

ISSN 2181-922X

LANGUAGE & CULTURE

UZBEKISTON O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER
LINGVISTIKASI

2024 Vol. 2 (6)

www.compling.tsuull.uz

ISSN 2181-922X

O‘ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2024 Vol. 2 (6)

compling.tsuull.uz

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti

Bosh muharrir:

Botir Elov

Bosh muharrir o'rinbosari:

Shahlo Hamroyeva

Mas'ul kotib:

Oqila Abdullayeva

Tahrir kengashi

Shuhrat Sirojiddinov (O'zbekiston), Eshref Adali (Turkiya), [Viktor Zaxarov] (Rossiya), Vladimir Benko (Slovakiya), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat O'rxun (Turkiya), Suyun Karimov (O'zbekiston), Abduvali Qarshiyev (O'zbekiston), Muxammadjon Musayev (O'zbekiston), Kamoliddin Shukurov (O'zbekiston), O'tkir Hamdamov (O'zbekiston), Tal'at Zuparov (O'zbekiston), Bahodir Mo'minov (O'zbekiston), Faxriddin Nurullayev (O'zbekiston), Zulxumor Xolmanova (O'zbekiston), Muqaddas Abdurahmonova (O'zbekiston), Elova Dilrabo (O'zbekiston), Ruhillo Alayev (O'zbekiston), Rasuljon Atamuratov (O'zbekiston), Malika Abdullayeva (O'zbekiston), Mannon Ochilov (O'zbekiston), Xolisa Axmedova (O'zbekiston), Zilola Xusainova (O'zbekiston), Uldona Abdurahmonova (O'zbekiston).

Jurnal haqida ma'lumot

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi – Oliy attestatsiya komissiyasi ilmiy nashrlar ro'yxatidagi “O'zbekiston: til va madaniyat” akademik jurnalining ilovasi hisoblanib, unda professor-o'qituvchilar, doktorantlar, stajor-tadqiqotchilar, mustaqil izlanuvchilar, magistrantlarning kompyuter lingvistikasi, jumladan, tabiiy tilga ishlov berish (NLP), o'zbek tilining formal grammatikasi, korpus lingvistikasi, mashina tarjimasi, nutqni qayta ishlash tizimlari, intellektual tizimlar, kompyuter leksikografiyasi hamda lingvistik ontologiyalar kabi sohalarga oid tadqiqotlari nashr qilinadi.

Jurnal ilovasi bir yilda to'rt marta chop etiladi.

O'zbek, turk, rus va ingliz tillarida yozilgan maqolalar qabul qilinadi.

Jurnalda kitoblarga yozilgan taqrizlar, adabiyotlar sharhi, konferensiyalar hisobotlari va tadqiqot loyihalari natijalari ham e'lon qilinadi.

Mualliflar fikri tahririyat nuqtayi nazaridan farq qilishi mumkin.

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi 2023-yildan chiqa boshlagan.

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti. O'zbekiston, Toshkent, Yakkasaroy tumani, Yusuf Xos Hojib ko'chasi, 103-uy.

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

Alisher Navoi'i Tashkent State University of the Uzbek Language and Literature

Chief editor:	Botir Elov
Deputy editor-in-chief:	Shahlo Hamroyeva
Responsible secretary:	Oqila Abdullayeva

Editorial board

Shukhrat Sirojiddinov (Uzbekiston), Eshref Adali (Turkiye), [Viktor Zakharov] (Russia), Vladimir Benko (Slovakia), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat Orhun (Turkey), Suyun Karimov (Uzbekistan), Abduvali Karshiyev (Uzbekistan), Mukhammadjon Musayev (Uzbekistan), Kamoliddin Shukurov (Uzbekistan), O'tkir Hamdamov (Uzbekistan), Tal'at Zuparov (Uzbekistan), Bahadir Mo'minov (Uzbekistan), Fakhridin Nurullayev (Uzbekistan), Zulkhumor Kholmanova (Uzbekistan), Muqaddas Abdurakhmonova (Uzbekistan), Elova Dirlabo (Uzbekistan), Ruhillo Alayev (Uzbekistan), Rasuljon Atamuratov (Uzbekistan), Malika Abdullayeva (Uzbekistan), Mannon Ochilov (Uzbekistan), Kholisa Akhmedova (Uzbekistan), Zilola Khusainova (Uzbekistan), Uldona Abdurakhmonova (Uzbekistan).

Information about the magazine

"Uzbekistan: language and culture. "Computer Linguistics" series is an appendix of the academic journal "Uzbekistan: Language and Culture" in the list of scientific publications of the Higher Attestation Commission, in which computer linguistics, including natural language processing (NLP) of professors-teachers, doctoral students, intern-researchers, independent researchers, master's students, researches related to formal grammar of the Uzbek language, corpus linguistics, machine translation, speech processing systems, intelligent systems, computer lexicography and linguistic ontologies are published.

The magazine supplement is published four times a year.

Articles written in Uzbek, Turkish, Russian and English languages are accepted.

The journal also publishes book reviews, literature reviews, conference reports, and research project results.

The opinion of the authors may differ from the editorial point of view.

"Uzbekistan: language and culture. "Computer Linguistics" series has been published since 2023.

Tashkent State University of Uzbek Language and Literature named after Alisher Navoi. Yusuf Khos Hajib street, 103, Yakkasaray district, Tashkent, Uzbekistan.

