



O'ZBEK-TURK TILLARIDA OMONIMIYANING AVTOMATIK DISAMBIGUATSIYASI

Abdurashidova Odina A'zam qizi
mustaqil tadqiqotchi

Автоматическая дизамбигуация
омонимии в узбекском и турецком языках

Абдурашидова Одина Азамовна
независимый исследователь

Automatic Disambiguation of Homonymy in Uzbek and Turkish

Abdurashidova Odina A'zam qizi
Independent Researcher

Annotatsiya: Ushbu tadqiqotda o'zbek-turk tillaridagi omonim birliklarni parallel korpus asosida avtomatik aniqlash va kontekst, morfologik hamda semantik belgilar orqali ularning ma'nosini disambiguatsiya qilish modeli taklif etiladi.

Аннотация: В данном исследовании предлагается модель автоматического выявления омонимичных единиц в узбекском и турецком языках на основе параллельного корпуса, а также их семантической дизамбигуации с использованием контекстуальных, морфологических и семантических признаков.

Abstract: This study proposes a model for the automatic identification of homonymous units in Uzbek and Turkish based on a parallel corpus, as well as their semantic disambiguation using contextual, morphological, and semantic features.

Kalit so'zlar: O'zbek tili, turk tili, omonimiya, omonim birliklar, parallel korpus, avtomatik aniqlash, semantik disambiguatsiya, morfologik tahlil, kontekstual tahlil, tabiiy tilni qayta ishlash (NLP).



Ключевые слова: Узбекский язык, турецкий язык, омонимия, омонимичные единицы, параллельный корпус, автоматическое выявление, семантическая дизамбигуация, морфологический анализ, контекстуальный анализ, обработка естественного языка (NLP).

Keywords: Uzbek language, Turkish language, homonymy, homonymous units, parallel corpus, automatic identification, semantic disambiguation, morphological analysis, contextual analysis, natural language processing (NLP).

Kirish. Bugungi kunda tabiiy tilni qayta ishlash (Natural Language Processing - NLP) texnologiyalarining rivojlanishi turli tillardagi leksik va semantik noaniqliklarni avtomatik aniqlash masalasini dolzarb tadqiqot yo'nalishlaridan biriga aylantirmoqda. Xususan, bir xil shaklga ega bo'lgan, ammo turli ma'nolarni ifodalovchi leksik birliklarni kontekst asosida farqlash avtomatik tarjima, elektron lug'atlar, axborot qidiruvi, matnlarni avtomatik tahlil qilish va til o'rgatish tizimlari uchun muhim ahamiyatga ega.

Omonimiya hodisasi inson tomonidan tabiiy tilni tushunishda odatda kontekst yordamida avtomatik tarzda hal qilinadi. Biroq kompyuter tizimi uchun ayni bir leksik birlikning qaysi ma'noda ishlatilganini aniqlash murakkab masala hisoblanadi. Masalan, o'zbek tilidagi **ot** birligi turli kontekstlarda hayvon nomini yoki boshqa semantik birlikni ifodalashi mumkin. Shuningdek, turk tilidagi **yüz** shakli kontekstga qarab turli ma'nolarni anglatishi mumkin. Demak, leksik shaklning o'zi yetarli axborot bermaydi; uning atrofidagi kontekst va boshqa lingvistik belgilarni hisobga olish zarur. Mazkur muammo kompyuter lingvistikasida **Word Sense Disambiguation (WSD)**, ya'ni so'z ma'nosini kontekst asosida avtomatik aniqlash masalasi doirasida o'rganiladi.

O'zbek va turk tillaridagi mental omonimiya qatorlariga bag'ishlangan tadqiqotlarda ham mazkur tillarda shakl va ma'no



munosabatlarining o'ziga xosligi ko'rsatilgan. Bu holat omonim birliklarni qiyosiy va kognitiv-semantik jihatdan tadqiq etish uchun nazariy asos yaratadi.

Word Sense Disambiguation jarayonini quyidagi model orqali ifodalash mumkin:

Kontekst + noaniq birlik → tegishli semantik ma'no

Masalan: **U yangi ot sotib oldi.** Bu yerda **ot** hayvon ma'nosida.

Boshqa kontekstda: **Uning oti Ali.** Bu yerda birlik ism yoki nom ma'nosini anglatadi. Demak, kompyuter tizimi **ot** shaklining o'ziga emas, balki uning kontekstual xususiyatlariga e'tibor qaratishi kerak.

Metodologiya. Tadqiqotda **qiyosiy-tavsifiy, korpus, kontekstual, morfologik va semantik tahlil** metodlaridan foydalanish taklif etiladi. Avtomatik disambiguatsiya jarayoni quyidagi bosqichlardan tashkil topadi: **Parallel korpus → preprocessing → tokenizatsiya → lemmatizatsiya → morfologik tahlil → omonim kandidatlarini aniqlash → gaplarni alignment qilish → kontekstual tahlil → turkcha ekvivalentni aniqlash → semantik disambiguatsiya.**

Muhokama. O'zbek-turk parallel korpusida omonimiyani avtomatik disambiguatsiya qilishda eng muhim masalalardan biri - bir xil shakldagi birliklarning kontekstga qarab turli ma'nolarni ifodalashidir. Quyidagi misollar ushbu jarayonni yaqqol ko'rsatadi.

