для практики: Обсуждаются практические последствия выводов для коммерческих банков, включая рекомендации по оптимизации методов инкассации и бухгалтерского учета.

Последствия для исследований: Определены возможности для будущих исследований, таких как продольные исследования, межстрановые сравнения и новые тенденции в технологиях и регулировании.

IX. **Заключение**. В этом разделе суммированы основные выводы, вклад и последствия исследования:

Краткое изложение результатов: Кратко изложены основные результаты исследования, включая понимание организационных структур, практики бухгалтерского учета, проблем и решений, связанных со сбором денежных поступлений и ценностей в коммерческих банках.

Вклад в знания: подчеркнут вклад исследований в область банковских операций, бухгалтерского учета и управления рисками.

Заключительные замечания: представлены заключительные размышления о значении исследования и его потенциальном влиянии на отраслевую практику и научный дискурс.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Рамазанова Д. Т (2020). Учет коммерческими банками кассовых операций. Проблемы современных методов интеграции и поиск новых подходов (с. 162-164).
2. Н. М. Муратбекова и Ч. Х. Нуралиева (2018). Совершенствование учета кассовых операций коммерческого банка. Как указано на страницах 126-130 в книге М. Рыскулбекова «Атындагы Кыргызский экономический университет Кабарлары».
3. А. Т. Мамажонов, И М. С. Хожибоев (2021). Учет денежных средств в банках – необходимость и важность. Прикладная и теоретическая наука, 123-126.

4. Kimani, J. G., Karungu, R. M., Kirui, K., and Nyangau, B. O. (2014). How Does Working Capital Management Affect the Risk of Liquidity? Kenya's Listed Commercial Bank Case. Research Journal of Accounting and Finance, 5(10), 34–49.
5. Chen, L. (2019). 'Innovations in Cash Logistics and Transport in Chinese Commercial Banks', Journal of Banking Technology, 11(2), pp. 45-60.
6. Garcia, M., and Lopez, R. (2021). 'Compliance and Regulatory Challenges in Cash Management', American Banking Review, 33(4), pp. 110-125.
7. Jackson, H., and Lee, S. (2023). 'Impact of Digitalization on Cash Operations in Australian Banks', Journal of Digital Banking, 7(1), pp. 75-89.
8. Müller, F., and Schmidt, H. (2022). 'Accounting Practices for Cash and Valuables in European Banks', European Journal of Finance, 28(6), pp. 502-519.
9. Patel, R., and Kumar, A. (2021). 'Risk Management in Cash Collection and Handling', Indian Banking Journal, 26(3), pp. 34-50.
10. Smith, J., and Johnson, K. (2020). 'Efficient Cash Management Practices in Commercial Banks', Journal of Banking Efficiency, 14(4), pp. 200-215.
11. Thompson, Emily. «Technological Advancements and Their Impact on Cash Management in Banking Institutions.» University of Oxford, 2022.
12. Martinez, David. «An Analytical Examination of Risk Mitigation in the Physical Handling of Currency within the Banking Sector.» Harvard University, 2021.
13. Khan, Aisha. «Cash Operations in the Era of Digital Banking: A Paradigmatic Transition.» University of Toronto, 2023.
14. Greenfield, Robert. «Cashless Society: The Future of Banking.» Cambridge University Press, 2020, 1st edition.
15. Patel, Susan. «Modern Banking Operations: Cash, Compliance, and Technology.» Wiley Finance, 2021, 1st edition.

ALISHER NAVOIY IJODINING ERKIN VOHIDOV SHE'RIYATIDA O'RNI

Suyarova Gulruxsor Nurilla qizi

https://doi.org/10.5281/zenodo.23340346

Annotatsiya. *Ushbu maqolada Erkin Vohidov ijodida ham Alisher Navoiy shaxsi va she'riyatidan ta'sirlanish izlari tadqiq qilinadi. Navoyining ishq haqida bitganlari Erkin Vohidov ijodida qay shaklda akslanishi ko'rib chiqiladi.*

Kalit so'zlar*: peshso'z, adabiy-estetik tafakkur, mavzu, g'oya uslub, makon va zamon, tadqiq, talqin, tahlil*

Аннотация. *В данной статье исследуются следы влияния личности и поэзии Алишера Навои на творчество Эркина Вахидова. Рассматривается, в какой форме стихи Навои о любви находят свое отражение в произведениях Эркина Вахидова.*
Ключевые слова: *предисловие, литературно-эстетическое мышление, тема, идея, стиль, пространство и время, исследование, интерпретация, анализ.*

Abstract. *This article examines the traces of influence of Alisher Navoi's personality and poetry on the creative works of Erkin Vohidov. It explores how Navoi's writings on love are reflected in Erkin Vohidov's poetry.*

Keywords: *preface, literary-aesthetic thinking, theme, idea, style, space and time, research, interpretation, analys*

Alisher Navoiy ijodi asrlar davomida ijodkorlarni ruhlantirib kelgan. Xususan, Erkin Vohidov ijodida ham Alisher Navoiy shaxsi va she'riyatidan ta'sirlanish izlari ko'rinadi. Erkin Vohidov she'rlarida Navoiy baytlari mag'zini shoirona nigoh bilan chaqib, hikmatlarini elga taqdim etgan. Alisher Navoyining

Aylagin jondin judo
Etguncha jonondin judo.

misralaridan ilhomlangan shoir "Hijron yuki" she'rida uni peshso'z etib oladi. Navoyining hijron iztiroblaridan ta'sirlangan shoir she'rda oshiqning ma'shuqa visolidan bebahra o'tgan har lahzani yil bilan o'lchangan ko'ngil iztiroblarini go'zal tashbehbar bilan ifodalaydi. Shoir o'z she'rlarida Navoyiga ko'p murojaat etadi. Uning shaxsi va ijodini barcha uchun ibrat deb biladi. "Turkiston bozori" she'rida butun turkiy xalqlar birligini orzu qilgan shoir:

Barcha tillarda Navoiy
She'ridan devon bo'lur

deya Navoyining qardosh millatlar uchun birdek sevimli ekaniga urg'u beradi. Shuningdek, hazrat haqidagi "O'zbek Navoyini o'qimay qo'ysa" she'rida millat Navoyini o'qimay qo'ysa, ma'rifatdan yiroqlashishi, komil bo'lmagan elning yurti ulug' bo'lmasligini ta'kidlaydi. O'xshatish, tazod va boshqa badiiy san'atlar vositasida Navoyini tanimagan millat fojiasini

bayon etadi. Agar millat uni tushunsa, bilsa, o‘zbek o‘zligini bekam anglashini, chinakam Navoiy avlodi bo‘lib olamga tanilishini orzulaydi. Alisher Navoiy she’riyatiga ixlos, uning go‘zal g‘azallari Erkin Vohidovni ham aruz olamiga yetaklagan bo‘lsa ne ajab: “... xalqimiz Navoiy va Fuzuliylar, Hofiz va Bedillarning go‘zal va teran g‘azallarini biladi. Uning aruz vazniga didi ana o‘shalarni o‘qib tarbiya topgan” [2, 121]. Erkin Vohidovning nazira yo‘lidagi dastlabki mashqlari, Alisher Navoyining “Aylagach” radifli g‘azaliga bog‘lagan muxammasida go‘yoki o‘quvchini tadrijiy tarzda Navoiy hikmatiga tayyorlaydi. Birinchi banddagi “Misli oykim ko‘rk ochur sayri shabiston aylagach” misrasi Navoyining “Husni ortar yuzda zulfin anbarafshon aylagach” misrasiga kirish vazifasini o‘tagan: tun qorong‘usida oy ko‘rkamlashganidek, yuzidagi tundek qora zulfini yig‘gan yorning ham husni ortadi. Navoyining mantiqiy izchil baytlaridagi ayni holatni saqlashga intilgan shoir keyingi bandlarda ham mutanosiblikni saqlab qoladi. Ikkinchi bandda oshiqning ma’shuqaga murojaatida ham tadrijiylik kuzatiladi. Banddagi “xino”, “gul”, “qon” so‘zlari tanosib vositasida oshiq holini ochib beradi. She’rda Erkin Vohidovning misralari Navoiy baytining mazmuniga hamohanglik kasb etadi. “Har qachonkim kemaga ul oy safar raxtin solur” deb boshlanuvchi g‘azaliga bitilgan muxammasida Erkin Vohidovning bu yo‘nalishdagi tajribasi oshgani seziladi.