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

MUNDARIJA

Фархад Мирзаев Рамиз, Гюнель Новрузова Сиявуш

Компьютерное моделирование основные инструменты
исследования6

Xolisa Axmedova, Elbek Malikov

Bilinga asoslangan yondashuvlar asosida omonimiyani
farqlash.....15

Xolisa Axmedova, Shohnazar Sultonov

Statistik usullar yordamida polifunksional soʻzlarni
semantik farqlash.....26

Oqila Abdullayeva, Oʻgʻiloy Bozorqulova

Jahon tilshunosligida treebanklar tasnifi.....43

Zilola Xusainova, Surayyo Yangibayeva

Til korpusi turlari.....54

Shaxinabonu Mansurova

Son soʻz turkumini grammatik pos teglashning
lingvistik modellari.....64

Botir Elov, Zilola Xusainova, Sarvinoz Qosimova

Katta til modellari.....78

Oqila Abdullayeva, Fotima Oʻtkirova

Jahon tilshunosligida Dependancy Parsingga oid
tadqiqotlar.....92

CONTENT

Farhad Mirzayev Ramiz, Gunel Novruzova Siyavush

Computer simulation basic research

tools.....13

Xolisa Axmedova, Elbek Malikov

Differentiating knowledge-based

Homonymy.....24

Xolisa Axmedova, Shohnazar Sultonov

Semantic differentiation of polyfunctional words using

statistikal methods.....40

Oqila Abdullayeva, O'g'iloy Bozorqulova

The classification of treebanks in world linguistics.....52

Zilola Xusainova, Surayyo Yangibayeva

Types of language corpus.....62

Shaxinabonu Mansurova

Linguistic methods of grammatical pos tagging

of the number word group.....76

Botir Elov, Zilola Xusainova, Sarvinoz Qosimova

Large language models.....90

Oqila Abdullayeva, Fotima O'tkirova

Dependancy Parsing in world

linguistics.....104

STATISTIK USULLAR YORDAMIDA POLIFUNKSIONAL SO‘ZLARNI SEMANTIK FARQLASH

Xolisa Axmedova¹
Shohnazar Sultonov²

Annotatsiya. Ushbu maqola statistik usullar yordamida o‘zbek tilidagi polifunksional so‘zlarning semantik farqlanishini tahlil qilishga bag‘ishlangan. Katta hajmdagi matn korpusi ma’lumotlari asosida polifunksional so‘zlarning turli kontekstlarda qo‘llanishi tahlil qilinadi va ularning ma’no jihatidan farqlanishi aniqlanadi. Maqolada qo‘llanilgan statistik usullar, natijalar va xulosalar batafsil bayon etiladi. Tadqiqot natijalari o‘zbek tili semantik analizatorini ishlab chiqish va takomillashtirishda muhim ahamiyatga ega.

Kalit so‘zlar: *Semantik tahlil, tabiiy tilni qayta ishlash polifunksional so‘zlar, statistik usullar, polifunksionallik, tasniflash parametrlari.*

Kirish

Tabiiy tilga ishlov berish sohasida semantik analiz alohida ahamiyatga ega. Matnlarni semantik analiz qilish ham o‘z navbatida ikki muhim qismga bo‘linadi: Sentiment analiz qilish va so‘z ma’nosini aniqlash.

Sentiment analiz matnlarga hissiy bo‘yoqdorlik beradi.

So‘z ma’nosini aniqlash so‘zning kontekstdagi ma’nosini aniqlab beradi. So‘z ma’nosini aniqlash elementlarining har biri gap ma’nosini aniqlashda muhim ahamiyatga ega. Polifunksional so‘zlar ham ana shunday elementlardan biri. Leksik birliklarning polifunksionalligi muammosi zamonaviy tilshunoslikning dolzarb

¹Axmedova Xolisxon Ilxomovna – texnika fanlari bo‘yicha falsafa doktori (PhD), Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti.
E-pochta: xolisa9029@mail.ru
ORCID: 0000-0002-9828-1650

²Sultonov ShohnazarAbdurashid o‘g‘li – Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti magistranti.
E-pochta: sultonovshonazar@mail.com
ORCID: 0009-0008-1676-6736

muammolaridan biridir. Polifunksional soʻz masalasi jahon tilshunosligida rus, ingliz, tatar, xitoy tillarida koʻp uchraydi. Polifunksionallik tushunchasi baʼzi manbalarda koʻp funksiyalilik koʻrinishida uchraydi. Koʻpgina xorijiy lingvistik asarlarda koʻp funksiyalilik omonimiya hodisalari bilan bogʻliq holda koʻrib chiqiladi.

Polifunksional soʻz masalasi jahon tilshunosligida toʻliq yechim topmagan boʻlib, uning avtomatik aniqlanishi, korpusda polifunksionallikning bartaraf etilishi, lingvistik filtri va modeli borasida biror tadqiqot uchratmadik. Oʻzbek tilidagi polifunksional soʻzlarni semantik farqlashda tadqiqotchi Sh.Gulyamova “Polifunksional soʻzlarni lingvistik modellash tirishni turkumlar kesimida aniqlab olish kerak”, – deydi [Gulyamova, 2020. 47].

Polifunksional soʻzlar turli kontekstlarda turli vazifalarni bajaradi.

1-jadval: Polifunksional soʻzlar

Soʻz	Maʼnosi	Soʻz turkumi
Bemahal	Kech tun payti	Ot
	Nomuvofiq, nomaqbul, oʻrinsiz payt.	Ravish
	Juda erta, yoshlik paytda; odatdagi vaqtdan oldin.	
Baʼzan	Hamma vaqt ham emas, har zamon-har zamonda, ahyon-ahyonda, goh-gohida.	Ravish
	Ayrim hollarda, baʼzi vaqtlarda.	
	Gapning uyushiq boʻlaklarini bogʻlash uchun qoʻllanadi.	Bogʻlovchi
...	...	...

Bemahal soʻzi ishtirok etgan quyidagi gaplarda uning vazifalarini yaqqol koʻrish mumkin.

1. *Bemahal boʻlib qolganidan, oʻsha yerda tunab qoladigan boʻldilar.*

2. *Birovning uyiga boqma, bemahal əshikni qoqma!*

3. *Bemahal boʻlmasin hech kimga oʻlim,*

Otim Avaz, Buldur qassob uliman..

1-gapda ot soʻz turkumi vazifasida kelgan boʻlsa, 2- va 3-gaplarda ravish soʻz turkumi vazifasida bajarayotgan boʻlsa ham maʼnolari bir-biriga uzviy bogʻlangan. Bemahal soʻzi ravish soʻz turkumini oʻzida ham ikki xil maʼnoni ifodalashi mumkin. Garchi, soʻzning uchta maʼnosi mavjud boʻlsa-da, u 2 ta soʻz turkumi doirasida polifunksionallikni ifodalaydi. Ravish soʻz turkumi vazifasini

ifodalaganda aynan qaysi bir ma'noni ifodalashini aniqlash kerak. Bu esa polifunksionallikni aniqlashdagi yana bitta muhim vazifalardan biridir.

