



## MORFOLOGIK NOANIQLIKNI YECHISH YO‘LLARINING TAHLILI

**Elov Botir Boltayevich,**  
Texnika fanlar doktori, dotsent  
[elov@navoiy-uni.uz](mailto:elov@navoiy-uni.uz)  
ToshDO‘TAU

**Islomova Barno Xabibullo qizi,**  
tayanch doktoranti  
ToshDO‘TAU

**Annotatsiya.** Ushbu maqolada o‘zbek va jahon tillaridagi morfologik noaniqlik (ambiguity) hodisasi, uning turlari va kelib chiqish sabablari tadqiq etiladi. Morfologik noaniqlik bir so‘z shaklining bir necha grammatik ma’noga ega ekanligi bo‘lib, bu tabiiy tilni qayta ishlash (NLP) tizimlarida asosiy to‘siqlardan biri hisoblanadi. Tadqiqot davomida sinxron tahlil va korpus lingvistikasi usullaridan foydalanilgan.

**Kalit so‘zlari:** *morfologik noaniqlik, tabiiy tilni qayta ishlash, katta hajmdagi ma’lumotlar, mashinali o‘qitish, neyron tarmoqlar, gibridd tizimlar.*

**Abstract.** In this this article is studied the phenomenon of morphological ambiguity in Uzbek and world languages, its types and causes. Morphological ambiguity is the fact that a word form has several grammatical meanings, which is one of the main obstacles in natural language processing (NLP) systems. In his research is used synchronous analysis and corpus linguistics methods.

**Keywords:** *morphological ambiguity, natural language processing, big data, machine learning, neural networks, hybrid systems.*

**Kirish.** Morfologik noaniqlik tilshunoslikning eng murakkab va qiziqarli muammolaridan biridir. Birgina leksik birlikning kontekstdan tashqarida turli morfologik kategoriyalarga mansub bo‘lishi nutqni tushunish va kompyuter tahlilida qiyinchiliklar tug‘diradi. Morfologik noaniqlik (ambiguity) nafaqat nazariy tilshunoslik, balki zamonaviy texnologiyalar va kundalik muloqot uchun ham o‘ta dolzarb muammo hisoblanadi.



Hozirgi kunda kompyuterlar inson tilini tushunishi uchun morfologik tahlil birinchi bosqich hisoblanadi. Agar dastur “olma” so‘zini *meva (noun)* yoki fe‘l (*verbning inkor shakli*) ekanligini ajrata olmasa, tarjima mantiqsiz chiqadi. Foydalanuvchi “yoz” deb qidiruv berganda, tizim faslni yoki harakatni nazarda tutayotganini kontekst orqali aniqlashi shart.

Katta hajmdagi ma’lumotlar bazasida so‘zlarni guruhlashda morfologik noaniqlik xatoliklarga sabab bo‘ladi. So‘zni lug‘at shakliga keltirishda (masalan, “ot” so‘zini hayvon yoki “otmoq” fe‘lining asosi sifatida ko‘rish) aniqlik darajasi axborot tizimlarining sifatini belgilaydi.

Til doimo tejamkorlikka kam shakl bilan ko‘p ma’noni ifodalashga intiladi. Morfologik noaniqlik tildagi polisemiya (ko‘p ma’nolilik) va omonimiyaning natijasi hisoblanadi. Ushbu hodisani o‘rganish tilning grammatik qurilishi qanchalik moslashuvchan ekanligini va qaysi nuqtalarda nutqiy chalkashliklar (shovqinlar) yuzaga kelishini ko‘rsatib beradi [3:1-2].

Kundalik muloqotda ko‘pincha noaniqlik sezilmaydi, chunki inson miyasi kontekstni bir zumda tahlil qiladi. Biroq, birgina qo‘shimchaning (masalan, “-ma” inkor yoki ot yasovchi) noto‘g‘ri talqin qilinishi huquqiy oqibatlarga olib kelishi mumkin. Adabiyotda morfologik noaniqlikdan ataylab badiiy san’at (masalan, tajnis) yaratishda foydalaniladi.

Bugungi kunda sun’iy intellekt va mashina tarjimasi rivojlanayotgan bir davrda, so‘zlarni to‘g‘ri lemmalash (aslini topish) va teglash (POS-tagging) juda muhim.

