

LINGVISTIKANING ISTIQBOLLARI: RAQAMLI TEXNOLOGIYALAR VA SUN'IY INTELLEKT DAVRIDA TILSHUNOSLIK YO'NALISHLARI

O'ralov Abbosbek Sayibnazar o'g'li

TDYU qoshidagi Namangan akademik litseyi o'qituvchisi

Email: sardorclash.197@gmail.com

<https://doi.org/10.5281/zenodo.22827371>

ANNOTATSIYA: Mazkur maqolada zamonaviy lingvistikaning rivojlanish tendensiyalari va istiqbolli yo'nalishlari tahlil qilinadi. Kompyuter lingvistikasi, korpus lingvistikasi, kognitiv lingvistika, psixolingvistika, neyrolingvistika va sotsiolingvistika kabi yo'nalishlarning rivojlanish imkoniyatlari ko'rib chiqiladi hamda bu jarayon o'zbek tilshunosligi misolida - xususan, o'zbek tili milliy korpusini yaratish tajribasi orqali - amaliy tarzda ko'rsatiladi. Maqolada, shuningdek, yetakchi sun'iy intellekt til modellarining o'zbek tilini qayta ishlashdagi joriy cheklavlari ham ko'rsatib o'tiladi.

Kalit so'zlar: PISA, lingvistika, kompyuter lingvistikasi, korpus lingvistikasi, sun'iy intellekt, tabiiy tilni qayta ishlash, kognitiv lingvistika, neyrolingvistika, sotsiolingvistika, o'zbek tili milliy korpusi, raqamli texnologiyalar.

Kirish.

Til insoniyat tarixidagi eng muhim ijtimoiy hodisalardan biri bo'lib, xalqning tarixiy xotirasi, ma'naviy merosi va milliy o'zligini aks ettiruvchi murakkab tizimdir. Lingvistika tilning paydo bo'lishi, rivojlanishi, ichki tuzilishi va jamiyatdagi o'rnini ilmiy jihatdan tadqiq etadi. Dastlab fonetika, leksikologiya, morfologiya va sintaksis doirasida rivojlangan bo'lsa, bugungi kunda uning tadqiqot doirasi sezilarli darajada kengaygan [3]. XXI asr raqamli texnologiyalarning tilshunoslik oldiga yangi vazifalar qo'yishi bilan xarakterlanadi. Elektron matnlar, ijtimoiy tarmoqlardagi yozishmalar hamda audio-video materiallar lingvistik tadqiqotlar uchun katta ma'lumotlar manbaiga aylanmoqda, sun'iy intellektning rivojlanishi esa inson tilini kompyuter orqali tushunish va qayta ishlash imkoniyatlarini kengaytirmoqda [4].

Ushbu maqolaning maqsadi zamonaviy lingvistikaning asosiy istiqbolli yo'nalishlarini tizimli tahlil qilish va ularni o'zbek tilshunosligining amaliy tajribasi - xususan, milliy korpus yaratish jarayoni - misolida ko'rsatishdan iborat. Bu yondashuv maqolani sof tavsifiy sharhdan

farqlab, aniq holat tahliliga (*case study*) asoslangan xulosalar chiqarish imkonini beradi.

Kompyuter lingvistikasi va tabiiy tilni qayta ishlash. Sun'iy intellekt tizimlari katta hajmdagi matnlarni tahlil qilib, so'zlar o'rtasidagi munosabatlarni aniqlashi, matn yaratishi va tarjima qilishi mumkin. Tabiiy tilni qayta ishlash texnologiyalarining rivojlanishi natijasida kompyuterlar inson tilining nafaqat grammatik tuzilishini, balki mazmuni va kontekstini ham tushunishga tobora yaqinlashmoqda [4].