Sh.Gulyamova polifunksionallik haqidagi o'zbek tilida mavjud nazariy qarashlarni umumlashtirib, semantik analizator uchun polifunksional so'zlarni modellashtirish borasida bir qator ishlarni amalga oshirgan. Tadqiqotchi 183 ta polifunksional so'zlarni so'z turkumlari kesimidagi jami 21 ta guruhga ajratgan [Gulyamova, 2020. 105]. Ushbu maqolada biz ajratilgan guruhlarga polifunksional so'zlarni biriktirib chiqdik.(2-jadval)

2-jadval: Polifunksional so'zlar so'z turkumlari doirasida uchrashiga ko'ra guruhlari va soni

№	So'z turkumlari	Soni
1	Yuklama yoki ravish	1
2	Bog'lovchi yoki yuklama	2
3	Ot yoki ko'makchi	1
4	Ot yoki undov	1
5	Undov yoki modal so'z	1
6	Modal yoki yuklama	2
7	Bog'lama yoki ot	1
8	Sifat yoki undov	1
9	Ravish yoki undov	1
10	Sifat yoki ko'makchi	6
11	Olmosh yoki ravish	3
12	Ravish yoki ko'makchi	21
13	Mustaqil fe'l va ko'makchi fe'l	26
14	Bog'lovchi va modal so'z	6
15	Ko'makchi bog'lovchi	1
16	Yuklama ko'makchi	1
17	Ot yoki ravish	5
18	Ravish yoki bog'lovchi	1
19	Ot yoki sifat	72
20	Ravish yoki modal so'z	4
21	Sifat yoki ravish	16
22	Sifat yoki modal	3
23	Mustaqil va yordamchi fe'l	3
24	Ravish yoki ko'makchi yoki modal so'z	1
25	Sifat yoki ravish yoki ko'makchi	1
26	Sifat yoki ravish yoki yuklama	1

2-jadvaldan koʻrinadiki, polifunksional soʻzlarni ikkita soʻz turkumi doirasida 23 ta guruhga ajratib, uchta soʻz turkumi doirasida 3 ta guruhga ajratib oʻrganish mumkin.

Jahonda kontekstda uchragan polifunksional soʻzlarni semantik farqlash masalasi uchramasa ham, baʼzi turkiy tillarda bu masalani yechish boʻyicha harakatlarni koʻrish mumkin. Polifunksionallikni aniqlash deganda, soʻzning kontekstda, asosan, ularni qaysi soʻz turkumi vazifasini bajarib kelganini aniqlashni tushunish mumkin.

Polifunksionallikni aniqlashning mavjud yondashuvlari anʼanaviy deterministikaga boʻlinadi, yaʼni lokal va global sintaktik tahlil, sintaktik lugʻat va ehtimolliklarga asoslanadi. Grammatik va statistik hodisalardan foydalangan holda korpusdagi soʻzlarning belgilari asosida polifunksionallik hodisasi aniqlanadi. Polifunksionallikni muvaffaqiyatli chiqishi uchun aniq natijalar kerak. Jahon kompyuter lingvistikasida gaplarni semantik tahlil qilishda, asosan, 3 ta usuldan foydalanilgan:

Qoidalarga asoslangan usul – soʻzlarni til grammatikasi qoidalari yordamida semantik tahlil qilish nazarda tutilgan. Oʻzbek tilida qator qoidalar ishlab chiqilgan. Ushbu qoidalarga asoslangan holda baʼzi turdagi polifunksional soʻzlarni farqlovchi matematik modellari X. Axmedova tomonidan taklif etilgan [Elov va boshqalar, 2024. 421-431]. “Hosil qilingan matematik modellar asosida alohida modullar yozilishi zarur boʻlib, ular kontekstga va omonim soʻzlarning oʻziga, ularning soʻz turkumiga yoki parametrlariga qarab faqat maʼlum hollarda oʻz isbotini topadi. Ushbu usulning mohiyati shundan iboratki, baʼzi hollarda kontekst tahlili gapning bir qismini sintaktik tuzilishini va uning yordamida soʻz shakllarini tushunishga yordam beradi”- deya taʼkidlaydi tadqiqotchi.

Polifunksional soʻzlar ham omonimlarga oʻxshab birikuvchilar asosida farqlanishi mumkin. Ammo faqat bu metodning oʻzi barcha polifunksional soʻzlarni farqlab, model tuzish uchun yetarli emas. Chunki baʼzi soʻzlarning soʻz turkumi kontekstda aniqlanadi.

Yordamchi, alohida olingan soʻzlar yoki ular bilan mustaqil soʻzlar orasidagi polifunksionallik hodisasi kontekstda reallashadi. Bir soʻz ham bogʻlovchi, ham koʻmakchi boʻlib, gapda aynan qaysi funksiyani bajarib kelishi ularning pozitsiyasi, birikuvchisi hamda vazifasiga bogʻliq.

Masalan, *aniq* ham sifat, ham modal soʻz boʻlib (gapda kirish soʻz vazifasini bajaradi) keladi.

Aniq sif. – eshitish, ko'rish, tushunish jihatdan yaxshi, to'la ifodali. Masalan: *Saodatxon orziqib javob kutar, Tojiboy aka bo'lsa, aniq javob bermay, gapni aylantirar edi.* (S.Zunnunova, Olov) *Ko'pchilik tomonidan ma'qullangan amaliy, aniq takliflar kiritildi.* (Gazetadan)

Aniq mod. – fikrning chinligini bildiradi; albatta, shubhasiz. *Xolmirza aka ichidan "shu g'ayrating bo'lsa, ikki yilda Xirmontepadan aniq ayriyamiz", deb o'ylaydi.* (A.Qodiriy, G'irvonlik Mallavoy) *Endi kimga og'iz solsang, g'iring demay tegishi aniq.* (S.Ahmad, Yulduz)

A.Kaplan tomonidan kontekstni minimal bartaraf etishni o'rganish bo'yicha olib borilgan tajriba polifunksional so'zlarni bartaraf etish uchun ham qo'l keladi. Omonimlarni modellashtirish masalasida bu tajriba xususida batafsil to'xtalgan edi [Gatiavlin va boshqalar, 2019. 237].

Polifunksionallik hodisasi butun jumla tarkibi bilan birgalikda namoyon bo'ladi. Masalan:

1. *Cholning bo'yin tomirlari bo'rtib chiqqani, iyaklari orqaga tortilgani traktor fonarida **aniq** ko'rinib turardi* (S.Ahmad, Ufq. 83-b.).

2. *Manavini to'ldirib, adreslarni **aniq** qilib yoz* (S.Ahmad, Ufq. 104-b.).

3. *Pastga qulashi **aniq** edi* (S.Ahmad, Ufq. 133-b.).

4. *O'zi daryoga tushib ketishi **aniq*** (S.Ahmad, Ufq. 150-b.).

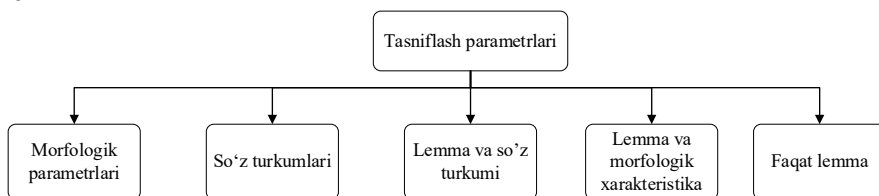
5. *Tashqarida bo'layotgan shovqin-suron, ketmon va belkuraklarnng ovozi, toshga qadalayotgan lomlarning qarsillashi, ketmon peshlayotgan temirchilarning jarang-jurungi taxta devor tirqishidan o'tib Belyavskiyga **aniq** eshilib turardi* (S.Ahmad, Ufq. 156-b.).