Ayriliq oni yaqindir, kema yo‘l bongin cholur,
Vah meni tashlab firoqqa, yor yiroqqa yo‘l olur.
Jon borib jono bilan, sohilda bir jismim qolur,
Har qachonkim kemaga ul oy safar raxtin solur,
Mavjlig‘ daryo kabi oshufta ko‘nglum qo‘zg‘olur.

Oldingi muxammasi ma’noga yetaklovchi vazifasini o‘tagan bo‘lsa, ushbu muxammasda Navoiy misralari bilan mazmun-ma’no uyg‘unligi ko‘rinadi. “To muhabbat dashti bepoyonida ovoramen” deb boshlanuvchi g‘azalida Erkin Vohidovning uslubi yanada yaqqolroq ko‘rinadi. Oldingi ikki muxammasdan farqli ravishda ushbu muxammasga birmuncha ijodiy yondashilganini ko‘rishimiz mumkin. Birinchi baytida Navoiy bepoyon muhabbat dashtida ovvora, har balo kelsa-da, ishq oshubidan bechoraligini aytib yozg‘irsa, Erkin Vohidov ma’noni yanada kuchaytirib “ko‘ngil”, “ko‘z”, “jon” tashxislaridan foydalanib, oshiqning bechora holini yanada kuchaytirgan. Shuningdek, uchinchi bandda ham oh dudida bir uchqun kabi ko‘rinib, qo‘nimsiz yulduzga mengzalgan oshiqni shoir Karbalo dashti ohulari yoniga eltadi. Laylivash dardidan Majnun bilan bir jon, bir tan oshiq holini xusni ta’lil san’atida aks ettiradi. “Ko‘zung ne balo qaro bo‘lubtur” g‘azalida Navoiy qo‘llagan ruju’ san’ati Erkin Vohidov muxammasining asosi bo‘lgan

Ishqing me nga to jafo bo‘lubtur,
Ko‘nglumga jafo safo bo‘lubtur,
Qoshing, ne jazoki, yo bo‘lubtur,
Ko‘zung ne balo qaro bo‘lubtur,
Kim jong‘a balo qaro bo‘lubtur.

Keyingi bandlarda ham ma’no misrama-misra oshib, bo‘yoqlari yanada quyuqlashib boradi. Maqta’da Navoyining ishq haqida bitganlari ishq ahli orasida qo‘shiqqa aylanishi aytilsa, Erkin Vohidov ayni misralarning izohini ishqning ikki jahon saodati, umrning asl ma’nosi,

shoirning to‘g‘ri yo‘lga olib boruvchi tilagi deb biladi. Erkin Vohidovning hazrat Naviy g‘azallariga bog‘lagan muxammaslarida u xoh tadrijiy, xoh izohlovchi, yoinki hamohanglikda bo‘lsin, benazir so‘z kashshofining ijodiga cheksiz muhabbati aks etadi. Bu mehr-muhabbat Erkin Vohidov ijodining boshqa qirralarida ham yaqqol namoyon bo‘ladi. “Nizomiy panjasiga panja urmoq” mashaqqatini his etgan sohir ijod sahibining ohanrabosi o‘zidan keyingi she’riyat muhiblarini ham hamisha ijod olamiga ruhlantirib turadi.

FOYDALANILGAN ADABIYOTLAR

1.Vohidov E. Tanlangan asarlar. – T.: Sharq, 2016.
2. Вохидов Э. Изтироб. – Т.: Ўзбекистон, 1992.
3. Vohidov E. Shoir-u she’r-u shuur. – T.: Yosh Gvardiya, 1987.
4. Vohidov E. Iztirob. – T.: O‘zbekiston, 1992.
5. Sultonov U. Cho‘lponning adabiy-estetik qarashlari: Filol. fan. nomzodi... diss. – T., 1995; – 128 b.
6. Сент-Бёв Ш. – О. Литературные портреты. – М.: 1970.
Mirzayev T., Mirvohidov R. To quyosh sochgayki nur. – T.: O‘zbekiston NMIU, 2016.
7. Vohidov E. Saylanma. I jild: Ishq savdosi (She’rlar va dostonlar). – Toshkent: "Sharq" NMAK, 2018. – 384 b. (Erkin Vohidov asarlarining so‘nggi rasmiy qayta nashri)
8. Qurolov M. Erkin Vohidov va jahon adabiyoti an’analari // "O‘zbek tili va adabiyoti" jurnali. – Toshkent, 2021. – № 4. – B. 62–67. (Erkin Vohidovning jahon adabiyoti bilan aloqasi bo‘yicha aniq ilmiy maqola)
9. Rasulov A. Badiiy muloqot va estetika masalalari. – Toshkent: Muharrir nashriyoti, 2019. – 210 b. (Adabiy-estetik qarashlar va badiiy tafakkurga bag‘ishlangan monografiya

THE IMPORTANCE OF PREVENTIVE MEASURES IN CREATING SAFE WORKING CONDITIONS

O'roqov Azizbek Usmonaliyevich

https://doi.org/10.5281/zenodo.23367032

Abstract. *Work-related injury and disease remain a heavy public health burden, and by several measures that burden is growing. This study examined why preventive measures, rather than reactive responses, should anchor efforts to create safe working conditions. A secondary-data design combined three strands: global burden estimates from the International Labour Organization and peer-reviewed modeling; U.S. Census of Fatal Occupational Injuries data for 2015–2024; and a structured synthesis of systematic reviews, meta-analyses, and field experiments on preventive interventions and their economic return. About 2.93 million deaths in 2019 were attributed to work, close to nine in ten of them from disease rather than from accidents. In the United States, the fatal injury rate stayed between 3.3 and 3.7 per 100,000 full-time workers all decade, and more than a third of the 2024 deaths involved hazards for which engineering controls have long existed. Intervention evidence was strongest for enforcement backed by penalties; randomized inspections cut injuries by 9.4% and injury costs by 26% with no detectable job loss. Training reliably changed behavior but showed weak links to health outcomes. Prevention appears to work best when it acts early, targets hazards and organizations rather than individual vigilance alone, and is tracked through leading indicators.*

Keywords: *occupational safety, preventive measures, safe working conditions, hierarchy of controls, safety climate, return on prevention*

Introduction

Workplace safety is often judged through accident investigations, compensation, and revised procedures. Yet the International Labour Organization (ILO, 2023) attributed approximately 2.93 million deaths to work-related causes in 2019, alongside 395 million nonfatal occupational injuries. These figures invite a different question: which decisions could have prevented exposure in the first place? Preventive measures encompass more than warnings or personal protective equipment. The hierarchy of controls prioritizes eliminating hazards and redesigning work before relying on workers to avoid danger. Reason's (1997) account of organizational accidents likewise directs attention to latent conditions, including poor maintenance, unreasonable workloads, and inadequate supervision. Safety-climate research connects workers' perceptions of management priorities with safety outcomes (Christian et al., 2009; Zohar, 1980), while ISO 45001 sets out a framework for continuous improvement (International Organization for Standardization [ISO], 2018). Together, these approaches locate reliable protection before an incident.