Shu bilan birga, sun'iy intellekt oldida hal qilinishi kerak bo'lgan muammolar ham mavjud: mashinalar kinoya, metafora, yashirin ma'no va dialekt kabi pragmatik xususiyatlarni har doim ham to'g'ri anglay olmaydi. Bu muammo kam resursli tillar - jumladan, o'zbek tili - uchun ayniqsa dolzarb. 2025-yil avgust oyida e'lon qilingan milliy sun'iy intellekt til modeli loyihasi doirasida ta'kidlanishicha, hozirgi yetakchi til modellari (ChatGPT, Claude, Gemini) o'zbek tilini atigi 60–70 foiz aniqlik bilan qayta ishlay oladi [14].

Korpus lingvistikasi. Korpus lingvistikasi katta hajmdagi real til materiallarini elektron shaklda to'plash va ularni statistik hamda lingvistik jihatdan tahlil qilishga asoslanadi. Korpus yordamida ma'lum bir so'zning qanday kontekstlarda ishlatilishi, uning ma'no variantlari va grammatik shakllari aniqlanishi mumkin [4].

O'zbek tili milliy korpusi: amaliy holat tahlili. O'zbek tilshunosligida korpus lingvistikasining rivojlanishi nazariyadan amaliyotga o'tgan aniq misol sifatida ko'rsatilishi mumkin. "2020–2030-yillarda o'zbek tilini rivojlantirish va til siyosatini takomillashtirish konsepsiyasi"da davlat tilining zamonaviy axborot texnologiyalariga integratsiyalashuvi va o'zbek tili milliy korpusini yaratish ustuvor vazifa sifatida belgilangan [10].

Ushbu yo'nalishdagi dastlabki ilmiy asos N.Z. Abdurahmonovanning tadqiqotlarida qo'yilgan bo'lib, muallif o'zbek tili elektron korpusini yaratish g'oyasini ilgari surib, uning konseptual asoslarini ishlab chiqqan; natijalar 2021-yilda himoya qilingan doktorlik dissertatsiyasida yoritilgan [6].

Sh. Hamroyevaning mualliflik korpusini tuzish mezonlariga bag'ishlangan tadqiqoti bilan bir qatorda, keyingi bosqichda amaliy dasturiy ishlar boshlangan - natijada foydalanuvchilarga ochiq bo'lgan uzbekcorpora.uz portali ishga tushirilgan [11]. Yaratilgan korpus infratuzilmasi asosida bir qator amaliy tadqiqotlar olib borilmoqda: Yodgorov (2023) korpus matnlarini avtomatik qayta ishlash usullarini tavsiflagan [7].

Elov va Xudayberganov (2024) o'zbek tili korpusi matnlarini POS-teglash (so'z turkumlarini avtomatik belgilash) usullarini ishlab chiqqan [9].

Kasimova (2024) esa milliy korpus materiallari asosida o'zbek tilidagi darajalanish hodisasini tahlil qilgan [8].

Ushbu misol shuni ko'rsatadiki, lingvistik nazariya (korpus tuzish tamoyillari) bilan amaliy dasturiy ishlanma (veb-portal, teglash algoritmlari) o'rtasidagi bog'liqlik ta'minlangandagina yo'nalish haqiqatan istiqbolli bo'lib qoladi. Ayni paytda, bunday korpuslarning hajmi va janr xilma-xilligi xalqaro (masalan, Rossiya Milliy korpusi yoki Leipzig Corpora Collection) andozalar bilan solishtirganda hali

cheklangan bo'lib, bu - quyida ko'rib o'tiladigan - resurs yetishmovchiligi muammosining aniq ko'rinishidir.

Kognitiv va neyrolingvistika. Kognitiv lingvistika til va inson tafakkuri o'rtasidagi munosabatlarni o'rganadi; "Vatan", "oil", "erkinlik" kabi tushunchalar nafaqat umumiy ma'noga, balki har bir xalqning tarixiy-madaniy tajribasiga xos xususiyatlarga ham ega [5].

