

Elov Botir, Axmedova Xolisxon

# O'ZBEK TILI SEMANTIK ANALIZATORINING MATEMATIK MODELLARI

*Monografiya*

MACHINE  
LEARNING

MATH

STATS

NLP

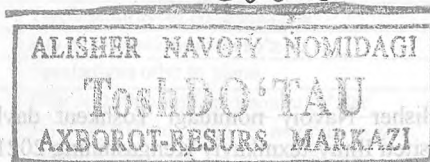
Toshkent - 2021

**ELOV BOTIR, AXMEDOVA XOLISXON**

# **O‘ZBEK TILI SEMANTIK ANALIZATORINIG MATEMATIK MODELLARI**

*Monografiya*

36659



**“Farovon yurt”**

**Toshkent – 2021**

UO'K 364.3(575.1)

364.3(575.1)

O' 26

O'zbek tili semantik analizatorinig matematik modellari: Monografiya [Matn].  
-Toshkent: "Farovon yurt", 2021. - 108 b.

KBK 364.3(575.1)

*Mazkur monografiyada o'zbek tilidagi omonim, polisemantik va polifunksional so'zlarni semantik tahlil qiluvchi tizim uchun matematik modellar haqida so'z yuritiladi. Omonim va polifunksional so'zlarni semantik tahlil qilishda ularni so'z turkumlari doirasida uchrashiga ko'ra guruhlariga ajratish, bir so'z turkumi doirasidagi omonim so'zlar va polisemantik so'zlarni aniqlashda semantik konkordans, Yashirin Markov modellaridan foydalanish samarali natija berishi asoslanadi. Kitob talabalar, magistrantlar, tadqiqotchilar hamda sohaga qiziquvchilar uchun mo'ljallangan.*

**Mas'ul muharrirlar:**

Filologiya fanlari doktori, professor: Laylo Raupova

**Taqrizchilar:**

Texnika fanlari doktori, professor: Dilnoza Muxammadiyeva

Filologiya fanlari doktori, dotsent: Shahlo Hamroyeva

Monografiya Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti Ilmiy-texnikaviy Kengashining 2021 yil 29 noyabrdagi 8-sonli Qaroriga asosan chop etildi.

ISBN 978-9943-6174-9-0

© "Farovon yurt" 2021.

## Monografiyada qoʻllanilgan belgilashlar

№	Belgi	Izoh
1.	<i>N</i>	Ot soʻz turkumidagi soʻzlar
2.	<i>V</i>	Fel soʻz turkumidagi soʻzlar
3.	<i>Adj</i>	Sifat soʻz turkumidagi soʻzlar
4.	<i>Num</i>	son soʻz turkumidagi soʻzlar
5.	<i>Pr</i>	Olmosh soʻz turkumidagi soʻzlar
6.	<i>Adv</i>	Ravish soʻz turkumidagi soʻzlar
7.	∨	Dizyunksiya amali, “yoki” maʼnosini anglatadi
8.	⊕	Soʻzlarning oʻzaro birikishini anglatadi
9.	<i>H<sub>N,V</sub></i>	Ot va fel orasidagi omonimlik hodisasi
10.	<i>aff<sub>s,i</sub><sup>N</sup></i>	Otga qoʻshiluvchi sintaktik va lugʻaviy shakl yasovchi qoʻshimcha
11.	<i>aff<sub>s,j</sub><sup>V</sup></i>	Felga qoʻshiluvchi sintaktik va lugʻaviy shakl yasovchi qoʻshimcha
12.	<i>H<sub>num,N</sub></i>	Ot va son orasidagi omonimlik hodisasi
13.	<i>W<sup>dec</sup></i>	Oʻnlik soʻzlar
14.	<i>H<sup>num</sup></i>	Son omonimligi
15.	<i>aff<sup>ord</sup></i>	Tartib sonlarni bildiruvchi qoʻshimchalari
16.	<i>W<sub>k</sub><sup>N</sup></i>	Raqam soʻzlar
17.	<i>N<sub>f</sub><sup>c</sup></i>	Sanaladigan otlar toʻplami
18.	<i>V<sup>num</sup></i>	Son omonimga birikuvchi fellar
19.	<i>H<sup>N</sup></i>	ot soʻz turkumidagi omonim
20.	<i>aff<sup>CS</sup></i>	Kelishik qoʻshimchalar
21.	<i>H<sub>pv</sub></i>	Predikativ soʻzning omonimligi
22.	<i>V<sup>H</sup></i>	Felning harakat nomi shakli
23.	<i>V<sup>pv</sup></i>	Predikativ soʻz
24.	<i>H<sub>N,Aux</sub></i>	Koʻmakchi va ot oʻrtasidagi omonimlik
25.	<i>aff<sup>N</sup></i>	Otga qoʻshiluvchi qoʻshimchalar
26.	<i>W</i>	Soʻz
27.	<i>W<sub>j</sub><sup>Aux</sup></i>	Koʻmakchi omonimni birikishi mumkin boʻlgan soʻzlar
28.	<i>H<sup>E</sup></i>	Undov soʻzning omonimligi
29.	<i>V<sup>NE</sup></i>	ot soʻz turkumidagi omonim soʻz bilan birikaoladigan feʼllar toʻplami.
30.	<i>H<sub>num,Adj</sub></i>	Son va sifat soʻz turkumlari orasidagi omonimlik
31.	<i>H<sup>num</sup></i>	Son soʻz turkumidagi omonim soʻzlar
32.	<i>aff<sup>ord</sup></i>	Tartib bildiruvchi qoʻshimchalar
33.	<i>W<sub>k</sub><sup>N</sup></i>	k ta ot soʻz turkumidagi soʻz
34.	<i>N<sub>f</sub><sup>c</sup></i>	Sanaladigan otlar toʻplami
35.	<i>H<sup>Adj</sup></i>	Sifat soʻz turkumidagi omonim soʻzlar
36.	<i>H<sub>V,Aux</sub></i>	Fel ∨ koʻmakchi orasidagi omonim soʻzlar
37.	<i>W<sup>T</sup></i>	Makon va zamoni ifodalovchi soʻzlar
38.	<i>H<sup>Aux</sup></i>	Koʻmakchi omonim soʻzlar
39.	<i>H<sup>V</sup></i>	Fel soʻz turkumidagi omonim soʻzlar
40.	<i>aff<sub>k</sub><sup>V</sup></i>	Feʼl omonimga qoʻshiluvchi baʼzi qoʻshimchalar
41.	<i>H<sub>Adj,m</sub></i>	Sifat ∨ modal soʻzlarning omonimligi
42.	<i>H<sub>Adj,y</sub></i>	Sifat ∨ feʼl soʻz turkumlari orasidagi omonim soʻzlar
43.	<i>aff<sup>Ad,d</sup></i>	Sifat darajalarini ifodalovchi qoʻshimchalar toʻplami
44.	<i>H<sub>Adv,p</sub></i>	Ravish ∨ yuklama orasidagi omonimlik



45.	$H^{Adv}$	Ravish so'z turkumidagi omonim so'z
46.	$W^{Adv}$	Ravish so'z turkumidagi omonim so'zlar bn birikadigan so'zlar
47.	$aff_i^{adv}$	Ravishga qo'shilishi mumkin bo'lgan qo'shimchalar
48.	$H^{imit}$	Taqlid so'zlar orasidagi omonim so'zlar
49.	$aff_i^{CS}$	Kelishik qo'shimchalar to'plamining qism to'plami
50.	$V_i^{imit}$	Taqlid so'z bilan birikadigan fellar
51.	$N_k^{imit}$	Kaqlid so'z bilan brika oladigan ot
52.	$V_k^{qz}$	Qo'shma fe'ning 2-komponenti
53.	$H_{adj,adv}$	Sifat v ravish so'z turkumlari orasidagi omonim so'zlar
54.	$W_i^{adv,b}$	Ravishning chap tomonidan birikuvchi so'zlar
55.	$H^{adv}$	Ravish omonimligi
56.	$W_i^{adv,a}$	Ravishning o'ng tomonidan birikuvchisi
57.	$H_{N,V,Pr}$	Ot v fel v olmosh orasidagi omonimlik
58.	$H^{Pr}$	Olmosh omonimligi
59.	$W_k^{Pr}$	Olmosh omonim bilan o'ng tomonidan birikishi mumkin bo'lgan so'zlar
60.	$H_{Adj,v}$	Sifat v fe'l omonimligi
61.	$H_{N,Adv,v}$	Ot v ravish v fel so'z turkumlari orasidagi omonimlik
62.	$H^{Adv}$	Ravish omonimligi
63.	$aff_i^{adv}$	Ravishga qo'shiluvchi qo'shimchalar
64.	$aff_i^v$	Fe'l omonim so'zga qo'shiluvchi qo'shimchalar
65.	$H_{num,N,V}$	Son v ot v fel omonimligi
66.	$aff_j^{agg}$	Jamlovchi qo'shimchalar to'plami
67.	$H_{V,N,imit}$	Fel v ot v taqlid so'z orasidagi omonimlik
68.	$H_{E,v,Adv}$	Undov v fel v ravish so'z turkumlari orasidagi omonimlik
69.	$H_{N,Adj,v,Adv}$	Ot v sifat v fel v ravish so'z turkumlari orasidagi omonimlik
70.	$P$	Polisemantik so'z
71.	$W_i^b$	Polisemantik so'zning chap tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
72.	$W_j^a$	Polisemantik so'zning o'ng tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
73.	$T$	Boolean tipidagi kattalik
74.	$Pf_{adj,N}$	Sifat v ot so'z turkumidagi polifunksionallik
75.	$Pf^N$	Ot so'z turkumidagi polifunksional so'z
76.	$W_i^b$	Polisemantik so'zning chap tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
77.	$W_j^a$	Polisemantik so'zning o'ng tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
78.	$T$	Boolean tipidagi kattalik
79.	$Pf_{adj,N}$	Sifat v ot so'z turkumidagi polifunksionallik
80.	$Pf^N$	Ot so'z turkumidagi polifunksional so'z
81.	$W_i^b$	Polisemantik so'zning chap tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
82.	$W_j^a$	Polisemantik so'zning o'ng tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
83.	$T$	Boolean tipidagi kattalik
84.	$Pf_{adj,N}$	Sifat v ot so'z turkumidagi polifunksionallik
85.	$Pf^N$	Ot so'z turkumidagi polifunksional so'z
86.	$W_i^b$	Polisemantik so'zning chap tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
87.	$W_j^a$	Polisemantik so'zning o'ng tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
88.	$T$	Boolean tipidagi kattalik
89.	$Pf_{adj,N}$	Sifat v ot so'z turkumidagi polifunksionallik
90.	$Pf^N$	Ot so'z turkumidagi polifunksional so'z
91.	$W_i^b$	Polisemantik so'zning chap tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
92.	$W_j^a$	Polisemantik so'zning o'ng tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami
93.	$T$	Boolean tipidagi kattalik
94.	$Pf_{adj,N}$	Sifat v ot so'z turkumidagi polifunksionallik
95.	$Pf^N$	Ot so'z turkumidagi polifunksional so'z

