

KOMPYUTER LINGVISTIKASIDA OMONIMLIK HODISASI TADQIQI

Alijonova Fayruza

Mirzo Ulug‘bek nomidagi O‘zbekiston Milliy universiteti

I bosqich magistratura talabasi

Email: fayruzaalijonova@gmail.com

Abdulahobov G‘iyosiddin

Farg‘ona Davlat universiteti o‘qituvchisi

Email: abduvaxobov2018@gmail.com

ANNOTATSIYA

Maqolada o‘zbek tili elektron korpusida omonimlar bilan ishlash jarayonida yuzaga keladigan muammo va kamchiliklarni aniqlash, ularning kelib chiqish sabablarini o‘rganish, jahon va o‘zbek kompyuter lingvistikasida omonimlar bilan bog‘liq tadqiqotlar va ularning tahlili yoritilgan. O‘zbek tili elektron korpus menejeri uchun omonimlar bilan bog‘liq masalalarni o‘rgangan holda ularning dasturiy bazasini ishlab chiqish korpus menejeri uchun muhim masalalardan biri hisoblanadi. Korpus uchun omonimlarning dasturiy ta‘minotini ishlab chiqish jarayonida amalga oshirish lozim bo‘lgan vazifalar va ularning yechimlari maqolada keltirib o‘tilgan.

Kalit so‘zlar. Korpus, baza, model, token, lug‘at, modellashtirish, algoritim, qidiruv interfeysi

STUDY OF THE PHENOMENON OF HOMONYMY IN COMPUTATIONAL LINGUISTICS

ABSTRACT

The article describes the problems and shortcomings in the process of working with homonyms in the electronic corpus of the Uzbek language, the study of their causes, research and analysis of homonyms in world and Uzbek computer linguistics. One of the most important issues for the Uzbek language electronic corpus manager is to develop their software base by studying the issues related to homonyms. The tasks to be performed in the process of developing homonymous software for the corpus and their solutions are given in the article.

Keywords. Corpus, base, model, token, dictionary, modeling, algorithm, search interface

ИССЛЕДОВАНИЕ ЯВЛЕНИЯ ОМОНИМИИ В КОМПЬЮТЕРНОЙ ЛИНГВИСТИКЕ

АННОТАЦИЯ

В статье описаны проблемы и недостатки в процессе работы с омонимами в электронном корпусе узбекского языка, изучения их причин, исследования и анализа омонимов в мировой и узбекской компьютерной лингвистике. Одним из наиболее важных вопросов для руководителей электронных корпусов узбекского языка является развитие их программной базы путем изучения вопросов, связанных с омонимами. В статье приведены задачи, которые предстоит решать в процессе разработки одноименного программного обеспечения для корпуса, и пути их решения.

Ключевые слова. Корпус, база, модель, токен, словарь, моделирование, алгоритм, поисковый интерфейс

Matnlar korpusi («corpus» – lotincha «tana» degan ma’noni anglatadi) – elektron holda saqlanadigan ma’lum til birliklari bo‘lib, ular tilshunoslar uchun turli xil muammolarni hal etish uchun tatbiq etishda va turli yo‘nalishdagi tadqiqotlar uchun zaruriyatga qarab turli shakllarda tuziladi. Til birliklarining qay tarzda saqlanishiga qarab maxsus dasturlar yordamida har bir kerakli so‘z yoki so‘z birikmasi uchun darhol uning qo‘llanishi bo‘yicha misollar topilishi, imlo bo‘yicha variantlari, sinonimik qatorlari topilishi mumkin. Matnlar korpusiga oid ilmiy tadqiqotlar salmog‘ining ko‘payishi natijasida tilshunoslikda korpus lingvistikasi yo‘nalishi shakllandi. Kompyuterda yaratilgan birinchi matnlar korpusi Braun korpusi (БК, inglizcha Brown Corpus, BC) hisoblanadi, u 1961-yilda Braun universitetida yaratilgan, har biri 2000 so‘zli 500 ta matn fragmentini o‘z ichiga oladi. 1970-yillarda 1 mln so‘zni o‘z ichiga olgan matnlar korpusi asosida rus tilining chastotali lug‘ati yaratildi. 1980-yillarda Shvetsiyaning Upsala universitetida ham rus tilida matnlar korpusi yaratildi. Buyuk Britaniyada Ingliz tili Banki (Bank of English) hamda Britaniya Milliy Korpusi (British National Corpus, BNC), Rossiyada Rus tilining mashina fondi (Машинный фонд русского языка) hamda Rus tilining Milliy korpusi (Национальный корпус русского языка) loyihalari ishlab chiqildi. Masalan, Rus tilining Milliy korpusi hajmi hozirgi kunda 149 mln so‘zdan iborat. Keyingi yillarda internet tizimining rivojlanishi virtual matnlar korpusining yuzaga kelishiga olib keldi. Ya’ni internetdagi qidiriv saytlari, elektron kutubxonalar, virtual ensiklopediyalar korpus vazifasini bajarmoqda. Korpusning janri va tematik rangbarangligi internetdan foydalanuvchining