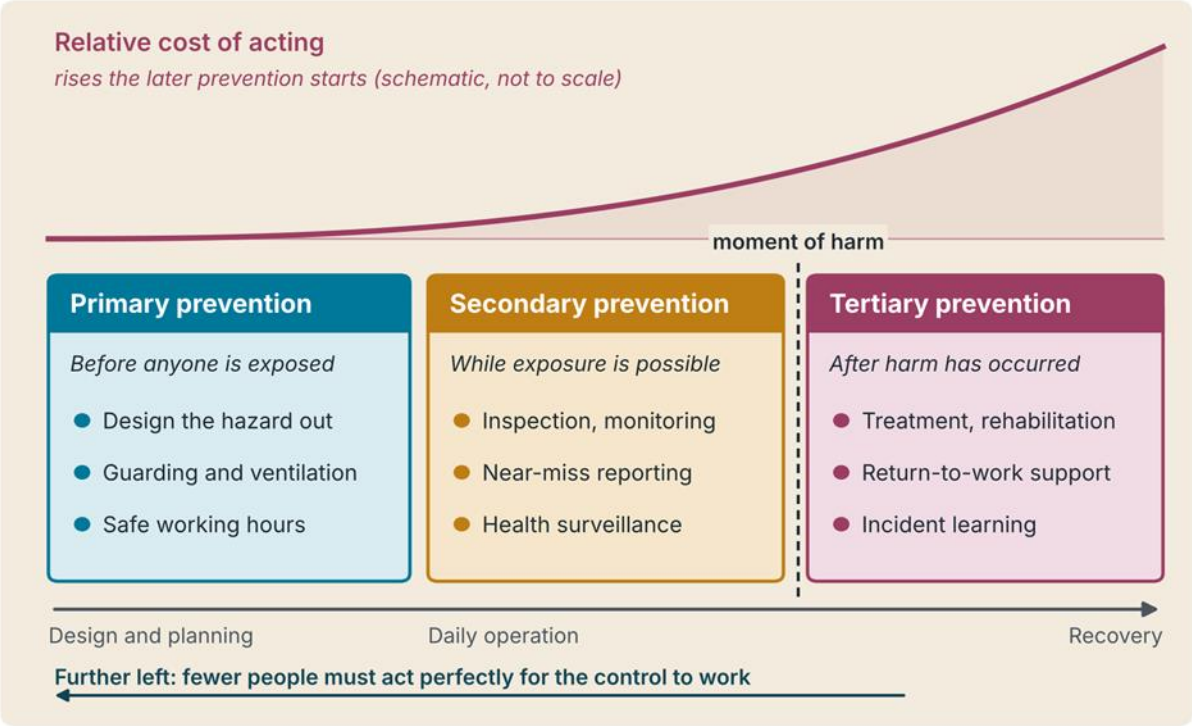
Prevention as a Matter of Timing

Prevention can be understood in terms of when action occurs. Primary prevention removes or reduces hazards before exposure; secondary prevention identifies emerging risks or early harm; and tertiary prevention limits the consequences of an injury or illness. Figure 1 places these approaches along the life of a hazard. Redesigning equipment before installation can protect many workers without depending on repeated individual decisions, whereas responding

after an accident cannot undo the original exposure. The diagram is illustrative, not a cost model, and does not discount rehabilitation. Organizations should accordingly prioritize planning, procurement, and work design.

Figure 1

Prevention Horizons Across the Life of a Workplace Hazard



Note. Schematic developed by the author from the public health classification of primary, secondary, and tertiary prevention. The cost curve illustrates relative direction only and is not drawn from data.

Gaps Addressed by This Study

Research on occupational safety commonly evaluates training, enforcement, management systems, and working conditions separately. This can obscure whether improved awareness or fewer reported injuries reflect stronger controls. Moreover, fatal injuries are counted more routinely than work-related diseases, even though disease accounts for most occupational deaths (Takala et al., 2024). The study combines global burden estimates, a decade of U.S. fatal-injury data, and intervention evaluations. Its purpose is not to establish a new causal effect but to identify where the existing evidence most convincingly supports earlier and more effective prevention.

Method

Design

This study used a descriptive secondary-data design, combining estimates of occupational burden, national fatal-injury trends, and published evaluations of prevention (Figure 2). The strands were compared, not statistically pooled, because their definitions and outcomes differ. No human participants were recruited or assigned interventions.

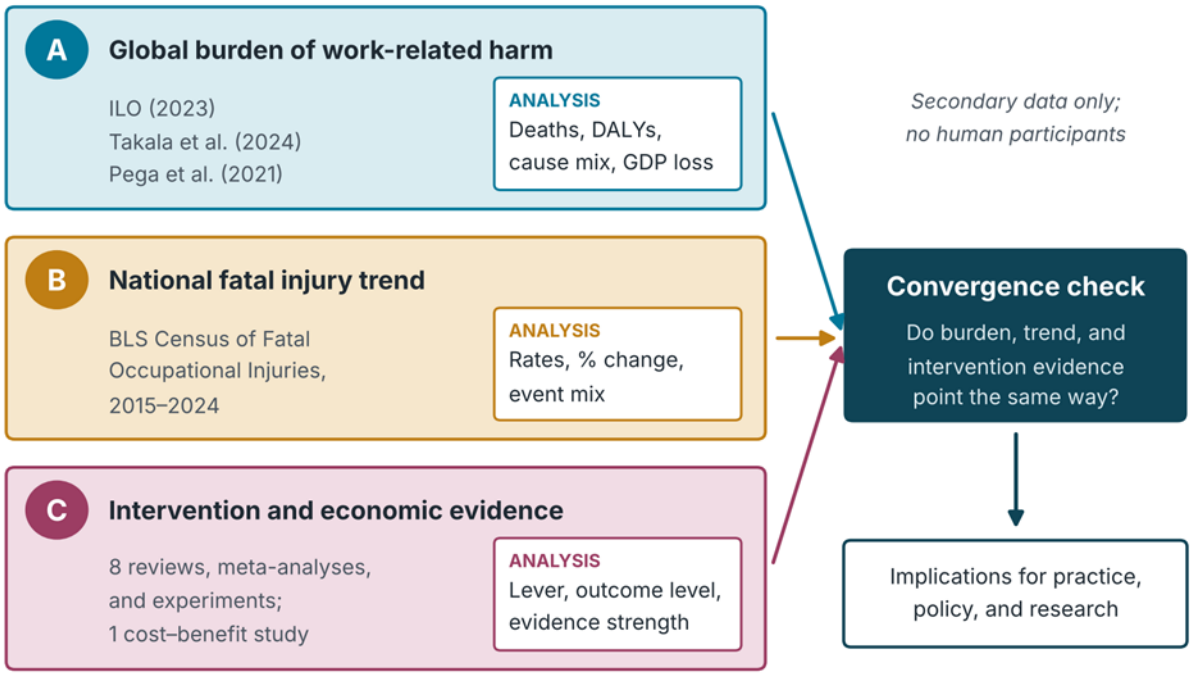
Data Sources

Global burden estimates were drawn from the ILO's (2023) report for 2019, Takala et al.'s (2024) analysis of 181 countries, and the WHO–ILO assessment of deaths attributable to long working hours (Pega et al., 2021). Fatal-injury counts, event types, and rates for 2015–2024

were taken from the U.S. Bureau of Labor Statistics (BLS, 2026) Census of Fatal Occupational Injuries. The U.S. series offers consistent definitions and verified records, but is not globally representative. Figure 2 summarizes this combined approach.

Figure 2

Analytic Design of the Study



Note. DALYs = disability-adjusted life years; GDP = gross domestic product; BLS = U.S. Bureau of Labor Statistics.

Selection of Intervention and Economic Evidence

The intervention synthesis prioritized peer-reviewed systematic reviews, meta-analyses, and controlled or natural-experiment studies that assessed regulation, training, management systems, organizational climate, or job design. Eligible sources reported at least one relevant outcome—knowledge, behavior, injury, illness, or cost. Eight intervention publications and one international economic study were retained for focused comparison. The purposive selection supports a focused synthesis, not an exhaustive review.

Analysis

For the BLS series, I compared annual counts and rates, calculated the ten-year mean, and examined changes by event type. A separate descriptive grouping identified 2024 fatalities arising from falls, contact with equipment, fires, electrical exposure, or oxygen deficiency—hazards with established engineering or design responses. Transportation events were excluded despite preventable risks. Interventions were assessed by measure and outcome, distinguishing knowledge gains from injury or cost reductions. Evidence-strength assessments were retained as reported by the respective review authors rather than independently reassigned.

Results

Scale and Composition of Work-Related Harm

The global evidence indicates that occupational harm is predominantly a disease burden. Of an estimated 2.93 million work-related deaths in 2019, approximately 2.6 million resulted from disease and 330,000 from accidents (ILO, 2023). Circulatory disease and cancer accounted

for 912,000 and 843,000 deaths, respectively (Takala et al., 2024). Long working hours were independently associated with an estimated 745,194 deaths from ischemic heart disease and stroke in 2016 (Pega et al., 2021). Table 1 also documents 180 million work-attributable disability-adjusted life years and an estimated economic loss equivalent to 5.8% of global gross domestic product (Takala et al., 2024). Differing methods complicate trends but not the significance of occupational disease. Long-term exposure therefore belongs within safety management.