## KIRISH

Jahon tilshunosligida 1930-1940-yillarda amaliy tilshunoslik, eng avvalo, chet tilini o'qitish jarayoni, uni o'qitish metodikasi, grammatikaning o'ziga xos xususiyatlarini tasvirlash orqali tushunilgan. Rossiyada 1950-yillardan boshlab amaliy tilshunoslik kompyuter texnologiyalari va avtomatik axborotni qayta ishlash tizimlari bilan bog'langan. 1960-yillarda maxsus ilmiy yo'nalish sifatida shakllangan amaliy tilshunoslik tildan foydalanishning amaliy masalalarini yechishga qaratilgan. Jumladan, jahon tilshunosligida XX asrning 60-yillarida paydo bo'lgan korpus tilshunosligi nutq amaliyoti uchun yangi metodologik asosni, tadqiqotning ustuvor yo'nalishlarini yaratdi. Til korpusi va u bilan bog'liq bo'lgan fan, ya'ni korpus lingvistikasi XX asrning oxiri XXI asr boshida tilshunoslik va til o'qitish metodikasiga shiddat bilan kirib kelgan yangi yo'nalishdir. Til faoliyatining zamonaviy shartlari nafaqat nutq namunalarini, balki alohida xronologik holatlar, nutq amaliyotini tizimlashtirishni o'rganishni o'z ichiga oladi.

Dunyo tilshunosligida XIX asr oxiri XX asrning boshlarida korpuslar lingvistik tadqiqotlar o'tkazish yoki ko'pincha amaliy muammolarni yechish uchun (masalan, til birliklarining chastotasini hisoblash uchun) yaratila boshlandi. Korpus tilshunosligi uchun kompyuterlarning ixtiro qilinishi va keng qo'llanilishi bilan rivojlanishning yangi bosqichi boshlanib, yaratilgan korpuslar mavjudlaridan nafaqat saqlash formatida, balki hajmlarida ham farq qilishini tahlil etish dolzarb masalaga aylandi. Kompyuter lingvistikasi sohasida turli maqsadli korpuslar yaratish, mavjud korpuslarni mukammallashtirish, boyitish, kengaytirish kabilar yechimini kutayotgan masalalardan biridir.

Kompyuter lingvistikasini kompyuter vositalaridan foydalanish sohasi sifatida dasturlar, ma'lumotlarni qayta ishlash texnologiyalarini tashkil etish va muayyan sharoitlarda til faoliyatini modellashtirish uchun, shuningdek, tilshunoslikda tilning kompyuter modellari va unga oid fanlar doirasida qo'llash mumkin.

Til juda murakkab ta'lim bo'lganligi tufayli kompyuter tilshunosligida avtomatik matn tahlili; avtomatik matn sintezi; avtomatik lug'atlarni yaratish va qo'llab-quvvatlash; avtomatlashtirilgan axborot-qidiruv tizimlarini yaratish; mashina

tarjimai; avtomatik til o'rganish tizimlarini yaratish; lingvistik ma'lumotlar bazalarini yaratish; nazariy va amaliy tilshunoslik masalalarini yechishning dasturiy vositalarini ishlab chiqish kabi turli yo'nalishlar rivojlanib bormoqda.

O'zbekistonda mustaqillik yillarida kompyuter va korpus lingvistikasi bo'yicha sanoqli tadqiqotlar amalga oshirilgan. Binobarin, "Bugungi globallashuv davrida har bir xalq, har qaysi mustaqil davlat o'z milliy manfaatlarini ta'minlash, bu borada, avvalo, o'z madaniyatini, azaliy qadriyatlarini, ona tilini asrab-avaylash va rivojlantirish masalasiga ustuvor ahamiyat qaratishi tabiiy. O'zbek tilining xalqimiz ijtimoiy hayotida va xalqaro miqyosdagi obro'-e'tiborini tubdan oshirish, unib-o'sib kelayotgan yoshlarimizni vatanparvarlik, milliy an'ana va qadriyatlarga sadoqat, ulug' ajdodlarimizning boy merosiga vorislik ruhida tarbiyalash, mamlakatimizda davlat tilini to'laqonli joriy etishni ta'minlash"<sup>1</sup> maqsadida korpus lingvistikasi istiqbolli ilmiy yo'nalish sifatida uning nazariy muammolarini yechishni taqozo etadi. Bu esa fanimiz oldida turgan dolzarb vazifalardan biri – o'zbek tili birliklari avtomatik tahlilida omonimlik, ko'p ma'nolilik va polifunksionallikni farqlashning matematik modellari va semantik tahlil qilish algoritmlarini yaratishni kun tartibiga qo'yadi.

Dunyo tilshunosligida axborot-qidiruv tizimida semantik analizatorning o'rni, matnga avtomatik ishlov berish, matnning avtomatik analiz/sintez dasturlari, semantik analizatorlar tadqiqiga oid olib borilgan izlanishlardan quyidagi ilmiy natijalar olingan: lingvistik analizator tushunchasi, uning turlari tavsiflangan, semantik analiz va sintezning matnning avtomatik tahlilidagi ahamiyati yoritilgan (Boston University, New York University, AQSh; Uralskiy Federalniy Universitet, Rossiya); lingvistik parser, sintaktik parseming ahamiyati va ulardan foydalanish usullari aniqlangan (Uralskiy Federalniy Universitet, Rossiya); matnga avtomatik ishlov berish texnologiyalari o'rganilgan (University of Tokyo, Yaponiya, Australian National University, Avstraliya, Oxford University Language Centre, Angliya); Morfo-Analizator, SegmMorf, Lintez, Psi-An kabi tahlil tizimlarining

<sup>1</sup> O'zbekiston Respublikasi Prezidenti Shavkat Mirziyoyevning "O'zbek tilining davlat tili sifatidagi nufuzi va mavqeiini tubdan oshirish chora-tadbirlari to'g'risida"ga farmon // www.xabar.uz

asoslari ishlab chiqilgan (Xarkov davlat pedagogika universiteti, Ukraina; Qozon federal universiteti (Rossiya, Tatariston Respublikasi); matnga avtomatik morfologik va sintaktik ishlov berish modeli va dasturlari (Xarkov davlat pedagogika universiteti, Ukraina; Baki Dövlət universiteti, Ozarbayjon); turkiy tillar kompyuter lingvistikasida tabiiy tilning kompyuter ishlovi, sintaktik-semantik analizator tuzish masalalari tadqiq qilingan (Tatariston Respublika Fanlar akademiyasi amaliy tilshunoslik va semiotika instituti), Qozon federal universiteti (Rossiya, Tatariston Respublikasi), V.Vernardskiy nomidagi Tavriya akademiyasi Qrim federal universiteti (Rossiya, Qrim Respublikasi), o'zbek-ingliz tili mashina tarjimasining lingvistik ta'minoti, til korpusining lingvistik bazasini tuzish tamoyillari ishlab chiqilgan; o'zbek-ingliz parallel korpusi instrumentariylarini tuzish asoslari, o'zbek milliy korpusida idiomalar bazasini yaratish, o'zbek tili morfoanalizatori va uning lingvistik ta'minotini yaratish masalasi ishlab chiqilgan (Mirzo Ulug'bek nomidagi O'zMU, Alisher Navoiy nomidagi ToshDO'TAU, Al-Xorazmiy nomidagi TATU, TerDU, BuxDU).

Jahon tilshunosligida tabiiy tilga avtomatik ishlov berish, analizatorlar yaratish bo'yicha, jumladan, axborot-qidiruv tizimi tahlil dasturlari, zamonaviy axborot texnologiyalari vositalarida ishlov berishni yo'lga qo'yish uchun til korpuslari, semantik analizator hamda ularning lingvistik asoslarini yaratish kabi yo'nalishlarda tadqiqotlar olib borilmoqda.

O'zbek tili leksikologiyasida ilk bor leksik dubletlar haqida ma'lumot berildi. Keyinchalik M. Mirtojiev, I. Qo'chqortoyev, A. Hojiyev, B. Isabekov, R. Shukurovlarning<sup>2</sup> leksik omonim, sinonim va antonimlar tadqiqiga bag'ishlangan maxsus asarlari maydonga keldi. Bu asarlar o'zbek tili leksikologiyasining yanada takomillashuvi uchun xizmat qildi. Korpus lingvistikasida omonimiya, polisemiya va polifunksionallik muammosini yechish, omonim birliklarni teglash va matnni avtomatik o'qish jarayonida omonimiyani bartaraf etish masalasiga oid qator tadqiqotlar vujudga kelgan.

---

<sup>2</sup> Nurmonov A. O'zbek tilshunosligi tarixi.-Toshkent: O'zbekiston, 2002.-177 b-B.80.

