

# ЛИНГВИСТ



(Илмий мақолалар тўплами)

I

ЎЗБЕКИСТОН РЕСПУБЛИКАСИ ОЛИЙ ВА ЎРТА МАХСУС  
ТАЪЛИМ ВАЗИРЛИГИ

МИРЗО УЛУҒБЕК НОМИДАГИ ЎЗБЕКИСТОН МИЛЛИЙ  
УНИВЕРСИТЕТИ

# ЛИНГВИСТ

(Илмий мақолалар тўплами)

II

Тошкент  
«Akademnashr»  
2011 йил

ISBN 978-9943-397-27-9



9 789943 397279

**for+ot** (олмош) → учун

**for+пайт равиши** → давомида.

Шунда **for** кўмакчи вазифасида от ёки олмошдан кейин таржима қилинади.

Демак, таржима дастури учун лингвистик таъмин яратишда содда таҳлил босқичларидан бошлаб, мураккаб кўринишларини ҳам кўриб чиқиш керак бўлади. Шунингдек, бошқа грамматик воситалар ҳам шу тарзда ишланса мақсадга мувофиқ деб ҳисоблаймиз.

### **Фойдаланилган адабиётлар:**

1. Raymond Murphy/ English grammar in USE. Cambridge University press 1985.
2. Гаппоров М. Инглиз тили грамматикаси. – Тошкент, 2006.
3. Исоков У. English-Uzbek dictionary. – Toshkent, 2008.

**Абжалова М, магистрант**

### **Лингвистик таҳлил дастурлари ва матнни таҳрир қилиш**

XXI асрнинг илк йилларида тилшунослик соҳасида компьютердан фойдаланиш натижасида янги, замонавий йўналиш бўлган компьютер тилшунослиги юзага келди. Ушбу мақолада берилган маълумотлар компьютер тилшунослиги муаммоларининг бири билан боғлиқ. Бизга аёнки, компьютер тилшунослигининг ўзида ҳам бир неча йўналишлар мавжуд, улар ичида ўзбек тилидаги матнларни компьютер воситасида таҳлил қилиш ва таҳрир қилиш дастурининг лингвистик таъмини ўзига хос ўрин тутди. Мақолада рус, инглиз, немис, француз каби тиллардаги сўзларни, гапларни, матнни компьютерда ишлашга мўлжалланган дастурларда ёзаётганда ёки тармоқ ичида (Интернет) қидирувга берилганда уларни морфологик, синтактик, семантик жиҳатдан таҳлил қилувчи дастурлар борасида маълумотлар берилди. Бундан кўзланган мақсад жаҳон тилшунослиги томонидан тан олинган ва айни дамда фойдаланилаётган асосий лингвистик программалар ҳақида илмий-назарий фактларни, уларнинг ишлаш даражасини, таҳлил ёки таҳрир давомидаги имкониятларини ўрганиш.

Лозанна (Швейцария) университетининг ва Париждаги Олий амалий мактабнинг профессори Муртазо Маҳмудян таъкидлаганидек, «тил, ўз таърифига кўра, инсоният тажрибасининг барча соҳаларини қопламоғи лозим»<sup>1</sup>. Шунинг назарда тутилган ҳолда бугунги кунда илм-фаннинг турли соҳаларида, кундалик ҳаётда компьютердан фойдаланиш жараёнида ўзбек тилида матн ёзилаётганда матннинг тўғридан-тўғри таҳририни таъминлаш масаласи ўзбек тилшуносларининг олдида ечимини қутаётган масала бўлиб турибди.

Ўзбек тилидаги матнларни (биринчи кадамда) Microsoft Office Word дастурига киритиш жараёнида матнни автоматик таҳрир қилишга эришиш, бунинг учун дастурнинг лингвистик таъминини яратиш мақсадида бугунги кунда кенг қўлланишда бўлган дастурлар чуқур ўрганилди. Уларнинг амалий аҳамияти қисқа шарҳ тарзида берилди.

Лингвистик таҳлил дастурлари жаҳон тилшунослигида қуйидагича гуруҳланади:

1. Лингвистик технологиялар ва системалар<sup>2</sup>

2. Матнни лингвистик таҳлил қилиш дастурлари (морфология, синтаксис)

3. Матнни статистик таҳлил қилиш утилитлари

Қуйида фақат биринчи гуруҳ таркибига кирувчи дастурлар борасида лингвистик маълумотлар берилди.

