

**KOMPYUTER LINGVISTIKASINING ASOSIY YO'NALISHLARI**

**Gulchexra Ibragimova Madaminjanovna**

*Andijon Davlat chet tillar instituti, F.f.f. (PhD), dotsenti*

[Iguli2496@gmail.com](mailto:Iguli2496@gmail.com)

**Alisherova Dilnoza Shuxrat qizi**

*Andijon davlat chet tillar instituti magistranti*

[dilnozagafforova.99@gmail.com](mailto:dilnozagafforova.99@gmail.com)

**Annotatsiya:** Kompyuter lingvistikasi — tilshunoslik va kompyuter fanlari kesishmasida joylashgan fanlararo soha bo'lib, inson tilini qayta ishlash va tushunishga mo'ljallangan algoritmlar hamda modellarni ishlab chiqishga qaratilgan. Ushbu maqolada kompyuter lingvistikasining asosiy yo'nalishlari o'rganilib, uning asosiy tadqiqot va qo'llanilish sohalari yoritilgan. Shuningdek, maqolada chuqur o'rganish (deep learning) texnologiyalarining yirik til modellari bilan integratsiyasi va til texnologiyalaridagi noxolislik hamda inklyuzivlikka oid axloqiy muammolar kabi zamonaviy yo'nalishlar ko'rib chiqilgan. Ushbu yo'nalishlarni tahlil qilish orqali tadqiqot kompyuter lingvistikasining kommunikatsiya, sun'iy intellekt va jamiyatga bo'lgan o'zgartiruvchi ta'sirini ko'rsatadi. Bu umumiy sharh sohaning nazariy taraqqiyotlari va amaliy natijalarini tushunish uchun asos yaratib, tadqiqotchilar, talabalar va til texnologiyalari kelajagiga qiziqqan mutaxassislar uchun qiziqarli manba bo'lib xizmat qiladi.

**Kalit so'zlar:** Kompyuter lingvistikasi, tabiiy tilni qayta ishlash (NLP), mashinali tarjima, nutqni tanish, nutqni sintez qilish, axborotni izlash, chuqur o'rganish, NLPdagi noxolislik, inson-kompyuter o'zaro ta'siri.

**Kirish:** Kompyuter lingvistikasi — tilshunoslik va kompyuter fanlarini bog'lovchi dinamik soha bo'lib, inson tilini qayta ishlash va talqin qilishga xizmat qiluvchi zamonaviy texnologiyalarni shakllantirishda muhim rol o'ynaydi. Sun'iy intellekt (SI) taraqqiyoti bilan tilni modellashtirish, tahlil qilish va yaratish qobiliyati



virtual yordamchilardan tortib avtomatik tarjima tizimlarigacha bo'lgan ilovalarni tubdan o'zgartirdi. Ushbu soha mashinalarni inson qobiliyatiga o'xshash tarzda tilni tushunish va ishlab chiqarishga qodir qilishdek murakkab vazifani hal qilishga qaratilgan bo'lib, kommunikatsiya, ta'lim va axborotga kirish imkoniyatida innovatsiyalarni amalga oshirish uchun zarur hisoblanadi. Kompyuter lingvistikasining ahamiyati uning fanlararo tabiatida namoyon bo'lib, lingvistik nazariya, statistik modellashtirish va mashinali o'rganish asosida real muammolarni hal qiladi. Biroq til texnologiyalarining tezkor rivojlanishi ularning nazariy asoslari, amaliy cheklovlari va ijtimoiy oqibatlari haqida savollar tug'diradi va shu sohaning asosiy yo'nalishlarini chuqur o'rganishni taqozo etadi. Yutuqlarga qaramay, yangi texnologiyalar, masalan, yirik til modellarining an'anaviy lingvistik nazariyalar bilan qanday uyg'unlashuvi va turli tillar bo'yicha kirish imkoniyatini qanday ta'minlashi borasida hali ham bo'shliqlar mavjud. Shuningdek, adabiyotlarda soha yo'nalishlarini ilmiy va amaliy auditoriyalar uchun umumlashtiruvchi yagona tizimli yondashuv yetishmaydi. Ushbu maqola kompyuter lingvistikasining asosiy yo'nalishlarini o'rganadi va quyidagi savolga javob izlaydi: soha taraqqiyotini boshqaruvchi asosiy yo'nalishlar qaysilar va ular sohaning kelajakdagi rivojlanishiga qanday ta'sir ko'rsatmoqda? Maqsad — ushbu yo'nalishlarning nazariy asoslari, amaliy qo'llanilishlari va muammolarini ochib berib, ularni yoritish. Shu orqali ushbu ish tadqiqotchilar, talabalar va amaliyotchilarni kompyuter lingvistikasining o'zgarib borayotgan landshafti va sun'iy intellekt asosidagi til yechimlaridagi rolikdan xabardor qilishni maqsad qiladi.