O'zbek kompyuter lingvistikasi yo'nalishida amalga oshirilgan sanoqli tadqiqotlarda kompyuter xotirasi omonim birliklarni "tanishi" va "o'qiy olishi"ga mo'ljallangan tahlil dasturlarini yaratish borasida harakatlar mavjud; o'zbek tilidagi omonimlarni teglash muammolari bo'yicha ayrim mulohazalar, omonimiyani aniqlash algoritmini tuzish bo'yicha dastlabki harakatlar amalga oshirilgan. Tilshunos tadqiqotchilar tomonidan tavsiya qilingan lingvistik modellar asosida semantik analizator yaratishga dastlabki qadam qo'yilmoqda.

akademiyasi amaliy (Rossiya, Tatariston Respublikasi) V. V. Ermakovich nomidagi Tavriya akademiyasi (Grim federal universiteti) (Rossiya, Qrim Respublikasi) o'zbek-ingliz tili mashina ta'limining lingvistik ta'minoti, til korpusining lingvistik bazasini tuzish, tashviihlan ishlab chiqilgan; o'zbek-ingliz parallel korpusi instrumentariyalarini tuzish asosari, o'zbek milliy korpusida idiomalar bazasini yaratish, o'zbek tili notranskript va uning lingvistik ta'minoti yaratish masalasi ishlab chiqilgan (Mizco Ulug'bek nomidagi O'xMU, Alisher Navoiy nomidagi ToshDU, TATU, TerDU, BukDU).

Tabiiy tilda avtomatik ishlov berish, analizator yaratish bo'yicha, jumladan, eksport-qiruv tizimi tahrir dasturini, xamonaviy eksport texnologiyalar vositalarida ishlov berishni yo'lga qo'yish uchun til korpuslari, semantik analizatorlarda ulaning lingvistik asoslarini yaratish kabi yo'nalishlarda tadqiqotlar olib borilmoqda.

O'zbek tili leksikologiyasida ikki bo'lsak doblatlar papida matnlot berilib, Keyinchalik M. Mirziyoyev, I. Qo'chqorov, A. Haliyev, B. Isakov, R. Shukurovning leksik omonim, sinonim va antonimlar tadqiqatiga bag'ishlangan nazariy asarlari tayyorlangan. Bu asarlar o'zbek tili leksikologiyasining yanada takomillashtirish uchun xizmat qildi. Korpus lingvistikasida omonimiyat, polissemiya va polissemiklik nazariyasini yechish, omonim birliklarni teglash va matn avtomatik o'qish jarayonida omonimiyatni beqaror etish masalasiga oid dastur tadqiqotlarida mavjud kelgan.



## **I BOB. Semantik munosabatlar va ularning matn tahlilida qo'llanilishi**

### **1.1. Tabiiy tillardagi matnlarni tahlil qilish tizimlari**

Tabiiy tilni qayta ishlash (Natural Language Processing (NLP)) – bu sun'iy intellekt va matematik tilshunoslikning umumiy yo'nalishi bo'lib, tabiiy tillardagi matnlarni kompyuterda analiz va sintez qilish muammolarini o'rganadi. Ushbu muammoni hal qilish inson va kompyuter o'rtasida o'zaro aloqaning yanada qulay shaklini yaratishni anglatadi.

Tabiiy tilni avtomatik qayta ishlash muammosi keyingi yarim asrdan ortiq davr mobaynida dolzarb bo'lib qolmoqda. Muammoning murakkabligi va aniq g'oyaning yo'qligi uni hal qilish yo'llarining qiyinligini bildiradi. Lingvistik analizatorlar esa matnga avtomatik ishlov berish vositalari sifatida alohida ahamiyat kasb etadi. Lingvistik analizatorni quyidagi turlarga bo'lib o'rganish maqsadga muvofiq:

1. Morfologik analizator
2. Sintaktik analizator
3. Semantik analizator

Katta guruhlar o'z ichida lemmatizator, morfoanalizator, stemmer, parser, orfokorrektorlarni oladi. Bu analizatorlarning har biri o'zining ishlash prinsipi, maqsadi, vazifasiga ega.

#### **1.1.1. Morfologik analizatorlar.**

Morfologik analizator – lug'atdagi alohida so'z/so'z shakllarni taqqoslovchi va so'zlarning grammatik xususiyatlarini topuvchi algoritmlar majmui. Morfologik analizator morfologik lug'atlar yordamida tahlil qilingan jumladagi so'zlarning ketma-ketligini, morfologik xususiyatini aniqlaydi. Turkiy tillar bo'yicha morfologik analizator yaratish masalasi o'tgan asrning 60-yillaridan buyon kun tartibidan o'rin olgan. O'sha davrdan boshlab ko'plab tadqiqotlar yuzaga chiqdi. Xususan, tatar tili morfologik modeli bo'yicha tahlil 1990-yillarning o'rtalaridan boshlab amaliyotga tatbiq etila boshlandi. O'sha davrda tatar tilining morfologik analizatori uch modelga asoslangan: generativ, paradigmatic va ikki bosqichli model. Tatar tilining dastlabki morfologik analizatori ikki fayldan iborat bo'lgan.

Birinchi fayl –qoidalar fayli (Rules), ikkinchi fayl- leksik birliklardan iborat leksikon (o‘zak va affiksial morfemlardan tashkil topgan) hamda ularning morfotaktik qoidalari izohi ham tavsiflanadi. Leksikon o‘z navbatida ichki leksikondan iborat. Ichki leksikonning tuzilishi graflar bilan bog‘liq. Ushbu tahlil FST (finite state transducer), ya‘ni chekli avtomat transyutiri texnologiyasi orqali amalga oshirilgan. Turkiy tillarni o‘z ichiga olgan morfologik analizator<sup>3</sup>, sintezator (boshqird, gagauz, turk, qozoq, tatar, qoraqalpoq, qirg‘iz, qrim-tatar, qo‘miq, chuvash, o‘zbek, uyg‘ur, tuva, saxa, no‘g‘ay) va mashina tarjimasiga (boshqird, turk, qozoq, tatar, qoraqalpoq, qirg‘iz, qrim-tatar, qo‘miq) mo‘ljallangan platformadan foydalanish mumkin. Ushbu platformada to‘liq morfologik tahlil qilish imkoniyati chegaralangan, uni yanada takomillashtirish zarur.

Dunyo va turk tilshunosligida morfologik analizator tuzish masalalari keng tadqiq etilgan. Jumladan, A.V.Diybo, A.K.Asiryana, A.V.Sheyrovich, A.R.Gatiatullin, A.M.Bashirov, P.V.Jeltov, N.A.Israilova, P.S.Bakasova, M.G.Malkovskiy, A.S.Starostin, M.G.Malkovskiy, A.S.Starostin, D.H.Younger, Ye.A.Kanevskiy, N.V.Kolpakova, N.A.Israilova, P.S.Bakasova, S.Bird, E.Klein, E.Loper, M.Korobov, A.D.Moskvina, D.Orlova, P.V.Panicheva, O.A.Mitrofanova, O.N.Lyashkevskaya, I.Astafyeva, A.Bonch-Osmolovskaya, A.Gareyshina, Yu.Grishina, V.Dyachkov, M.Ionov, A.Koroleva, M.Kudrinskiy, A.Lityagina, Ye.Luchina, Ye.Sidorova, S.Toldova, S.Savchuk, S.Koval va K.K.Boyarskiy kabi bir qator olimlar tadqiqotida batafsil yoritilgan<sup>4</sup>. O‘zbek tilida ham avtomatik

<sup>3</sup> <http://turkic.apertium.org>

<sup>4</sup> Асирян А.К. Сравнение инструментов морфологической разметки / Intellectual potential of the XXI century 2017 // <http://www.sworld.education>; Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков // Филология и культура. Philology and culture. 2014. №2(36). – С. 20-26.; Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словарей на базе структурнофункциональной модели тюркской морфемы / Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». –Казань, 2017. – 327 с.; Гатиатуллин А. Р., Баширов А.М., Осипов Г.С., Смирнов И.В., Шелманов А.О. Методы лингвистического анализа текстов на татарском языке и их применение в поисковой системе Eхastus // Труды ИСА РАН. Том 66. 1/2016. – 18-25 с.; Желтов П.В. Разработка морфологического анализатора чувашского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Казань, 2017. – 327 с.; Желтов П.В. Создание национального корпуса чувашского языка: проблемы и перспективы // Современные проблемы науки и образования. – 2015. – № 1-1.; URL: <http://www.science-education.ru/ru/article/view?id=19046> (дата обращения: 18.01.2021); Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Казань, 2017. – 327 с.; Малковский М.Г., Старостин А.С. Модель синтаксиса в системе морфо-синтаксического анализа «TR

morfologik tahlilning ayrim masalalari yoritilgan ishlar mavjud<sup>5</sup>. Sh.Hamroyeva tomonidan esa o'zbek tili morfoanalizatori va uning lingvistik ta'minotini yaratish masalasi tadqiqot obyekti sifatida o'rganilgan. Tadqiqotda lingvistik ta'minot vositalari tavsifi, ularning morfo-analizator va til korpuslari yaratishdagi ahamiyati; morfologik analiz/sintez instrumentariylari; lingvistik ta'minotning tuzilishi, tarkibi va vazifalari asoslanib, o'zbek tili morfologik analizatorining lingvistik ta'minoti masalasi hal qilingan<sup>6</sup>. N.Abdurahmonova o'zining "O'zbek tili korpusini morfologik tglashda FST texnologiyasi tatbiqi" maqolasida morfo-analizatorni yaratishda FST(finite state transducer) texnologiyasining qulayligi va kamchiliklari haqida mulohaza yuritgan. Dunyo tilshunosligida bir nechta morfologik analizator yaratilgan bo'lib, ularning asosiy vazifasi so'zni asos va qo'shimchalarga ajratadi. Ilg'or analizatorlar ko'proq ehtimollik usullari asosida ishlay oladi. Sintaktik, grammatik usulni o'z ichiga olgan analizator, parser deb ataladi va asosan, sayt hamda internet manbalarini tahlil qilish uchun ishlatiladi.<sup>7</sup> Zamonaviy tilshunoslikda so'zlarning asosini aniqlash stimmer deyiladi. Zamonaviy