6. *Sizdan loyihaga **aniq** amal qilishingizni so'raymiz* (S.Ahmad, Ufq. 170-b.).

Ko'rinadiki, 1-2-5-6-gaplarda *aniq* sifat turkumiga mansub bo'lsa, 3-4-gaplarda modal so'z bo'lib kelgan. Bu holatda ularni birikuvchilar orqali aniqlash oson emas. Shuning uchun butun gap mazmunidan chiqib yondashish kerak. Gapning umumiy mazmuni orqali so'zning qanday funktsiya bajarishi reallashadi.

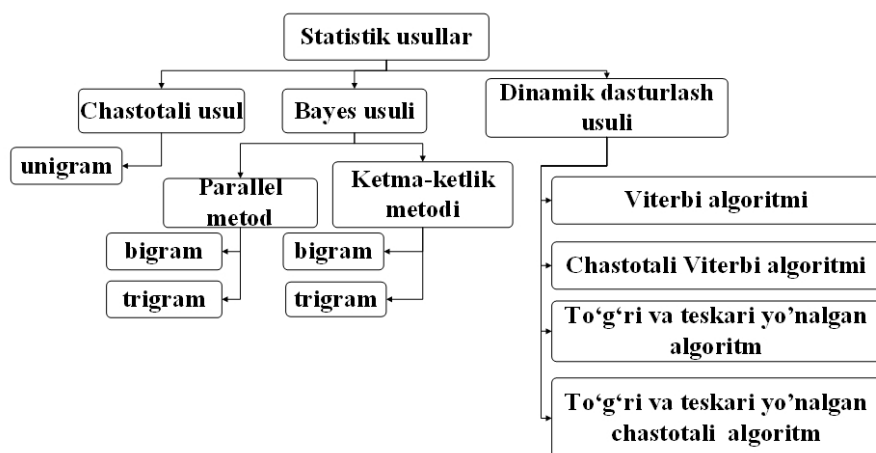
"O'zbek tili omonimlarining izohli lug'ati"da ham mustaqil, ham ko'makchi fe'l maqomida kela oladigan so'zlar omonim sifatida ko'rsatilgan. Izlanishlar asosida shuni ta'kidlash kerakki, biz bunday kategoriyalarga ega so'zlarga omonim emas, balki polifunksional so'z sifatida qaraymiz [Gulyamova, 2022. 182-185].

Statistik maʼlumotlarga asoslangan yondashuv – masalani yechishda soʻzlarning grammatik parametrlarini tasniflash orqali foydalaniladi. Bu parametrlar turli tabiiy tillarda turlicha tanlab olinadi. Masalan, rus tilida morfologik omonimiyani aniqlashda soʻz turkumi, soʻzning qaysi jins(rod)ga oidligi, birlik yoki koʻplik shakli, lemma, lemma va soʻz turkumi, faqat lemmasi, polifunksionalligi kabi parametrlar ajratib olingan. Bu usul yordamida oʻzbek tilidagi omonim soʻzlarni semantik farqlashda foydalanilgan [Axmedova, 2023, 78-82]. Ushbu maqolada polifunksional soʻzlarni tasniflash parametrlari yordamida semantik farqlash masalasini koʻrib chiqamiz. Bu parametrlarni oʻzbek tili uchun quyidagicha ajratish mumkin:



1-rasm. Tashniflash parametrlari

Soʻzning semantik tahlilining aniqlik darajasi yuqori boʻlmasa, tasniflanish parametrlarini kengaytirish mumkin. Ajratib olingan parametrlar yordamida statistik usullardan foydalanish mumkin. Ushbu usullarning asosiy vazifasi kontekst tarkibini n-grammlarga ajratish, yaʼni kontekstdagi kirish soʻzining birikuvchilarini aniqlab, baholash metodlari yordamida baholashdan iborat [Рахматуллаев, 1984. 25]. Baʼzi statistik usullarda soʻz va uning qoʻshimchalari orqali qaror qabul qilinsa, baʼzilarida soʻzning kontekstdagi semantik valentliklari yordamida xulosa qilinadi. Bundan kelib chiqadiki, statistik maʼlumotlarga asoslangan usullar qaror qabul qilish parametrlariga koʻra bir necha guruhlariga boʻlinadi. (2-rasm) Chastotali usulda unigram, yaʼni soʻzning oʻzi orqali semantik tahlil qilinadi, Bayes usulida esa soʻzning bigram va trigrammalaridan foydalaniladi. Dinamik dasturlash usulida esa n-grammalardan foydalaniladi.



2-rasm. Statistik usullar iyerarxiyasi

Ushbu metodlar orqali baholashda kontekstni tokenizatsiya va teglashtirish jarayonlari olib boriladi. O'zbek tilida tokenizatsiya, stemming va lemmatizatsiya masalalari Z.Xusainova tomonidan yechilgan [Xusainova, 2024]. O'zbek tili leksik birliklarini POS teglash ham muhim masalalardan biridir. O'zbek tili leksik birliklarini POS teglash bo'yicha B.Elov, N.Xudoyberganovlarning ilmiy maqolalarini uchratdik [Elov va boshqalar, 2024]. Tabiiy tilning leksik birliklarini POS teglashda turli yondashuvlarni qo'llash mumkin. Bu usullarning deyarli barchasida joriy so'zning n-gramlaridan foydalaniladi. Leksik birliklarni POS teglash orqali qisman polifunksionallikni bartaraf etish mumkin. Ya'ni shunday polifunksional so'zlar borki, ular turli kontekstlarda turli so'z turkumi vazifasini bajaradi. Bunday polifunksional so'zlarni POS teglash orqali aniqlash mumkin.

Dastlab, polifunksional so'zlarni semantik farqlashdagi muammolarni aytib o'tsak:

– *Faqat turli so'z turkumlari vazifasini bajaruvchi polifunksional so'zlarni aniqlash. Kasal* - polifunksional so'z bo'lib, kontekstda ikkita so'z turkumi vazifasini bajarishi mumkin, ya'ni

1. Anzirat buvi ***kasal*** qizini parvarishlash uchun har kuni uzoq qishloqdan qatnaydi.

2. Shifokorxonaga olib kelingan ***kasalni*** ko'rib hayratlandim.

1-gapdagi kasal so'zi sifat so'z turkumi vazifasini, 2-gapda esa ot so'z turkumi vazifasini bajargan. Xuddi shunday uchta so'z turkumini ifodalab 3 ta vazifada keladigan polifunksional so'zlarni ham uchratish mumkin.

– *Har bir so'z turkumi doirasida bir nechta vazifalarni anglatuvchi polifunksional so'zlarni semantik farqlash. Misol uchun, ba'zan* so'zi (3-jadvalga qarang).