Table 1

Burden and Economic Indicators Drawn From Secondary Sources

Indicator	Estimate	Year	Source
Global			
Work-related deaths, all causes	2.93 million	2019	ILO (2023)
Deaths from work-related disease	2.6 million	2019	ILO (2023)
Deaths from occupational accidents	330,000	2019	ILO (2023)
Non-fatal occupational injuries	395 million	2019	ILO (2023)
Deaths from work-related circulatory disease	912,000	2019	Takala et al. (2024)
Deaths from work-related malignant neoplasms	843,000	2019	Takala et al. (2024)
Deaths attributable to working ≥55 hr/week	745,194	2016	Pega et al. (2021)
Work-attributable DALYs	180 million	2019	Takala et al. (2024)
Economic loss from work-related harm	5.8% of GDP	2019	Takala et al. (2024)
United States			
Fatal occupational injuries	5,070	2024	BLS (2026)
Fatal injury rate per 100,000 FTE workers	3.3	2024	BLS (2026)
Share of fatalities from transportation incidents	38.2%	2024	BLS (2026)
Firm level			
Return on prevention (benefit–cost ratio)	2.2	—	Bräunig & Kohstall (2013)

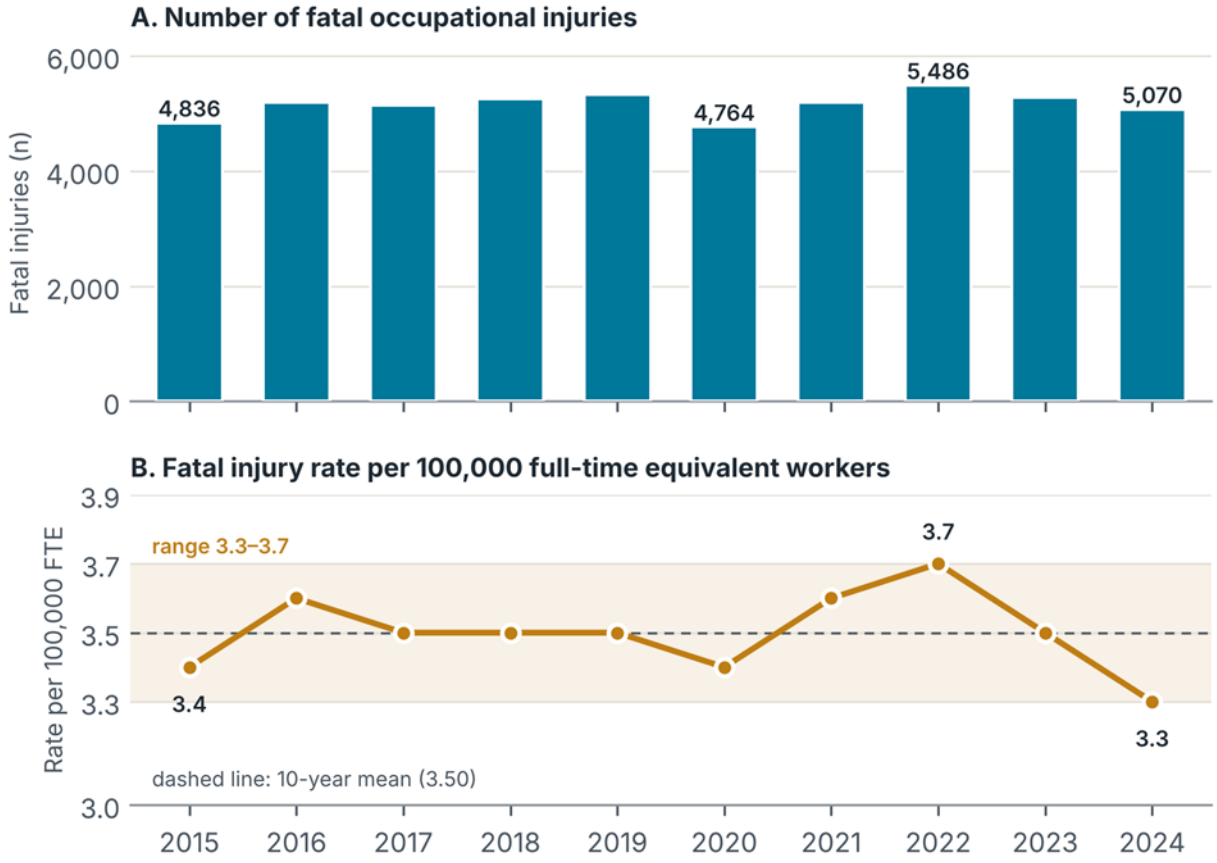
Note. ILO = International Labour Organization; DALYs = disability-adjusted life years; GDP = gross domestic product; BLS = U.S. Bureau of Labor Statistics; FTE = full-time equivalent. Estimates from different producers rely on different methods and should not be summed. The dash indicates that the estimate draws on a multi-country company survey rather than a single reference year.

A Decade of Flat Fatal Injury Rates

The U.S. recorded 5,070 fatal occupational injuries in 2024, down from 5,283 in 2023; the corresponding rate declined from 3.5 to 3.3 per 100,000 full-time equivalent workers (BLS, 2026). Figure 3 places this improvement in a longer perspective. From 2015 through 2024, the annual rate remained between 3.3 and 3.7, averaging 3.50, while the 2024 death count exceeded that of 2015. Pandemic disruptions limit interpretation of the 2020 decline. The broadly stable rate shows limited progress without establishing its causes.

Figure 3

Fatal Occupational Injuries in the United States, 2015–2024



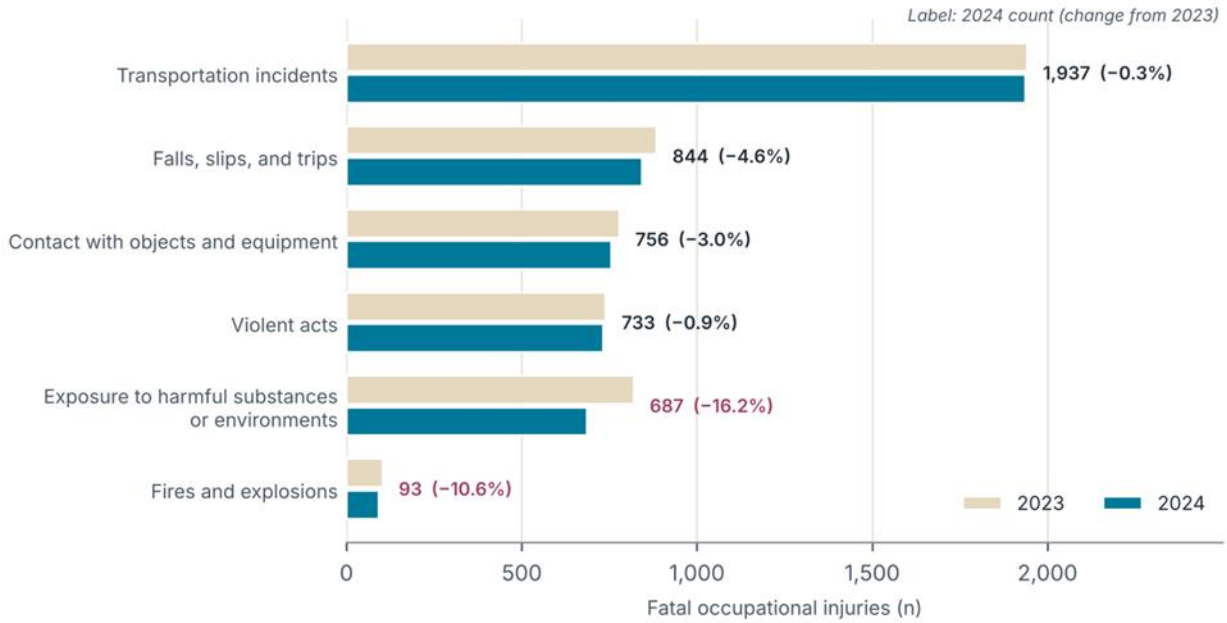
Note. Panel A shows counts and Panel B shows rates on separate scales to avoid a dual-axis display. FTE = full-time equivalent. Data from the Census of Fatal Occupational Injuries (U.S. Bureau of Labor Statistics, 2026); chart drawn by the author.

Mechanisms Behind the 2024 Fatalities

Transportation incidents caused 1,937 U.S. occupational deaths in 2024, or 38.2% of the total. Falls, slips, and trips accounted for 844 deaths; contact with objects or equipment for 756; violent acts for 733; and exposure to harmful substances or environments for 687 (BLS, 2026). Figure 4 compares the principal event categories with 2023. Much of the annual decline occurred in the exposure category, including fewer drug and alcohol overdoses, while falls and equipment-contact deaths changed relatively little. Under a narrow classification, 1,858 deaths (36.6%) involved hazards addressed by established controls, including fall protection, machine guarding, isolation, and confined-space testing. This identifies opportunities rather than proven preventable cases.