EETON» //Тр. междунар. Конф. Диалог 2006. – М., 2006. – С. 481-492.; Younger D.H. Recognition and parsing of context-free languages in time n3. Information and Control 10 (2). – P.189-208.; Каневский Е.А., Колпакова Н.В. К вопросу построения морфологического анализатора // Компьютерная лингвистика и интеллектуальные технологии. 1999. Т. 2. С. 98–106; Израилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т. 2. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 327 с.; Bird S., Klein E., Loper E. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. Beijing, 2009.; Korobov M. Morphological Analyzer and Generator for Russian and Ukrainian Languages // Analysis of Images, Social Networks and Texts: 4th International Conference, AIST 2015, Yekaterinburg, Russia, April 9–11, 2015. Communications in Computer and Information Science, Springer, 2015. P. 320–332.; Москвина А.Д., Орлова Д., Паничева П.В., Митрофанова О.А. Разработка ядра синтаксического анализатора для русского языка на основе библиотек NLTK // Компьютерная лингвистика и вычислительные онтологии. Труды XIX Международной объединенной научной конференции «Интернет и современное общество». Санкт-Петербург, 22–24 июня 2016 г. СПб, 2016. С. 44–54.; Ляшевская О.Н., Астафьева И., Бонч-Осмоловская А., Гарейшина А., Гришина Ю., Дьячков В., Ионов М., Королева А., Кудринский М., Литягина А., Лучина Е., Сидорова Е., Толдова С., Савчук С., Коваль С. Оценка методов автоматического анализа текста: морфологические парсеры русского языка // Компьютерная лингвистика и интеллектуальные технологии. 2010. № 9 (16). С. 318–326; Каневский Е.А., Боярский К.К. Морфолого-лексический анализатор и классификация текста // Прикладная лингвистика в науке и образовании. Материалы V международной научно-практической конференции. Санкт-Петербург, 2010. С. 157–163.

<sup>5</sup>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.

<sup>6</sup> Hamroyeva SH. O'zbek tili morfologik analizatorining lingvistik ta'minoti.–GlobeEdit,2020; Mengliyev B. Hamroyeva SH. Turkiy tillardagi morfologik analizator xususida // O'zbek tilini dunyo miqyosida keng targ'ib qilish bo'yicha hamkorlik istiqbollari // Xalqaro ilmiy-amaliy anjuman materiallari. – Toshkent, 2020, oktabr. – B. 360-364.

<sup>7</sup> Е.В. Ерискина, Д.С. Курушин, Д.В. Яруллин, Д.Ю. Короткин. Объектно-ориентированная модель морфологического анализатора русскоязычного текста // Пермский национальный исследовательский политехнический университет, Пермь-2018

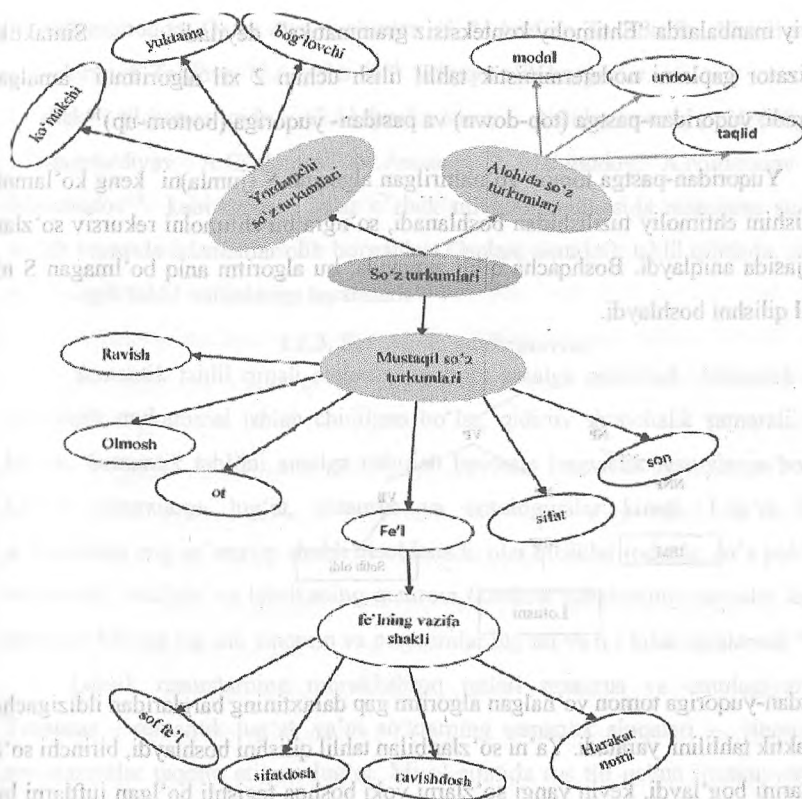
tilshunoslikda bir qancha morfo-analizatorlar ishlab chiqilgan. Ular, asosan, lug'atlarga asoslanadi. Quyidagi jadvalda asosiy morfologik analizatorlarni qiyosiy tahlili keltirilgan.

Asosiy morfo-analizatorlar<sup>8</sup>

Nomi	Tabiiy tillar	Lug'at boyligi	Natijasi
Rumor	Rus tili	125 ming so'z	So'z va uning so'z shakllari
MyStem	Rus tili	Lug'atlarsiz	So'z va uning morfologik xususiyatlari
Morphology	Rus, ingliz va nemis tillari	115 ming so'z	So'z, uning so'z shakllari, o'zagi va dastlabki shakllar
IsPELL	Rus, ingliz va nemis tillari va bosh.	137.2 ming so'z	Imlo tekshiruvini so'z va uning so'z shakllari

Ushbu morfo-analizatorlar orasida aniqlik darajasi eng yuqori bo'lgani MyStem tizimi hisoblanadi. Bu yagona analizator o'rnatilgan lug'atlarga asoslangan evristik usullarsiz ishlaydi. Bu tizim Python dasturlash tilining Pymystem kutubxonasi foydalangan holda yaratilgan. Pymystem kutubxonasi esa tabiiy tildagi so'zlarning morfologik xususiyatlari o'rnatilgan. MyStem analizatorining rus tili grammatikasiga moslashtirilgan. Agar bu tizimni o'zbek tiliga moslashtirsak, o'zbek tilidagi so'z turkumlari va ularning xarakteristikasi quyidagicha ierarxiya qilinadi.

<sup>8</sup> Сравнение и создание морфологических анализаторов в NLTK // Habr URL: <http://habr.com/post/340404/> (дата обращения: 24.11.2018).



1-rasm: Oʻzbek tilidagi soʻz turkumlarining xarakteristik ierarxiyasi

Tabiiy tillarning grammatik xususiyatidan kelib chiqib soʻzlar morfologik tahlil qilinadi.

### 1.1.2. Sintaktik analizatorlar

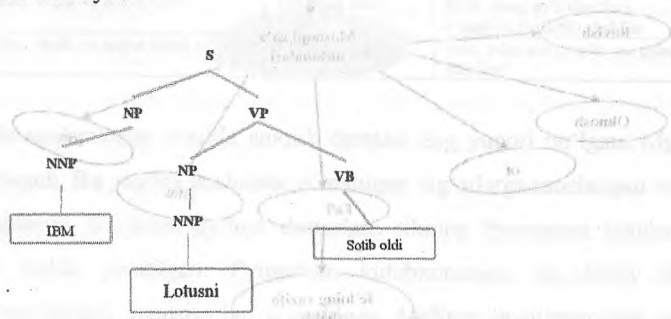
**Sintaktik analizator**- bu kirish tilini xususiyatlaridan kelib chiqib kiritilgan gapni sintaktik tahlil qiluvchi kompyuter dasturidir. Baʼzi manbalarda “Dasturning ifodani oʻqish va sintaktik tahlil qilishi boʻlimi” deya taʼriflanadi. Xomskiy grammatikalarining ierarxiyasida tillarni 4 guruhga ajratilgan. Ularni ichida eng qiziqarli kontekstsiz grammatika va tillar. Kontekstsiz grammatika tushunchasi

<sup>9</sup> [https://cpp.com.ru/shildt\\_spr\\_po\\_c/29/2905.html](https://cpp.com.ru/shildt_spr_po_c/29/2905.html)



xorijiy manbalarda “Ehtimoliy kontekstsiz grammatika” deyiladi. Sintaktik analizator gaplarni nodeterministik tahlil tilish uchun 2 xil algoritmni amalga oshiradi: yuqoridan-pastga (top-down) va pastdan- yuqoriga (bottom-up)<sup>10</sup>.

Yuqoridan-pastga tomon yo‘naltirilgan algoritm S (jumla)ni keng ko‘lamni tuzilishini ehtimoliy tuzilishidan boshlanadi, so‘ngra bu ehtimolni rekursiv so‘zlar darajasida aniqlaydi. Boshqacha qilib aytganda, bu algoritm aniq bo‘lmagan S ni tahlil qilishni boshlaydi.



Pastdan-yuqoriga tomon yo‘nalgan algoritm gap daraxtining barglaridan ildizigacha sintaktik tahlilini yaratadi. Ya‘ni so‘zlar bilan tahlil qilishni boshlaydi, birinchi so‘z juftlarini bog‘laydi, keyin yangi so‘zlarni yoki boshqa tegishli bo‘lgan juftlarni bu juftlarga bog‘laydi. Bog‘lanish jarayoni asta-sekin boshlang‘ich noterminal S ga keladi - ya‘ni gaplardagi barcha so‘zlar bir tuzilishga bog‘lanadi.