3-jadval: *Baʻzan* soʻzining maʼnolari

Ravish	Hamma vaqt ham emas, har zamon-har zamonda, ahyon-ahyonda, goh-gohida	1
	Ayrim hollarda, baʼzi vaqtlarda	2
Bogʻlovchi	Gapning uyushiq boʻlaklarini bogʻlash uchun qoʻllanadi	3

3-jadvalda keltirilganidek, *baʻzan* soʻzi ikkita soʻz turkumini ifodalasa-da, 3 ta vazifani bajarishi mumkin. Xuddi shunday, turli soʻz turkumlari vazifasida koʻplab maʼnolarni anglatuvchi polifunksional soʻzlarni ham uchratish mumkin. *Boʻlmoq* soʻzi ana shunday soʻzlar jumlasidandir.

4-jadval: *Boʻlmoq* soʻzining maʼnolari

Mustaqil feʼl	Yuz bermoq, voqe boʻlmoq.	1
	Kelmoq, boshlanmoq; oʻtmoq, kechmoq (fasl, davr, payt va sh. k. haqida).	2
	Baholanmoq, biror narxga olinmoq yoki sotilmoq.	3
	Amalga oshmoq, bitmoq. Boʻladigan ish. Boʻlmaydigan ish.	4
	Taxt holga kelmoq, tayyor boʻlmoq, bitmoq.	5
	Yetarli boʻldi deb hisoblamog; toʻymoq.	6
	(ch. k. soʻz bilan) koʻchma Toʻymoq; koʻrib-bilib xafsalasi qaytmoq; joniga tegmoq.	7
	Tugatmoq, oxiriga yetkazmoq.	8
	Yaramog; toʻgʻri, muvofiq kelmoq.	9
	Ega boʻlmoq; bor, mavjud boʻlmoq.	10
	Biror joyga borib, tashrif buyu-rib, maʼlum vaqt turmoq, qoʻnmoq, yashamoq.	11
	Biror-bir tadbirda ishtirok etmoq, qatnashmoq, hozir boʻlmoq.	12
	Oʻsmog, bitmoq (oʻsimlik, hosil haqida).	13
	Biror darajaga yetmoq, oʻz nasliga tortmoq.	14
	(ch. k. soʻz bilan) Dunyoga kelmoq, tugʻilmoq.	15
	(ch. k. soʻz bilan) Kelib chiqmoq, tugʻilmoq.	16
	Biror harakat yoki narsaning yetarli boʻlganini, meʼyordan oshganini, qoniqqanlikni, mamnunlikni ifodalaydi.	17
	Biror harakat, tadbirni maʼqullamoq, unga qoʻshilmoq.	18
	Sifat turkumidagi soʻzlar bilan birikib, qoʻshma feʼl hosil qiladi, mas, hayron boʻlmoq, xursand boʻlmoq, kir boʻlmoq.	19
	-(a)r, -(a)digan, -(y)digan, -moqchi affiksli feʼllar bilan qoʻshilib, istak, mayl maʼnolarini bildiradi.	20
	-day, -dek affiksli feʼllar bilan qoʻshilib, ish-harakatning jarayoni, holatini bildiradi.	21

	-(i)b affiksli ravishdoshdan keyin ko'makchi fe'l vazifasida kelib, ish-harakatning tugallanganligini bildiradi.	22
	Bo'lishsiz shakllarda kelib, ish-harakatning amalga oshirilishi qiyinligini yoki mumkin emasligini bildiradi.	23
	Ot kesim tarkibida bog'lama vazifasida keladi.	24
	Butunni (sindirib, qirgib, yorib yoki biror boshqa yo'l bilan) bo'laklarga ajratmoq, bo'lak-bo'lak qilmoq.	25
	Ma'lum qismlarga, guruhlariga ajratmoq, taqsimlamoq.	26
	mat. Bo'luv amalini bajarmoq, yechmoq.	27
	O'rtada taqsim qilmoq, ulashmoq.	28
	Tashkilot, muassasa, korxona va sh. k. ni kichikroq mustaqil qismlarga, bo'limlarga ajratmoq.	29
	ko'chma Gap, so'z, suhbat va sh. k. ni to'xtatib yoki kesib qo'ymoq.	30
	Shaxsning xayoli, fikri, o'yi va b. ga xalaqit bermoq, uni chalg'itmoq.	31
Bog'lovchi	Bo'lmoq I fl. shart shakli.	32
	zidl. bog'l. Esa.	33

Agar polifunksional so'z ikki yoki uch so'z turkumi doirasida bittadan ma'noga ega bo'lsa, bu muammoni statistik usullar yordamida aniqlash nisbatan osonroq. 3-jadvalda *bo'lmoq* so'zining 33 ta vazifasi keltirilgan. Bunday holatda so'zning vazifasini aniqlash uchun statistik hamda mashinali o'qitish yondashuvlaridan foydalanish maqsadga muvofiq. X.Axmedova o'zbek tilidagi omonim so'zlarni semantik tahlil qilishda statistik ma'lumotlarga asoslangan yondashuvning Chastotali usulidan foydalangan [X.Axmedova, 2022. 160]. Biz ham o'zbek tilidagi polifunksional so'zlarni kontekstda aniqlashda ushbu usuldan foydalandik. Chastotali usuldan fodalalanishdan avval korpus ma'lumotlaridan *ba'zan* polifunksional so'zlar ishtirok etgan gaplarni ajratib oldik.

Ba'zan so'zi ishtirok etgan gaplarni namuna sifatida keltiramiz.

1. *Ha, ba'zan shunday sho'xlik ham qilib turish kerak. (P.Tursun, O'qituvchi)*

2. *Ba'zan Gulnorning yuragi toshib, butun dardini, yaralarini Unsinga ochishni istaydi. (Oybek, Tanlangan asarlar)*

3. *Yo'lchi ba'zan yolg'iz, ba'zan o'z yoniga birovni hamroh qilib, qancha jar, qancha qo'rg'on va quloqlarni kezib, suv keltirar edi. (Oybek, Tanlangan asarlar)*

1 va 2-gaplardagi *ba'zan* polifunksional so'zi ravish, 3-gapda esa bog'lovchi vazifasini bajarmoqda. Xuddi shunday, o'zbek tili ta'limi korpusi ma'lumotlari orasidan joriy so'z qatnashgan gaplar

