

ZAMONAVIY NUTQIY MODELLARDA TABIIY TILNI QAYTA ISHLASH VA MENTAL GRAMMATIKA MASALALARI

<https://doi.org/10.70728/phil.v05.i02.021>

Xo'jaaxmatova Qunduzoy Baxtiyor qizi
qunduzxojaaxmatova@gmail.com
SDVMCHBU Toshkent filiali

Annotatsiya: Ushbu maqolada zamonaviy nutqiy modellar va ularda tabiiy tilni qayta ishlash jarayonini lingvistik jihatdan baholash va undagi muammolarni aniqlash, shuningdek, Lingvistika sohasida insonlarning ong bilan bo'g'liq bo'lgan til ko'nikmalarini egallashga qaratilgan asosiy yo'nalishlaridan biri sifatida Psixolingvistikaning insonlarda tilni qayta ishlash hamda sun'iy intellektda bu jarayonning ishlash prinsiplarini tahlil qilinishi kabi bir qancha masalalar yoritib berilgan. Xususan bu borada inson tili va sun'iy intellektda nutq jarayoning hosil bo'lishi va nutq kontekstida mental grammatikaning funksional xususiyatlariga ham alohida urg'u berib o'tilgan.

Kalit so'zlar: nutq modellari, psixolingvistika, sun'iy intellekt, mental grammatika TTS(text to speech), STT(Speech to text) tizimlari, ChatGPT, Gemini

NATURAL LANGUAGE PROCESSING AND MENTAL GRAMMAR ISSUES IN MODERN SPEECH MODELS

Kho'jaakhmatova Kunduzoy Bakhtiyor qizi
kunduzxojaaxmatova@gmail.com
Samarkand State University of Veterinary Medicine, Animal Husbandry and
Biotechnology Tashkent branch

Abstract: This article covers a number of issues, such as modern speech models and the linguistic assessment of the natural language processing process in them and the identification of problems in it, as well as the analysis of the principles of language processing in humans and artificial intelligence in the field of Psycholinguistics, one of the main directions in Linguistics aimed at acquiring language skills related to consciousness. In particular, special emphasis is placed on the formation of the speech process in human language and artificial intelligence and the functional features of mental grammar in the context of speech.

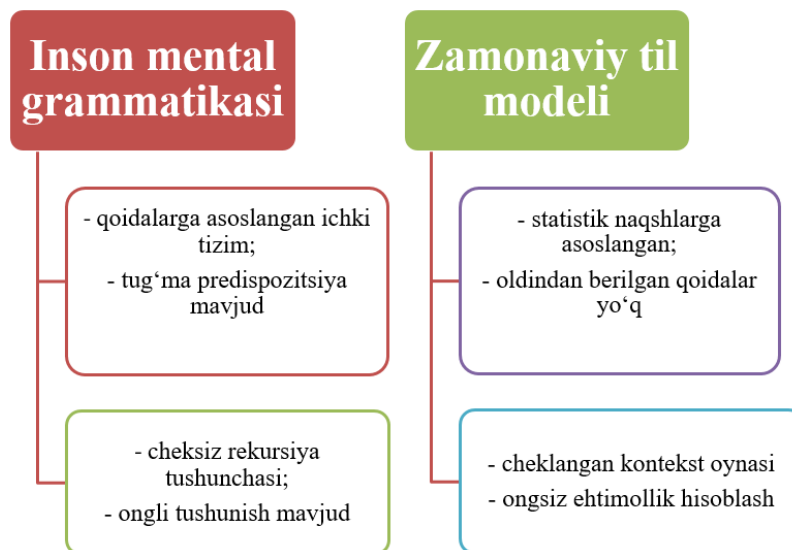
Keywords: speech models, psycholinguistics, artificial intelligence, mental grammar TTS (text to speech), STT (Speech to text) systems, ChatGPT, Gemini

