# O‘ZBEKISTONDA KOMPYUTER LINGVISTIKASI YO‘NALISHINING PAYDO BO‘LISHI

**Temirova Munisa Soatali qizi**
Andijon davlat Universiteti magistranti
E-mail: temirovamunisa.97@gmail.com
**Hasanova Shaxnoza Kalonboy qizi**
Andijon davlat universiteti magistranti
E-mail: shahnozaxasanova412@gmail.com
Doi: https://doi.org/10.5281/zenodo.20320498

**Annotatsiya:** Ushbu maqolada O‘zbekistonda kompyuter lingvistikasi yo‘nalishining shakllanishi va rivojlanish bosqichlari tahlil qilinadi. Kompyuter lingvistikasi lingvistika va axborot texnologiyalarining kesishgan sohasi sifatida tilni avtomatik qayta ishlash, modellashtirish va tahlil qilish masalalarini o‘rganadi. Tadqiqotda ushbu yo‘nalishning paydo bo‘lishiga ta’sir qilgan asosiy omillar, jumladan raqamli texnologiyalar rivoji, til resurslarini elektronlashtirish ehtiyoji va sun’iy intellektga bo‘lgan talab ko‘rib chiqiladi. Shuningdek, o‘zbek tilining korpus lingvistikasi, mashina tarjimasi va tabiiy tilni qayta ishlash tizimlaridagi o‘rni yoritiladi. Maqolada sohaning nazariy asoslari bilan birga amaliy yo‘nalishlari ham tahlil qilinib, mavjud muammolar va istiqbollar ko‘rsatib beriladi. Natijada kompyuter lingvistikasining O‘zbekistondagi ilmiy va amaliy ahamiyati asoslab beriladi.

**Kalit so‘zlar:** kompyuter lingvistikasi, korpus lingvistikasi, tabiiy tilni qayta ishlash, mashina tarjimasi, o‘zbek tili, raqamli til resurslari, sun’iy intellekt, morfologik tahlil, til texnologiyalari, avtomatik tizimlar.

**Аннотация.** В данной статье анализируются этапы формирования и развития компьютерной лингвистики в Узбекистане. Компьютерная лингвистика рассматривается как междисциплинарная область, объединяющая лингвистику и информационные технологии, и изучающая вопросы автоматической обработки, моделирования и анализа языка. В исследовании рассматриваются основные факторы, повлиявшие на развитие данного направления, включая цифровизацию языковых ресурсов, развитие информационных технологий и рост интереса к системам искусственного интеллекта. Особое внимание уделяется корпусной лингвистике узбекского языка, системам машинного перевода и технологиям обработки естественного языка. Также анализируются теоретические и прикладные аспекты, существующие проблемы и перспективы развития. В результате обосновывается научная и практическая значимость компьютерной лингвистики в Узбекистане.

**Ключевые слова:** компьютерная лингвистика, корпусная лингвистика, обработка естественного языка, машинный перевод, узбекский язык, языковые ресурсы, искусственный интеллект, морфологический анализ, языковые технологии, автоматические системы.

**Abstract.** This article analyzes the formation and development stages of computational linguistics in Uzbekistan. Computational linguistics is considered an interdisciplinary field that combines linguistics and information technologies, focusing on the automatic processing, modeling, and analysis of language. The study examines key factors influencing the emergence of this field, including the digitalization of language resources, the development of information technologies, and the growing demand for artificial intelligence systems. Special attention is

given to corpus linguistics of the Uzbek language, machine translation systems, and natural language processing technologies. Both theoretical foundations and applied aspects are discussed, along with existing challenges and future prospects. As a result, the scientific and practical significance of computational linguistics in Uzbekistan is justified.

**Keywords:** computational linguistics, corpus linguistics, natural language processing, machine translation, Uzbek language, language resources, artificial intelligence, morphological analysis, language technologies, automated systems.