### Лингвистик технологиялар ва системалар

Номи, муаллиф(лар)и	Шарҳлар
AskNet	Семантик савол-жавоб тарзидаги қидирув системаси <u>AskNet</u> ва ишлов бериш инструментлари рус ва инглиз тилларидаги матнларни лингвистик таҳлил қилишни тўлиқ амалга оширади. Лингвистик таҳлил қилиш модули ўз зиммасига <u>морфологияни</u> (лугатли ва лугатсиз), <u>синтаксис</u> , <u>семантикани</u> олади. Текст сарлавҳасининг семантик модули мавжуд.
Интелл Сервис	Программалари махсулотлар тўла версиялар билан корпоративли (уюшмалли), сайтли ва шахсий (махсус) қидирув тизимлари тақлиф қилинган. Интернет бўйича савол-жавобли қидирув метақидирувли

<sup>1</sup> Маҳмудиян М. Лингвистика. М., 1985. с.173.

<sup>2</sup> info@krugosvet.ru

	<p>системалар базасида (<a href="http://www.asknet.ru">www.asknet.ru</a>) амалга оширилади. Ҳозирда аналитик қидирув системаси AQUA ишлаб чиқарилмоқда. Бу тизим матнни автоматик текширишда семантик нуқтаи назардан саволларга жавоб топади. Программалар ва SDK савдо-сотикда кенг ёйилмоқда. Лингвистик таҳлил даражалари: графематик, морфологик, синтактик, семантик.</p>
<p><b>Russian Context Optimizer</b></p> <p>RCO</p>	<p>Рус тилидаги матнларни синтактик ва морфологик жиҳатдан таҳлил қилувчи йиғма (комплексли) дастур. Унинг лингвистик таҳлил модулларига қуйидагилар киради: морфологик луғатлар, графем луғатлар, синтактик қоидалар, семантик классификаторлар, синонимлар луғати. Ишланманинг асослари морфологик ва фактографик таҳлил модуллари томонидан татбиқ этилган. Морфологик луғатнинг 15000 лексема ҳажмисидан фойдаланилган. Дастурлар Windows, Oracle, SDK учун ишлаб чиқилган ва у Windows ҳамда Unix платформаларида ишлайди. Дастурлар ва SDK савдо-сотикда тарқалмоқда. Лингвистик таҳлил даража-лари: графематик, морфологик, синтактик, семантик.</p>
<p><b>Ontos</b></p> <p>Ontos</p>	<p>Ушбу дастур матнли ҳужжатларни таҳлил қилишга, аннотациянинг тузилишини текширишга, келтирилганларга (<u>OntosMiner</u>, <u>LightOntos for Workgroups</u>, <u>Ontos SOA</u>, <u>TAIS Ontos</u>) ишлов беришга мўлжалланган. Дастур алгоритмлари графематик, морфологик ва семантик таҳлилга қаратилади. Системалар морфологик луғатлар ва семантик картрижлардан асосий тиллар (инглиз, немис, француз, рус) учун фойдаланади. Матнда йўл қўйилган хатоларнинг фактик тўғри жавоблари (вариантлари) кўрсатилади, графаларнинг боғланиши бериледи. Программалар Java.net да амалга оширилади ва улар Microsoft SQL Server да ишлатилади. Лингвистик таҳлил даражалари: графематик, морфологик, семантик — тематик синонимлардан фойдаланиш даражасида.</p>
<p><b>Google Mini</b></p> <p>Google Desktop</p>	<p>Дастурли аппарат комплекси <u>Google Mini</u> ва <u>Google Desktop</u> программалари тармоқ ичида, яъни Интернетда компьютер воситасида хабарларни қидиришга мўлжалланган. Асосий тилларда (рус, инглиз ва бошқ.) калит сўзлар бўйича морфологик</p>

