

## TILSHUNOSLIK VA SUN'IY INTELEKT: RAQAMLI TRANSFORMASIYA VA YANGI ILMIY PARADIGMALAR

***G'aniyeva Lobar Sa'dulla qizi, f.f.b.f.d. (PhD)***

*Oriental universiteti katta o'qituvchisi*

0009-0000-9438-0603

[ganiyevalobar29@gmail.com](mailto:ganiyevalobar29@gmail.com)

***Iskandarova Muslima Baxodir qizi***

*Oriental universiteti talabasi*

**Annotatsiya:** Ushbu maqolada tilshunoslik fani va sun'iy intellekt (SI) texnologiyalarining o'zaro integratsiyasi, Katta til modellari (LLM) davrida tilshunoslar oldida turgan yangi konseptual vazifalar va tadqiqot yo'nalishlari tahlil qilinadi. Tadqiqotda tilning matematik modellarini yaratishda lingvistik nazariyalarning o'rni, kognitiv tilshunoslikning SI arxitekturasidagi ahamiyati va o'zbek tilini raqamlashtirishdagi fundamental muammolar ko'rib chiqiladi. Maqola yakunida zamonaviy tilshunos mutaxassislar uchun zarur bo'lgan yangi kompetensiyalar va istiqboldagi ilmiy yo'nalishlar bo'yicha xulosalar berilgan.

**Kalit so'zlar:** *Sun'iy intellekt, NLP (Tabiiy tilni qayta ishlash), katta til modellari (Large Language Models LLM), Kognitiv lingvistika, gallutsinatsiya kompyuter lingvistikasi, korpus lingvistikasi, transformator modellari, malumotlar generatsiyasi, ontologik baza.*

**Аннотация:** В данной статье анализируется интеграция лингвистической науки и технологий искусственного интеллекта (ИИ), а также новые концептуальные задачи и направления исследований, стоящие перед лингвистами в эпоху больших языковых моделей (LLM). В исследовании рассматриваются роль лингвистических теорий в создании математических моделей языка, значение когнитивной лингвистики в архитектуре ИИ и фундаментальные проблемы цифровизации узбекского языка. В завершение статьи представлены выводы о новых компетенциях, необходимых для современных специалистов-лингвистов, и перспективных направлениях научных исследований.

**Ключевые слова:** Искусственный интеллект, NLP (Обработка естественного языка), большие языковые модели (LLM), когнитивная лингвистика, галлюцинация, компьютерная лингвистика, корпусная лингвистика, модели-трансформеры, генерация данных, онтологическая база.

**Abstract:** This article analyzes the integration of linguistic science and artificial intelligence (AI) technologies, as well as the new conceptual tasks and research directions facing linguists in the era of Large Language Models (LLMs). The study examines the role of linguistic theories in developing mathematical language models, the importance of cognitive linguistics in AI architecture, and the fundamental challenges of digitalizing the Uzbek language. The article concludes with insights into the new competencies required for modern linguistic experts and future scientific research trajectories.

**Keywords:** Artificial Intelligence, NLP (Natural Language Processing), Large Language Models (LLM), Cognitive Linguistics, Hallucination, Computational Linguistics, Corpus Linguistics, Transformer Models, Data Generation, Ontological Base.