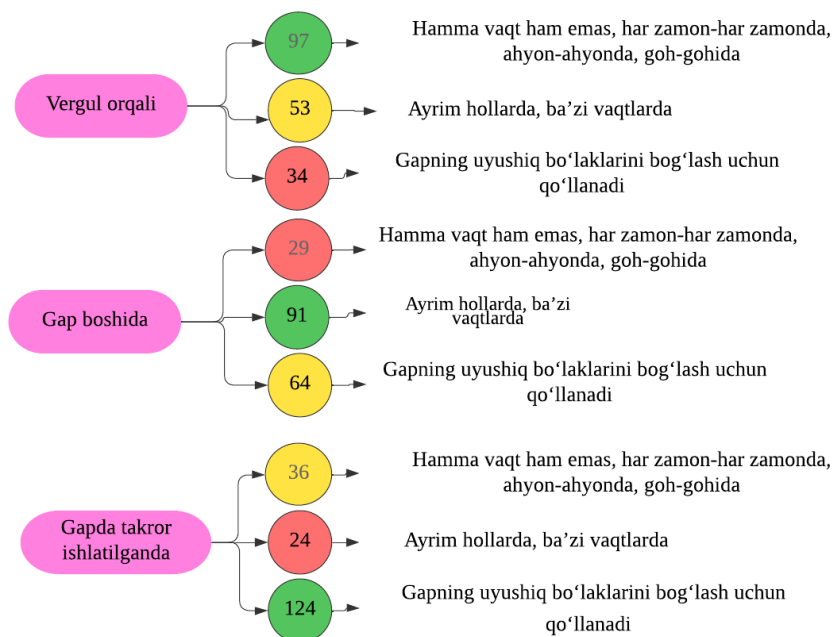
ajratib olinib tahlil qilindi. Oʻzbek tili taʼlimiy korpusidan *baʼzan* soʻzi qidirilganda 183 ta oʻrinda uchradi. Ularni tahlil qilib chiqdik va soʻzning semantik maʼnolarini farqlovchi omillarni belgilab oldik (4-jadval).

4-jadval: *Baʼzan* soʻzini semantik farqlash omillari

Nº	Farqlovchi omillar	Semantik maʼnosi
1	Vergul orqali	1
2	Gap boshida	2
3	Gapda takror ishlatilishi	3

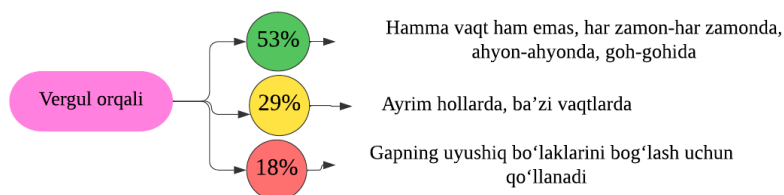
4-jadvaldan koʻrish mumkinki, *baʼzan* soʻzidan oldin vergul belgisi qoʻyilganda, ravish soʻz turkumi vazifasini bajarib, “*Ayrim hollarda, baʼzi vaqtlarda*” degan maʼnoni anglatishi mumkin. Gap boshida kelganida koʻp hollarda “*Hamma vaqt ham emas, har zamon-har zamonda, ahyon-ahyonda, goh-gohida*” maʼnosini anglatishi mumkin. Agarda *baʼzan* soʻzi gapda ikki marta qatnashgan boʻlsa, bu soʻzni bogʻlovchi vazifasini bajaryapti deyish mumkin.

Oʻzbek tili taʼlimiy korpusi maʼlumotlaridan *baʼzan* soʻzi qatnashgan gaplar ajratib olinib semantik teglandi hamda 4-jadvalda keltirilgan tasniflash parametrlari boʻyicha ajratib chiqildi. Natijalarni quyidagi rasmda koʻrish mumkin.



3-rasm. *Baʼzan* soʻzining tasniflash parametrlari boʻyicha olingan statistik maʼlumotlari

3-rasmda keltirilgan statistik ma'lumotlar asosida Chastotali usulning qiymatlarini hisoblaymiz. Chastotali usulda har bir tasniflash parametri bo'yicha aniqlangan gaplarning soni umumiy gaplar soniga bo'linadi. Ya'ni, biz ba'zan so'zi ishtirok etgan 183 ta gapni tasniflash parametrlari asosida guruhlariga ajratdik. Ularning har birini sonini 183 ga bo'lish orqali yuqori ustunlikka ega bo'lgan ma'nolarni aniqlaymiz. Hisoblash orqali quyidagi natijalar oldik.



4-rasm. 1-parametr bo'yicha olingan natijalarning foiz qiymati

Ushbu parametr bo'yicha yangi kiritilgan gapni semantik tahlil qilamiz. Misol uchun :

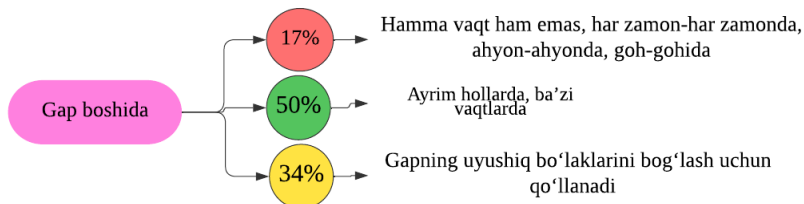
*Ha, **ba'zan** shunday sho'xlik ham qilib turish kerak.*

Bu gapda polifunksional so'z **ba'zan** so'zi bo'lib, o'zidan oldin vergul belgisi qo'yilgan. 4-rasmdagi statistik ehtimolliklarga ko'ra ushbu gapdagi ba'zan so'zining ma'nosi 53% aniqlik bilan "*Hamma vaqt emas, har zomon-har zamonda ahyon- ahyonda*", goh gohida degan ma'noni anglatadi.

Navbatdagi gapni keltiramiz.

*Me'mor **ba'zan** shunday uzoq xayollarga borardi.*

Bu gapni 2-tasniflash parametri bo'yicha semantik tahlil qilish mumkin. (5-rasm).



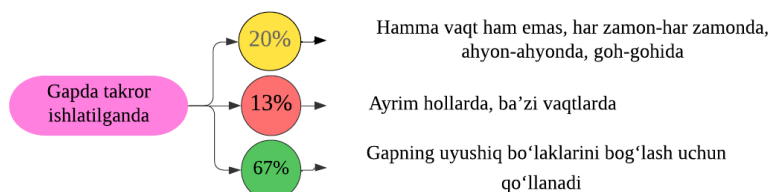
5-rasm. Ba'zan so'zining 2-parametri boyicha foiz ko'rsatkichlari

Ushbu gapdagi polifunksional so'z 2-tasniflash parametri bo'yichi 50% aniqlik bilan ravish so'z turkumi vazifasini bajarib, "*Ayrim hollarda, ba'zi vaqtlarda*" degan ma'noni anglatadi.