qiziqishlariga bog‘liq. Masalan, ilm-fan doirasida Wikipedia katta hajmdagi matnlar korpusi sifatida ishlatilmoqda¹⁴

Tilshunoslikning ilk va o‘rta bosqichlarida omonimiya hodisasining paydo bo‘lishi, taraqqiyoti va yondosh hodisalarga munosabatlari o‘rganilgan bo‘lsa, zamonaviy lingvistikada uning korpusda berilishi masalasi kun tartibiga qo‘yildi. Bu vazifani bajarish uchun mavjud lug‘atlar axborot banki vazifasini o‘taydi. Korpus lingvistikasida omonimlik muammosini yechish, omonim birliklarni teglash va matnni avtomatik o‘qish jarayonida omonimiyani bartaraf etish masalasi (“snyatiye omonimii”) ga oid qator tadqiqotlar vujudga kelgan. Jumladan, G. I. Kustova, O. N. Lyashevskaya, Y. V. Paducheva, Ye. V. Raxilina, B. P. Kobritsov, T. I. Reznikova, V. V. Kukanova, A.A. Kretovlar shu masalalar yechimiga bag‘ishlangan qator ishlarni e‘lon qilishgan¹⁵.

Hozirda 50 milliondan ortiq so‘zni o‘zida jamlagan, Nilufar Abduraxmonova va uning ilmiy jamoasi tomonidan tadqiqot ishlari olib borilayotgan <http://uzbekcorpus.uz/> o‘zbek tili uchun nisbatan mukammal korpus hisoblanadi. 2018-yildan buyon ushbu korpus ustida tadqiqot ishlari amalga oshirib borilmoqda. Korpus lingvistikasiga doir tadqiqot ishlariga qaramay korpus hali ham o‘zining kamchiliklariga ega, xususan, interfeysda omonimlar bog‘liq muammolar hali to‘laligicha o‘z yechimini topmadi. Ushbu maqolada ayrim kamchiliklar va ularning yechimlari yuzasidan fikr yuritiladi.

Lug‘atlar bilan ishlash, so‘zlarni morfoanaliz qilish, tokenlarga ajratish jarayonida omonimlar bilan bog‘liq muammolar yuzaga keladi. Masalan, qidiruv jarayonida foydalanuvchi qidiruv interfeysiga *tur* so‘zini kiritisa, asosi *tur* bo‘lgan barcha so‘zlar topiladi. Ammo, ushbu so‘z turli xil ma‘nolarni ifodalishini hisobga olsak, foydalanuvchi uchun bu holatda qiyinchilik tug‘iladi.

Shularni hisobga oladigan bo‘lsak, korpus menejeri uchun omonimlarning mukammal bazasini va to‘g‘ri ishlash algoritmini yaratish to‘g‘risida yechimlar xususida fikr yuritimiz.

Kompyuter intellektual sistemasidagi omonimiya masalasini tadqiq etish uchun quyidagi vazifalarni amalga oshirish lozim bo‘ladi:

- o‘zbek tilidagi badiiy asarlardan, izohli lug‘atlardan va omonimlar lug‘atidagi omonim so‘zlarni ajratib olish;
- omonimlarning tarqalish darajasini aniqlash;
- omonimiyaning sistem xarakterini ochib berish;
- omonimlarning simmetriya/asimmetriya (disimmetriya) jihatlarini ko‘rsatib berish;
- omonimlarni so‘z turkumlariga ko‘ra tasniflash;

¹⁴ Rahimov A. Kompyuter lingvistikasi asoslari. Toshkent: Akademnashr, 2011, . –B 86.