Neyrolingvistika esa tilning inson miyasida qanday shakllanishi va qayta ishlanishini o'rganadi; bu yo'nalishdagi tadqiqotlar nutq buzilishlarini aniqlash va reabilitatsiya usullarini takomillashtirishda muhim ahamiyat kasb etadi [4].

Sotsiolingvistika va ijtimoiy tarmoqlar tili. Til jamiyatdan tashqarida mavjud bo'la olmaydi - jamiyatdagi siyosiy, iqtisodiy va texnologik o'zgarishlar tilda ham aks etadi. Globallashuv natijasida ko'ptillilik va tillararo aloqa kuchaymoqda, yoshlar nutqida esa yangi so'zlar, qisqartmalar va internet leksikasi ko'paymoqda [3].

Qisqa yozishmalar, emoji, aralash tilli nutq va yangi internet iboralari zamonaviy sotsiolingvistika uchun muhim tadqiqot materialidir; bu jarayonlarni muntazam kuzatish alohida korpus va metodologiya talab qiladi.

Mashinaviy tarjima va terminologiya. Tillar o'rtasidagi aloqa hajmining ortishi tezkor tarjimaga bo'lgan ehtiyojni kuchaytirmoqda, biroq sifatli tarjima uchun faqat lug'aviy ekvivalentlarni topish yetarli emas - kontekst, uslub va frazeologik ma'no ham hisobga olinishi kerak [4].

O'zbek tilining mashinaviy tarjimadagi imkoniyatlarini oshirish uchun sifatli parallel korpuslar (o'zbek-ingliz, o'zbek-rus) yaratish zarur - bu vazifa yuqorida ko'rsatilgan milliy korpus loyihasining tabiiy davomi sifatida qaralishi mumkin. Shu bilan bir qatorda, informatika, biotexnologiya va sun'iy intellekt kabi sohalarda yangi terminlar soni ortib

bormoqda, bu esa yagona terminologik me'yorni shakllantirish zaruratini kuchaytiradi.

Amaliy yo'nalishlar: lingvistik ekspertiza va ta'lim.

Lingvistik ekspertiza huquq, OAV va reklama sohalarida til materiallarini tahlil qilishga xizmat qiladi; internetdagi axborot hajmining ortishi bilan matn muallifligini aniqlash va tahdid mazmunidagi ifodalarni aniqlash kabi vazifalarda sun'iy intellekt dan dastlabki tahlil vositasi sifatida foydalanish imkoniyatlari kengaymoqda. Ta'lim sohasida esa nutqni tahlil qiluvchi dasturlar va o'quvchining xatolariga moslashadigan individual ta'lim tizimlari istiqbolli hisoblanadi.

Muammoli jihatlar. Yuqorida ko'rib o'tilgan o'zbek tili milliy korpusi tajribasi zamonaviy lingvistikaning kamida uchta tizimli muammosini yaqqol namoyon etadi. Birinchidan, resurs yetishmovchiligi - mavjud korpus hajmi va janr xilma-xilligi xalqaro andozalardan orqada qolmoqda. Ikkinchidan, ma'lumot sifati muammosi - sun'iy intellekt tizimlari sifatsiz yoki noto'g'ri teglangan ma'lumotlardan foydalansa, natija ham kutilgan darajada bo'lmaydi; aynan shuning uchun POS-teglash kabi aniq metodologik ishlar (Elov, Xudayberganov, 2024) alohida qimmatga ega. Uchinchidan, tilning tez o'zgarishi - lug'atlar va me'yoriy manbalar internet leksikasidagi o'zgarishlarni tezkor aks ettira olmaydi, bu esa korpusni muntazam yangilab turishni talab qiladi.