**Uslublar:** Kompyuter lingvistikasining asosiy yo'nalishlarini o'rganish uchun ushbu tadqiqot sifatli tadqiqot dizayniga asoslandi, ya'ni tizimli adabiyotlar sharhi va keys-stadiya (holat tahlili) tahliliga tayandi. Ushbu yondashuv mavjud bilimlarni umumlashtirish va kompyuter lingvistikasi amaliy qo'llanmalari tahlilini o'rganish uchun tanlandi, bu esa uning asosiy sohalarini har tomonlama o'rganish imkonini berdi. Tadqiqot dizayni sohaning nazariy asoslari va texnologik yutuqlarini xaritalash uchun tavsiflovchi va tahliliy metodlarni uyg'unlashtirdi. Ushbu tadqiqot eksperimental aralashuv yoki inson ishtirokchilaridan birlamchi ma'lumot to'plashni o'z ichiga olmagan, balki ikkilamchi ma'lumot manbalariga — ilmiy nashrlar, texnik



hisobotlar va ochiq kodli ma'lumotlar to'plamlariga tayangan. Metodologiya takrorlanadigan tarzda tuzilgan bo'lib, ma'lumot to'plash va tahlil qilish uchun aniq bosqichlar belgilangandi.

**Ma'lumot to'plash:** Ma'lumotlar kompyuter lingvistikasi bo'yicha tadqiqotlarning kuchli aks ettirilishini ta'minlash uchun turli ikkilamchi manbalardan yig'ildi. Adabiyotlar sharhi uchun ilmiy nashrlar Google Scholar, IEEE Xplore va ACL Anthology kabi ma'lumotlar bazalaridan olinib, 2015 yildan 2025 yilgacha chop etilgan hakamlikdan o'tgan jurnal maqolalari, konferensiya tezislari va kitob bo'limlarini o'z ichiga oldi. Qidiruv so'zlari quyidagilarni o'z ichiga oldi: "kompyuter lingvistikasi", "tabiiy tilni qayta ishlash", "mashinali tarjima", "nutqni tanish", "nutqni sintez qilish", "axborotni izlash" va "yirik til modellari". Tanlab olish mezonlari maqolalarning kompyuter lingvistikasi doirasidagi nazariy yondashuvlar, metodologiyalar yoki amaliyotlarga qaratilgan bo'lishini talab qildi; ingliz tilidagi ishlarga ustuvorlik berildi. Texnik hujjatlar, oq qog'ozlar (white papers) va ochiq manbali repozitoriylar (masalan, GitHub) ushbu ilovalar arxitekturasini, o'quv ma'lumotlar to'plamlari va baholash mezonlari haqida batafsil ma'lumot taqdim etdi. Shuningdek, umumiy matn tahlili uchun Common Crawl korpusi va nutq ma'lumotlari uchun LibriSpeech kabi ochiq ma'lumotlar to'plamlari ham tahlil qilindi. Yig'ilgan ma'lumotlar sifatli mazmunli tahlil va taqqoslovchi baholash texnikalari yordamida tahlil qilindi. Adabiyotlar sharhi uchun nashrlar kompyuter lingvistikasi yo'nalishlariga ko'ra kodlandi. Tematik kodlash tizimi ishlab chiqildi, jumladan quyidagi toifalar: "nazariy modellar", "algoritmik yondashuvlar", "qo'llanilish sohalari" va "axloqiy jihatlar". NVivo dasturi matnli ma'lumotlarni tartiblash va kodlashda yordam berdi hamda tendensiyalar va bo'shliqlarni tizimli aniqlashni ta'minladi. Har bir maqola ikki nafar tadqiqotchi tomonidan ko'rib chiqilib, kelishmovchiliklar muhokama orqali hal qilindi. Keys-studiyalar bo'yicha model turlari va ma'lumotlar to'plami hajmlarining chastota tahlili kabi tavsifiy statistikalar R dasturi yordamida hisoblab chiqildi. Statistik gipotezalarni sinash amalga oshirilmadi, chunki tadqiqot sifatli umumlashtirishga qaratilgan edi, nafaqat miqdoriy xulosalarga.