Zamonamizning har sohada rivojlanib borayotganligi ayniqsa, raqamli texnologiya va axborot tizimlarining insonlarga dunyo bo'ylab cheksiz imkoniyatlar hamda qulayliklar yaratib berayotganligida o'z isbotini topmoqda. Bunga misol qilib sun'iy intellektni keltirishimiz mumkin. Sun'iy intellekt bu- tizimlashtirilgan, katta ma'lumotlar hamda matnlar asosida ishlaydigan tizmdir. Bugungi kunda uning turli xil ko'rinishlari mavjud va u turli maqsadlarda foydalanishga qaratilgan. Sun'iy intellekt deyarli barcha sohalarda, ta'lim, tibbiyot, business, moliya hamda boshqalarda keng qo'llanilib kelinmoqda. Ularning ichida eng keng tarqalgani ChatGPT, Gemini, Deepseek va boshqalardir. Bu tizimlar orqali biz matnli, audioli, rasm va video ko'rinishidagi ma'lumotlarni qabul qilib olishimiz mumkin. Tabiiyki savol tug'iladi, xo'sh ma'lumotlar ba'zasidan qanday foydalaniladi va buning uchun unda til ko'nikmalari mavjudmi? Har qanday tizimda u sizga kerakli ma'lumotni kerakli formatda topib berishi uchun siz undan ona tilingizda yoki boshqa chet tillarida savol so'rashingiz mumkin. Biroq savolni va sizga berayotgan javoblarni qay tarzda sizga uzatadi? Yoki tahlil qiladi? Bu tabiiyki har qanday sun'iy intellekt tizimida Tabiiy tilni qayta ishlash bosqichi orqali o'rganiladi va tahlil qilinadi. Tabiiy tilni qayta ishlash jarayonida zamonaviy nutqiy modellar datda sun'iy neyron tarmoqlarga, ayniqsa transformer arxitekturasiga asoslanadi. Ular matnni kichik bo'laklarga ajratib, uning kontekstual muvofiqligini, gaplarning grammatik hamda semantik jihatdan tuzilishini modellashtiradi. Nutqiy modellar, shu boisdan, zamonaviy tilshunoslikning yangi bosqichi sifatida qaralmoqda hamda ular korpus lingvistikasi, psixolingvistika, til tipologiyasi, amaliy tilshunoslik kabi sohalarda o'zining ilmiy va amaliy ahamiyatini ko'rsatmoqda. Tabiiy tilni qayta ishlash orqali ulardagi nutqni psixolingvistik, amaliy tilshunoslik jihatdan uchraydigan muammolarni aniqlash va tahlil qilish imkoniyatiga ega bo'lish mumkin. Tabiiy tilni qayta ishlashda ko'zlangan asosiy maqsad bu-kompyuterlarga tizimlashtirilgan matnni tushuntirish hamda mazmunan to'g'ri yondashishni o'rgatishdir. Tabiiy tilni qayta ishlashning bir qancha yo'nalishlari mavjud; mashina tarjimasini, nutqni tanish, kayfiyat /tuyg'ularni tahlil qilish, savollarga javob berish, matnni referatlash yoki xulosalash, Chatbot, matn tasniflari, belgilarni aniqlash va boshqalar. Tabiiy tilni qayta ishlash orqali raqamli texnologiyalar har qanday elektron hamda og'zaki matnni tinglaydi, aniqlaydi va tushunadi. Til modellarini lingvistik jihatdan tahlil qilish deganda, ularning til birliklari bilan qanday ishlashi va lingvistik qoidalarga qay darajada mos kelishi o'rganiladi. Bu holatda Mental grammatika tushunchasi haqida so'z yuritish zarur. Mental grammatika bu inson ongi orqali shakllangan ichki til qoidalari bo'lib o'zaro muloqot jarayonida so'zlarni anglash, grammatik jihatdan birlashtirish orqali gaplarni hosil qilish, shuningdek uni kontekstga mos ravishda qo'llay olishdir. Bu haqida Noam Chomskiyning umumiy grammatika nazariyasida ham insonning biror bir tilda

muloqot qilishi tug'ma grammatik qobiliyatga egaligi va tilni ichki qoidalar asosida o'zlashtirishi haqida tushunchalar berib o'tilgan. Shunga ko'ra mental grammatikani muloqot jarayonida tabiiy tarzda inson tomonidan shakllanishini quyidagi xususiyatlarga qarab bilish mumkin. Fonologik xususiyatga ko'ra so'z hosil qilishda tovushlar tizimi orqali to'g'ri talaffuz qilish, urg'u, intonatsiya hamda fonetik o'zgarishlar inson ichki til qoidalari orqali avtomatik amalga oshadi va bu ko'pincha ona tilida suhbatlashganda yaqqol o'z aksini topadi. Morfologik xususiyatga ko'ra esa so'zlarning yasaliishi hamda so'z yasovchi qo'shimchalarni to'g'ri qo'llash va uni grammatik konteksta tahlil qilish qobiliyatini beradi. Masalan: o'qimoq -o'qidi, kitob-kitobxon so'zlarida bir nechta so'zlarning bitta asosdan so'z yasaliishi yoki zamonda to'g'ri qo'llash inson ongi orqali avtomatik amalga oshadi. Sintaktik xususiyatga ko'ra gap qurilishi, soda va murakkab gaplarni hosil qilish gaplarda harakat holat va harakat bajaruvchisini tanish va amaliy tarzda har qanday yangi murakkab yoki soda gaplarni tuzish hodisalarini o'z ichiga oladi. Masalan: o'zbek tilida gap strukturasi quyidagicha; Ega+to'ldiruvchi+kesim bo'lsa ba'zi tillarda Ega+Kesim+ To'ldiruvchi shaklida keladi. Semantik xususiyatiga keladigan bo'lsak bunda har qanday tildagi ko'p ma'nolilik, o'xshatish, kinoya, hazil kabi kontekstga mos gaplar tuzishni o'z ichiga oladi. Zamonaviy nutqiy modellarning hozirgi ko'rinishlarida, masalan, ChatGPT, Gemini, SunoAI, Detektor, Gamma, Alice va boshqalarda tabiiy tilni qayta ishlash jarayoni turli bosqichlarni o'z ichiga oladi. Bu tizimlashtirishning birinchi bosqichi tokenizatsiya bo'lib, bunda matn kichik birliklarga-tokenlarga(ya'ni so'z, so'z birikma yoki belgilar) ajratiladi. Har bir token raqamli identifikatorga aylantiriladi. Keyingi bosqich Embedding(vektorlashtirish) bo'lib, har bir token ko'p o'lchamli vektorga o'tkaziladi. Bu vektor tokenlarni semantik ma'no, sintaktik ro'l, kontekstga bog'liqlik jihatdan o'rganadi, ya'ni chuqur o'rganish (Deep learning) orqali. Nutqiy modellar transformer arxitekturasiga asoslanadi. Nutqiy modellar juda katta matn korpuslarida o'qitiladi, bunda asosiy vazifa keying tokenni bashorat qilishdir. Genetarsiya ya'ni matn yaratish bosqichida tokenlar oldingi berilgan tokenlarga asoslanib ehtimollik taqsimotini hisoblaydi va eng ehtimolga yaqin bo'lgan tokenni tanlaydi, jarayon esa ketma-ketlikda takrorlanadi. Demak bundan bilish mumkinki, inson nutq tizimida tilni qayta ishlash jarayoni ongli, biologik, hamda ichki grammatik qoidalariga asoslangan holda ishlasa, modellarda esa algoritm, hisoblash, ehtimollik asosida kechadi. Zamonaviy Katta Til Model(LLM)lar Ashish Vaswani va hamkorlari tomonidan taklif qilingan Transformer arxitekturasiga asoslanadi. Bu arxitektura sintaktik daraxtlarni ochiq ko'rinishda qurmaydi (ya'ni parse tree yasamaydi), lekin ularni yashirin fazoda (hidden space) implisit tarzda kodlaydi. Sun'iy intellekt (SI) tizimlarida til va mantiqning o'zaro munosabati dolzarb masala bo'lib, haqiqiy intellektual tizim yaratishda bu ikki omilning integratsiyasi muhim ahamiyatga ega. Til va mantiqning integratsiyasi masalasi shuni anglatadiki, tabiiy tilni tushunish (matnda nima