Figure 4

Fatal Occupational Injuries in the United States by Event or Exposure, 2023 and 2024



Note. Categories follow the Occupational Injury and Illness Classification System used by the Census of Fatal Occupational Injuries. Labels give the 2024 count and the percentage change from 2023; changes of 10% or more are shown in a darker accent. Data from the U.S. Bureau of Labor Statistics (2026); chart drawn by the author.

What the Intervention Evidence Shows

Table 2 compares evidence on distinct preventive measures. Enforcement provides the clearest injury-related findings. Levine et al. (2012) studied 409 randomly inspected California establishments and 409 matched controls; inspections were followed by 9.4% fewer injuries and a 26% reduction in injury costs, without detectable losses in employment or firm performance. Across 43 studies, Tompa et al. (2016) found stronger evidence for inspections that resulted in citations or penalties than for visits without such consequences. Credible oversight can influence organizational choices, although penalties alone cannot ensure safety.

Table 2

Selected Evidence on the Effectiveness of Preventive Measures

Source	Design	Preventive lever	Principal finding
Levine et al. (2012)	Natural field experiment; 409 inspected and 409 matched control establishments	Randomized government inspection	Injury rate fell 9.4% and injury costs 26%; no detectable effect on employment, sales, credit ratings, or firm survival
Tompa et al. (2016)	Systematic review; 43 studies	Regulatory enforcement	Strong evidence that inspections with citations or penalties reduce injuries; strong evidence that inspections without penalties do not
Robson et al. (2007)	Systematic review; 13 studies of adequate quality	OSH management systems	Results generally positive, none negative; evidence judged insufficient to recommend for or against
Robson et al. (2012)	Systematic review; 22 studies	OSH training	Strong evidence of change in work behavior; insufficient evidence of effects on symptoms, injuries, or illness
Burke et al. (2006)	Meta-analysis	Training method	More engaging methods yielded greater knowledge gains and larger reductions in accidents, illnesses, and injuries
Christian et al. (2009)	Meta-analysis	Safety climate; person factors	Knowledge and motivation most strongly related to safe behavior; group safety climate most strongly related to accidents and injuries
Beus et al. (2010)	Meta-analysis	Safety climate	Injuries predicted later climate more strongly than climate predicted injuries; management commitment the most robust predictor
Nahrgang et al. (2011)	Meta-analysis; 203 samples, N = 186,440	Job demands and resources	Supportive environment the most consistent resource, and risks and hazards the most consistent demand, across industries

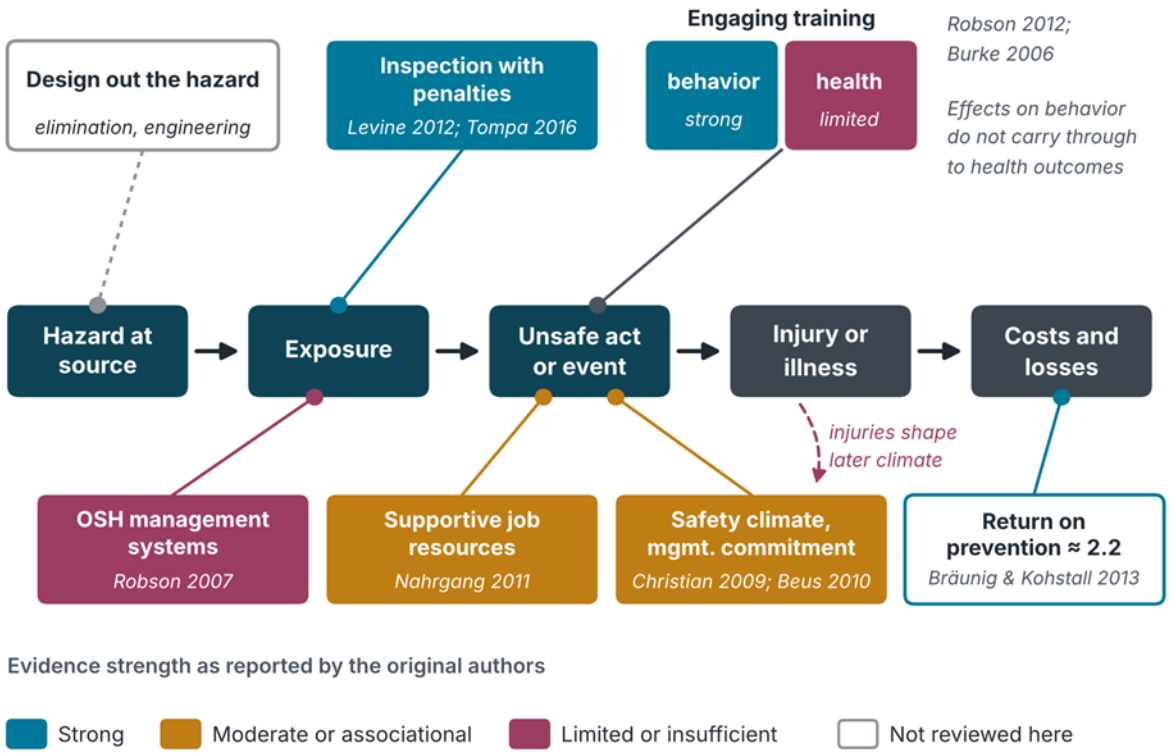
Note. OSH = occupational safety and health. Findings are summarized as reported by the original authors; strength-of-evidence wording is theirs.

Interventions aimed at individuals produce a more qualified picture. Robson et al. (2012) found strong evidence that occupational safety training changes behavior but insufficient evidence that it reduces injuries or illness; Burke et al. (2006), however, reported better outcomes from interactive and practice-based instruction than from passive teaching. Organizational context also matters. Christian et al. (2009) associated knowledge and motivation

with safe behavior, and group safety climate with injuries. Across 203 samples, Nahrgang et al. (2011) connected job resources and hazards with safety. Beus et al. (2010) cautioned that the relationship is reciprocal: previous injuries may shape later perceptions of safety climate more strongly than climate predicts future injuries. Management-system interventions generally showed promising but methodologically insufficient evidence (Robson et al., 2007). Figure 5 locates these measures along the hazard-to-harm pathway without creating a new evidence ranking.

Figure 5

Preventive Levers Positioned Along the Causal Chain From Hazard to Loss



Note. Author's synthesis of the sources in Table 2 and Bräunig and Kohstall (2013). Shading reflects evidence strength as stated by each source's own authors and was not re-graded. Elimination and engineering controls are shown for completeness but were outside the reviewed evidence. The dashed arrow reflects the finding of Beus et al. (2010) that injuries predict later safety climate. OSH = occupational safety and health.

The Economic Case

Prevention also has a financial rationale, although estimates require care. A multicountry company study reported an average benefit–cost ratio of approximately 2.2 for occupational safety investments (Bräunig & Kohstall, 2013). Because it partly reflects respondent judgments, this is not a guaranteed return. Nevertheless, it is consistent in direction with the observed reduction in injury costs after randomized inspections (Levine et al., 2012) and the substantial economic losses associated with occupational harm globally (Takala et al., 2024). Economic benefits reinforce, but do not replace, the ethical case.

Discussion

Considered together, the evidence challenges the idea that safe working conditions are primarily achieved by asking individuals to be more careful. Global estimates reveal a large burden of occupational disease, the U.S. fatal-injury series shows little sustained improvement, and intervention studies give comparatively strong support to enforceable organizational measures. These findings neither establish universal preventability nor quantify overall effectiveness. They emphasize work design, resources, exposure limits, and oversight before harm.

Why Progress Has Stalled

Several features of present practice may explain the distance between safety commitments and outcomes. Injury counts are visible and quickly available, whereas illnesses related to exposure, excessive hours, or chronic strain can take years to appear. Such risks can be underestimated. Documented training may be favored over costly engineering changes. Yet improved knowledge is not equivalent to demonstrated injury reduction (Robson et al., 2012). Finally, enforcement loses force when inspections are detached from meaningful follow-up (Tompa et al., 2016). These remain interpretations, not tested causal explanations.

Implications for Practice and Policy

For employers, prevention should begin at design and procurement, when hazards can be eliminated or modified. Practical training remains essential (Burke et al., 2006), but cannot replace guards or safe schedules. Overtime policies deserve attention given the disease burden associated with long hours (Pega et al., 2021). Leading indicators include resolved hazard reports, maintenance, worker involvement, and corrective actions. Regulators can reinforce these practices through targeted inspections, credible sanctions, and stronger occupational-disease surveillance. ISO 45001 (ISO, 2018) provides a management structure, but certification alone does not establish safer conditions (Robson et al., 2007). Long-term health effects of organizational changes also require research.