Morfologik noaniqlikni (disambiguation) yechish bu matndagi ko‘p ma’noli so‘z shaklining aynan qaysi morfologik kategoriyaga tegishli ekanligini kontekstga qarab aniqlash jarayonidir [4:2-12]. Bugungi kunda ushbu muammo lingvistik va texnik darajada quyidagi usullar orqali hal qilinadi:



– Qoidalarga asoslangan usullar tilshunoslar tomonidan ishlab chiqilgan qat'iy grammatik qoidalarga tayanadi. Masalan, “*olma*” so‘zidan oldin “*qizil*” yoki “*shirin*” sifati kelsa, tizim uni ot (meva) deb teglaydi. Agar undan oldin “*uni*” olmoshi kelsa yoki gap buyruq ohangida bo‘lsa, fe‘l deb hisoblaydi. Noaniqlik darajasi juda yuqori bo‘lgan so‘zlar uchun maxsus lug‘atlar yaratiladi [6:52-78].

– Statistik va ehtimollik usullarda so‘zning qaysi turkumga oidligi katta matnlar to‘plami (korpus) asosida hisoblanadi [2:5-9]. Tizim so‘zlar zanjirini tahlil qiladi va “*A so‘zidan keyin B turkumi kelish ehtimoli qancha?*” degan savolga javob izlaydi. Masalan: O‘zbek tilida egalik qo‘shimchasidan keyin kelishik qo‘shimchasi kelish ehtimoli 99% ni tashkil etadi.

– Mashinali o‘rganish va neyron tarmoqlari hozirgi kunda eng samarali yechim hisoblanadi [2:1-5]. So‘zlar ko‘p o‘lchovli fazoda vektor sifatida tasvirlanadi. Ma‘nodosh va grammatik yaqin so‘zlar bir-biriga yaqin joylashadi. Ushbu usulga asoslangan BERT, GPT modellari gapni shunchaki chapdan o‘ngga o‘qimaydi, balki gapdagi barcha so‘zlarning bir-biriga bog‘liqligini bir vaqtda tahlil qiladi. Bu “*olma*” so‘zining ma‘nosini gapning boshidagi yoki oxiridagi uzoq bog‘lanishlar orqali ham aniqlash imkonini beradi.

Gibrid tizimlar eng yuqori aniqlik (98 foizdan yuqori) beradigan yechim [5:2-6]. Bunda ham qat'iy lingvistik qoidalar, ham neyron tarmoqlarining imkoniyatlari birlashtiriladi. Dastlab qoidalar orqali aniq xatolar aniqlanadi, so‘ngra murakkab holatlar sun‘iy intellekt tomonidan yechiladi. Bunda matn so‘zlarga ajratiladi va morfologik tahlil orqali so‘zning barcha mumkin bo‘lgan variantlari ajratiladi. Masalan, “*olma*” = *meva / ol+ma*. Keyingi bosqichda kontekstni tahlil qilib, noto‘g‘ri variantlar o‘chiriladi. Oxirida eng yuqori ehtimollikka ega bo‘lgan bitta variant qoldiriladi.

Quyidagi jadvalda morfologik noaniqlikni yechish usullarining texnik imkoniyatlari, amaliy samaradorligi va kamchiliklari keltirilgan.

*1-jadval. Morfologik noaniqlikni bartaraf etish usullarining qiyosiy tahlili*

<b>Tahlil mezonlari</b>	<b>Qoidalarga asoslangan usullar</b>	<b>Statistik usullar</b>	<b>Neyron tarmoqlari</b>
<b>Ishlash prinsipi</b>	Tilshunoslar tomonidan yozilgan qat'iy qoidalarga asoslanadi.	Matnlar korpusidagi so'zlar ketma-ketligi ehtimolligi.	Vektorli hisoblash va sun'iy intellekt o'z-o'zini o'rganishi.
<b>Kontekstni ko'ra olish</b>	Faqat qoidada ko'rsatilgan qisqa masofani ko'radi.	Cheklangan (odatda faqat o'zidan oldingi 1-2 so'zni) so'zlarga asoslangan.	Global, ya'ni butun gap yoki matnni bir vaqtda tahlil qiladi.
<b>O'zbek tili uchun samarasi</b>	Qo'shimchalar ketma-ketligi qat'iy bo'lgani uchun samarasi yuqori.	Katta hajmdagi belgilangan korpus talab etilganligi bois samarasi o'rta darajada.	Juda ko'p ma'lumot va quvvat talab qilsada natijasi yuqori hisoblanadi.
<b>Xatolik ustida ishlash</b>	Qoidaning xato ishlaganini ko'rish va tuzatish oson.	Statistik xatolikni tuzatish uchun qayta o'qitish kerak.	Ichki hisoblashlar asosida amalga oshirilganligi uchun uni aniqlash murakkab.
<b>Yangi so'zlarga munosabat</b>	Dasturga qo'shilmaguncha ularni tanimaydi.	O'xshashlik asosida taxmin qilishi mumkin.	Kontekst orqali yangi so'z turkumini aniqlay oladi.