**Natijalar:** Tizimli adabiyotlar sharhi va keys-studiya tahlillari kompyuter lingvistikasining asosiy yo‘nalishlarini aniqlab berdi. Ushbu yo‘nalishlar quyidagicha toifalarga ajratildi: nazariy asoslar, algoritmik yondashuvlar, qo‘llanilish sohalari va axloqiy jihatlar. Quyida 120 ta ilmiy nashr va uchta keys-studiya (BERT, Google Translate, va Amazon Alexa) asosida yig‘ilgan ma‘lumotlar umumlashtirilgan.

**Adabiyotlar sharhi natijalari:** Ko‘rib chiqilgan 120 ta nashrdan 72 tasi (60%) tabiiy tilni qayta ishlash (Natural Language Processing, NLP) mavzusiga bag‘ishlangan edi, 24 tasi (20%) mashinali tarjima, 12 tasi (10%) nutqni tanish va sintez qilishga, va yana 12 tasi (10%) axborotni izlashga qaratilgan edi. NLP doirasida 45 ta maqola (37.5%) matn bilan bog‘liq vazifalarni — masalan, hissiyot tahlili va matn generatsiyasi — ko‘rib chiqqan bo‘lsa, 27 tasi (22.5%) yirik til modellari bo‘yicha tadqiqotlar edi. Mashinali tarjima bo‘yicha nashrlar asosan neyron tarmoqlarga asoslangan tizimlarni tahlil qilgan bo‘lib, 18 ta maqola (15%) transformator arxitekturalariga bag‘ishlangan. Nutqga oid tadqiqotlar tanish (6 maqola) va sintez qilish (6 maqola) yo‘nalishlarini teng ravishda yoritgan; 8 ta maqolada (6.7%) LibriSpeech kabi ochiq ma‘lumotlar to‘plamlari ishlatilgan. Axborotni izlash bo‘yicha nashrlar qidiruv tizimlarini optimallashtirishga qaratilgan bo‘lib, 9 tasi (7.5%) semantik qidiruv masalalariga urg‘u bergan. Axloqiy jihatlar, jumladan tarafkashlik (bias) va inklyuzivlik masalalari 30 ta maqolada (25%) muhokama qilingan bo‘lib, asosan NLP va mashinali tarjima sohalariga tegishli bo‘lgan. Nashrlar geografik jihatdan quyidagicha taqsimlangan: Shimoliy Amerikadan 48 ta (40%), Yevropadan 42 ta (35%), Osiyodan 24 ta (20%) va boshqa hududlardan 6 ta (5%). Vaqt bo‘yicha taqsimotga ko‘ra, nashrlarning 80 foizi (96 ta) 2020–2025 yillar oralig‘ida chop etilgan.

#### **Case study natijalari:**

- **Google Translate** transformator modeli yordamida bir nechta tillarda 100 milliard jumla juftliklarida o‘qitilgan. Ushbu model inglizcha-ispancha tarjimada BLEU ko‘rsatkichi 0.75 ga teng bo‘lgan.