**KIRISH.** Raqamli texnologiyalar asrida tilshunoslik fani o‘zining klassik tavsiflash metodikasidan voz kechib, konstruktiv va bashorat qiluvchi xarakterga ega bo‘lgan yangi bosqichga o‘tdi. Sun‘iy intellekt, xususan, tabiiy tilni qayta ishlash (NLP - Natural Language Processing) sohasidagi rivojlanishlar tilga nafaqat ijtimoiy hodisa, balki yuqori darajada tartiblangan matematik struktura sifatida qarashni taqozo etmoqda [Jurafsky Martin, 2023:118]. Agar XX asr tilshunosligi tilning ichki qonuniyatlarini o‘rganish bilan cheklangan bo‘lsa, XXI asr tilshunosligi ushbu qonuniyatlarni algoritmlarga ko‘chirish va “mashina tafakkuri” ni shakllantirish vazifasini o‘z zimmasiga oldi. Bu jarayon tilshunoslar oldiga nafaqat nazariy, shuningdek, texnologik hamda til va tafakkur, til va jamiyat munosabatida yuz beradigan lingvistik hodisalarni sun‘iy intellektda tilning matematik modellarini yaratish masalasini ham qo‘ymoqda. Lingvistikaning eng muhim yo‘nalishlari – semantika, pragmatika, diskurs analizi – SI uchun hamon murakkab. “Ko‘p ma’nolilik”ni to‘g‘ri aniqlash, kinoyani tushunish, sarkazm yoki urf-odatga xos poetic ifodalarni yechish, har xil ma’no pardalari (denotativ, konnotativ, me’yoriy, turlicha registr va uslub) bilan ishlashda inson tafakkuridagi ijodiy yondashuv hal qiluvchi omil bo‘lib qolmoqda. Shu sababli kelajakda ayni shu murakkab sohalarda lingvistlarning ilmiy izlanishlari qat’iy talab etiladi [Mengliyev B., Hamroyeva Sh., 2025:442].

SI modellarining, xususan neyron tarmoqlarning ishlash prinsipi inson miyasining tilni o‘zlashtirish mexanizmlariga asoslangan. Biroq, zamonaviy **Large Language Models (LLM)** modellari tilni qanday “tushinishi” masalasi hali ham ilmiy bahslarga sabab bo‘lmoqda. Negaki turli nutqiy vaziyatlarda yuzaga keladigan jarayonlarni, xususan,

his-tuyg‘u, kinoya, ko‘p ma‘nololik, omonimiya hodisalarini SI tabiiy intellektdan farli ravishda to‘g‘ridan to‘g‘ri tanimaydi. Zamonaviy SI modellarining asosi bo‘lgan “Distributiv semantika” g‘oyasi (ya‘ni, so‘zning ma‘nosi uning atrofidagi so‘zlar bilan belgilanadi) bevosita Ferdinand de Sossyur va Zellig Harris kabi tilshunoslarning nazariyalariga borib taqaladi. Ferdinand de **Sossyur** tilni shaxmatga qiyoslaydi, ya‘ni shaxmatning o‘z strukturasi va sistemasi bo‘lgani kabi tilda ham bu o‘z aksini topgan. Til strukturasi o‘z sistemalariga ega. Til birliklarining vazifalari aniq, xuddi shahmat donalaridek, ammo ularni qaysi paytda qanday harakat qilishi aniq belgilanmagani va bu jarayonda ham inson tafakkuri ishlashini hisobga olsak, bu vaziyat tafakkurni to‘laligicha sun‘iylashtirish masalasini yuzaga keltiradi. Mashina so‘zlarni ko‘p o‘lchovli vektor maydonida joylashtirish orqali ular o‘rtasidagi semantik masofani hisoblaydi. Ammo bu jarayon har doim ham mukammal natija bermaydi. Tabiiy intellektga ehtiyoj qoldiradi.

Kognitiv tilshunoslikda til va tafakkur munosabati markaziy o‘rin tutadi. SI modellarida "Attention" (e‘tibor) mexanizmining joriy etilishi inson nutqni qabul qilayotganda kalit so‘zlarga urg‘u berish jarayonini matematik modellashtirish imkonini berdi (Vaswani,2017:48). Bu tilshunoslikdagi "aktual bo‘linish" nazariyasining algoritmik ko‘rinishidir. Bu kabi algoritmlarning ishlashi samarali deb qaralsada, nutqning aktual bo‘linishida ma‘nolar hamisha mutanosib kelmaydi. Bu kabi jarayonlar tilshunoslar oldiga bir qancha muammolar va vazifalarni belgilaydi. Texnologik taraqqiyot tilshunoslik sohasida yangi kasbiy ehtiyojlarni keltirib chiqardi. Bugungi kunda tilshunos mutaxassis, bir qancha funksional vazifalarni bajarishi shart. Jumladan:

**Sifatli lingvistik ma‘lumotlar generatsiyasini yaratish.** SI modellari “toza” va tizimlashtirilgan ma‘lumotlarga muhtoj. Tilshunoslar endilikda nafaqat matn yig‘uvchi, balki **data annotator** (ma‘lumotlarni lingvistik teglovchi) sifatida ish olib borishi natijasida matnlarni morfologik, sintaktik va semantik jihatdan to‘g‘ri markirovka qilish, ya‘ni algoritmlarning aniqligini belgilovchi asosiy omillarni belgilab olishdan iborat.

**Etik nazorat va Lingvistik filtrlash.** SI modellarida uchraydigan gender kamsitishlari, ijtimoiy stereotiplar va noto‘g‘ri ma‘lumotlar (gallutsinatsiyalar) bevosita til korpuslaridagi nuqsonlar natijasidir. Algoritmardagi bu kabi lingvistik xatolarni (bias) aniqlash va neytrallashtirish metodikasini ishlab chiqish juda zarur (Bender & Koller, 2020). Ayniqsa o‘zbek tilida bu ishlar jadallikda amalga oshirilishini talab qiladi. Negaki, bugungi kunga qadar foydalanilayotgan ma‘lumotlar algoritmlarida xatoliklar yo‘qori darajada.

O‘zbek tili agglütinativ qurilishga ega bo‘lgani sababli, uning uchun yaratiladigan SI modellarida o‘ziga xos murakkabliklar mavjud.

**1-jadval. Til strukturasiidagi asosiy muammolar va ularning yechimi.**

Muammo	Sabab	Yechim
<b>Morfologik tarkib</b>	Birgina o‘zakdan minglab so‘z shakllari hosil bo‘lishi.	Morfologik analizatorlar va stemmerlarni mukammallashtirish.
<b>Resurs tanqisligi</b>	Internet segmentida o‘zbek tilidagi ilmiy-falsafiy matnlarning kamligi.	Milliy korpusni (uzbekcorpus.uz) boyitish va ochiq ma’lumotlar bazasini yaratish.
<b>Alifbo muammosi</b>	Kirill va lotin yozuvlarining parallel qo‘llanilishi.	Universal transliteratsiya va unifikatsiya algoritmlarini joriy etish.

**2-jadval. Til sistemasidagi masalalar.**

Muammo turi	Lingvistik tavsif	SI uchun texnologik yechim	Tilshunosning yangi vazifasi
<b>Gallutsinatsiya</b>	Haqiqatga mos kelmaydigan uydirma faktlarni generatsiya qilish.	<b>RAG (Retrieval-Augmented Generation)</b> va Fact-checking algoritmlari.	Tashqi bilimlar bazasini (Verifikatsiya bazasi) shakllantirish.
<b>Polisemiya va Omonimiya</b>	So‘zlarning ko‘p ma’noiligi va shakldoshligi (masalan, <i>ot</i> , <i>tut</i> ).	<b>Contextual Embeddings</b> (BERT, RoBERTa modellari).	Semantik teglash (Sense disambiguation) korpuslarini yaratish.
<b>Pragmatik xatolik</b>	Matn osti ma’nosini, piching va kontekstni tushunmaslik.	<b>RLHF (Reinforcement Learning from Human Feedback).</b>	Modelni insoniy muloqot ssenariylari asosida baholash va “tarbiyalash”.

