



MORFOLOGIK TAHLIL MUAMMOLARI VA ULARNING YECHIMLARI

Islomova Barno Xabibullo qizi
tayanch doktoranti
ToshDO‘TAU

Annotatsiya. Mazkur maqola o‘zbek tili morfologik tahlilining nazariy asoslari va uning tabiiy tilni qayta ishlash (NLP) sohasidagi zamonaviy muammolariga bag‘ishlangan. Maqolada tilshunoslikning klassik metodlari va raqamli texnologiyalar o‘rtasidagi bog‘liqlikni tahlil qilinib, agglyutinativ tillarga xos bo‘lgan omonimiya, o‘zak variativligi va murakkab so‘z shakllari tahlilidagi to‘siqlar ko‘rsatilgan. Tadqiqotda chekli holatlar texnologiyasi (FST), neyron tarmoqlari va katta til modellarining (LLM) o‘zbek tili morfologiyasini avtomatlashtirishdagi o‘rni yoritilgan. Maqolada morfologik tahlil sifatini oshirish uchun lingvistik qoidalar va statistik modellarni birlashtirgan gibril yondashuv hamda milliy annotatsiya standartini joriy etish zarurligi ilmiy asoslangan hamda o‘zbek tilining global raqamli ekotizimdagi nufuzini oshirishi dalilangan.

Kalit so‘zlari: *morfologiya, NLP, UzMorph, katta til modellari, gibril yondashuv.*

Abstract. This article is devoted to the theoretical foundations of morphological analysis of the Uzbek language and its modern problems in the field of natural language processing (NLP). The article analyzes the relationship between classical methods of linguistics and digital technologies, and shows the obstacles in the analysis of homonymy, stem variation, and complex word forms inherent in agglutinative languages. The study highlights the role of Finite State Technology (FST), neural networks, and Large Language Models (LLM) in automating the morphology of the Uzbek language. The article argues that the need for a hybrid approach combining linguistic rules and statistical models and the introduction of a national annotation standard to improve the quality of morphological analysis is



scientifically justified and will increase the prestige of the Uzbek language in the global digital ecosystem.

Keywords: *morphology, NLP, UzMorph, big language models, hybrid approach.*

Kirish. Morfologiya tilning sathlararo iyerarxiyasida markaziy o'rinni egallaydi. U fonologiya va sintaksisni bog'lovchi ko'priq vazifasini o'taydi. Morfologik tahlilning asosiy maqsadi so'zning ichki tuzilishini, morfemalarning birikish qonuniyatlarini va so'z shakllarining grammatik ma'nolarini aniqlashdan iborat.

Bugungi kunda, global raqamlashtirish jarayonida morfologik tahlil faqat nazariy tilshunoslik doirasidan chiqib, tabiiy tilni qayta ishlash (NLP) sohasining eng muhim qismiga aylandi. Biroq, o'zbek tili kabi boy morfologik imkoniyatlarga ega tillarda so'z o'zgarishi va so'z yasalishi tizimidagi murakkabliklar tahlil jarayonida bir qator muammolarni keltirib chiqarmoqda.

O'zbek tili morfologiyasining poydevori o'tgan asrning o'rtalarida yaratilgan fundamental tadqiqotlarga borib taqaladi.

G.Abdurahmonov va A.Rustamovlar “Qadimgi turkiy til” [2:15-60] asarida morfologik tahlilning tarixiy asoslari o'rganilgan bo'lib, unda o'zbek tilidagi qo'shimchalarning genezisi va ularning vaqt o'tishi bilan qanday transformatsiyaga uchragani tahlil qilingan.

“Hozirgi o'zbek adabiy tili” [5:5-72] asari morfologiyani tizimli o'rganishning boshlang'ich nuqtasi bo'lib, so'z turkumlarining klassik tasnifini va morfemalarning tahlil usullarini belgilab bergan. Shuningdek, mualliflar tomonidan an'anaviy morfologik tahlildan farqli o'laroq, so'zlarni substansial (mohiyatli) tahlil qilish taklif etilgan. Unda morfologik omonimiya va polifunksionallik muammolari yangicha talqin qilingan.