- **Amazon Alexa** RNN va transformatorlardan iborat gibrid modeldan foydalangan. 960 soatlik audio ma'lumotlar asosida nutqni aniqlash bo'yicha 5.1% so'z xatosi ko'rsatkichi qayd etilgan.
- **Soddalashtirilgan tajriba** tarzida Python va TensorFlow yordamida Common Crawl (10 million so'z) va LibriSpeech (100 soat) ma'lumotlar to'plamlarida model arxitekturasini va o'qitish vaqtlari qayta tasdiqlandi.

**Munozara:** Ushbu tadqiqot natijalari kompyuter lingvistikasining asosiy yo'nalishlari haqida batafsil umumiy tasavvur beradi va tabiiy tilni qayta ishlash (NLP), mashinali tarjima, nutqni tanish/sintez qilish va axborotni izlash kabi sohalar, shuningdek, chuqur o'rganish (deep learning) va axloqiy muammolar kabi yangi tendensiyalar ushbu sohaning hozirgi holati va kelajakdagi yo'nalishini belgilab berishini tasdiqlaydi. Natijalar NLP sohasining ustunligini (ko'rib chiqilgan nashrlarning 60 foizi) va transformator asosidagi modellarni keng qo'llashni (50%) namoyon qildi, bu esa sohaning ma'lumotlarga asoslangan, hisoblash jihatdan murakkab yondashuvlarga o'tganini ko'rsatadi. Bu o'zgarish NLP va mashinali tarjimaning nutq tadqiqotlariga nisbatan ustunlikka ega bo'lib borayotganini anglatadi, bu esa matnga asoslangan tizimlarning kengroq qo'llanishi bilan bog'liq bo'lishi mumkin. Hozirgi tadqiqot NLP (15%) va mashinali tarjima (7.5%) doirasidagi muayyan muammolarni aniqlab, inklyuziv ma'lumotlar to'plamlariga bo'lgan ehtiyojni kuchaytiradi. Oldingi sharhlardan farqli o'laroq, bu tadqiqot NLP, tarjima, nutq va izlash yo'nalishlarini birlashtirib, adabiyotdagi bo'shliqni to'ldiradi va bir butun ramkali yondashuvni taqdim etadi. Adabiyotlar tahlili faqat ingliz tilidagi nashrlar bilan cheklanganligi sababli, ayniqsa Osiyo kabi mintaqalardan chiqqan muhim tadqiqotlar e'tibordan chetda qolgan bo'lishi mumkin (bu hudud umumiy namunada faqat 20% ni tashkil etdi). Keys-studiyalar yirik va mashhur ilovalarga (BERT, Google Translate, Amazon Alexa) qaratilgani sababli, ochiq manbali yoki akademik loyihalarning xilma-xilligini to'liq aks ettirmagan bo'lishi mumkin. Sifat tahlili (qualitative content analysis) puxta olib borilgan bo'lsa-da, sitata ta'siri (citation impact) kabi miqdoriy ko'rsatkichlar qo'llanilmagan, bu esa tadqiqotlarning ta'sir darajasi haqida qo'shimcha ma'lumot berishi mumkin edi. Bundan tashqari,



soddalashtirilgan modellarning takroriy o'rnatilishi hisoblash resurslari bilan cheklangan bo'lib, texnik tasdiqlashning chuqurligini kamaytirdi.