Limitations

The analysis has important limits. Global disease estimates depend on modeling assumptions, and figures from different producers cannot be combined as though they arose from one census. The U.S. injury series does not capture global risk or occupational disease. The grouping of fatalities amenable to engineering controls also required judgment and cannot establish preventability case by case. Finally, intervention studies were selected purposively, with the reviews' own assessments of evidence quality retained. The synthesis identifies patterns, not pooled causal effects or precise returns.

Conclusion

Safe working conditions cannot be produced through post-incident responses alone. The evidence emphasizes the continuing burden of occupational disease, persistent fatal-injury risks, and the influence of organizational decisions. Effective prevention combines engineering controls, manageable schedules, credible oversight, and worker participation, with training as support rather than substitute. Its success should be measured by reduced exposure and harm, not simply documented compliance.

REFERENCES

1. Beus, J. M., Payne, S. C., Bergman, M. E., & Arthur, W., Jr. (2010). Safety climate and injuries: An examination of theoretical and empirical relationships. *Journal of Applied Psychology*, *95*(4), 713–727. https://doi.org/10.1037/a0019164
2. Bräunig, D., & Kohstall, T. (2013). *Calculating the international return on prevention for companies: Costs and benefits of investments in occupational safety and health* [Final report]. International Social Security Association.
3. Burke, M. J., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R. O., & Islam, G. (2006). Relative effectiveness of worker safety and health training methods. *American Journal of Public Health*, *96*(2), 315–324. https://doi.org/10.2105/AJPH.2004.059840
4. Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology*, *94*(5), 1103–1127. https://doi.org/10.1037/a0016172
5. International Labour Organization. (2023). *A call for safer and healthier working environments*. International Labour Office.
6. International Organization for Standardization. (2018). *Occupational health and safety management systems — Requirements with guidance for use* (ISO Standard No. 45001:2018).
7. Levine, D. I., Toffel, M. W., & Johnson, M. S. (2012). Randomized government safety inspections reduce worker injuries with no detectable job loss. *Science*, *336*(6083), 907–911. https://doi.org/10.1126/science.1215191
8. Nahrgang, J. D., Morgeson, F. P., & Hofmann, D. A. (2011). Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology*, *96*(1), 71–94. https://doi.org/10.1037/a0021484
9. Pega, F., Náfrádi, B., Momen, N. C., Ujita, Y., Streicher, K. N., Prüss-Üstün, A. M., Descatha, A., Driscoll, T., Fischer, F. M., Godderis, L., Kiiver, H. M., Li, J., Magnusson Hanson, L. L., Rugulies, R., Sørensen, K., & Woodruff, T. J. (2021). Global, regional, and national burdens of ischemic heart disease and stroke attributable to exposure to long working hours for 194 countries, 2000–2016: A systematic analysis from the WHO/ILO Joint Estimates of the Work-related Burden of Disease and Injury. *Environment International*, *154*, Article 106595. https://doi.org/10.1016/j.envint.2021.106595
10. Reason, J. (1997). *Managing the risks of organizational accidents*. Ashgate.
11. Robson, L. S., Clarke, J. A., Cullen, K., Bielecky, A., Severin, C., Bigelow, P. L., Irvin, E., Culyer, A., & Mahood, Q. (2007). The effectiveness of occupational health and safety management system interventions: A systematic review. *Safety Science*, *45*(3), 329–353. https://doi.org/10.1016/j.ssci.2006.07.003
12. Robson, L. S., Stephenson, C. M., Schulte, P. A., Amick, B. C., III, Irvin, E. L., Eggerth, D. E., Chan, S., Bielecky, A. R., Wang, A. M., Heidotting, T. L., Peters, R. H., Clarke, J. A., Cullen, K., Rotunda, C. J., & Grubb, P. L. (2012). A systematic review of the effectiveness of occupational health and safety training. *Scandinavian Journal of Work, Environment & Health*, *38*(3), 193–208. https://doi.org/10.5271/sjweh.3259
13. Takala, J., Hämäläinen, P., Sauni, R., Nygård, C.-H., Gagliardi, D., & Neupane, S. (2024). Global-, regional- and country-level estimates of the work-related burden of diseases and

accidents in 2019. *Scandinavian Journal of Work, Environment & Health, 50*(2), 73–82. https://doi.org/10.5271/sjweh.4132

14. Tompa, E., Kalcevich, C., Foley, M., McLeod, C., Hogg-Johnson, S., Cullen, K., MacEachen, E., Mahood, Q., & Irvin, E. (2016). A systematic literature review of the effectiveness of occupational health and safety regulatory enforcement. *American Journal of Industrial Medicine, 59*(11), 919–933. https://doi.org/10.1002/ajim.22605
15. U.S. Bureau of Labor Statistics. (2026, February 19). *National census of fatal occupational injuries in 2024* (News Release No. USDL-26-0230). U.S. Department of Labor. https://www.bls.gov/news.release/cfoi.nr0.htm
16. Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology, 65*(1), 96–102. https://doi.org/10.1037/0021-9010.65.1.96

Mundarija / Contents

1	**Abdiraimova Shaxnoza Abdugappar qizi,** THE IMPORTANCE OF THE FAMILY INSTITUTION IN PROMOTING HEALTHY LONGEVITY: MODERN SOCIOLOGICAL APPROACHES	4
2	**Abdullayeva Nasima Murodullayevna,** DIALEKTAL KORPUS YARATISHNING ZAMONAVIY TAJRIBALARI ASOSIDA QASHQADARYO SHEVASI LINGVISTIK BAZASINI SHAKLLANTIRISH	10
3	**Abdullayeva Shaxnoza Raximjon qizi,** AXBOROT TEXNOLOGIYALARI TERMINLARINI ONTOLOGIK TASNIFLASHDAGI MUAMMOLAR VA ULARNING YECHIMLARI	17
4	**Abdunazarova Maftuna Faxriddinovna,** BADIIY MULOQOTDA IMPLIKATURA INTERPRETATSIYASI VA PRAGMATIK TO'SIQ MUAMMOSI (ERKIN A'ZAM ASARLARI MISOLIDA)	21
5	**Abduraxmanova Azoda Djo'rayevna, Solixov Murod Botir o'g'li,** SUN'IY INTELEKT TEXNOLOGIYALARI YORDAMIDA ERGANOMIKANI TAHLIL QILISH ORQALI KORXONALARDA MEHNAT XAVFSIZLIGINI TA'MINLASH	27
6	**Akhmedova Anorkhon Nasibali kizi, Bokijonova Sayyorakhon Tukhtamurod kizi,** DEVELOPING WRITTEN TRANSLATION COMPETENCE THROUGH AI-ASSISTED LEARNER PROFILING AND DIFFERENTIATED INSTRUCTION	31
7	**Nafisa Amanlikova,** O'ZBEK XALQ DOSTONLARINING PARALLEL KORPUSINI YARATISH VA UNING LINGVISTIK AHAMIYATI	42
8	**Sabina Amonova,** SINONIMLAR O'QUV LUG'ATLARIDA IZOHLASH MUAMMOSI VA MULTIMEDIALI YONDASHUV	46
9	**Durdona Anvarova,** KORPUSGA ASOSLANGAN OMONIMLAR LUG'ATINI YARATISHDA OMONIMLARNI TASNIFLASH VA LOYIHALASHTIRISH MASALALARI	51
10	**Umida Askarova,** KORPUS TAHLILI ASOSIDA NLP TEXNOLOGIYALARINING RIVOJLANISHI VA LINGVOSTATISTIK TADQIQOTLARDAGI AHAMIYATI	57
11	**Saodat Boysariyeva,** O'ZBEK TILIDA OLMOSHLARNING REFERENSIAL XUSUSIYATLARI VA DISKURSDAGI KOMMUNIKATIV VAZIFALARI	62
12	**Джуманиязова Интизор Давлатназар кизи,** ЛИНГВИСТИЧЕСКИЕ ИНДИКАТОРЫ ДЛЯ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ И ПОИСКА АФОРИЗМОВ	69
13	**Ergashev Sirojiddin,** K-MEANS KLASTERLASH ALGORITMIGA NEYRON TARMOQ ARXITEKTURASINI QO'LLASH	74
14	**G'anijonova Shaxlo Dilmurod qizi,** FRAZEOLOGIK BIRLIKLARNI KORPUS UCHUN TEGLASH MASALASIDA ZAMONAVIY YONDASHUVLAR VA TADQIQOTLAR SHARHI	78
15	**С. С. Гулямов, С. С. Гулямов,** СОЦИАЛЬНО ОРИЕНТИРОВАННАЯ ЦИФРОВАЯ ТРАНСФОРМАЦИЯ ЭКОНОМИКИ ЦЕНТРАЛЬНОАЗИАТСКИХ СТРАН	84
16	**Ibodullayeva Kunduz Rustam qizi,** CORPUS-BASED LINGUOSTATISTICAL ANALYSIS OF UZBEK SCIENTIFIC DISCOURSE: LEXICAL UNITS AND DISCOURSE MARKERS	103
17	**Iskandarova Aybibi Rakhimbayevna,** ELEKTRON LUG'AT BAZASI UCHUN SO'Z BIRIKMASI – FRAZEOLOGIK BIRLIK JUFTLIGINING IDENTIFIKATSIYA NAMUNALARI	109
18	**Iskandarova Iroda Bakir qizi,** O'ZBEK TILI NUTQ SINTEZATORI	116
19	**Adizova Nilufar Istamovna,** TURIZM LEKSIKASINING N-GRAM MODELI VA GIPONIM-GIPERONIMIK MUNOSABATLARI	120
20	**Jamolova Sarvinoz Shuhratjon qizi,** BREND NOMLARINING LINGVISTIK MOHIYATI	126
21	**Jo'raeva Zilola Obidjon qizi,** AVTOMATIK NUTQNI TANISH TEXNOLOGIYASINING ZAMONAVIY YONDASHUVLARI	131
22	**Madina Jo'raqulova,** KORPUS MATNLARINI SINTAKTIK TEGLASHNING AHAMIYATI	137
23	**Kamolova Ozodaxon Muzaffar qizi,** NAZARIY VA AMALIY DARS MASHG'ULOTLARINING UYG'UNLIGI HAMDA TA'LIM YO'NALISHLARIDA JORIY ETILISHI	142
24	**Kanayeva Ozoda Esirgapovna,** NOMLANGAN BIRLIKLARNI ANIQLASH TIZIMLARINING NAZARIY ASOSLARI VA EVOLYUTSIYASI	148
25	**Sevinch Kenjayeva,** A CONTEXT-AWARE NOISY-CHANNEL MODEL FOR AUTOMATIC SPELLING CORRECTION OF UZBEK USER-GENERATED TEXTS	154
26	**Sanobarkhon Khabibullaeva,** O'ZBEK TILI PARSING TIZIMI UCHUN SODDA GAPLARNING SINTAKTIK MODELLARI: TADQIQOTLAR SHARHI	161
27	**Khajimuratov Abdukakhkhor Abdumutalovich,** OLIY TA'LIM MUASSASALARIDA TADBIRKORLIK EKOTIZIMINI SHAKLLANTIRISH VA UNING MINTAQA IQTISODIY O'SISHIGA TA'SIRINI TIZIMLI MODELLASHTIRISH	164
28	**Mukhlisa Khamidova,** DIGITAL TRANSFORMATION AND GOVERNANCE OF HIGHER EDUCATION INSTITUTIONS IN UZBEKISTAN: INSTITUTIONAL CAPACITY, DATA INFRASTRUCTURE AND INTERNATIONAL COMPETITIVENESS	169
29	**Begoyim Kholbekova,** POETRY TRANSLATION AND THE PROBLEM OF EQUIVALENCE	179