Navbatdagi gapni keltiramiz:

*Ba'zan Qoratomyning ishxonasida, **ba'zan** choyxonada tunab yurdi.*

Bugapda ba'zan so'zi ikki marta qatnashgan bo'lib 3-tasniflash parametri yordamida aniqlash mumkin. (6-rasm).



6-rasm. Ba'zan so'zining 3-parametri bo'yicha ma'nosini aniqlash

Demak, keltirilgan gapdagi *ba'zan* polifunksional so'zi ikki marta uchraganini hisobga olib, 3-tasniflash parametriga mos keladi va 67 % aniqlik bilan "Gapning uyushiq bo'laklarini bog'lash uchun qo'llanadi".

Yuqorida keltirilgan ketma-ketliklar kabi boshqa guruhga mansub polifunksional so'zlarni semantik farqlovchi parametrlarni ajratish mumkin. Bu parametrlarni ajratib olish uchun nisbatan kattaroq hajmdagi statistik kuzatuvlar kerak bo'ladi. Keltirilgan hisob-kitoblar asosida olingan natijalar yuqori aniqlikka ega emas, albatta. Ammo natijalarni yanada yuqori qilish mumkin. Buning uchun katta hajmdagi dataset olib uni tahlil qilish, tasniflash parametrlarini yanada mukammal qilib tanlash lozim. Undan tashqari yuqori aniqlikdagi natijalarni olish uchun boshqa zamonaviy yondashuvlardan foydalanish zarur.

Xulosa

Semantik analizatorni ishlab chiqish uchun dastavval semantik analiz va sintezning matnning avtomatik tahlilidagi ahamiyatini yoritish muhim. Avtomatik semantik tahlil kompyuter lingvistikasining aktual va shu bilan birga murakkab vazifalaridan biridir. U inson aql-zakovati bilan bog'liq bo'lgan har qanday tizimni yaratishning zaruriy bosqichidir.

Dunyo tilshunosligida morfologik, semantik va sintaktik analizator bo'yicha keng tadqiqotlar olib borilmoqda. Semantikani rasmiylashtirish o'ta murakkab vazifa. Muammoning tub murakkabligi shundaki, mazmun (ma'no)ni o'rganishda tildan tashqariga chiqib, tashqi olamga, til doirasidan tashqarida yotuvchi predmet va tasavvurlar tasnifiga murojaat qilishga to'g'ri keladi.

Avtomatik semantik analiz va sintez kompyuter lingvistikasining murakkab vazifalaridan biri bo'lib, matnning avtomatik tahlilidagi ahamiyati katta. Axborot-qidiruv tizimida

semantik analiz va sintez jarayonini mukammallashtirish, semantik qidiruv va uning muammolarini yechish semantik analizator vazifasini aniqlashtirishga, kelajakda SemA (semantik analizator) tuzishga yo'l ochadi.

Foydalanilgan adabiyotlar

- Abdurahmonova N. Automatic morphological analyze for English-Uzbek system // Известия Кыргызский государственный технический университет им. И.Раззакова. Теоретический и прикладной научно-технический журнал. № 2 (38) – Бишкек, 2016. – С. 12-18.
- Mengliyev B., Hamroyeva Sh. Turkiy tillardagi morfologik analizatorlar xususiyada// O'zbek tilini dunyo miqyosida keng targ'ib qilish bo'yicha hamkorlik istiqbollari // Xalqaro ilmiy-amaliy anjuman materiallari. – Toshkent, 2020, oktabr. – B. 360-364.
- Hamroyeva Sh. O'zbek tili morfologik analizatorining lingvistik. – GlobeEdit, 2020.
- Elov B. Hamroyeva Sh. Morfologik analizatorni yaratish usullari // Til va madaniyat 2022. 4-son.
- Elov B., Hamroyeva Sh., Axmedova X. Methods for Creating a Morphological Analyzer, 2023. Doi:10.1007/978-3-031-27199-1_4.
- Ерискина Е.В., Курушин Д.С., Яруллин Д.В., Корюкин Д.Ю.Объектно-ориентированная модель морфологического анализатора русскоязычного текста // Пермский национальный исследовательский политехнический университет, Пермь-2018.
- Сравнение и создание морфологических анализаторов в NLTK // Habr URL: habr.com/post/340404/ (дата обращения: 24.11.2018).
- Зо Зон Су. Разработка синтаксических анализаторов языков программирования с учетом контекстных условий// тема диссертации и автореферата по ВАК РФ 05.13.11, кандидат физико-математических. 1985.
- Буштедт В.А. Модель принятия решения на основе синтаксического анализа в задачах обработки патентной информации//2011 – кандидат технических наук.
- Сокирко А, Толдова С. Сравнение эффективности двух методик снятия лексической и морфологической

- неоднозначности для русского языка (скрытая модель Маркова и синтаксический анализатор именных групп) // Интернет-Математика, 80–94. Retrieved from <http://hdl.handle.net/10995/1391>.
- Maxmudov N., Nurmonov A. O'zbek tilining nazariy grammatikasi. Toshkent, 1995, 37-58-betlar.
- Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. Автоматическая обработка текстов на естественном языке и анализ данных. - НИУ ВШЭ, 2017. – 269 с.
- Каневский Е.А., Боярский К.К. Семантико-синтаксический анализатор SemSin // Компьютерная лингвистика и интеллектуальные технологии, 2012;
- Gulyamova Sh. Polifunksionalva ko'p ma'noli so'zlarni farqlashda lokalizatsiya hodisasi//“O'zbek amaliy tilshunosligi va lingvodidaktika masalalari” mavzusidagi respublika ilmiy-amaliy konferensiya materiallari to'plami. – Toshkent, 2021.
- Elov B., Abdurakhmonova N., Axmedova X., Iskandarova A., Fattaxova D. Designing Processes and Models of Semantical Differentiation for Polyfunctional Words in the Uzbek Contexts// Communications in Computer and Information Science 2145 CCIS, pp. 305-316. 2024.
- Гулямова Ш.К. Полифункционализм хусусида // Ўзбекистонда илмий-амалий тадқиқотлар мавзусидаги онлайн конференция. 15-август. № 19. – Тошкент, 2020. – Б. 46-48.
- Elov B., Axmedova X., Abdullayeva O., Farkhodovich K.B. Modeling of Models and Processes that Differentiate Semantically Polyfunctional Words in the Context of the Uzbek Language// Lecture Notes in Networks and Systems 1002 LNNS, pp. 421-431. 2024.

SEMANTIC DIFFERENTIATION OF POLYFUNCTIONAL WORDS USING STATISTIKAL METHODS

Xolisa Axmedova¹
Shohnazar Sultonov²

Abstract. This article is devoted to the analysis of the semantic differentiation of polyfunctional words in the Uzbek language using statistical methods. The study analyzes the use of polyfunctional words in different contexts based on a large text corpus and identifies their semantic differentiation. The article describes in detail the statistikal methods used, results and conclusions. The results of the study will be of great importance in the creation and improvement of semantic analyzers of the Uzbek language. The article also provides suggestions for future research.