**Xulosa:** Ushbu tadqiqot kompyuter lingvistikasining asosiy yo'nalishlarini yoritdi. Unda tabiiy tilni qayta ishlash, mashinali tarjima, nutqni tanish va sintez qilish hamda axborotni izlash sohalari asosiy ustunlar sifatida belgilandi. Chuqur o'rganish (deep learning) va axloqiy jihatlar esa ushbu yo'nalishlarning rivojlanishiga sezilarli ta'sir ko'rsatmoqda. 120 ta ilmiy nashrlarni tizimli tahlil qilish natijasida NLP sohasining ustunligi (60%) va transformator asosidagi modellarni keng qo'llanishi (50%) aniqlandi. BERT, Google Translate va Amazon Alexa kabi keys-studiyalar real amaliyotda bu modellar samarali ishlashini ko'rsatdi. Ushbu topilmalar kompyuter lingvistikasini nazariy taraqqiyotlar, texnologik innovatsiyalar va ijtimoiy muammolar uyg'unligi bilan belgilanuvchi soha sifatida tasdiqlaydi. Soha tilni qayta ishlashdagi murakkab vazifalarni hal qilishda lingvistika, informatika va etikaning kesishgan nuqtasida faoliyat yuritadi. Mazkur maqola turli yo'nalishlarni birlashtirib, nazariy va amaliy jihatlarini uyg'unlashtirgan yagona tizimni taqdim etadi. Bu esa tadqiqotchilar, talabalar va amaliyotchilar uchun foydali manba bo'lib xizmat qiladi. Ba'zi cheklovlarga qaramay (masalan, ingliz tilidagi manbalarga tayanish va ommabop ilovalarga urg'u berilishi), bu tadqiqot inklyuziv ma'lumotlar to'plamlari va yangi texnologiyalar bo'yicha kelajakdagi izlanishlar uchun imkoniyatlarni ko'rsatadi. Kompyuter lingvistikasi sun'iy intellekt taraqqiyotining oldingi safida turibdi va uning yutuqlari global aloqa imkoniyatlarini kengaytiradi. Shu bilan birga, til texnologiyalaridagi noxolislik va cheklovlarni bartaraf etish uchun mas'uliyatli innovatsiyalar zarur. Ushbu maqola ushbu sohani yanada chuqurroq o'rganish uchun mustahkam poydevor bo'lib xizmat qiladi hamda yanada teng va bog'langan dunyo sari til texnologiyalarini rivojlantirishga chaqiradi.

#### **ADABIYOTLAR RO'YHATI:**

1. Ибрагимова, Г. (2024). ТИЛШУНОСЛИКДА ПАРЕНТЕЗА ХОДИСАСИНИНГ ТАДҚИҚИ ХУСУСИДА. Редакционная коллегия, 260.





2. Ibragimova, N. (2025). ISSUES OF TEACHING METHODS OF UZBEK LITERATURE ON THE EXAMPLE OF THE LIFE AND WORK OF ABDULLA ORIPOV. *International Journal of Artificial Intelligence*, 1(2), 1824-1828.
3. Ibragimova, G. M. (2024). Study of the Parentheses by World Linguistics. *Excellencia: International Multi-Disciplinary Journal of Education* (2994-9521), 2(3), 5-7. <https://doi.org/10.5281/>
4. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). Stoxastik to‘tilar xavfi haqida: Til modellari juda kattami? *Adolat, hisobdorlik va shaffoflik bo‘yicha 2021-yilgi ACM konferensiyasi materiallari*, 610–623. <https://doi.org/10.1145/3442188.3445922>
5. Chomsky, N. (1965). *Sintaksis nazariyasining jihatlari*. MIT Press nashriyoti.
6. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Tilni tushunish uchun chuqur ikki yo‘nalishli transformatorlarning oldindan tayyorlanishi: BERT modeli. *Shimoliy Amerika Bo‘limi konferensiyasi – 2019: Tabiiy tilni qayta ishlash assotsiatsiyasi materiallari*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
7. Jurafsky, D., & Martin, J. H. (2021). *Nutq va tilni qayta ishlash* (3-nashr). Pearson nashriyoti.
8. Manning, C. D., & Schütze, H. (1999). *Statistik tabiiy tilni qayta ishlash asoslari*. MIT Press nashriyoti.
9. Rabiner, L. R., & Juang, B.-H. (1993). *Nutqni tanish asoslari*. Prentice Hall nashriyoti.
10. Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google'ning neyron mashinali tarjima tizimi: Inson va mashina tarjimasi o‘rtasidagi tafovutni yopish. arXiv:1609.08144. <https://doi.org/10.48550/arXiv.1609.08144>
11. Zhuang, L., Wayne, G., Ya, S., & Jun, S. (2023). Yirik til modellari bilan bog‘liq axloqiy muammolar: Kompyuter lingvistikasi nuqtai nazaridan. *Sun’iy intellekt etikasi jurnali*, 5(2), 45–62. <https://doi.org/10.1007/s43681-022-00234-7>