Dunyo kompyuter lingvistikasi sohasida qator sintaktik analizatorlar yaratilgan bo‘lib, ular shu ikkita algoritm asosida sintaktik tahlilni amalga oshiradi. PetitParser<sup>11</sup>, ЭТАП-3<sup>12</sup>, Bison<sup>13</sup> kabi ko‘plab sintaktik analizatorlardan

<sup>10</sup> [http://www.solarix.ru/for\\_developers/docs/rules.shtml](http://www.solarix.ru/for_developers/docs/rules.shtml)

<sup>11</sup> С.В. Чубейко, А.Н. Цуриков, В.С. Палагута. Анализ синтаксического разбора текста с помощью парсер-комбинаторов // Ростовский государственный университет путей сообщения, Ростов-на-Дону // Инженерный вестник Дона, №4 (2018) [ivdon.ru/magazine/archive/n4y2018/S335](http://ivdon.ru/magazine/archive/n4y2018/S335)

<sup>12</sup> Дружков К.Ю. Цели и ЛЕГКЕ PARSER OF ETAP-3 LINGUISTIC PROCESSOR: EXPERIMENTS ON RANKING SYNTACTIC

HYPOTHESES // ИИПИ РАН, Москва

<sup>13</sup> <https://ita.sibstis.ru/sites/csc.sibstis.ru/files/courses/trans/LabWork4.pdf>

foydalanilmoqda. Qator dunyo olimlari N. V.Arefev, Zo Zon Su, V.A.Bushtedt, A.Sokirko, S.Toldova, N.A.Valenda<sup>14</sup> tadqiqot olib borganlar.

O'zbek tilshunosligida G.Abdurahmonov, A.Sulaymonov, X.Xoliyorov, J.Omonturdiyev, A.G'ulomov, M.Asqarov, N.Maxmudov, A.Nurmonov, N.Maxmudov<sup>15</sup> kabi qator olimlar o'zbek tili grammatikasida matnlarni sintaktik tahlili borasida izlanishlar olib borganlar. Gaplarni sintaktik tahlil qilishda, avvalo, morfologik tahlil natijalariga tayaniladi.

### 1.1.3. Semantik analizatorlar

Semantik tahlil orqali semantik qidiruv amalga oshiriladi. Semantik tahlil qanchalik mukammal ishlab chiqilgan bo'lsa, qidiruv shunchalik samarali natija beradi. Semantik tahlilni amalga oshirish bevosita lingvistik resurslarga bog'liq. Leksik resurslarga lug'at, tezaurus va ontologiyalar kiradi. Lug'at leksik axborotning eng an'anaviy shakli hisoblanadi: ular birliklar (odatda, so'z yoki so'z birikmasi), tuzilishi va leksikaning qamrovi (konkret sohalarning atamalar lug'ati, umumiy leksika lug'ati, sinonim va paronimlar lug'ati va h.) bilan farqlanadi<sup>16</sup>.

Leksik resurslarning murakkabroq turlari tezaurus va ontologiyalardir. Tezaurus – semantik lug'at, ya'ni so'zlarning semantik aloqalari – sinonimlar, assosiasiyalar taqdim etilgan lug'at. Misol sifatida rus tili uchun ijtimoiy-siyosiy leksikani qamrab oluvchi axborot-qidiruv RuTez tezaurusini keltirish mumkin<sup>17</sup>. Ontologiya tushunchasi tezaurus tushunchasi bilan chambarchas bog'liq<sup>18</sup>.

<sup>14</sup> Арефьев, Николай Викторович. Методы построения и использования компьютерных словарей сочетаемости для синтаксических анализаторов русскоязычных текстов//тема диссертации-ВАК РФ 05.13.11. 2012; 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; Бушгедт, Владислав Андреевич. 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>

<sup>15</sup> Abduraxmonov G., Sulaymonov A., Xoliyorov X., Omonturdiyev J. Hozirgi o'zbek adabiy tili. Sintaksis.// Toshkent: O'qituvchi, 1979, 18-39-betlar.; G'ulomov A., Asqarov M. Hozirgi o'zbek adabiy tili. Sintaksis. //Toshkent: O'qituvchi, 1965, 27-53-betlar; Maxmudov N, Nurmonov A. O'zbek tilining nazariy grammatikasi. Toshkent, 1995, 37-58-betlar; Maxmudov N. Semantiko-sintaksicheskaya assimetriya v prostom predlozenii uzbekskogo yazyka ADD. Tashkent, 1984, s 45.

<sup>16</sup> 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.

<sup>17</sup> De Smet W., Moens M.F. Cross-language linking of news stories on the web using interlingual topic modelling // Proceedings of the 2Nd ACM Workshop on Social Web Search and Mining. – SWSM '09. – New York, NY, USA: ACM, 2009. – Pp. 57–64.

<sup>18</sup> Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. – М.: Наука. 1986.

Ontologiya – turli vazifalar uchun ko‘p foydalanishga qaratilgan muayyan bilim sohasi tushunchalari, obyektlar majmui. Ontologiyalar tildagi mavjud so‘z birikmalari asosida yaratilishi mumkin. Bu holda ular lingvistik ontologiya deb ataladi. Xuddi shunday lingvistik ontologiya WordNet<sup>19</sup> tizimi hisoblanib, u inglizcha so‘zlarni o‘z ichiga olgan katta leksik manba. Unda ot, sifat, fe‘l, ravish hamda ularning bir necha turdagi semantik bog‘lanishlari keltirilgan.

Semantik komponent ontologiya asosida amalga oshirilishi mumkin. Ushbu yondashuvni amalga oshirish uchun matn resurslariga mos tushunchalarga ta‘rif berish; har bir sinf uchun matn resurslari tarkibining formal strukturasi aniqlash kabilar zarur.

Semantik tahlil formal semantik yondashuv asosida ham amalga oshirilishi mumkin. Bu holda gapning qiymati matematik formula yordamida ifodalanadi. Semantik komponent matnning semantik va sintaktik tahlil natijalariga asoslanib, har bir gapning sintaktik tuzilishini tasvirlovchi belgiga ega bo‘ladi<sup>20</sup>.

Qidiruv tizimlari SQ (semantik qidiruv)ni amalga oshirishda juda ko‘p muammolarga duch keladi. Jumladan, qidiruv so‘rovini kiritishda foydalanuvchi aynan nimani nazarda tutganini, ya‘ni turli kontekstlarda so‘z yoki iboraning bir qancha ma‘nolari mavjudligini aniqlash. Semantik qidiruv kompyuter yo‘li bilan hal qilib bo‘lmaydigan muammolarni yechishda yordam berolmaydi. Semantik qidiruv nafaqat kontekstni, balki boshqa omillar tahlilini ham inobatga oladi. Data Science Blogathon nashri o‘zining “Step by Step Guide to Master NLP - Semantic Analysis” nomli ilmiy maqolasida semantik tahlil qilishning qadamlarini sanab o‘tgan. Unga ko‘ra:

1. Semantik tahlilning birinchi tarkibiy qismi so‘z ma‘nosini ajratish bo‘lib, unda biz alohida so‘zlarning ma‘nosini o‘rganamiz. Bu komponentni leksik-semantika deb nomlaganlar<sup>21</sup>

<sup>19</sup> <http://wordnet.princeton.edu>

<sup>20</sup> Гаршина В.В., Богоявленская Ю.А. Разработка лингвистического парсера русского языка // <http://www.vestnik.vsu.ru/pdf/analiz/2012/02/2012-02-29.pdf>

<sup>21</sup> Chirag Goyal-June 23,2021// <https://www.analyticsvidhya.com/blog/2021/06/part-9-step-by-step-guide-to-master-nlp-semantic-analysis/>

2. Alohida soʻzlarning kombinatsiyalarini, brikuvchilarini oʻrganish. Bunda otli brikmalar, feʼlli brikmalar, frazalar, lokatsiya va kollakatsiya hodisalarining natijalari nazarda tutilgan.
3. Semantik takomillashtirish algoritmlarini matn maʼlumotlari namunalari bilan taʼminlash orqali Machine learning algoritmlari yordamida avtomatik ravishda semantik tahlilni amalga oshirish mumkin, qidiruv tizimi va uning oldingi natijalari asosida aniq prognozlar qilishga oʻrgatishimiz mumkin.
4. Semantik analiz elementlarini alohida oʻrganish lozimligi taʼkidlanadi.

– Giponimiya<sup>22</sup>

– Omonimiya

– Polisemiya

– Sinonimiya

– Antonimiya

5. Semantik tizimni loyihalashtirish. Tizimning asosi boʻlgan maʼlumotlarni tizimlashtirish, yaʼni:

– Obyektlarni yaratish

– Izohlar

– Bogʻlanishlar

– Strukturalar, masalan, ot, feʼl brikmalarining strukturalari.

6. Semantik tahlil maʼnoni ifodalash uchun quyidagi yondashuvlardan foydalanadi –

– Birinchi tartibli predikatlar mantiqi

– Semantik tarmoqlar

– Formalar

– Qoidalarga asoslangan arxitektura

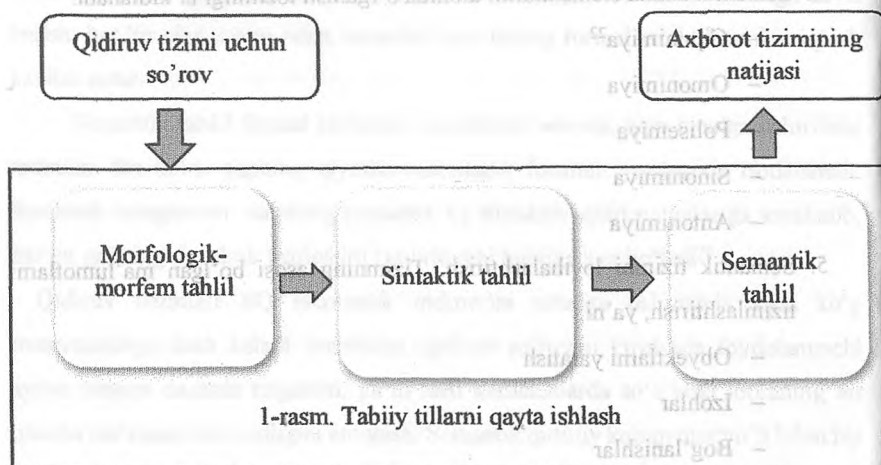
– Case grammatikasi<sup>23</sup>

<sup>22</sup> Chirag Goyal-June 23,2021// <https://www.analyticsvidhya.com/blog/2021/06/part-9-step-by-step-guide-to-master-nlp-semantic-analysis/>

<sup>23</sup> Charles J. Fillmore, "A Private History of the Concept Frame." *Concepts of Case*, ed. by René Dirven and Günter Radden. Gunter Narr Verlag, 1987// <https://www.thoughtco.com/case-grammar-linguistic-theory-1689744>

— Konseptual grafik

tizimi yaratiladi. Tabiiy tillarni qayta ishlash jarayonida morfologik, sintaktik va semantik tahlil ketma-ketlik bilan amalga oshiriladi. Kirish tilidagi matnlarni semantik tahlil qiluvchi tizimning natijasi morfologik va sintaktik tahlilga asoslanadi. Bunday tizimlarning tahlil jarayoni quyidagi chizmada yaqqol namoyon bo'ladi.