deyilayotganini anglash) ko‘pincha mantiqiy xulosa chiqarish bilan chambarchas bog‘liq.

Bu farqni qisqacha qilib quyidagi tartibda ko‘rsatish mumkin:



Insonlar tilda mantiqiy fikrlarini ifodalaydilar: hukmlar, xulosalar, sabab-oqibat munosabatlari tilda aks etadi. Sun‘iy intellektga odamlar va boshqa tirik organizmlarning tabiiy aqli qarshi qo‘yiladi. Ammo barcha robotlarda ham sun‘iy intellekt mavjud emas, chunki faqat aqlgina fikr yuritish, o‘rgatish, mustaqil qaror qabul qilish imkoniga ega [1]. Sun‘iy intellekt tizimi matndagi mazmunni to‘laonli anglab, undan to‘g‘ri xulosa chiqara olishi uchun til modellari va mantiqiy inferensiya o‘rtasida muvofiqlik bo‘lishi kerak. Hozirgi vaqtda til va mantiq integratsiyasiga erishish uchun neyron va ramziy yondashuvlarni birlashtirish g‘oyalari ilgari surilmoqda. SRI International tadqiqotchilari ta’kidlaganidek, amaliy SI tizimlarida bilimga asoslangan (ramziy) va ma’lumotga asoslangan (statistik) usullarni uyg‘un qo‘llash talab etiladi - faqat mantiqiy qoidalarga tayanuvchi tizimlar juda tor doiradagi muammoni hal eta olsa, faqat ma’lumotdan o‘rganuvchi tizimlar “chuqur” mantiqiy inference qilolmaydi. Shunday ekan, tilni tushunuvchi modellarni mantiqiy qobiliyatlar bilan boyitish, yoki aksincha, mantiqiy tizimlarni tabiiy til bilan “ovqatlantirish” sun‘iy intellektni yangi bosqichga olib chiqishi kutilmoqda.

Foydalanilgan adabiyotlar

1. Q.B.Xo‘jaaxmatova, “Mental Grammatikaning Bilingvlar Nutqini O‘stirshdagi Funktsional Xususiyatlari”, Italy, 2025.
2. Абдурахмонов Қ. Сунъий интеллект: кеча, бугун, эртага // Янги Ўзбекистон газетаси, 2022 йил, 27 сентябрь, 198-сон
3. Brown T. va boshq. Language Models are Few-Shot Learners. NeurIPS (arXiv preprint), 2020.
4. Wolfram S. What Is ChatGPT Doing ... and Why Does It Work? StephenWolfram.com blogi, 2023.

5. Wei va boshq., Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, arXiv, 2022
6. To'rayev B.O. Sinergetik tafakkur va sun'iy intellekt // "Falsafa va insoniyat: o'tmish, hozirgi zamon, kelajak" ilmiy-nazariy anjuman. Alfraganus universiteti. 2024. 23-bet.
7. IBM Research. "Neuro-symbolic AI." research.ibm.com maqolasi, 2023.