

# TILSHUNOSLIKDAGI ZAMONAVIY YO'NALISHLAR: MUAMMO VA YECHIMLAR



ONLAYN

XALQARO ILMIY-AMALIY

KONFERENSIYA

Andijon - 2020

<http://conference.adu.uz>

## TATBIQIY TILSHUNOSLIK MASALALARI VA KOMPYUTER LINGVISTIKASI

### ЛИНГВИСТИК ДАСТУРЛАР УЧУН ОМОНИМЛАРНИ ТАҲЛИЛ ҚИЛИШ ТЕХНОЛОГИЯСИ

*Абжалова М.А., ф.ф.д., (PhD),  
Навоний давлат кончилиқ институти*

**Аннотация.** Мазкур мақолада ўзбек тилидаги матнларни автоматик таҳрир ва таҳлил қилиш дастури орқали ўзбекча матнларда учрайдиган омоним шаклларни таҳлил қилиш технологияси ҳақида сўз боради. Матнда учрайдиган омошакллар лингвистик корпусга боглиқ ҳолда рус ва хориж тилишунослигида алоҳида тадқиқ қилинади. Ўзбек Миллий корпусининг яратилмаганлиги боис лингвистик дастурлар учун мазкур технология ишлаб чиқилди.

**Калим сўзлар:** лингвистик дастур, омоним, корпус, модел, тег, ID.

**Annotation.** This article discusses the technology for analyzing homonyms found in Uzbek texts using software for automatic editing and analysis of texts in Uzbek. In Russian and foreign linguistics, depending on the language corpus, methods of analysis of homonymous words are studied. This technology was developed for linguistic programs for the analysis of texts of the Uzbek language, in which the Uzbek national corpus has not yet been created.

**Key words:** linguistic program, homonym, corpus, model, tag, ID.

Маълумки, омонимлар ёзилиши бир хил фонемалардан иборат шакллар ҳисобланади. Шу боис омонимияда ифода плани етакчи [1]. Омоним ҳодисада ясама лексема ва туб лексема, шунингдек, уларга грамматик маъно ифодаловчи аффикс кўшилганда воқе бўладиган шаклий тенг келиш ҳам қамраб олинади. Масалан, *терим* сўзшакли *тер* – феъл лексемасига *-им* лексема ясовчисининг бирикиши натижасида ҳосил қилинган лексеманинг бирлик, бош келишиқ формасига тенг.

Аниқроқ бўлиш учун, бир неча турдаги омонимларни ажратиб кўрсатиш мумкин:

- 1) нутқнинг бир қисми ёки луғатдаги шакли (лемма) фарқланади;
- 2) баъзи морфологик хусусиятлар, масалан, келишиқ, эгалиқ ёки сонда (бой морфологияли тилларда мавжуд) фарқланади;
- 3) бир-биридан фақат мазмун жиҳатдан фарқланади (бу ҳолат семантик фарқланишда учрайди).

Тилнинг морфологик сатҳини ўзида қамраган универсал тизим (корпус) мавжуд эмаслиги сабабли кўпгина тилларда турланган сўзшакллари ёки сўз туркумлари доирасидаги сўзларнинг аффикслар билан бирикишида лингвистик таъминотга киритилган морфологик луғатларга мурожаат қилинади. Шу боис турли хил луғатлар учун омонимлик (ноаниқлик) тушунчаси бироз фарқ қилиши мумкин.

Морфологик луғатлар билан бир қаторда корпус, деб аталадиган тўлиқ матнлар тўплами мавжуд [2]. Корпус таъминотининг шакллантирилишига қараб омоним шаклларнинг нафақат морфологик, балки семантик ва синтактик хусусиятлари ҳам намоён бўлади [3]. Ахборот олишнинг қулай имкониятлари яратилиши натижасида корпус тил ҳақидаги мукамал маълумот манбаига айланади, масалан, омоним шаклларнинг энг кенг тарқалган тури статистикасини билиш мумкин.

Бугунги кунда лингвистик тадқиқот ва амалий топшириқлар ечими учун тил корпуслари замонавий тилшуносликнинг инкор этиб бўлмас иш қуролига айланди, чунки корпус турли лингвистик топшириқларни ечишга хизмат қилади [4,5,6].