1-rasm. Tabiiy tillarni qayta ishlash

Xulosa o'rnida aytish mumkinki, avtomatik semantik tahlil va qidiruv kompyuter lingvistikasining dolzarb hamda murakkab vazifalaridan biri sifatida baholanadi. Ko'rib o'tilganidek, semantik analiz va sintez matnning avtomatik tahlilidagi ahamiyati katta.

### 1.2. Matnlarni semantik analiz qilishning model va metodlari

Dunyo kompyuter tilshunosligida semantik analiz va analizator borasida bir qator olimlar ilmiy izlanish olib borishgan. A.V.Tuzov, N.A.Mixail, A.V.Sokirko, N.A.Schlaefer, A.V.Mochalova, D.Sh.Suleymanov, A.R.Gatiatullin,



Ye.A.Kanevskiy, K.K.Boyarskiy<sup>24</sup> kabi olimlarning tadqiqotlarini ana shunday ishlar sifatida qayd etish mumkin. Ammo semantik analizator tuzish yo'llari, maqsad va vazifalari maxsus tadqiqot obyekti bo'lmagan. Dunyoning qator olimlari tomonidan tabiiy tillar uchun semantik analizator yaratish nazariyalari ishlab chiqilgan bo'lsa-da, shu paytgacha mukammal semantik analizator yaratilmagan. Semantik analizni amalga oshiruvchi tizim turli tillarda turlicha model va algoritmlarga asoslanadi. Bu tabiiy tillarning grammatikasiga bog'liq. Yuqorida sanab o'tilgan elementlarning har birini modellashtirish orqali mukammal tizim yaratish mumkin. Joriy tadqiqod ishida omonim, polisemantik va polifunksional so'zlarni semantik analiz qiluvchi tizimni yaratish uchun matematik model, algoritim, ma'lumotlar bazasi strukturasi ishlab chiqiladi. Omonimlikni aniqlash ko'plab kompyuter dasturlari uchun foydalidir. Ayniqsa, kompyuter lingvistikasi, xususan qidiruv tizimlarini ishlash tezligi, ma'lumotlar hajmining kamayishi va so'rovlarning kamayishi kabi imkoniyatlar yaratadi. Omonimiyani aniqlashning mavjud yondoshuvlari an'anaviy deterministikaga bo'linadi, ya'ni lokal va global sintaktik tahlil, sintaktik lug'at va ehtimolliklarga asoslanadi. Grammatik va statistik hodisalardan foydalangan holda korpusdagi so'zlarning belgilari asosida omonimiya hodisasi aniqlanadi. Omonimiyani muvaffaqiyatli chiqishi uchun aniq natijalar kerak. Bundan tashqari ma'lumotlar hajmining kattaligi ularni qayta ishlashda sekinlashtiradi, bu esa **kombinator portlash** deb ataladi. Bunday algoritmlarga

<sup>24</sup> Тузов, В.А. Компьютерная семантика русского языка / В.А. Тузов. – СПб.: Изд-во С.-Петербурга, 2004. – 400 с. Н.А. Михаил // Использование семантических связей в системах анализа текстов на естественном языке: дис. ... канд. техн. наук: 05.13.11 // Назаров Михаил Александрович, М., 2002; Сокирко, А.В. Семантические словари в автоматической обработке текста (по материалам системы ДИАЛИНГ): дис. ... канд. техн. наук: 05.13.17 / Сокирко Алексей Викторович. – М., 2001. – 120 с.; Мочалова, А.В. Алгоритм семантического анализа текста, основанный на базовых семантических шаблонах с удалением / А.В. Мочалова // Научно-технический вестник информационных технологий, механики и оптики. – 2014.– № 5. – С. 126 –132.; Mathematical model of an ontological-semantic analyzer using basic ontological-semantic patterns / A.V.Mochalova, V.A.Mochalov // Lecture Notes in Artificial Intelligence, Proceedings of 15th Mexican International Conference on Artificial Intelligence. – 2016. – P. 53–66.; Семантический анализ. Некоторые проблемы автоматической обработки текста / А.В.Мочалова, В.П.Захаров, В.А.Кузнецов, В.А.Мочалов. – Петрозаводск: ПИН, 2016. – 60 с.; Семантический анализатор русскоязычного текста для вопросно-ответной системы // ДИССЕРТАЦИЯ на соискание ученой степени кандидата технических наук. – Петрозаводск, 2017. – 128 с.; Сулейманов Д.Ш., Гатиатуллин А.Р. Структурно-функциональная компьютерная модель татарских морфем. — Казань: Фэн, 2003. — 220 с. – С. 152.; Каневский Е.А., Боярский К.К. Семантико-синтаксический анализатор SemSin // Компьютерная лингвистика и интеллектуальные технологии. 2012; Боярский К.К., Каневский Е.А. Предсинтаксический модуль в анализаторе SemSin // Интернет и современное общество. Санкт-Петербург, 2013. С. 280–286.

misol sifatida lingvistik protsessor komponenti ETAP va lingvistik analizator DIALING olish mumkin. Birinchi, ETAP da sintaktik analizning filtrlash metodidan foydalaniladi, bunda soʻzlar orasidagi bogʻlanishlar, soʻz brikmalari tahlil qilinadi va berilgan matn ichida oʻchiriladi(filtrlanadi). Bu jarayon soʻzlar orasidagi bogʻlanishlar daraxt koʻrinishiga kelguniga qadar davom etadi. DIALING- tizimi turli soʻz brikmalarining kombinatsiyalariga asoslangan holda analiz qiladi. ENGTWOL-ingliz tili morfologik analizatori barcha sintaktik qoidalarga asoslangan holda gap tuzilishidagi oldingi va keyingi soʻzlar orqali analizni amalga oshiruvchi lokal sintaktik metod. Yuqorida keltirilgan algoritmlarning barchasi deterministik yaʼni qoidalarga asoslangan holda analizni amalga oshiradi. Ushbu noaniqlikni aniqlashning yana bir yoʻli korpusda berilgan misollar yordamida olingan ehtimollikdan foydalanishdir. Ushbu yondoshuvning namunasi sifatida M Xarst ism (ot)lardagi noaniqlikni bartaraf etish uchun omograflar, lokal kontekstlarga asoslangan katta korpus namunasini ishlab chiqqan.

Polisemantik soʻzlarni ayrim grammatik va leksik xususiyatlari kontekstlarda jamlangan boʻladi. Bu bosqich "boshqariladigan taʼlim" deyiladi. Algoritm soʻzning noaniq ishlatilishi uchun bir xil kontekstlarni yaratadi va toʻgʻri maʼnolarni tanlash orqali "oʻz-oʻzini oʻqitish" ni amalga oshiradi. Turli soʻzlar uchun algoritmnining aniqligi 73% dan 100% gacha edi. D.Yarovski tomonidan tuzilgan shunga oʻxshash tizim taxminan 96% aniqlikka ega.

Yana bir keng tarqalgan ehtimollik yondashuvi Yashirin-Markov modelidan(HMM) foydalanishga asoslangan algoritm. Algoritmning asosiy gʻoyasi jumladagi har bir soʻz uchun quyidagi funksiyaning qiymatini maksimal darajada oshiruvchi Grammatik tegni tanlashdan iborat:

$$P(\text{soʻz}|\text{tag}) * P(\text{tag}|\text{oldingi n ta teg})$$

Bu yerda,  $P(\text{tag}|\text{oldingi n ta teg})$  - shartli ehtimollik (korpusda hisoblangan), oldindan aniqlangan n ta teg bilan joriy tegning paydo boʻlish ehtimolligi,

$P(so'z|tag)$  –shartli ehtimollik (korpus ma'lumotlari yordamida hisoblanadi) so'zning Grammatik xususiyatlari asosida aniqlangan tegi. HMM modeli murakkab hisoblashlarga ega bo'lsa ham, amalda turli xil soddalashtirishlarga ega. Ingliz tili grammatikasi uchun 96% aniqlik bilan so'zlarning ma'nolarini farqlaydi. Ushbu modelni ingliz tili bilan solishtirganda rus tiliga qo'llash qiyin bo'lishi mumkin, rus tilidagi so'z shakllanishi va so'z o'zgarishining boyligini hisobga olgan holda belgilangan korpuslarni juda katta hajmda bo'lishini talab qiladi.