Xulosa va takliflar. Lingvistikaning kelajakdagi rivojlanishi an'anaviy tadqiqot metodlari bilangina cheklanib qolmaydi: raqamli texnologiyalar, katta ma'lumotlar va sun'iy intellektning rivojlanishi fanning mazmuni va metodologiyasini tubdan boyitmoqda. O'zbek tili milliy korpusi misolida ko'rsatilganidek, bu jarayon nazariy tadqiqotlar (Abdurahmonova, Hamroyeva) bilan amaliy dasturiy ishlanmalar (Yodgorov; Elov, Xudayberganov) o'rtasidagi izchil hamkorlik natijasida haqiqiy samara beradi.

Shu asosda quyidagi aniq takliflarni ilgari surish mumkin:

Mavjud o'zbek tili milliy korpusining hajmini va janr tarkibini xalqaro andozalar (masalan, Rossiya Milliy korpusi yoki Leipzig Corpora Collection) darajasiga yaqinlashtirish;

O'zbek-ingliz va o'zbek-rus parallel korpuslarini maqsadli tarzda kengaytirish, chunki bu bevosita mashinaviy tarjima sifatiga ta'sir qiladi;

Yetakchi til modellarining o'zbek tilidagi 60–70 foizlik aniqlik ko'rsatkichini oshirish bo'yicha milliy loyihalarni ilmiy jamoatchilik ishtirokida kengaytirish;

POS-teglash va darajalanish kabi mikrostrukturaviy tadqiqotlarni (Elov, Xudayberganov, 2024; Kasimova, 2024) boshqa grammatik hodisalarga ham tatbiq etish.

Lingvistikaning istiqboli, xulosa qilib aytganda, uning texnologik taraqqiyotning faol ishtirokchisiga aylanish qobiliyatida - ya'ni nazariy bilim bilan amaliy raqamli ishlanmalarni birlashtira olish qobiliyatida namoyon bo'ladi.

Foydalanilgan adabiyotlar ro'yxati:

1. Saussure, F. de. Course in General Linguistics. - London: Peter Owen, 1960.
2. Chomsky, N. Syntactic Structures. - The Hague: Mouton, 1957.
3. Crystal, D. The Cambridge Encyclopedia of Language. - Cambridge: Cambridge University Press, 1997.
4. Jurafsky, D., Martin, J. H. Speech and Language Processing. - 3rd ed. draft. - Pearson, 2023.
5. Lakoff, G., Johnson, M. Metaphors We Live By. - Chicago: University of Chicago Press, 1980.
6. Abdurahmonova, N.Z. O'zbek tili elektron korpusini yaratishning lingvistik-nazariy asoslari: filol. fanlari dokt. ... diss. - Toshkent, 2021.
7. Yodgorov, U. S. Methods of processing the Uzbek language corpus texts // International Journal of Open Information Technologies. - 2023. - Vol. 11, No. 12. - P. 143–151.
8. Kasimova, M. B. O'zbek tili milliy korpusida darajalanish hodisasining ilmiy tahlili // ilmiy maqola. - 2024.
9. Elov, B., Xudayberganov, N. O'zbek tili korpusi matnlarini POS teglash usullari // Computer Linguistics: problems, solutions, prospects. - 2024. - Vol. 1, No. 1.
10. O'zbekiston Respublikasi Prezidentining "2020–2030 yillarda o'zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish konsepsiyasini tasdiqlash to'g'risida"gi Farmoni. - Toshkent, 2020.
11. O'zbek tili milliy korpusi: rasmiy portal. - URL: uzbekcorpora.uz.
12. Yusupov, O'. Tilshunoslikka kirish: o'quv qo'llanma. - Toshkent, 2019.
13. O'zbek tilining izohli lug'ati: 5 jildlik / A. Madvaliyev tahriri ostida. - Toshkent: O'zbekiston Milliy ensiklopediyasi, 2006–2008.
14. Milliy sun'iy intellekt til modeli loyihasi (2025-yil avgust) va o'zbek tilini qayta ishlash muammolari tahlili // Worldly Knowledge. Yil tadqiqotchisi - 2025. - URL: wordlyknowledge.uz.