Корпус муайян электрон манбага лексик, морфологик, грамматик, семантик белгилар асосида ишлов берилганлиги сабабли лингвистик тадқиқотларда фойдаланиш учун жуда қулай имкониятни беради. Ш.Хамроева таъкидлаганидек, тил корпуси электрон кутубхонадан фарқли равишда муайян тилни тадқиқ қилиш, ўрганиш ва ўргатиш учун зарурий, фойдали ва қизиқарли матнлар тўпланишини назарда тутди [5].

Турли корпуслардаги омонимлик турлича бўлади. Бу морфологиянинг нафақат ҳар хил моделларидан фойдаланилишига, балки матнлар услуги ва маълумотлар миқдорига ҳам нисбатан боғлиқдир. Рус тили корпусига бирлаштирилган матнлар таркибида омонимликка эга сўзшакллар кўп учрайди. Шу боис хорижий ва рус тилшунослигида омонимликни бартараф этиш масаласи (“снятия омонимии”) алоҳида тадқиқ қилинади [6,7,8].

Грамматикага оид манбалардан маълумки, омонимияни йўқ қилиш (яъни ноаниқликни бартараф этиш)нинг деярли барча усуллари икки гуруҳга бўлинади [8]:

1. Грамматик меъёрларга асосланган усуллар. Ўз навбатида, улар қуйидагича гуруҳланади:

- технологияларнинг аралашувисиз қўлда киритилган методлар;
- қоидаларни автоматик яратиш усуллари.

2. Статистикага асосланган усуллар.

Ушбу гуруҳларнинг ҳар бирида афзаллик ва камчиликлар мавжуд. Шундай ҳолат бўладики, ҳар иккала гуруҳга тегишли методлар комбинациясидан юзага келган метод яхшироқ натижани кўрсатиши мумкин. Таъкидлаш жоизки, манбалари кўрсатилган изланишларда омонимлик ҳодисаси корпус лингвистикасига боғлаб тадқиқ қилинган. Ваҳоланки, кўплаб тилларнинг миллий корпуси ҳануз яратилган эмас. Шунга биноан Брилл усули, яширин Марков модели, моделлар модификацияси каби методларидан миллий корпусига эга бўлмаган тиллар матнидаги омонимларни таҳлил қилишда қўллаб бўлмайди.

Ўзбек миллий корпусининг яратилмаганлиги эътиборга олиниб, ўзбек тилидаги омоним шаклларни таҳлил қилишда N-gramга боғлиқ равишда омоним-таҳлил технологияси ишлаб чиқилди. Бунинг учун ўзбек тилидаги омонимлар лингвистик таъминотга ID билан киритилди, бу жиҳат контекстда омонимларнинг грамматик шаклланишини таҳлил қилиш имкониятини беради. Таҳлил аниқлигини ошириш мақсадида омонимшаклнинг туркуми аниқланди ва унга рамзий қиймат берилди. Шундан сўнг омонимлар базаси (List of homonym) сўз бирикмалари модели (Models of CompWords)га боғланди (List of homonym=> Models of CompWords). Натижада киритилган M – омоним билан келган N – сўзга боғлиқ равишда ЛТдаги M – омонимнинг ID рақамига тўғри келувчи Q – аффиксларни олади. Яъни, M[id]+Q[id].

Мазкур усулни мисол орқали таҳлил қиламиз:  $енг_1 \rightarrow N/от$  (кийимнинг қўлни қоплаб турадиган қисми),  $енг_2 \rightarrow V/феъл$  (ғолиб келмоқ, устун келмоқ),  $енг_3 \rightarrow V/феъл$  (ейишга ундовчи буйруқ феъли) омонимшакллар матнда синтактик модулда яратилган сўз бирикмалари моделидаги сўз туркумлари рамзларига мос равишда ўзидан кейин келган сўз (биграмма – икки белги кетма-кетлиги) билан синтактик боғланишда қабул қиладиган грамматик воситаси таҳлил қилинади. Демак, бу ерда N ва V билан боғлиқ моделларга мурожаат қилинади:  $N_{\text{com-ни}} + V$

(енгини кесмоқ),  $N_{\text{сом-га}} + V$  (енгига қадамоқ),  $\text{Pron}_{\text{ни(дир)}} + V$  (ҳаммани енгмоқ),  $\text{Pron}_{\text{дан(дир)}} + V$  (мендан енгилмоқ),  $\text{Adj}_{\text{дан}} + V$  (Vh) (кучсиздан енгилмоқ) ва ҳ.к.