30	**Zebo Gulboyevna Kodirova,** A REVIEW OF SELECTED STUDIES ON THE DEVELOPMENT OF AN ARABIC-UZBEK SEMANTIC SEARCH SYSTEM	183
31	**Kuvondikova Gavhar Isomiddinovna,** O'QUV KORPUSINI SHAKLLANTIRISH METODIKASI VA CHASTOTALI LUG'AT TUZISH TAMOYILLARI	188
32	**Maxkamova Mohinur Mirodil qizi,** "O'ZBEK VA TURK TILLARIDA AVTOMATIK MORFOLOGIK TAHLIL VA LEMMATIZATSIYA: MASHINA TARJIMASI UCHUN ZAMONAVIY YONDASHUVLAR"	194
33	**Мирмахмудова Ирода Ирисматовна,** САТ-ВОСИТАЛАР ВА КАТТА ТИЛ МОДЕЛЛАРИ ЁРДАМИДА ТАРЖИМА ТРАНСФОРМАЦИЯЛАРИНИ АВТОМАТИК АНИҚЛАШ: КОМИССАРОВ ТАСНИФИ МИСОЛИДА	199
34	**Dildora Mirvaliyeva,** RASMIY HUJJATLAR TILIDA QO'LLANILADIGAN LEKSIK VA GRAMMATIK BIRLIKLARNING O'ZIGA XOSLIGI	204
35	**Mirzayeva Faxriya Shavkat qizi,** O'ZBEK TILIDAGI MATNLARDA XATOLIKLARNI ANIQLASHDA DAMERAU–LEVENSHTEYN MASOFASI VA MARKOV ZANJIRI MODELLARINING QIYOSIY TAHLILI	208
36	**Mirzayeva Maftuna Baxtiyor qizi,** OYBEKNING "NAVOIY" ROMANIDA PORTRET BIRLIKLARINING LEKSIK-SEMANTIK MUNOSABATLARI	214
37	**Muxammadiyeva Kamola Doston qizi,** O'ZBEKCHA-RUSCHA PARALLEL KORPUSLARI ASOSIDA ALIGNER TIZIMINI YARATISHDA NLP TEXNOLOGIYALARI	220
38	**Elmira Nazirova, Nilufar Madetbayeva,** IBORALARNI MATNDAN ANIQLASH VA IKKI TILLI EKVIVALENTLIGINI BAHOLASHDA ONTOLOGIYAGA ASOSLANGAN MATEMATIK MODEL	225
39	**Normamatova Mashhura,** MUHAMMAD YUSUF MUALLIFLIK KORPUSI: SHAKLLANTIRISH TAMOYILLARI VA LINGVOSTATISTIK TAHLIL	231
40	**Nazokat O'rinboyeva,** TABIIY TILNI QAYTA ISHLASH VA BILIMLAR GRAFLARI: INTEGRATSIYALASHGAN YONDASHUV VA ZAMONAVIY TENDENSIYALAR	238
41	**O'rinboyeva Roza Toliboy qizi,** NUTQ KORPUSINI ANNOTATSIYALASHNING LINGVISTIK ASOSLARI: SAMARQAND VILOYATI SHEVASI MISOLIDA	241
42	**Aziz O'temisov,** FOMA DASTURIY MUHITIDA QARAQALPOQ TILIDAGI SON SO'Z TURKUMINI MODELLASHTIRISH	248
43	**Odilova Olmosxon Dilshodjon qizi,** HUQUQIY HUJJATLAR ASOSIDA NER TIZIMINI YARATISH VA MATNLARNI TASNIFLASH	253
44	**Olimjonova Ziynat Bobomurod qizi,** RAQAMLI MAKONDA SEVGI FENOMENI: ER-XOTIN MUNOSABATLARIDA HISSIY YAQINLIK VA ISHONCH DINAMIKASI	259
45	**Po'latova Gulxayyo G'ofurjon qizi,** O'ZBEK TILI TEZAURUSIDA SIFAT LEKSEMALARINING LINGVISTIK VA METODOLOGIK XUSUSIYATLARI	263
46	**Qadirova Mokhira Bakhodir qizi, Saparova Mokhira Fayzullayevana,** THE DISTINCTIVE FEATURES OF THESAURUS DICTIONARIES	267
47	**Qurbonnazarova Aisla O'ktamjonovna,** DIALEKTAL MATERIALLARNI YIG'ISHNING ZAMONAVIY METODLARI VA ULARNI SAMARQAND SHEVASI KORPUSINI YARATISHDA QO'LLASH	271
48	**Rahimov Usmonjon Ergashevich , Sirojiddinova Muxarram Raximjon qizi ,** BADIIY DISKURSDA FRAZEOLOGIK BIRLIKLARNING LINGVOTERAPEVTIK TA'SIRI (CHO'LPONNING "KECHA VA KUNDUZ" ROMANI MISOLIDA)	279
49	**Safiye Kırlangıç,** O'ZBEK VA TURK RAQAMLİ NUTQİDA YANGİ LEKSİK BİRLİKLAR: QİYOSİY TAHLİL	285
50	**Sattorova Nafisa Isomidinovna,** ERNEST HEMENGUEY VA OYBEK PERSONAJLARIDA "SADOQAT" KONSEPTINI TARKIBLANISHI	291
51	**Sharipov Qudratbek Quraltayevich,** KORXONALARDA YASHIL INVESTITSIYALARNI MOLIYALASHTIRISH SAMARADORLIGINI BAHOLASHNING EKONOMETRIK MODELI	296
52	**Sirojiddinov Ziyoviddin Sharofiddin o'g'li, Hakimov Muftax Hamidovich,** QOIDAVIY KKT FORMALIZMI VA YIRIK TIL MODELLARI INTEGRATSIYASI: ARXITEKTURA VARIANTLARI TASNIFI VA BAHOLASH MEZONLARI	300
53	**Sobirova Madinaxon Raxim qizi,** XITOY VA O'ZBEK TILLARIDAGI EKOTURIZM TERMINLARINING STRUKTUR XUSUSIYATLARI VA RAQAMLI TERMINOLOGIK BAZANI SHAKLLANTIRISH MASALALARI	306
54	**Jumagul Sulaymonova,** O'ZBEK TILIDA SENTIMENT TAHLILNING NAZARIY VA LINGVISTIK MASALALARI	309
55	**Султонова Сайёра Холмирзаевна,** ДИАГНОСТИКА КАК ИНСТРУМЕНТ ФОРМИРУЮЩЕГО ЭКСПЕРИМЕНТА: АВТОРСКАЯ МЕТОДИКА И ПЛАТФОРМА «ДИАГКОМП-РУС»	315
56	**Суюнова Мохинур Илхом кизи,** КУЛЬТУРНАЯ ОБУСЛОВЛЕННОСТЬ РЕЧЕВЫХ АКТОВ В УЗБЕКСКОМ ЯЗЫКЕ: КОРПУСНАЯ ВЕРИФИКАЦИЯ	321
57	**Z. Taxirov,** USLUBIYAT VA NUTQ MADANIYATI	328
58	**Go'zal S, Tojiyeva,** RAQAMLI TA'LIM MUHITIDA BO'LAJAK PEDAGOGLARNING KASBIY RIVOJLANISHINI BAHOLASHNING INTELLEKTUAL TEXNOLOGIYALAR	333
59	**Kamola Toshtemirova,** GENERATIONAL DIGITAL KNOWLEDGE GAP IN UZBEK ORGANIZATIONS	338