Mustaqillik yillarida morfologik tahlilga funksional-semantik va kognitiv nuqtai nazardan yondashish kuchaydi.

Akademik A.Nurmonovning [4:12-28] morfofonologiya sohasidagi ishlari morfologik tahlilning fonetik sath bilan bog'liqligini ochib beradi. U o'zakdagi tovush o'zgarishlarining tahlil jarayoniga ta'sirini ilmiy asoslab bergan.

Hozirgi kunga kelib esa, mazkur masalalar kompyuter lingvistikasi va NLP tizimlari asosida amalga oshirilmoqda. Ya'ni, morfologik tahlilning eng dolzarb muammolari hozirda axborot texnologiyalari bilan bog'liq holda o'rganilmoqda.

K.Beesley va L.Karttunenlar tomonidan yozilgan manbada morfologik tahlilni avtomatlashtirishning matematik asoslari yoritilgan. O'zbek tili uchun yaratilayotgan morfologik analizatorlar (masalan, UzMorph) [6:1-2] aynan shu asardagi “Chekli holatlar texnologiyasi”ga tayanadi.

M.Abdurahmonova [1:1-5] o'zbek tili morfologik tahlilini algoritmizatsiyalash muammolarini tadqiq etgan. Unda agglyutinatив tillar uchun stemmer va lemmatizatorlar yaratishning ilmiy muammolari tahlil qilinadi.

Hozirgi raqamli texnologiyalar rivojlangan davrda ayrim manbalar so'z yasashini alohida sath sifatida ko'rsatsa, boshqalari uni morfologiyaning tarkibiy qismi deb hisoblaydi.

B.Elov va boshqalar tomonidan yozilgan “Katta til modellari” [3:1-4] nomli ilmiy maqola o'zbek tilshunosligida sun'iy intellekt va neyron tarmoqlarining qo'llanilishi borasidagi eng yangi tadqiqotlardan biri hisoblanadi. Ushbu maqola o'zbek tili uchun milliy LLM yaratishning yo'l xaritasi bo'lib xizmat qiladi va sohadagi texnik muammolarga lingvistik yechimlar taklif etadi.

Adabiyotlar tahlili asosida ushbu yo'nalishdagi asosiy muammolarni quyidagi tartibda keltirish mumkin:

Texnik muammolar. O'zbek tili kabi agglyutinatив tillarda har bir so'z shaklini lug'atga kiritish imkonsiz. Shuning uchun chekli holatli morfologiya

qo'llaniladi. Unda tizim so'zni o'zakdan boshlab, qo'shimchalarni ma'lum bir “yo'l” bo'ylab tahlil qiladi. Agar qo'shimcha noto'g'ri tartibda kelsa (masalan, kelishikdan keyin yasovchi qo'shimcha), tizim uni xato deb hisoblaydi.

Semantik muammolar. Morfologik tahlildagi ko'p ma'nolilikni faqat qoidalar bilan yechib bo'lmaydi. Bunda so'zning faqat o'zini emas, balki chap va o'ng tomonidagi qo'shni so'zlarni ham tahlil qilish talab etiladi.

Misol, “U kitobni olmadi” (fe'l) vs “Savatda qizil olma bor” (ot).

Transformer modeli “savatda” va “bor” so'zlari asosida “olma” so'zining morfologik statusini 99% aniqlikda teglaydi.

Strukturaviy muammolar. O'zbek tilida morfologik tahlilning yagona standartining yo'qligi tadqiqotlarni qiyinlashtiradi. O'zbek tili milliy korpusi doirasida yagona morfologik teglash tizimini joriy etish lozim. Bu tizimda har bir qo'shimcha universal xalqaro kodlar bilan teglanishi shart.

Quyidagi jadvalda morfologik tahlil muammolari va ularning yechimlari keltirilgan.