**Kirish.** Kompyuter lingvistikasi zamonaviy tilshunoslikning eng tez rivojlanayotgan yo‘nalishlaridan biri bo‘lib, u tabiiy tilni kompyuter vositalari yordamida tahlil qilish, qayta ishlash va modellashtirish bilan shug‘ullanadi. Ushbu fan lingvistika, matematika, informatika va sun’iy intellektning kesishgan nuqtasida shakllanib, bugungi kunda nafaqat nazariy, balki amaliy jihatdan ham muhim ahamiyat kasb etmoqda. Dunyo miqyosida kompyuter lingvistikasi XX asrning ikkinchi yarmidan boshlab mustaqil ilmiy yo‘nalish sifatida shakllangan bo‘lsa, O‘zbekistonda bu soha nisbatan keyinroq, ammo tezkor rivojlanish bosqichiga kirgan yo‘nalishlardan biri hisoblanadi. O‘zbekistonda kompyuter lingvistikasi fanining paydo bo‘lishi bevosita axborot texnologiyalarining kirib kelishi, til resurslarini raqamlashtirish ehtiyoji va ta’lim tizimida zamonaviy texnologiyalardan foydalanish zarurati bilan bog‘liq. Dastlabki bosqichlarda asosiy e’tibor o‘zbek tilini kompyuter tizimlarida ifodalash, matnlarni kodlash, elektron lug‘atlar yaratish va oddiy avtomatlashtirilgan lingvistik tahlil vositalarini ishlab chiqishga qaratilgan. Keyinchalik esa bu yo‘nalish kengayib, korpus lingvistikasi, mashina tarjimasi, morfologik va sintaktik tahlil tizimlari kabi murakkab ilmiy-amaliy yo‘nalishlarni o‘z ichiga oldi.

O‘zbekiston ilmiy maktabida kompyuter lingvistikasining shakllanishida mahalliy tilshunoslar va axborot texnologiyalari mutaxassislarining hamkorligi muhim rol o‘ynadi. O‘zbek tilining grammatik tuzilishi, agglutinativ xususiyati va so‘z yasalish tizimi kompyuter modellashtirish uchun o‘ziga xos murakkabliklarni yuzaga keltiradi. Shu sababli, dastlabki tadqiqotlar asosan tilning formal modeli va algoritmik ifodalanishini ishlab chiqishga qaratildi. Bu jarayonda o‘zbek tilini avtomatik tahlil qilish uchun lingvistik qoidalar bazasi, morfologik tahlil algoritmlari va elektron resurslar yaratish yo‘lga qo‘yildi. Mustaqillikdan keyingi davrda esa kompyuter lingvistikasining rivojlanishi yanada jadallashdi. Lotin yozuviga o‘tish, davlat tilining raqamli muhitda keng qo‘llanilishi, shuningdek, global axborot tizimlariga integratsiya jarayonlari ushbu sohaning ahamiyatini yanada oshirdi. Natijada o‘zbek tilining raqamli korpuslari, avtomatik tarjima tizimlari va intellektual lingvistik dasturlar yaratishga bo‘lgan ehtiyoj kuchaydi. Bu esa o‘z navbatida ilmiy izlanishlarning yangi bosqichini boshlab berdi.

Bugungi kunda O‘zbekistonda kompyuter lingvistikasi nafaqat filologiya, balki sun’iy intellekt va ma’lumotlar fanlari bilan ham uzviy bog‘langan holda rivojlanmoqda. Nutqni aniqlash tizimlari, matnlarni avtomatik tahrirlash, sentiment tahlili va mashina tarjimasi kabi yo‘nalishlar amaliyotda keng qo‘llanilmoqda. Shu bilan birga, o‘zbek tilining raqamli resurslarini boyitish, lingvistik korpuslarni kengaytirish va zamonaviy algoritmik modellarni ishlab chiqish dolzarb vazifalardan biri bo‘lib qolmoqda. Mazkur maqolada O‘zbekistonda kompyuter lingvistikasi yo‘nalishining shakllanishi, rivojlanish bosqichlari hamda uning ilmiy va amaliy ahamiyati tahlil qilinadi. Shuningdek, sohaning bugungi holati va istiqbollari haqida ilmiy xulosalar beriladi.