1. "Ot" birligi

1-misol: O'zbekcha: *U yangi **ot** sotib oldi.*

Turkcha: *O yeni bir **at** satın aldı.*

Bu kontekstda **ot** hayvon ma'nosida qo'llangan. Turkcha tarjimada uning **at** shaklida berilishi ushbu ma'noni tasdiqlaydi.

2-misol:

O'zbekcha: *Uning **oti** Ali.*

Turkcha: *Onun **adı** Ali.*



Bu yerda **ot** birlik sifatida ism yoki nom ma'nosida qo'llangan. Turkcha tarjimada **ad** leksemasining egalik shakli — **adı** ishlatilgan.

Demak:

ot → **at** = hayvon

ot → **ad/adı** = ism, nom

Mazkur juftlik parallel tarjimaning disambiguatsiya jarayonidagi ahamiyatini yanada yaqqol ko'rsatadi:

kul → **kül** = modda

kul → **gül** = kulmoq

Natijalar. Tadqiqot natijasida o'zbek-turk tillarida omonim birliklarni avtomatik aniqlash va semantik disambiguatsiya qilishda bir nechta asosiy lingvistik omillar aniqlandi.

Birinchidan, **leksik shaklning o'zi omonim birlikning ma'nosini aniqlash uchun yetarli emas**. Bir xil shakl turli kontekstlarda turli semantik qiymatga ega bo'lishi mumkin.

Ikkinchidan, **morfologik belgi** disambiguatsiyaning muhim komponenti hisoblanadi. Xususan, **yoz, yuz, olma, tut, kul** birliklarida ot va fe'l shakllarining farqlanishi ma'noni aniqlash imkonini beradi.

Uchinchidan, **kontekstual birliklar** omonim ma'noni aniqlashda muhim rol o'ynaydi. Masalan, *ot sotib olmoq, ot/ism, olma yeyish, kitobni olma* kabi birikmalar semantik talqinni aniqroq belgilaydi.

To'rtinchidan, **parallel tarjima ekvivalenti** qo'shimcha semantik signal sifatida ishlatilishi mumkin. Masalan: **ot** → **at**, **ot** → **ad**, **olma** → **elma**, **tut** → **dut**, **kul** → **kül**, **kul** → **gül** kabi mosliklar omonim birliklarning ma'nosini avtomatik guruhlash imkoniyatini beradi.

Beshinchidan, tarjima ekvivalenti har doim ham yetarli emas. **Yoz** → **yaz** kabi holatlarda ikki xil ma'no bir xil tarjima shaklida ifodalanishi mumkin. Shu sababli avtomatik model parallel tarjimani morfologik va kontekstual belgilar bilan birlashtirishi kerak.



Xulosa. Tadqiqotda omonim birliklarning ma'nosini aniqlashda **kontekst, morfologik xususiyat, sintaktik munosabat va parallel tarjima ekvivalentlari** muhim ahamiyatga ega ekanligi asoslandi.

O'zbek tilidagi **ot, yoz, yuz, olma, tut, kul** birliklari misolida olib borilgan tahlil bir xil shakldagi birliklarning kontekstga qarab turli ma'nolarni ifodalashini ko'rsatdi. Ularning turkcha ekvivalentlari esa ayrim hollarda semantik disambiguatsiya uchun qo'shimcha signal bo'lib xizmat qilishi mumkin. Tadqiqotning muhim natijalaridan biri shundaki, o'zbek-turk parallel korpusida omonimiyani avtomatik aniqlashni faqat lug'aviy yoki statistik yondashuv asosida amalga oshirish maqsadga muvofiq emas. **Lingvistik va kompyuter usullarini birlashtirgan gibrid model** samaraliroq yondashuv bo'lishi mumkin.

Taklif etilayotgan model quyidagi ketma-ketlikni nazarda tutadi:

parallel korpus → morfologik tahlil → omonim kandidatlarini aniqlash → kontekstual tahlil → alignment → tarjima ekvivalentini aniqlash → semantik disambiguatsiya.

Foydalanilgan adabiyotlar:

1. Torunoğlu Selamet D., Şentaş A., Eryiğit G. Gamified Crowd-sourcing for Word Sense Disambiguation of Turkish // ACM Transactions on Asian and Low-Resource Language Information Processing. – 2025. – Vol. 24, No. 11. – Article 130. – DOI: 10.1145/3767160.
2. Musurmankulova Sh. Theoretical Basis of the Formation of Uzbek-Turkish Parallel Corpus // Central Asian Journal of Literature, Philosophy and Culture. – 2023. – Vol. 4, No. 12. – DOI: 10.51699/cajlp.v4i12.1113.
3. Khamroeva Sh., Kholmonova I. Structure of the Uzbek-Turkish Parallel Corpus: Linguistic and Extralinguistic Annotation // 2025 10th International Conference on Computer Science and



Engineering (UBMK). – 2025. – B. 468–473. – DOI: 10.1109/UBMK67458.2025.11207056.

4. Xolmonova I.A. O'zbek-turk parallel korpusini yaratishning metodologik yondashuvlari // Buxoro davlat universiteti ilmiy axboroti. – 2025. – № 4 (121).

5. Abdurahmanova M. Ranks of Mental Homonymy in the Uzbek and Turkish Languages // Foreign Linguistics and Linguodidactics. – 2024. – Vol. 2, No. 3. – B. 273–279.

6. Keçeci Ç. A Study on the False Equivalence of Türkiye Turkish and Uzbek Turkish Idioms // Karadeniz International Scientific Journal. – 2024. – № 64. – B. 25–32. – DOI: 10.17498/kdeniz.1567818.