Keywords: *Semantic analysis, natural language processing, polyfunctional words, statistikal methods, polyfunctionality, classification parameters.*

References

- Abdurahmonova N. Automatic morphological analyze for English-Uzbek system // *Izvestiya Kirgizskiy gosudarstvennyy texnicheskiy universitet im. I.Razzakova. Teoreticheskiy i prikladnoy nauchno-texnicheskiy jurnal.* № 2 (38) – Bishkek, 2016. – S. 12-18.
- Mengliyev B., Hamroyeva Sh. Turkiy tillardagi morfologik analizatorlar xususiyada // *O'zbek tilini dunyo miqyosida keng targ'ib qilish bo'yicha hamkorlik istiqbollari* // *Xalqaro ilmiy-amaliy anjuman materiallari.* – Toshkent, 2020, oktabr. – B. 360-364.
- Hamroyeva Sh. *O'zbek tili morfologik analizatorining lingvistik.* – GlobeEdit, 2020.

¹ Axmedova Xolisxon Ilxomovna – doctor of philosophy of technical sciences (PhD). Tashkent State University of Uzbek Language and Literature named after Alisher Navoi.

E-mail: xolisa9029@mail.ru

ORCID: 0000-0002-9828-1650

² Sultonov Shohnazar Abdurashid o'g'li – Graduate student of Computer Linguistics. Tashkent State University of Uzbek Language and Literature named after Alisher Navoi.

E-mail: sultonovshonazar@mail.com

ORCID: 0009-0008-1676-6736

- Elov B. Hamroyeva SH. MORFOLOGIK ANALIZATORNI YARATISH USULLARI//Til va madaniyat 2022. 4-son.
- Elov, Botir & Mirdjonovna, Hamroyeva, Ilxomovna, Axmedova. (2023). Methods for Creating a Morphological Analyzer. Doi:10.1007/978-3-031-27199-1_4.
- Yeriskina Ye.V., Kurushin D.S., Yarullin D.V., Koryukin D.Yu.Obyektno-orientirovannaya model morfologicheskogo analizatora russkoyazychnogo teksta // Permskiy natsionalnyy issledovatel'skiy politexnicheskii universitet, Perm-2018
- Sravnenie i sozdanie morfologicheskix analizatorov v NLTK // Habr URL: habr.com/post/340404/ (data obraščeniya: 24.11.2018).
- Zo Zon Su. Razrabotka sintaksicheskix analizatorov yazikov programmirovaniya s uchetom kontekstnix usloviy//tema dissertatsii i avtoreferata po VAK RF 05.13.11, kandidat fiziko-matematicheskix. 1985.
- Bushtedt V.A. Model prinyatiya resheniya na osnove sintaksicheskogo analiza v zadachax obrabotki patentnoy informatsii // 2011 – kandidat texnicheskix nauk.
- Sokirko A,Toldova S. Sravnenie effektivnosti dvux metodik snyatiya leksicheskoy i morfologicheskoy neodnoznachnosti dlya russkogo yazika (skritaya model Markova i sintaksicheskii analizator imennyykh grupp) // Internet-Matematika, 80–94. Retrieved from <http://hdl.handle.net/10995/1391>
- Maxmudov N., Nurmonov A. O'zbek tilining nazariy grammatikasi. Toshkent, 1995, 37-58-betlar.
- Bolshakova Ye.I., Voronsov K.V., Yefremova N.E., Klishinskiy E.S., Lukashevich N.V., Sapin A.S. Avtomaticheskaya obrabotka tekstov na yestestvennom yazike i analiz dannix. - NIU VShE, 2017. – 269 s.
- Kanevskiy Ye.A., Boyarskiy K.K. Semantiko-sintaksicheskii analizator SemSin // Kompyuternaya lingvistika i intellektualnie texnologii. 2012.
- Gulyamova Sh. Polifunksionalva ko'p ma'noli so'zlarni farqlashda lokalizatsiya hodisasi // "O'zbek amaliy tilshunosligi va lingvodidaktika masalalari" mavzusidagi respublika ilmiy-amaliy konferensiya materiallari to'plami. – Toshkent, 2021.
- Elov B., Abdurakhmonova N., Axmedova X., Iskandarova A., Fattaxova D. Designing Processes and Models of Semantical Differentiation for Polyfunctional Words in the Uzbek Contexts // Communications in Computer and Information

Science 2145 CCIS, pp. 305-316. 2024.

Gulyamova Sh.K. Polifunksionallik xususida // – O'zbekistonda ilmiy-amaliy tadqiqotlar mavzusidagi onlayn konferensiya. 15-avgust. № 19. – Toshkent, 2020. – B. 46-48.

Boltayevich E.B., Tursunaliyevna A.M., Ilxomovna A.X., Xolmo'minovna A.O., Farkhodovich K.B. Modeling of Models and Processes that Differentiate Semantically Polyfunctional Words in the Context of the Uzbek Language// Lecture Notes in Networks and Systems 1002 LNNS, pp. 421-431. 2024.

Jurnal 2017-yil 26-oktyabrda O'zbekiston Respublikasi Matbuot
va axborot agentligi tomonidan 0936-raqam bilan ro'yxatdan o'tgan.

Jurnal O'zbekiston Respublikasi Oliy Attestatsiya Komissiyasi tomonidan
filologiya fanlari bo'yicha falsafa doktori (PhD) va fan doktori (DSc)
dissertatsiyalari asosiy ilmiy natijalari chop etilishi lozim bo'lgan ro'yxatga
kiritilgan (30.10.2021. № 308/6).

Tahririyatga kelgan maqolalar mualliflarga qaytarilmaydi.

Manzil: Toshkent shahri, Yakkasaroy tumani, Yusuf Xos
Hojib ko'chasi 103-uy.
Telefonlar: +99871 281-45-11, +99871 281-41-93.
Website: compling.tsuull.uz
E-mail: kompling@navoiy-uni.uz

Bosishga **.**.****-yilda ruxsat etildi.
Bichimi 70x100 1/16, Ofset bosma. "Cambria" garniturasida.
Shartli b.t. 7,51. Nashr b.t. 7,62.

"O'zbekiston: til va madaniyat" jurnali tahririyatida
tayyorlandi va sahifalandi.
"YASHNOBOD NASHR" bosmaxonasida chop etildi.
Adadi 300 nusxa. Buyurtma №2.
Bosmaxona manzili: Toshkent shahar Yashnobod tumani,
58-a harbiy shaharcha.