**Adabiyotlar tahlili.** Kompyuter lingvistikasi zamonaviy tilshunoslik va axborot texnologiyalarining kesishgan eng muhim yo‘nalishlaridan biri sifatida shakllanib, bugungi kunda o‘zbek tilshunosligida ham alohida ilmiy paradigma sifatida qaralmoqda. Mavzuga oid adabiyotlar tahlili shuni ko‘rsatadiki, O‘zbekistonda kompyuter lingvistikasining shakllanishi va rivojlanishi bir necha bosqichli jarayon bo‘lib, u dastlab nazariy lingvistika doirasida,

keyinchalik esa amaliy hisoblash tizimlari bilan integratsiyalashgan holda rivojlangan. Dastlabki nazariy manbalarda kompyuter lingvistikasining mohiyati, uning lingvistika, matematika va informatika bilan o‘zaro aloqasi yoritiladi. Ushbu manbalarda tilni formal modellashtirish, algoritmik yondashuvlar va lingvistik ma’lumotlarni strukturaviy tahlil qilish asosiy masala sifatida qaraladi. Ayrim tadqiqotlarda kompyuter lingvistikasining asosiy vazifasi sifatida tabiiy tilni mashina tushunadigan shaklga keltirish, ya’ni formalizatsiya jarayoni alohida ta’kidlanadi. Bu yo‘nalishdagi ishlarda til birliklarini matematik model sifatida ifodalash, morfologik va sintaktik tahlilni avtomatlashtirish g‘oyalari ilgari surilgan.

O‘zbek tilshunosligida kompyuter lingvistikasi bo‘yicha ilk yondashuvlar asosan elektron lug‘atlar yaratish, matnlarni avtomatik qayta ishlash va tilni kodlash masalalariga qaratilgan. Bu bosqichda tilshunoslar va dasturchilar hamkorligi muhim rol o‘ynagan bo‘lib, o‘zbek tilining agglutinativ xususiyati, so‘z shakllanishi va morfologik murakkabligi alohida ilmiy muammo sifatida ko‘rilgan. Adabiyotlarda aynan shu jihat o‘zbek tilini kompyuter tizimlariga moslashtirishda asosiy ilmiy qiyinchiliklardan biri sifatida qayd etiladi. Keyingi bosqichdagi ilmiy ishlar korpus lingvistikasi rivoji bilan chambarchas bog‘liq. Bu yo‘nalishdagi tadqiqotlarda o‘zbek tilining milliy korpuslarini yaratish, matnlar bazasini shakllantirish, lingvistik ma’lumotlarni avtomatik tahlil qilish masalalari keng yoritilgan. Korpus lingvistikasi doirasida so‘zlarning chastotasi, grammatik kategoriyalari, semantik bog‘lanishlari va matn strukturalarini raqamlashtirishga katta e’tibor qaratilgan. Ushbu adabiyotlarda korpus yaratish nafaqat lingvistik tadqiqotlar uchun, balki ta’lim va sun’iy intellekt tizimlari uchun ham muhim resurs ekani ta’kidlanadi.