1-jadval. Morfologik tahlil muammolari va ularning yechimlari

Muammo guruhi	Tavsifi	Yechimi
Lingvistik omonimiya	Bir xil shaklga ega qo'shimchalarning turli funksiyalarda kelishi.	Kontekstual tahlil (Disambiguatsiya): Neyron tarmoqlari yordamida so'zning gapdagi o'rnini aniqlash.
Agglyutinativ zanjir	O'zbek tilida bitta o'zakka 10 tagacha qo'shimcha qo'shilishi (so'z shakllari ko'pligi).	Finite State Transducers (FST): So'zni “matematik chekli holatlar” asosida qatlam-qatlam tahlil qilish.
O'zak variativligi	Tovush almashishlari (metateza, eliziya) natijasida o'zakning o'zgarishi.	Morfologik qoidalar bazasi: O'zak o'zgarishlarini standartlashtirilgan qoidalar asosida “lemmatizatsiya” qilish.
Ma'lumotlar tanqisligi	LLM va NLP modellari uchun sifatli markirovka qilingan matnlarning kamligi.	Kam miqdordagi ma'lumot bilan neyron tarmoqlarga asoslangan o'zi o'rganuvchi algoritmlarni tatbiq etish.

Bu kabi morfologik tahlil muammolarini yechish uchun Gibrid yondashuv eng samarali hisoblanadi. Bunda, lingvistik qoidalar tahlil uchun, statistik modellar

ehtimollarni hisoblash uchun, katta til modellari murakkab kontekstual tahlil uchun qo‘llaniladi.

Bu uch komponent birlashganida, o‘zbek tili morfologik tahlili 95-98% aniqlikka erishishi mumkin, bu esa mashina tarjimasini va sun‘iy intellekt sifatini tubdan yaxshilaydi.

Xulosa.

O‘zbek tili morfologik tahlilidagi muammolar faqat texnik emas, balki metodologik xarakterga ega. Ushbu muammolarni tizimli yechish orqali biz nafaqat tilshunoslikni rivojlantiramiz, balki o‘zbek tilining global raqamli ekotizimida (Google Translate, Siri, ChatGPT va h.k.) to‘laqonli va xatosiz faoliyat yuritishini ta‘minlaymiz. Kelajakdagi tadqiqotlar e‘tiborini “neyron tarmoqlari + lingvistik qoidalar” kombinatsiyasiga qaratish eng istiqbolli yo‘l hisoblanadi.

Shuningdek, morfologik tahlil endi faqat nazariy tilshunoslikning predmeti emas, balki yuqori texnologik ehtiyojga aylandi. Klassik tilshunoslikdagi muammolar (omonlar, o‘zak o‘zgarishi) aniqlikni kamaytirmoqda. Ushbu muammolarni yechish uchun lingvistik bilimlar bilan sun‘iy intellekt texnologiyalarini birlashtirgan gibrid modellardan foydalanish davr talabidir.

Bundan tashqari, morfologik tahlilning barqaror rivojlanishi uchun yagona milliy annotatsiya standarti zarur. Milliy korpus doirasida yagona morfologik teglar tizimini joriy etish barcha raqamli platformalar uchun universal zamin yaratadi.

Foydalanilgan adabiyotlar ro‘yxati

1. Abdurahmanova M., Olimova M. Korpus lingvistikasining ahamiyati // Computer linguistics: Problems, solutions, prospects. – 2022. – T. 1. – №. 1.
2. Abdurahmonov F., Rustamov A. Qadimgi turkiy til. – Toshkent: O‘qituvchi, 1982.
3. Elov B., Xusainova Z., Qosimova S. Katta til modellari // Uzbekistan: Language and Culture. – 2024. – T. 2. – №2.



4. Nurmonov A., Sobirov A., Qosimova N. Hozirgi o‘zbek adabiy tili. – Toshkent: Ilm-ziyo, 2010. –142 b.
5. Sayfullayeva R., Mengliyev B., Boqiyeva G., Qurbonova M., Yunusova Z., Abuzalova M. Hozirgi o‘zbek adabiy tili. – Toshkent, 2009.
6. Менглиев Д.Б. UzMorph [правообладатели: Менглиев ДБ (RU), Барахнин ВБ (RU)] // Свидетельство о государственной регистрации программы для ЭВМ. – 2021. – №2021662588.