<sup>3</sup> www.asknet.ru

Google	жиҳатдан текширувни амалга оширади. Лингвистик таҳлил имкониятлари: графематик, морфологик.
Yandex.Server Яндекс	Дастур (HTTP) сервери учун мўлжалланган, бир ёки бир неча сайтлар бўйича қидирувни ва рақамлашни амалга оширади. Қидирув рус, инглиз ва украин тилларининг морфологияси бўйича ишлайди. Қидирув натижаси ҳужжатлар рўйхати релевантлик ёки сана асосида тартибланиши ҳисобланади. Релевантликда ҳужжатдаги сўзларнинг миқдори, уларнинг частотали характеристикаси ва сўзлараро жойлашиши инобатга олинади. Матнларнинг синтактик ва семантик таҳлили ҳисобга олинмаган. Дастур версиялари Windows ва Unix учун мўлжалланган. Лингвистик таҳлил даражалари: графематик, морфологик.
Galaktika-ZOOM Галактика	Программа ҳужжатнинг маънога эга сўз ва сўз бирикмаларини аниқлайди, фойдаланувчи томонидан киритилган калит сўзлар ҳамда уларнинг синонимлари юзасидан текширади, қидирув ишини олиб боради. Шунингдек, ҳужжатда учрайдиган сўзларнинг частотаси бўйича ҳисобни формаллаштиради. Дастур рус тилидаги матнларнинг таҳлилини таъминлайди. Алгоритмлар морфологик ва статистик анализдан фойдаланишга асосланади. Лингвистик таҳлил даражалари: графематик, морфологик.
SyTech Арион	Система бир типли илмий матнларга аналитик жиҳатдан ишлашни автоматлаштиришни таъминлайди ва ҳисоботларни ҳар қил кўринишда тақдим этади. Аммо дастур фактографик хабарлар билан боғланган аниқ хулосаларни бермайди ва матнни синтактик ва семантик таҳлил қилмайди. Программа қурилма тизимини янги жойга ўрнатишда кўп вақтни ва қўл меҳнатини талаб этади, шунингдек, системага киритилмаган номаълум сўзларни автоматик равишда топиш имкони йўқ. Система табиий тиллар доирасида ишлашга ноқулай. Лингвистик таҳлил даражалари: графематик, морфологик.
Кластерловчи кидирув тизимлари	Метақидирувли тизимлар кидирув натижаларининг кластерланиши билан бирга фойдаланувчи

<p>Visimo Nigma<sup>4</sup></p>	<p>томонидан қидиришга берилган сўзларни матн ичидан ҳам топиш имконига эга. Аммо фақат кластер анализдан фойдаланиш ушбу типнинг метақидирувли тизимларига мавжуд афзалликларни бермайди. Қидирув сифатини ошириш, айниқса, табиий тиллардаги сўровларга ишлов беришни фақат синтактик ва семантик таҳлилдан фойдаланиш асосида амалга ошириш мумкин. Лингвистик таҳлил даражалари: графематик, морфологик.</p>
<p>Quintura Searchcrystal<sup>5</sup></p>	<p>Қидирув натижаларини график тарзда тақдим қилувчи метақидирувли система. Қидирув бир неча қидирув тизими бўйича амалга оширилади. Натижалар статистик мезонлар асосида кластерланади. Дастур морфологик таҳлил қилишдан фойдаланади. Масалан “юувчи” сўзи қидирувга берилганда Quintura кластерга “ажойиб энага” сўзини киритади, унинг ўзаги билан боғлиқ “юваётиб” сўзини ва қидириляётган “юувчи” сўзини беради.</p>

#### Фойдаланилган манбалар:

6. Апресян Ю.Д. Идеи и методы современной структурной лингвистики. – М., 1966.
7. Баранов А.Н. Введение в прикладную лингвистику. – М.: Эдиториал УРСС, 2001. – 357 – С.
8. Беляева Л.Н., Герд А.С., Убин И.И. Автоматизация в лексикографии // Прикладная лингвистика. - СПб., 1996.
9. Беляева Л.Н., Откупщикова М.И. Автоматический (машинный) перевод // Прикладная лингвистика. - СПб., 1996.
10. Исаев И.А. Опыт автоматизации лексикографических исследований. Система DIALEX // Слово Достоевского. – М., 1996.
11. Кибирк А.Е. Прикладная лингвистика // Очерки по общим и прикладным вопросам языкознания. – М., 1992.
12. Новое в зарубежной лингвистике. Вып. XXIV. – (Компьютерная лингвистика). – М.: Прогресс, 1989.
13. Панков И.П., Захаров В.П. Информационно-поисковые системы // Прикладное языкознание. – СПб., 1996.

<sup>4</sup> www.wikipedia.ru

<sup>5</sup> http://www.dialog-21.ru/Archive/2004/Koval.htm/