¹⁵ Gulyamova Sh. O‘zbek tili semantik analizatorining lingvistik asoslari. Aftoref. Fil.f.d..., Farg‘ona 2022. – B. 21.

- matnlarni qayta ishlashda omonimlarni saralash;
- omonimlar matritsasini tuzish;
- omonimlarni modellashtirish.

Yuqoridagi vazifalarni amalga oshirish orqali, birinchidan, o‘zbek tilshunosligidagi omonimiyaning ilmiy-nazariy masalasiga aniqlik kiritiladi, ikkinchidan, o‘zbek tilining kompyuter sistemasidagi omonimiya bazasi yaratiladi. Uchinchidan, o‘zbek tilini chet tili sifatida o‘qitish, lug‘atlar tuzishda metodik tavsiyalar ishlab chiqiladi. Avval barcha omonim so‘zlar, keyin ularning izohi ro‘yxatga kiritiladi. Anglatayotgan ma‘nolari bo‘yicha so‘zlar tartiblab chiqiladi.

Masalan: ot

- 1) *ot (hayvon) – tulpur – horse;*
- 2) *ot (kishining nomi) – ism – name;*
- 3) *ot (ish-harakat) – uloqtirmoq – throw;*
- 4) *ot (so‘z turkumi) – so‘z turkumi – noun¹⁶.*

Matnlar bilan ishlaganda omonimlar izohli lug‘atida qayd etilmagan ba’zi omonimlik hodisasi bilan bog‘liq so‘zlar ham muammoli holatlarni keltirib chiqaradi. So‘zlarga morfemalarning qo‘shiluvi bilan yuzaga chiqadigan omonimlik bu jihatdan fikrimizga isbot bo‘ladi. Masalan,

qor + tushum kelishigi = qorni // qorin + egalik qo‘shimchasi = qorni;
yaralmoq + sifatosh = yaralar // yara + ko‘plik qo‘shimchasi = yaralar.

Bu holatda asosga qo‘shimchalar qo‘shilishi natijasida omonimlik yuzaga kelmoqda. Ushbu holatda so‘zlarning turkumlarini va grammatik shakllarini aniqlash orqali omonimlar bir-biridan farqlanadi. Omonimlar bilan bog‘liq keyingi holatni ko‘rib chiqamiz:

o‘qi(fe‘l) 0 grammatik shaklda;
o‘qi=o‘q(ot) +i (egalik qo‘shimchasi).

Bu holatda omonimlik asos va qo‘shimcha qo‘shilishi natijasida yuzaga kelmoqda.

Matnning bosh harflar masalasi ham muayyan muammoni yuzaga keltiradi.
Sen bilmaysan, Asal judayam ajoyib. (1) Biznikiga karimalar kelishdi (2). Bizga ma‘lumki, Yer o‘z o‘qi atrofida aylanadi. Uyga kelib qaragam, dadam yerga bosh qo‘yib yotgan ekan (3)

1-misolda asal so‘zi ismni anglatyaptimi yoki shirinliknimi? 2-misoldagi karimalar so‘zi Karima ismli qiz bilan birgalikdagi shaxslarni anglatyapti, ammo avtomatik tahrir dasturi ism kichik harf bilan yozilgan deya so‘zni xato hisoblaydi. 3-misolning 1-gapida Yer sayyorasi, 2-gapda esa ona zamin haqida fikr yuritilgan. Lekin sintaktikoldi tahrir jarayonida 1-gapdagi Yer so‘zi Yer sayyorasi tarzida

¹⁶ Po‘latov A. Q. Kompyuter lingvistikasi. – Toshkent: Akademnashr, 2011. –B. 21 –22

yoziilmagani uchun ushbu soʻzni ona zamin haqida fikr yuritilayotganini inobatga olib, xato sifatida koʻrsatadi. Bu holda dasturning maʼlumotlar bazasiga quyidagi qoidalarni kiritish talab etiladi:

1. Turdosh otlar atoqli otlar vazifasida ham keladi yoki aksincha: turdosh ot atoqli ot. Ushbu qoidani muqim boʻlishi uchun morfologik tahrir modullarida oʻzbek adabiy tilidagi turdosh va atoqli otlar guruhlari toʻliq va aniq koʻrsatilishi lozim.
2. Atoqli otlarda shakl yasovchi qoʻshimchasi -lar qoʻshilganda atoqli otlar bosh harf bilan emas, kichik harf bilan yoziladi. Atoqli ot+lar= turdosh ot.
3. Har bir uslubning oʻziga xos xususiyatlari dastur taʼminotiga kiritiladi. Kiritilayotgan maʼlumotlar asosida yozilayotgan gap qaysi uslubga xosligi avtomatik tarzda aniqlanadi, natijada gapda qatnashayotgan soʻzlar imlosi tekshiriladi.¹⁷

Parallel korpus bilan ishlash jarayonida tarjimada omonimlar bilan bogʻliq birmuncha xato va kamchiliklar ham kuzatiladi. Birgina rus tilida koʻplab soʻz shakllari grammatik omonimlardir. Masalan, *bake* shakli infinitivda ham feʼl (*bake pies*) va nominative holatdagi ism (*Russian bake*). Omonimlarning qaysi biri maʼlum bir misolda ifodalanganini aniqlash uchun kontekstni tahlil qilish kerak. Korpus soʻz shakllarini grammatik belgilash avtomatik ravishda grammatik omonimiyani ajrata olmaydigan dastur yordamida amalga oshiriladi. Pechka shakllari dastur tomonidan omonim sifatida tavsiflanadi, yaʼni ularga bir vaqtning oʻzida bir nechta grammatik xususiyatlar beriladi.

Pech shaklidan har qanday foydalanish bunday korpusda feʼl sifatida ham, ism sifatida ham koʻrib chiqiladi va ushbu soʻrovda ham, boshqasida ham nashrga joylashtiriladi. Shunday qilib avtomatik ravishda etiketlangan korpusda ishlayotganda faqat soʻzsiz soʻzlarni oʻz ichiga olgan pechka soʻzi bilan misollar roʻyxatini olishning imkoni yoʻq, garchi siz grammatik xarakteristikasini “feʼl” ni oʻrnatgan boʻlsangiz ham. Foydalanuvchi misollar qatoridan infinitive bilan kerakli kontekstni qoʻlda tanlashi kerak boʻladi.

Bundan tashqari, chiqarilmagan omonimiyaga ega korpusda lugʻatda boʻlmagan soʻzlar uchun grammatik farazlar avtomatik ravishda hosil boʻladi. Shu sababli, RNC omonimiyasi olib tashlangan matnlardan misollarni izlash imkoniyatini beradi. Asosiy korpusning ushbu qismida har bir omonim shakl uchun tilshunos dasturiy tahlilga tegishli xususiyatlar toʻplamidan faqat bittasini tanlaydi, yaʼni berilgan kontekst uchun toʻgʻrisini.

Oʻzbek tilidagi matnlarda omonimlarni farqlashda kodlardan, simvollardan, turkumni bildiruvchi teglardan foydalanish talab etiladi. Buning uchun izohli

¹⁷ Poʻlatov A. Q. Kompyuter lingvistikasi. – Toshkent: Akademnashr, 2011

lugʻatlarda foydalaniladigan rim raqamlaridan teg sifatida foydalanish mumkin. Faqat rim raqamlarini qatʻiy tarzda belgilab olish kerak boʻladi:

I – ot, II – fʻel, III – sifat, IV – ravish.

Oʻzbek tilida ot-feʻl omonimligi koʻpligi darajasidaga koʻra birinchi oʻrinda turadi (tut I- daraxt, tut II – harakat). Ot-sifat omonimligi (chanqoq I “tashnalik” – chanqoq III “tashna”) yoki feʻl-sifat omonimligi (yot II – harakat, yot III – “begona”) keying oʻrinta turadi. Ot-ravish omonimligi uchramaydi. Sifat-ravish-feʻl omonimligi kam uchraydi: tik II harakat, tik III “qiya”, tik IV “toʻppa-toʻgʻri)

oshiq I - “ishq sohibi”;

oshiq I – “tana qismi”;

oshiq I – “oʻyin nomi”;

oshiq II – “shoshmoq”;

oshiq III – “ortiq”.