Rus tilshunosligida omonimlarni o'rganishga bag'ishlangan tadqiqotlar juda ko'p. Omonimiya hodisasi R.A.Budagov [1953], V.V.Broun [1958], L.A.Novikov [1961], A.G.Shyepin [1962], V.M.Gluxov [1970], I.V.Arnold [1973], N.G.Chausova (1990), B. A. Bobnev (2010) va boshqalarning ishlarida maxsus tadqiq etilgan. Rus tilida omonim so'zlarni farqlashda statistik usullardan foydalanilgan. Statistik usulning asosiy vazifasi n-grammga ajratish, ya'ni kontekstdagi kirish so'zining birikuvchilari aniqlab, baholash metodlari yordamida baholanadi<sup>25</sup>:

- Chastota metodi-
- Bayes metodi;
  - o Parallel metod;
  - o Ketma-ketlik;
- Dinamik dasturlash metodi;
  - o Viterbi metodi;
  - o Viterbi chastotali algoritmi;
  - o To'g'ri va teskari harakat algoritmi;
  - o Chastotali to'g'ri va teskari harakat algoritmi.

Ushbu metodlar orqali baholashda kontekstni tokenizatsiya va teglashtirish jarayonlari olib boriladi. Ko'plab usullar so'z turkumlarini teglashda joriy so'z va uning atrofidagi (qo'shni) so'zlaridan foydalaniladi. Rus tilidagi SemSin tizimida

<sup>25</sup> Рысаков С.В. Клышинский Э.С. Статистические методы снятия омонимии // Новые информационные технологии в автоматизированных системах. -2015

morfologik omonimiyani *olib tashlash* (*снятие омонимии*) masalasi yechilgan. Rus tilida bu termin *снятие омонимии* deb nomlangan, chunki semantik analizatorning elementlarini aniqlanganini matn ichidan ajratib olib tashlansa, qolgan elementlarni analiz qilish- tizimning ishlash tezligini oshiradi, algoritmlarning aniq va to'g'ri ishlashini ta'minlaydi.

Yevropa tillarida omonimiya bilan Sh.Balli [1955], B.Trinka [1967] Heinz Dieter Maas, Christoph Rösener, va Axel Theofilidis (Nemis tili) va boshqa olimlar shug'ullangan.

Omonimiya K.Axanov (qozoq tili) [1956], M.Mirtojiyev (o'zbek tili) [1964], A.X.Axyamov (boshqird tili) [1966], R.Bekdjanova (qirg'iz tili) [1967], B.S.Salimgareyeva (tatar tili) [1972], A.Geldimuradov (turkman tili) [1974], A.Memetov (qrim-tatar tili) [1977], A.A.Gasanov (ozarbayjon tili) [1980] kabi turkolog olimlarning tadqiq mavzusi bo'ldi. Bu turkolog olimlar A.I.Smiriskiy, V.V.Vinogradov, O.S.Axmanova va boshqa tilshunos olimlar ishlab chiqqan omonimiya nazariyasiga tayanadilar.

Tatar tilida ham boshqa agglutinative turkiy tillar guruhlarida bo'lgani kabi morfema tilshunoslikning ham semantik, ham sintaktik ma'lumot tashuvchi eng muhim ma'noli birligidir. TuganTel - tatar tili korpusida morfologik polisemiyani aniqlash algoritmi, ma'lumotlar bazasi ishlab chiqilgan. Gataullin Ramil Raisovich o'zining nomzodlik dissertatsiyasida tatar tilidagi omonimiya va polisemiyani olib tashlash bo'yicha model va algoritmlar haqida mulohaza yuritgan. Tatar tilida ham omonimiya va polisemiyani farqlashda Markov modellaridan foydalanish tavsiya qilingan<sup>26</sup>.

O'zbek tilining an'anaviy leksikologiyasida F.Kamolov, Z.Ma'rufov, Ya.Pinxasov, S.Usmonov, Sh.Rahmatullayev, A.Hojiyev<sup>27</sup> va boshqalarning nomi

<sup>26</sup> Гатауллин Рамиль Раисович. Методы, модели и программный инструментарий разрешения многозначности в текстах// дисс-ВАК РФ-05.13.11 канд.наук.-2019

<sup>27</sup> Камол Ф. Лексикология // Хозирги замон ўзбек тили. – Тошкент: Фан, 1957. – 437 б.; Маъруфов З. Хозирги замон ўзбек тили. – Тошкент: Фан, 1957. – 354 б.; Пинхасов Я.Д. Хозирги замон тили лексикаси. – Тошкент:





B.P.Kobrisov, T.I.Reznikova<sup>30</sup>, B.P.Kobrisov<sup>31</sup>, V.V.Kukanova<sup>32</sup>, A.A.Kretov<sup>33</sup>, A. Ye.Aleksandr<sup>34</sup>, Y.E. Yermolayeva<sup>35</sup> lar shu masalalar yechimiga bag'ishlangan qator ishlarni e'lon qilishgan.

O'zbek kompyuter lingvistikasi yo'nalishida amalga oshirilgan sanoqli tadqiqotlarda kompyuter xotirasi omonim birliklarni "tanishi" va "o'qiy olishi" ga mo'ljallangan tahlil dasturlarini yaratish borasida harakatlar mavjud; o'zbek tilidagi omonimlarni teglash muammolari bo'yicha ayrim mulohazalar, omonimiyani aniqlash algoritmini tuzish bo'yicha dastlabki harakatlar amalga oshirilgan. Tadqiqotchilar M.Abjalova<sup>36</sup>, D.Axmedova<sup>37</sup>, Sh.Hamroyeva<sup>38</sup>, O'.Xoliyorov<sup>39</sup>,

<sup>30</sup>Кустова Г.И., Ляшевская О.Н., Падучева Е.В., Рахилина Е.В. Семантическая разметка лексики в Национальном корпусе русского языка: принципы, проблемы, перспективы // Национальный корпус русского языка: 2003-2005. Результаты и перспективы. М.: Индикс, 2005. – С. 155-174; Ляшевская О.Н. Топологические классы имен в семантической разметке Национального корпуса русского языка // Труды международной конференции «Корпусная лингвистика-2008» (Санкт-Петербург, 6-10 октября 2008 г.). СПб.: С.-Петербургский гос. университет, Факультет филологии и искусств, 2008. – С. 276-284; Рахилина Е.В., Кобрицов Б.П., Кустова Г.И., Ляшевская О.Н., Шеманова О.Ю. Многозначность как прикладная проблема: семантическая разметка в Национальном корпусе русского языка // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2006». – М., 2006. – С. 445-450; Рахилина Е.В., Кустова Г.И., Ляшевская О.Н., Резникова Т.И., Шеманова О.Ю. Задачи и принципы семантической разметки лексики в НКРЯ // Национальный корпус русского языка: 2006-2008. Новые результаты и перспективы / отв. ред. В.А.Лукин. СПб.: Не-стор-История, 2009. – С.215-239; Ляшевская О.Н. Топологические классы имен в семантической разметке Национального корпуса русского языка // Труды международной конференции «Корпусная лингвистика-2008» (Санкт-Петербург, 6-10 октября 2008 г.). СПб.: С.-Петербургский гос. университет, Факультет филологии и искусств, 2008. – С. 276-284.

<sup>31</sup> Кобрицов Б.П. Модели многозначности русской предметной лексики: глобальные и локальные правила разрешения омонимии. Автореф. канд. филол. наук. – М.: РГГУ, 2004; Кобрицов Б.П., Ляшевская О.Н., Шеманова О.Ю. Поверхностные фильтры для разрешения семантической омонимии в текстовом корпусе // Кобозева И.М., Нариньяни А.С., Селегей В.П. (ред.), Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции Диалог-2005. – М., 2005.

<sup>32</sup> Куканова В.В. Принципы семантической разметки национального корпуса калмыцкого языка // [http://kalmcorp.ru/sites/default/files/kukanova\\_25.pdf](http://kalmcorp.ru/sites/default/files/kukanova_25.pdf)

<sup>33</sup> Кретов А.А. Анализ семантических помет в НКРЯ // <http://ruscorp.ru/sbornik2008/11.pdf>

<sup>34</sup> Аникин, Александр Евгеньевич // Опыт семантического анализа праславянской омонимии на индоевропейском фоне // тема диссертации и автореферата по ВАК РФ 10.02.03, кандидат филологических наук. – М.: 1983

<sup>35</sup> Y. E. Yermolayeva Семантическое варьирование диалектного слова в русских говорах Башкирии: В связи с проблемой разграничения полисемии и омонимии // тема диссертации и автореферата по ВАК РФ 10.02.01, кандидат филологических наук // Ермолаева, Юлия Александровна. -2000

<sup>36</sup> Abjalova M.A. O'zbek tilidagi matnlarni tahrir va tahlil qiluvchi dasturning lingvistik modullari (Rasmiy va ilmiy uslubdagi matnlar tahriri dasturi uchun): Filol. fan. bo'yicha falsafa doktori (PhD)...diss. – Farg'ona, 2019. – 164 b.

<sup>37</sup> Axmedova D. Atov birliklarini o'zbek tili korpuslari uchun leksik- semantik teglashning lingvistik asos va modellarini. Filologiya fanlari bo'yicha fal.dok-ri.diss. - Buxoro, 2020. – 163 b.

<sup>38</sup> Hamroyeva Sh. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari: Filol. fan. bo'yicha falsafa doktori (PhD)...diss. – Qarshi, 2018. – 250 b

<sup>39</sup> Xoliyorov O'. O'zbek tili ta'limiy korpusini tuzishning lingvistik asoslari. Filol. fan. bo'yicha fal. dok-ri. diss. avtoref. – Termiz, 2021. – B. 16. – 52 b

Sh.Gulyamova<sup>40</sup> kabi qator tadqiqotchilarning izlanishlari bunga yaqqol misol bo'ladi. Tilshunos tadqiqotchilar tomonidan tavsiya qilingan lingvistik modellar asosida semantik analizator yaratishga dastlabki qadam qo'yamiz.

<sup>40</sup>Gulyamova Sh.Q. Semantik analizatori uchun omonimlikni farqlash omillaening ayrim masalalari xususida // "O'zbekistonning umidli yoshlari" mavzusidagi 6-son Respublika ilmiy talabalar, magistrlar yosh tadqiqodchilar va mustaqil izlanuvchilar uchun onlayn konferensiyasining materiallari to'plami. – Toshkent: Tadqiqot, 2021. – B. 74-76.