Mashina tarjimasi va tabiiy tilni qayta ishlash (NLP) yo‘nalishidagi ilmiy manbalar esa kompyuter lingvistikasining amaliy imkoniyatlarini kengaytirgan. Bu yo‘nalishda o‘zbek tili uchun avtomatik tarjima tizimlarini yaratish, sintaktik va semantik tahlil algoritmlarini ishlab chiqish, shuningdek, matnni avtomatik tahrirlash tizimlarini yaratish masalalari o‘rganilgan. Adabiyotlarda chuqur o‘rganish asosida neyron tarmoqlar va mashinaviy o‘rganish usullarining lingvistik modellashtirishdagi roli ortib borayotgani qayd etiladi. Ayniqsa, so‘nggi yillardagi tadqiqotlar o‘zbek tilining raqamli resurslarini boyitish va sun’iy intellekt tizimlariga integratsiya qilishga qaratilganligi bilan ajralib turadi. Shuningdek, o‘zbek olimlari tomonidan yaratilgan o‘quv qo‘llanmalar va monografiyalarda kompyuter lingvistikasining nazariy asoslari, lingvistik modellashtirish tamoyillari va algoritmik yondashuvlar batafsil bayon etilgan. Ushbu adabiyotlarda lingvistik birliklarni formal tizim sifatida ko‘rish, ularni kompyuter dasturlari yordamida tahlil qilish va qayta ishlash usullari yoritiladi. Ayniqsa, til birliklarini tasniflash, affiksal tizimni avtomatlashtirish va lemmatizatsiya jarayonlari o‘zbek tili materialida keng tahlil qilingan.

Xorijiy ilmiy manbalarda esa kompyuter lingvistikasi asosan tabiiy tilni qayta ishlash, katta ma’lumotlar bilan ishlash va sun’iy intellekt tizimlari bilan bog‘liq holda rivojlanganligi ko‘rsatiladi. Ushbu adabiyotlarda chuqur o‘rganish (deep learning) modellari, semantik vektorlashtirish va til modellarining rivoji kompyuter lingvistikasining yangi bosqichini belgilab berganligi ta’kidlanadi. Bu yondashuvlar o‘zbek tiliga ham tatbiq etila boshlagan bo‘lib, adaptatsiya jarayoni hali ham davom etmoqda. Umuman olganda, adabiyotlar tahlili shuni ko‘rsatadiki, O‘zbekistonda kompyuter lingvistikasi dastlab nazariy tilshunoslik doirasida shakllanib, keyinchalik amaliy axborot texnologiyalari bilan integratsiyalashgan holda rivojlangan. Bugungi kunda esa bu yo‘nalish sun’iy intellekt, ma’lumotlar tahlili va raqamli til resurslari bilan uzviy bog‘liq holda yangi bosqichga chiqmoqda. Shu bilan birga, o‘zbek tilining to‘liq raqamli infratuzilmasini yaratish, korpuslarni kengaytirish va avtomatik tizimlarni takomillashtirish dolzarb ilmiy muammo sifatida qolmoqda.

**Tadqiqot muhokamasi.** O‘zbekistonda kompyuter lingvistikasi yo‘nalishining shakllanishi va rivojlanishi tilshunoslik va axborot texnologiyalari integratsiyasining natijasi sifatida qaraladi. Tadqiqot natijalari shuni ko‘rsatadiki, ushbu yo‘nalish dastlab nazariy

tilshunoslikning raqamli muhitga moslashuvi sifatida yuzaga kelgan bo‘lsa, keyinchalik u mustaqil ilmiy-amaliy soha darajasiga ko‘tarilgan. Muhokama jarayonida o‘zbek tilining strukturaviy xususiyatlari, texnologik imkoniyatlar va ilmiy yondashuvlar o‘rtasidagi o‘zaro bog‘liqlik asosiy omil sifatida namoyon bo‘ladi. Birinchi navbatda, o‘zbek tilining agglutinativ tabiati kompyuter lingvistikasini rivojlantirishda muhim metodologik masalalarni yuzaga keltiradi. So‘z shakllarining ko‘p qatlamli affiksal tuzilishi avtomatik tahlil jarayonini murakkablashtiradi. Tadqiqotlarda bu holat morfologik analizatorlar yaratishda maxsus algoritmik yondashuvlarni talab qilishi bilan izohlanadi. Shu sababli, o‘zbek tili uchun ishlab chiqilgan dastlabki tizimlar asosan qoidaviy modelga asoslangan bo‘lib, ular til birliklarini segmentatsiya qilish va grammatik kategoriyalarni aniqlashga qaratilgan.