Jadvaldan ko'rinib turibdiki, har bir usulning o'z o'rnini bor. Qoidalarga asoslangan usul o'zbek tili kabi agglyutinativ tillarda bazaviy morfologik analizator yaratish uchun poydevor hisoblanadi. Chunki so'z yasovchi va o'zgartiruvchi qo'shimchalar matematik aniqlikda bir-biriga ulanadi.

Statistik usullar lug'atlarda mavjud bo'lmagan so'zlar yoki dialektizmlar uchraganda “eng yaqin ehtimollik”ni topish uchun yordamchi vositadir.

Neyron tarmoqlari zamonaviy tarjimonlar va aqlli tizimlar uchun eng ma'qul yechim. U “olma” so'zining yonidagi so'zga emas, gapning oxiridagi “shirin ekan” yoki “uni qo'lingga olma” degan ma'noli qismlariga qarab noaniqlikni hal qiladi.

Hozirgi kunda eng yuqori natija morfologik qoidalar bilan neyron tarmoqlarini birlashtirgan gibrid tizimlar hisoblanadi. Bunda dastlab qoidalar orqali imkonsiz variantlar to'planadi, so'ngra qolgan noaniqliklar sun'iy intellekt usullari yordamida hal qilinadi.

Quyidagi jadvalda morfologik noaniqlikni bartaraf etish usullarining aniqligi, resurs sarfi, moslashuvchanligi va kontekst qamrovi mezonlari asosidagi tahlili keltirilgan.

*2-jadval. Morfologik noaniqlikni bartaraf etish usullarining aniqligi, resurs sarfi, moslashuvchanligi va kontekst qamrovi mezonlari asosidagi tahlili*

Mezon	Qoidalarga asoslangan usullar	Statistik usullar	Neyron tarmoqlari
Aniqligi	O'rta	Yaxshi	Juda yuqori
Resurs sarfi	Kam (faqat xotira)	O'rta	Juda ko'p
Til o'zgarishiga moslashuv	Qiyin (qo'lda amalga oshiriladi)	Avtomatik	Avtomatik va tez
Kontekst qamrovi	Cheklangan	Faqat qo'shni so'zlar	Butun matn

Agar o'zbek tili uchun mukammal tizim kerak bo'lsa, eng ma'qul yo'l tilimizning agglyutinativ tabiatidan (qo'shimchalar ko'pligidan) kelib chiqib, aniq morfologik qoidalarni yozish va murakkab ma'no noaniqliklarini hal qilishni BERT kabi modellarga topshirish kerak.

**Xulosa.** Tahlillar shuni ko'rsatadiki, morfologik noaniqlikni yechishda yagona va mutloq mukammal usul mavjud emas. Har bir usul o'zining ustunlik va kamchiliklariga ega. An'anaviy qoidalar tilning "skeleti"ni tushunish va grammatik xatolarni filtrlash uchun zarur. Statistik usullar tezkor va ommaviy matnlarni qayta ishlashda iqtisodiy jihatdan eng samarali hisoblanadi. Neyron tarmoqlari inson nutqiga xos bo'lgan nozik farqlarni ilg'ash va yuqori sifatli tarjima yaratishda yuqori natija beradi. Kelajakdagi eng to'g'ri yo'nalish lingvistik bilimlar bilan sun'iy intellektni birlashtirishdir. Faqatgina qoidalarga tayanib bo'lmaydi (til juda boy), shuningdek, faqat sun'iy intellektga tayanib ham bo'lmaydi (u mantiqiy xatolarga yo'l qo'yishi mumkin).

Morfologik noaniqlikni yechish nafaqat texnik vazifa, balki inson tafakkurini raqamlashtirish jarayonidir. O'zbek tili uchun ushbu muammoni hal qilish milliy kontentni jahon axborot maydoniga integratsiya qilishning eng asosiy kalitidir.

### Foydalanilgan adabiyotlar ro'yxati



1. Elov B. et al. Machine Learning Approaches to Sentiment Analysis in Classic and Modern Literature //2025 3rd International Conference on IoT, Communication and Automation Technology (ICICAT). – IEEE, 2025. – P. 1-7.
2. Vaswani A. et al. Attention Is All You need // Advances in Neural Information Processing systems. – 2017. – T. 30. – P. 5998-6008.
3. Jiménez-Preciado A. L. et al. The power of words from the 2024 United States presidential debates: A Natural Language Processing approach // Information. – 2024. – T. 16. – №. 1. –P. 1–12.
4. Manning C., Schütze H. Foundations of Statistical Natural Language Processing. – MIT Press, 1999. – 1 – 680 b.
5. Mukazhanov N. et al. Combined Method of Citation Sentiment Analysis: Linguistic Features and Machine Learning. // Journal of Information Science. – 2023. – P. 1–10.
6. Rahimov S., Umurqulov B., Eshonqulova A. Hozirgi o'zbek adabiy tili // Toshkent, 2003.