Oshiq omonimlari ishtirok etgan matnlarni kompyuter dasturi (google translater) da tarjima qilib koʻramiz:

1. *Oyqiz bu qizga uzoq tikilgach, uni ashula va dutor oshigʻi deb oʻyladi.*

Долгое время уставившись на девушку, медведь подумал, что она влюблена в пение и дутар.

Bu misolda “yoqtirmoq” maʼnosidagi oshiq boʻlmoq qoʻshma feʻli tarjimon dasturning omonimni farqlamasligi natijasida “sevmoq” maʼnosidagi soʻzga oʻzgartirilgan.

2. *Endi hovlilarning eshiklari ochilgan, bolalar choʻntaklarini yongʻoq, oshiq'larga toʻldirib, koʻchaga chiqib boshlagan edi.*

Теперь двери дворов были открыты, и дети стали выходить на улицы, наполняя карманы орехами и любовниками¹⁸.

Yuqoridagi misollardan koʻrinib turibdiki, avtomatik tarjima jarayonida ham omonimlar bilan bogʻliq muammolar talaygina. Omonimlar modelini ishlab chiqish va uni tarjima jarayoniga oʻzlashtirish hal qilinishi kerak boʻlgan masalalarni birmuncha kamaytiradi. Bu holatda omonimlar oladigan valentliklarni bazasini yigʻish bu masalani yechishga yordam beradi, yaʼni oshiq soʻzi misolida koʻradigan boʻlsak, oʻyin maʼnosida kelayotgan “oshiq” soʻzi “oʻyin” soʻzini valentlik sifatida qabul qiladi, yoki tana aʼzosi sifatidagi “oshiq” soʻzi tana aʼzolari bilan bogʻliq boʻlgan soʻzlarni valentlik sifatida qabul qiladi. Albatta, bu muammoni hal qilish hali ham asosiy vazifalar qatorida turmoqda.

Oldin taʼkidlanganidek, semantik analizatorning omonimlarni farqlashi uchun dastlab ularning lingvistik filtri va modellari ishlab chiqilishi lozim. Bu esa lingvistik bilimni, yondashuvni talab etadi. Zotan, tabiiy tilni qayta ishlash (NLP) sunʼiy aql

¹⁸ A. Rahmanova. Oʻzbek tili milliy korpusini yaratishda kompyuter usullari. Avtoreferat. Andijon 2022. B - 18

sintezi va matematik tilshunoslikning eng muhim va dolzarb yo‘nalishlaridan biridir. Bu o‘rinda asosiy harakat insonning tabiiy tilida ifodalangan matnni tahlil qilishi va kompyuterning ushbu ma‘lumotni “idrok etish”, qayta ishlash qobiliyatini topish hamda takomillashtirishga qaratilgan. Tabiiy tilni qayta ishlash – bu sun‘iy intellekt bilan bog‘liq vazifa, chunki u atrofdagi dunyo va inson hissiyotlarini yuqori darajada tushunishni o‘z ichiga oladi. Shubhasiz, ushbu muammolarni hal qilish tadqiqotchidan tilshunoslik hamda sun‘iy intellekt sohasida yetarlicha bilimni talab qiladi. Tadqiqot uchun Sh. Rahmatullayevning “Omonimlarning izohli lug‘ati” hamda “O‘zbek tilining izohli lug‘ati” (5 jildli) da mavjud omonimlar lingvistik baza bo‘lib xizmat qildi. Ularni turkumiga asoslanib, quyidagicha guruhlab chiqdik: 1) bir so‘z turkumi doirasidagi omonimlar; 2) ot // fe‘l hamda sifat // fe‘l turkumi orasida kuzatiladigan omonimlar; 3) turli turkum doirasida vujudga keladigan omonimlar. Ot // fe‘l hamda sifat // fe‘l turkumi doirasida uchraydigan omonim so‘zlar, asosan, morfologik omil asosida farqlanadi. MO bir qator morfologik vositalar orqali yuzaga chiqadi. Omonim so‘zni farqlashda morfologik va sintaktik omillar muhim o‘rin tutadi. Ot va fe‘l orasidagi omonimlikda morfologik; sifat va fe‘l orasidagi omonimlikda morfologik-sintaktik; faqat ot va faqat fe‘l turkumi doirasidagi omonimlikda sintaktik-morfologik; faqat sifat turkumi doirasidagi omonimlikda sintaktik; turli turkum doirasidagi omonimlikda esa morfologik, sintaktik, morfologik-sintaktik va sintaktik-morfologik omil yetakchilik qilishi aniqlandi.¹⁹