Ikkinchi muhim jihat sifatida korpus lingvistikasi masalasi ko‘rib chiqiladi. Tadqiqot natijalari shuni ko‘rsatadiki, o‘zbek tilining raqamli korpuslari shakllanishi kompyuter lingvistikasining eng muhim yutuqlaridan biridir. Korpuslar til birliklarini real kontekstda o‘rganish, chastota tahlili o‘tkazish va semantik munosabatlarni aniqlash imkonini beradi. Shu bilan birga, korpuslarning hajmi va sifati lingvistik modellar aniqligiga bevosita ta’sir ko‘rsatadi. O‘zbek tilidagi korpuslar hali ham kengaytirish va standartlashtirish bosqichida bo‘lib, bu jarayon tadqiqotlarning dolzarbligini oshiradi. Uchinchi jihat sifatida mashina tarjimasi va tabiiy tilni qayta ishlash tizimlarining rivoji tahlil qilinadi. Tadqiqotlar shuni ko‘rsatadiki, dastlabki bosqichlarda ushbu tizimlar asosan lug‘aviy va grammatik moslikka asoslangan bo‘lib, ular murakkab semantik munosabatlarni to‘liq qamrab olmagan. Keyingi yillarda esa statistik va neyron tarmoqlarga asoslangan yondashuvlar joriy etilishi natijasida tarjima sifati sezilarli darajada yaxshilangan. Biroq o‘zbek tili uchun universal va mukammal ishlaydigan model yaratish hali ham ochiq ilmiy muammo bo‘lib qolmoqda.

Tadqiqot muhokamasida shuningdek, lingvistik resurslarning yetishmasligi masalasi ham alohida o‘rin egallaydi. O‘zbek tilining raqamli lug‘atlari, annotatsiyalangan matnlar bazasi va semantik tarmoqlari hali to‘liq shakllanmagan. Bu holat sun’iy intellekt asosidagi tizimlarni o‘qitishda cheklovlar yaratadi. Shunga qaramasdan, so‘nggi yillarda ilmiy jamoalar tomonidan bu yo‘nalishda sezilarli ishlar amalga oshirilmoqda va yangi ma’lumotlar bazalari yaratilmoqda. Yana bir muhim masala sifatida kompyuter lingvistikasining amaliy qo‘llanilishi tahlil qilinadi. Tadqiqot natijalariga ko‘ra, ushbu yo‘nalish ta’lim tizimida, elektron ta’lim platformalarida, avtomatik imlo tekshirgichlarda va nutqni aniqlash tizimlarida keng qo‘llanilmoqda. Bu esa tilning raqamli muhitdagi funksional imkoniyatlarini kengaytiradi. Shu bilan birga, amaliy tizimlarning samaradorligi bevosita lingvistik modellar aniqligi va algoritmlarning optimallashtirilganligiga bog‘liq.

Muhokama jarayonida ilmiy yondashuvlar evolyutsiyasi ham e’tiborga olinadi. Dastlabki tadqiqotlarda qoidaviy yondashuv ustun bo‘lgan bo‘lsa, hozirgi bosqichda ma’lumotlarga asoslangan, statistik va mashinaviy o‘rganish metodlari yetakchilik qilmoqda. Bu o‘zgarish kompyuter lingvistikasining nazariy asoslarini ham sezilarli darajada kengaytirdi. Endilikda til faqat qoidalar majmuasi sifatida emas, balki katta hajmdagi ma’lumotlar tizimi sifatida ham qaralmoqda. Shuningdek, o‘zbek tilining raqamli transformatsiyasi jarayonida xalqaro tajriba bilan integratsiya masalasi ham muhim o‘rin tutadi. Xorijiy modellarni mahalliy til xususiyatlariga moslashtirish jarayoni bir qator metodologik muammolarni keltirib chiqaradi. Ayniqsa, morfologik boylik va so‘z tartibining nisbatan erkinligi kabi xususiyatlar universal modellarni to‘g‘ridan-to‘g‘ri qo‘llashni cheklaydi. Shu sababli adaptatsiya va lokalizatsiya jarayonlari alohida ilmiy yondashuvni talab qiladi. Umuman olganda, tadqiqot muhokamasi natijalari shuni ko‘rsatadiki, O‘zbekistonda kompyuter lingvistikasi jadal rivojlanayotgan, ammo hali to‘liq shakllanmagan kompleks ilmiy yo‘nalishdir. Uning kelajakdagi rivoji lingvistik resurslarning boyitilishi, zamonaviy algoritmlarning joriy etilishi va fanlararo hamkorlikning kuchayishiga bevosita bog‘liq. Bu esa o‘z navbatida o‘zbek tilining raqamli ekotizimda to‘liq va samarali ishlatilishini ta’minlashga xizmat qiladi.