Умуман, омоним-тахлил усули муайян омоним сўзга бошқа сўзнинг маълум грамматик кўрсаткич орқали боғланиши билан тушунтирилади ва бу жараён таҳлил имконияти ҳамда таҳлилнинг аниқлигини оширади.

Хулоса қилиб айтганда, дунё компьютер лингвистикасида омонимликни бартараф этиш усуллари ўрганилганида, бу хусусдаги тажрибадан фойдаланиб ўзбекча матнлардаги сўзшаклларнинг тегишли текшириш формуласи яратилди. Омонимликни бартараф этиш учун ҳар бир сўзни “таснифлаш” керак, яъни уни лемма – гап бўлаги ва морфологик хусусиятлар мажмуи билан таққослаш мумкин, улар қулайлик учун бир тегга қўшилади. Барча мумкин бўлган тегларни ўрганиш учун морфологик луғатдаги сўзларга тегишли ҳаволаларни топиш ёки MyStem [9] каби морфологик анализаторни ишлатиш етарли бўлиб, у сўз тегларини топишда ёрдам беради. Шундан сўнг бир нечта теглар орасидан фақат тегишли тегни танлаш керак бўлади.

Омоним сўзшаклларни таҳлил қилишда қўлланилган оптимал лингвистик усул матнларни таҳрир ва таҳлил қилиш, машина таржимаси, матнларни қайта ишлаш жараёнларида муҳим омил бўлади.

#### Адабиётлар

- 1.Раҳматуллаев Ш. Ўзбек тили омонимларининг изоҳли луғати. – Тошкент: Ўқитувчи, 1984. – Б.5.
- 2.[www.dialog-21.ru/media/2138/zakharov.pdf](http://www.dialog-21.ru/media/2138/zakharov.pdf) Захаров В.П. Корпуса русского языка.
- 3.Очиқ рус миллий корпуси сайти – <http://opencorpora.org>.
- 4.Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. – СПб, 2006. – С.26.
- 5.Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Фил.фан. бўйича фалсафа докт. (PhD) дисс. автореф. – Қарши, 2018.
- 6.Кобрицов Б.П. Модели многозначности русской предметной лексики: глобальные и локальные правила разрешения омонимии. Автореф... канд. филол. наук. Москва: РГГУ, 2004.
- 7.Зеленков Ю.Г., Сегалович И.В., Титов В.А. Вероятностная модель снятия морфологической омонимии на основе нормализующих подстановок и позиций соседних слов (Электрон ресурс). [http://www.dialog-21.ru/media/2444/zelenkov\\_segalovich.pdf](http://www.dialog-21.ru/media/2444/zelenkov_segalovich.pdf)
- 8.Рысаков С.В. Методы борьбы с омонимией. <http://samag.ru/archive/article/3059>.
- 9.<http://tech.yandex.ru/mystem> – MyStem морфологик таҳлил дастури сайти.

## ПАРЕМАЛАР ТАРКИБИДА ИШТИРОК ЭТГАН КИЙИМ-КЕЧАК

### МАЪНОСИДАГИ ФУНКЦИОНИМЛАР ҲАКИДА

*Турдиматова Ш.Т., Хужанд давлат университети тадқиқотчиси*

**Аннотация.** Мақолада тил луғавий сатҳ бирликлариаро лисоний муносабатларнинг функционимлар деб аталувчи тури ҳақида фикр юритилиб, уларнинг ўзбек тили паремалари таркибидаги лисоний амали таҳлил қилинади.

**Таянч сўз ва иборалар:** луғавий бирлик, лисоний муносабат(лар), полисемия, омонимия, омофон, антонимия, дублет, функционимлар ва б.

**1.** Анъанавий тилшуносликда луғавий сатҳ бирликлариаро лисоний муносабатлар улараро шаклдошлик ва маънодошлик хусусиятига кўра ўрганилар эди. Шу асосда луғавий бирликлар кўп маъноли(полисемия), омонимия, омограф, омофон(луғавий бирликларнинг шакл муносабатига кўра турлари); синонимия,