Yuqoridagi muammoli holatlarni hal etish uchun omonimlarning grammatik bazasi va modeli ishlab chiqiladi, ya‘ni omonimlarning turkumiga ko‘ra ular oladigan grammatik qo‘shimlar bazasi yig‘iladi. Yig‘ilgan baza va model asosida algoritm tuziladi.

Ushbu algoritm quyidagicha ko‘rinishda bo‘ladi:

- foydalanuvchi qidirilayotgan so‘zni interfeysga kiritadi;
- oynada maxsus so‘rovlar chiqadi, ya‘ni so‘zning turkumi, agar so‘z ot turkumida bo‘lsa, sanaladigan yoki sanalmaydigan; fe‘l bo‘lsa, o‘timli yoki o‘timsiz; tub yoki yasama(bog‘lama-bog‘lama)

So‘rovlar asosida qidirilganda foydalanuvchi aynan o‘ziga kerakli bo‘lgan so‘zni topish va ajratib olish imkoniyatiga ega bo‘ladi.

Access dasturi yordamida yig‘ilgan omonimlar bazasi quyidagi jadvalda ko‘rsatilgan. Jadvalda omonimlarning turkumi, izohi, badiiy asar va ilmiy matnlardan olingan misollari ko‘rsatilgan (Ushbu jadval SH. Rahmatullayevning “Omonimlarning izohli lug‘ati” asosida yig‘ildi).

¹⁹ Sh. Gulyamova. O‘zbek tili semantik analizatorining lingvistik asoslari. Avtoreferat. Farg‘ona 2022. B - 22

Код	Suz_uzbek	suz_turkum	izohi	misol
233	devon	Noun	[arabcha >ko'pligi	Devonda yonboshlab yotgan Abbosxon Saidiyning hurmat
234	devon	Noun	eski. Xon (amir) s	Devond a hech kim yo'q.[Majiddin] Abusaid Mirzo davri
235	dim	Noun	[<dam I3] ot. Suku	Og'zingizni parshion qilmay sabr qilib tursangiz, shu hafta
236	dim	Adj.	Sernam va issiq (h	Dimlamoq, dimikmoq.Bugun havo ertalabdan dim (dam).
237	dim	Adv.	folk. Mutlaqo, but	Dim ularga qaray ko'rma, chirog'im.«Ravshan».
238	differensial	Noun	[latyncha differen	1.
239	differensial	Noun	Matematika term	
240	differensial	Noun	Texnika termini. A	
241	differensial	Adj.	[<ruscha differen	Differensial geometriya. Differensial tenglama.
242	dov	Noun	Qimorda katta tik	Dov ketmoq (Qiyoslang: katta ketmok,).dov
243	dov	Adj.	[<dov I] sifat.Qo'r	Dov yurak.Ey qahramon, dov yurak, jasur ulug' xalq, Qudra
244	dov	Noun	[<dov II] ot.Dadill	Yuragim dov bermadi.[Hojiya]anchagacha yuragi dov: ber
245	dov	Noun	Talab, da'vo	(sh)moq.Unga hechqanday dovim yo'q. Menga nima dovir
246	dov	Adj.	Bo'ychan, katta	Dov yigit bo'lib qolibsan.[Yo'lchi]sersoqol, dov askar bilar
247	dov	Noun	dov-daraxt(o'siml	Hamroqul naridan-beri d o v-d astgohini yig'ishtirib jo'nac
248	dovul	Noun	Kuchli bo'ron	Dovulda qolmoq. Dovul ko'pchilik imoratlarningtomini uc
249	dovul	Noun	[<tojikcha davl< a	Dovulchi (dovul chaluvchi shaxs). Dov ul har vaqt mirshab
250	doira	Noun	arabcha — «aylan	Doira chizmoq. Yarim doyra.Doirasini hisoblab chiqmoq.C
251	doira	Noun	Dumaloq gardishg	Doirachi.Doira chalmoq. Doira chertmoq.Boloxonaning to
252	dol	Noun	Bolalarning o'yini	.«Kim do I?» — «Men dol».
253	dol	Noun	[arabcha a harfi]o	Bu so'z dol bilan e mae, zol bilan yozida-di. Qaddim dol b
254	dol	Adv.	ravish; eskirgan; s	Nishonning dol o'rtasiga tegizmoq. Dol vaqtida kelmos.Ol