60	**Toshtemirova Mohichehra G'olibjon qizi,** IJTIMOIY-MADANIY TIZIMLARDA O'ZBEK MILLIY LIBOSLARI DIZAYNINI RAQAMLI TRANSFORMATSIYA QILISH VA VIZUAL KODLARNI PARAMETRIK MODELLASHTIRISH METODOLOGIYASI	340
61	**Umarov Xusniddin Egamboyevich,** BO'LAJAK BOSHLANG'ICH SINF O'QITUVCHILARINI ILMIY-IJODIY FAOLIYATGA YO'NALTIRISH PEDAGOGIK MUAMMO SIFATIDA	344
62	**Umarova Dildora Nizom qizi,** INOVATSION TA'LIM SHAROITIDA MAKTABGACHA TA'LIM TAYYORLOV GURUHI BOLALARIDA ATROF MUHITGA MUNOSABATNI SHAKILLANTIRISH	349
63	**Urazaliyeva Mavluda Yangiboyevna,** AUDIOKORPUS MATERIALLARIGA ISHLOV BERISHNING TEXNOLOGIK BOSQICHLARI: O'ZBEK TILI MISOLIDA	353
64	**Xanxodjayeva Gavxar,** SUN'IY INTELLEKT TEXNOLOGIYALARINING ZAMONAVIY MUAMMOLARI: IJTIMOIY, AXLOQIY VA TEXNOLOGIK TAHLIL	357
65	**Xojimurodova Ominaxon Bobomurod qizi,** TALABALAR YOZMA ISHLARIDAGI GRAMMATIK XATOLARNI O'ZBEK TILIDA AVTOMATIK ANIQLASH VA TUZATISH TIZIMINING YETISHMASLIGI	362
66	**Xolmurotova Muxlisa Botirovna,** IJTIMOIY TIZIMLARNI MODELLASHTIRISH VA RAQAMLI TRANSFORMATSIYA: ILMIY TAJRIBALAR, TEXNOLOGIYALAR VA RIVOJLANISH ISTIQBOLLARI	366
67	**Munisa Xudoyberganova,** SEMANTIK KORPUS VA UNING NAZARIY ASOSLARI	368
68	**Yusupova Iroda Ravshanbek qizi,** O'ZBEK TILI TA'LIMIY LUG'ATLARINI RAQAMLI MUHITDA MODELLASHTIRISH	372
69	**Davronov Shakhbos Erkin ugli, Kakhkhorov Siddik K.,** MECHANISMS FOR DEVELOPING MEDICAL STUDENTS' PROFESSIONAL COMPETENCE THROUGH DIGITAL TECHNOLOGIES	378
70	**Davronov Shakhbos Erkin ugli, Kakhkhorov Siddik K.,** FACTORS INFLUENCING THE DEVELOPMENT OF STUDENTS' PROFESSIONAL COMPETENCE IN A DIGITAL LEARNING ENVIRONMENT	383
71	**D.R. Raxmonova,** MAKTABGACHA TA'LIM TASHKILOTI RAHBARINING SHAXSIY SIFATLARI VA EMOTSIONAL INTELLEKTINING KONFLIKTLARNI BOSHQARISHDAGI O'RNI	389
72	**Solihova Maftuna Sharoq qizi,** PEDAGOGLARDA RAQAMLI KOMPETENSIYANI SHAKLLANTIRISH SAMARADORLIGI	392
73	**Solihova Maftuna Sharof qizi,** TABIIY FANLARNI O'QITISHDA RAQAMLI TA'LIM MUHITINI TAKOMILLASHTIRISHNING PEDAGOGIK MEXANIZMLARI	395
74	**Zarnigor Komilova Xayriddin qizi,** RAQAMLI TRANSFORMATSIYA SHAROITIDA TIJORAT BANKLARI FAOLIYATI SAMARADORLIGINI OSHIRISH MEXANIZMLARI	398
75	**Комилова Зарнигор Хайриддин кизи,** ОРГАНИЗАЦИЯ СБОРА ДЕНЕЖНЫХ ПОСТУПЛЕНИЙ И ЦЕННОСТЕЙ В КОММЕРЧЕСКИХ БАНКАХ: СТРАТЕГИИ УЧЕТА И ОПТИМИЗАЦИИ ПРОЦЕССОВ	401
76	**Suyarova Gulruxsor Nurilla qizi,** ALISHER NAVOIY IJODINING ERKIN VOHIDOV SHE'RIYATIDA O'RNI	407
77	**O'roqov Azizbek Usmonaliyevich,** THE IMPORTANCE OF PREVENTIVE MEASURES IN CREATING SAFE WORKING CONDITIONS	410

O‘ZBEKISTON RESPUBLIKASI OLIY TA’LIM, FAN VA INNOVATSIYA-LAR VAZIRLIGI

MIRZO ULUG‘BEK NOMIDAGI O‘ZBEKISTON MILLIY UNIVERSITETI

EUROPEAN ASSOCIATION OF OR SOCIETIES (EURO)

“SCIENCE AND INNOVATION” XALQARO ILMIY JURNALI

“IJTIMOIY TIZIMLARNI MODELLASHTIRISH VA RAQAMLI TRANSFORMATSIYA: ILMIY TAJRIBALAR, TEXNOLOGIYALAR VA RIVOJLANISH ISTIQBOLLARI (MSBC–2026)” MAVZUSIDAGI 4-XALQARO ILMIY-AMALIY KONFERENSIYASI

MATERIALLAR TO‘PLAMI

2026-YIL 23-25-SENTYABR

COLLECTION OF MATERIALS

OF THE 4TH INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE ON THE TOPIC “MODELLING AND SIMULATION OF SOCIAL-BEHAVIOURAL PHENOMENA IN CREATIVE SOCIETIES (MSBC-2026)”

September 23-25, 2026

Published: 15.10.2026. Font: «Times New Roman».

LLC «Science and innovation»
License Mass Media №:1597 27.04.2022
License Publisher №:038864 15.09.2022
Address: 100121, Tashkent, Yakkasaroy, st. Sh. Rustaveli 150,
Business center “Vega”, 6th floor, Office №:2
www.scientists.uz, info@scientists.uz, +998933549654, +998993319654