Muammo turi	Lingvistik tavsif	SI uchun texnologik yechim	Tilshunosning yangi vazifasi
<b>Mantiqiy uzilish</b>	Matn qismlari o'rtasidagi mantiqiy bog'liqlikning yo'qolishi.	<b>Knowledge Graphs (Bilimlar grafigi)</b> va Ontologik modellashtirish.	Tushunchalar o'rtasidagi mantiqiy iyerarxiya xaritasini chizish.

O'zbek tilshunoslari oldida tilning **ontologik bazasini** (so'zlarning o'zaro mantiqiy bog'liqlik xaritasi) yaratishdek fundamental vazifa turibdi. Bu mashinaga o'zbekcha so'zlarni shunchaki ketma-ketlik emas, balki tushunchalar tizimi sifatida anglash imkonini beradi. Buning uchun qomusiy va umumfilologik lug'atlarning elektron bazasini yaratish, har bir sohaga tegishli terminlar ontologiyasini tayyorlash muhim. Fanning keyingi taraqqiyotida tilshunoslik va SI integratsiyasi asosida **neyrolingvistik tahlil** (Inson nutqiy faoliyati va SI algoritmlari o'rtasidagi izomorflikni o'rganish.), **raqamli dialektologiya** (SI yordamida hududiy dialektlarni avtomatik klassifikatsiya qilish va xaritalash), **avtomatik tarjima va interpretatsiya** (Madaniy kontekstni hisobga oluvchi, piching va ko'chma ma'nolarni yetkazib beruvchi neyron tarjimonlarni yaratish) hamda **psixolingvistik diagnostika** (Inson yozgan matn yoki nutq orqali uning psixologik holatini aniqlash algoritmlarini lingvistik asoslash.) kabi yo'nalishlar vujudga keladi.

**XULOSA.** Tilshunoslik fani bugungi kunda o'zining ikkinchi "uyg'onish" davrini yashamoqda. Sun'iy intellekt tilshunoslikni nazariy cheklanganlikdan qutqarib, uni muhandislik darajasiga olib chiqdi. Ammo o'zbek tili "kam resursli" (low-resource) tillar sirasiga kirgani sababli, SI modellarida o'zbekcha **gallutsinatsiyalar** (uydirmalar) ko'proq uchraydi. Buning asosiy sababi SI modelining o'zbekcha matnlarni yetarli darajada tahlil qilmagani va ko'pincha inglizcha yoki ruscha mantiqni o'zbek tiliga so'zma-so'z ko'chirganidir. Umuman olganda, SI modellari lingvistik nazariyalarsiz (xususan semantika va pragmatika) inson darajasidagi intellektga erisha olmaydi. O'zbek tili uchun milliy SI modellarini yaratish ilmiy, ijtimoiy-siyosiy va shu bilan birga strategik ahamiyatga ega.

### ADABIYOTLAR RO‘YXATI

1. Bender E. M., Koller A. Climbing towards NLU: On Meaning, Forma and Understanding in the Age of Data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020.
2. Chomsky N. (1957). *Syntactic Structures*. –The Hague: Mouton, 1953.
3. Jurafsky D. Martin J. H. *Speech and Language Processing*. –Stanford University, 2023.
4. Madvaliyev, A. *O‘zbek tili terminologiyasi va leksikografiyasi masalalari*. – Toshkent: O‘zbekiston milliy ensiklopediyasi, 2021.
5. Miller, G. A. (1995). WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11), 39–41.
6. Vaswani A., Shazeer N., Parmar N. Attention is All You Need. *Advances in Neural Information Processing Systems (NIPS)*, 2017.
7. Ziyodullayev, A. R. *Kompyuter lingvistikasi asoslari*. – Toshkent: Universitet nashriyoti, 2022.
8. Uzbek National Corpus. [uzbekcorpus.uz](http://uzbekcorpus.uz) (Elektron resurs).
9. Mengliyev B., Hamroyeva Sh. Sun’iy intellekt va zamonaviy tilshunoslik vazifalari. <https://inlibrary.uz/index.php/ijsci/article/view/130204>