**Xulosa.** O‘zbekistonda kompyuter lingvistikasi yo‘nalishining shakllanishi va rivojlanishi tilshunoslikning zamonaviy texnologiyalar bilan integratsiyalashuvi natijasi sifatida yuzaga kelgan murakkab va ko‘p bosqichli jarayon hisoblanadi. Tadqiqot natijalari shuni ko‘rsatadiki, ushbu soha dastlab tilni formal modellashtirish va elektron resurslar yaratish bilan bog‘liq bo‘lgan bo‘lsa, keyinchalik korpus lingvistikasi, mashina tarjimasi va tabiiy tilni qayta ishlash kabi yo‘nalishlar hisobiga sezilarli darajada kengaydi. O‘zbek tilining agglutinativ xususiyati kompyuter lingvistikasida alohida metodologik yondashuvlarni talab etadi va bu holat ilmiy izlanishlarning yo‘nalishini belgilab beradi. Shuningdek, mamlakatda raqamli til resurslarini yaratish, lingvistik korpuslarni kengaytirish va sun’iy intellekt asosidagi tizimlarni rivojlantirish dolzarb masala bo‘lib qolmoqda. Kompyuter lingvistikasining amaliy ahamiyati ta’lim, tarjima, nutqni aniqlash va avtomatik matn tahriri tizimlarida yaqqol namoyon bo‘lmoqda. Umuman olganda, ushbu yo‘nalishning istiqbollari lingvistik resurslarning boyitilishi va zamonaviy algoritmlarning joriy etilishiga bevosita bog‘liqdir.

## Foydalanilgan adabiyotlar

1. Jurafsky, D., & Martin, J. H. (2023). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. Pearson.
2. Manning, C. D., & Schütze, H. (2021). Foundations of statistical natural language processing. MIT Press.
3. Bird, S., Klein, E., & Loper, E. (2020). Natural language processing with Python. O’Reilly Media.
4. Tursunov, A. (2022). O‘zbek tilini kompyuter lingvistikasida modellashtirish muammolari. Filologiya masalalari.
5. Karimov, B. (2021). O‘zbek tilining korpus lingvistikasi asoslari. O‘zbek tilshunosligi tadqiqotlari.
6. Yo‘ldoshev, M. (2023). Raqamli til resurslarini yaratish va ularning ahamiyati. Zamonaviy tilshunoslik va axborot texnologiyalari.
7. Habibullayev, F. (2020). O‘zbek tilida avtomatik tarjima tizimlarini rivojlantirish. Ilmiy axborotnoma.
8. Jurayev, S. (2022). Kompyuter lingvistikasining nazariy asoslari. Til va texnologiya.
9. Xolmatov, I. (2021). Tabiiy tilni qayta ishlash va sun’iy intellekt. Axborot texnologiyalari va tilshunoslik.
10. Rustamov, K. (2023). O‘zbek tilining raqamli transformatsiyasi. Ilm-fan va innovatsiya jurnali.