Access dasturi asosida yaratilgan so'rovlar yordamida omonimlarni turkumlari bo'yicha ajratib olib, ularning modellari tuziladi.

Yuqoridagi vazifalar bajarilganda korpusda omonimlar bilan bog'liq muammolarning asosiy qismi bartaraf etiladi. Muammolar to'laqonli bartaraf etilishi uchun tilshunoslar va dasturchilar o'zaro hamkorlikda ish olib borishlari maqsadga muvofiq. Omonimlar bilan bog'liq masalalarni hal qilish kelajakda korpusning kamchiliklar va xatolarsiz ishlashini ta'minlaydi. Bu korpus, umuman kompyuter lingvistikasi, sun'iy intellekt rivoji uchun qo'yilgan eng muhim qadamlardan bo'lishi shubhasiz.

FOYDALANILGAN ADABIYOTLAR

1. Po'latov A. Q. Kompyuter lingvistikasi. – Toshkent: Akademyashr, (2011).
2. Rahimov. A. Kompyuter lingvistikasi asoslari. «Akademyashr» nashriyoti, (2011).
3. Gulyamova. Sh. O'zbek tili semantik analizatorining lingvistik asoslari. Avtoreferat. Farg'ona, (2022).
4. Rahmanova. A. O'zbek tili milliy korpusini yaratishda kompyuter usullari. Avtoreferat. Andijon, (2022).
5. N. Abdurakhmonova, U. Tuliyevev and A. Gatiatullin, "Linguistic functionality of Uzbek Electron Corpus: uzbekcorpus. uz," 2021 International Conference on Information Science and Communications Technologies (ICISCT), 2021, pp. 1-4, doi:

10. 1109/ICISCT52966. 2021. 9670043. Aripov, M. , Sharipbay, A. , Abdurakhmonova, N. , Razakhova B. : Ontology of grammar rules as example of noun of Uzbek and Kazakh languages. In: Abstract of the VI International Conference “Modern Problems of Applied Mathematics and Information Technology - Al-Khorezmiy 2018”, pp. 37–38, Tashkent, Uzbekistan (2018)
6. Abduraxmonova, N. Z. "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) il dis. aftoref. " (2018).
7. Abdurakhmonova N. The bases of automatic morphological analysis for machine translation. *Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta*. 2016;2 (38):12-7.
8. Abdurakhmonova N, Tuliyeu U. Morphological analysis by finite state transducer for Uzbek-English machine translation/*Foreign Philology: Language. Literature, Education*. 2018(3):68.
9. Abdurakhmonova N, Urdishev K. Corpus based teaching Uzbek as a foreign language. *Journal of Foreign Language Teaching and Applied Linguistics (J-FLTAL)*. 2019;6(1-2019):131-7.
10. Abdurakhmonov N. Modeling Analytic Forms of Verb in Uzbek as Stage of Morphological Analysis in Machine Translation. *Journal of Social Sciences and Humanities Research*. 2017;5(03):89-100.
11. Kubedinova L. Khusainov A. , Suleymanov D. , Gilmullin R. , Abdurakhmonova N. First Results of the TurkLang-7 Project: Creating Russian-Turkic Parallel Corpora and MT Systems. *Proceedings of the Computational Models in Language and Speech Workshop (CMLS 2020) co-located with 16th International Conference on Computational and Cognitive Linguistics (TEL 2020)*. 2020/11: 90-101
12. Abdurakhmonova N. Dependency parsing based on Uzbek Corpus. In *Proceedings of the International Conference on Language Technologies for All (LT4All) 2019*.