### 1.3.O'zbek tili semantik analizatorida matnlarni teglash usullari

So'nggi yillarda matnlarni semantik tahlil qilish muhim ahamiyat kasb etadi. Axborotning rivojlanishi Internet resurslarini haddan ortiq ko'p yuklanishi muammosini ko'paytiradi. XXI asr boshlarida Internetdagi sahifalar soni 4 milliarddan oshdi va har kuni 7 millionga ko'paymoqda. Strukturalanmagan ma'lumotlar bilan shug'ullanadigan foydalanuvchilar, ko'plab tashkilotlar va shaxslar tomonidan taqdim etilgan matnli axborotlar ma'lumotlarning katta qismini tashkil qiladi. Hozirgi vaqtda jahon bozorida matnli ma'lumotlarni semantika nuqtai nazaridan tahlil qiluvchi dasturiy mahsulotlar mavjud. Matnlarni sintaktik-semantik tahlil tarmog'ini yaratishga imkon beruvchi AOT va Semantic analyzer Group mahsulotlarini alohida ajratib ko'rsatish mumkin. Undan tashqari IBM kompaniyasi tomonidan ishlab chiqilgan IBM text Miner platformasi matnlarga izoh berish, tasniflash, klasterlash, kalit so'zlarni qidirish va tuzish uchun yordamchi dasturlarni o'z ichiga qamrab oladi. Rossiyaning Yandex News tizimi ham ma'lumotlarni avtomatik ravishda yangiliklarni guruhlash imkonini beradi, hujjatlar to'plami asosida syujetlar va maqolalardan tezislar tuzadi. InfoStream xizmatiga kirish esa hujjatlarning semantik yaqinligini hisobga olgan holda tezkor ma'lumotlar olish imkonini beradi. Bundan tashqari mobil agregator tizimlaridan biri bo'lgan Summary 2013-yil mart oyida Yahoo! tomonidan sotib olindi, ammo u rus, o'zbek tilidagi matnlarni qayta ishlash imkonini bermas edi.<sup>41</sup> Shunday qilib, hozirda mavjud bo'lgan dasturlar to'plami muammolarni to'liq hal qila olmaydi. Bu turli tabiiy tillarning grammatikasi, har xil turdagi matnlarni semantik tahlil qilishning murakkabligi, noaniqligi bilan izohlanadi. Ishlab chiqilayotgan yechim mavjud tizimlardan farqli o'laroq o'zbek tili semantik analizatorining debochasi desak, mubolag'a bo'lmaydi. Ushbu tizimda dastlab shakldosh so'zlar va ularni qurshovda aniqlash, ma'nolarini farqlash algoritmlari qo'llaniladi, ushbu masala o'z yechimini topgach polisemantik va polifunksional so'zlarni qurshovda farqlash algoritmlari yaratiladi va bu algoritmlar asosida dasturiy mahsulot ishlab chiqiladi. Yaratilajak

<sup>41</sup> Ладанова Е.О. 1, Ямашкин С.А. "Семантический анализатор для выделения фактов из текстовых сообщений".// Международный научно-исследовательский журнал, 2017.

axborot tizimi quyidagi funksional vazifalarni muvaffaqiyatli hal qilish imkonini beradi:

- O'zbek tilidagi mavjud omonim, polifunksional va polisemantik so'zlarni lug'at ko'rinishida namoyish etish;
- Matnda mavjud polifunksional va polisemantik so'zlarni aniqlash;
- Bir so'z turkumi va turli so'z turumlari doirasidagi omonim so'zlarni matnda aniqlash;
- Omonim, polifunksional va polisemantik so'zlar ishtirokidagi gaplardan namuna keltirish;
- Foydalanuvchilar uchun ro'yhatdan o'tish;
- Ro'yhatdan o'tgan foydalanuvchilar uchun tizimda mavjud bo'lmagan so'zlarni kiritish imkoniyati;
- Katta hajmdagi matnlardan omonim, polifunksional va polisemantik so'zlarni ajratish va ularni statistikasini namoyish etish.

Yuqorida sanab o'tilgan har bir funksional vazifaning bajarilishi lingvistik va matematik modellar, algoritmlar asosida amalga oshiriladi.

Tadqiqot ishini amalga oshirish rejasi quyidagi bosqichlarni o'z ichiga oladi:

- Matnlarni semantik tahlil qilish masalalarida mahalliy va xorijiy tajribalarini o'rganish va adabiy manbalarni sharhlash;
- Birinchi bosqichda, semantik lug'atlarni, matnli ma'lumotlarni tahlil qilish metodologiyasini ishlab chiqish;
- Omonim so'zlarni turkumlarga, ya'ni guruhlarga<sup>42</sup> ajratgan holda matematik modellar yaratiladi;
- Polifunksional va polisemantik so'zlarni semantik tahlil qiluvchi matematik model ishlab chiqiladi;
- Yaratilgan lingvistik modellar asosida farqlovchi omillar to'planadi va ma'lumotlar bazasiga kiritiladi;
- Matematik modellar asosida algoritmlar yaratiladi;

<sup>42</sup> Gulyamova Sh. Axmedova X. "O'zbek tili semantik analizatori uchun omonim so'zlar ma'lumotlar bazasini shakllantirish masalasi xususida" "So'z san'ati xalqaro jurnali": <http://dx.doi.org/10.26739/2181-9297-2021-3-102>

- Algoritmlar asosida dasturiy ta'minot yaratiladi;
- Yaratilgan dasturlar kompleksi sinovdan o'tkaziladi va taklif qilingan yechimning samaradorligini ya'ni katta hajmdagi ma'lumotlarni semantik tahlilining to'g'riligini tekshiriladi.

Tadqiqot ishining amalga oshirilishi o'zbek tili semantik analizatori yaratilishidagi dastlabki qadam desak mubolag'a bo'lmaydi. Yuqoridagi vazifalarning bajarilishi tilimizdagi qator muammolarni yechishga yordam beradi. Tabiiy tillarni qayta ishlash jarayoni (NLP-Natural languages processing) ning muammolarini ketma-ket modellashtirish zamon talabidir. Omonim so'zlarning qurshovda aniqlanishi bevosita so'z turkumlarini teglash muammosiga duch keladi. Ayniqsa, so'z turkumlarini teglash qadimiy va eng mashhur muammolardan hisoblanadi. So'z turkumlarini teglashtirishdagi asosiy maqsadimiz gap tuzilishini modellashtirishdan iborat. Masalan, gapni o'qish va qanday so'zlar ot, olmosh, fe'l, ravish va hokazo vazifasini bajarishini aniqlash nutq teglarining bir qismi hisoblanadi. *Korpus tilshunosligida so'z turkumlarini teglash, grammatik teglash yoki so'z turkumlariga ajratish deb ham ataladi, bu matn (korpus) dagi so'zni nutqning ma'lum bir qismiga mos keladigan tarzda belgilash, uning ta'rifi va kontekstidan kelib chiqqan holda, ya'ni o'zidan oldin va keyin kelgan va o'zaro bog'liq ibora, jumla yoki paragrafdagi so'zlar.*

Omonim so'zlarni semantik analiz qilish jarayonida biz Michayel Collins<sup>43</sup> ilmiy maqolalarida keltirilgan teglash usullariga murojaat qilamiz. Masalan, o'zbek tilidagi ikki so'z turkumi doirasidagi omonim so'zlarning ma'nolarini kontekstda aniqlash jarayonida Qoidalarga asoslangan teglash usuli ham, Stoxastik teglash usuli ham muhim ahamiyat kasb etadi.

**Qoidalarga asoslangan teglash.** So'z turkumlarini avtomatik teglash bu tabiiy tilni qayta ishlash jarayoni bo'lib, unda statistik metodlar bilan qoidalarga asoslangan usullardan ko'ra ko'proq muvaffaqiyatga erishilgan. Odatda qoidalarga asoslangan yondashuvlar noma'lum yoki noaniq so'zlarga teglar belgilash uchun

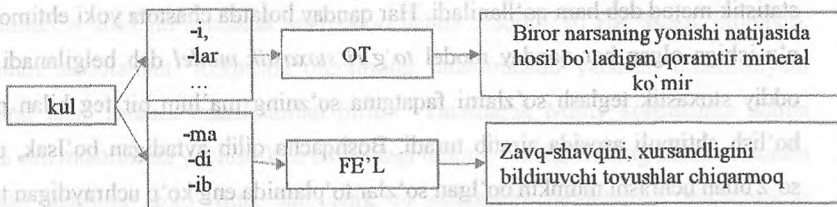
<sup>43</sup> Michayel Collins "Tagging with Hidden Markov Models" //2011



kontekstli ma'lumotlardan foydalanadi. So'zning lisoniy xususiyatlarini ajratish, oldingi va keyingi so'zini hamda boshqa jihatlarini tahlil qilish orqali amalga oshiriladi. Masalan, oldingi o'rindagi so'z kitob bo'lsa, u holda bu so'z ot bo'lishi kerak. Shu kabi boshqa so'z turkumlari uchun ham qoidalar keltirish mumkin:

Agar noma'lum X so'zidan oldin aniqlovchi va undan keyin ot kelsa, u holda bu so'z sifat deb belgilanadi.

**Kul** omonim so'zining xususiyatlari bunga misol bo'la oladi



Bu yerda **kul** so'zi ot so'z turkumiga mansub bo'lib kontekstda uchraganida tarkibida

-ning -lardan -chalarga -laringiz

-ga -larning -chalardan -laring

-dan -larga -chalarning -ng

-da -cha -chalarni -ingiz

-ni -chalar -lari -imiz

-larni -chalarni -larim -chalarim

-larga -chalarning -larimiz -chalari

kabi qo'shimchalar bo'lishi mumkin. **Kul** so'zi fe'l so'z turkumiga mansub bo'lganida esa

-b, -ma, -maydigan, -gach

-may, -maydi, -gin, -magach

-ydi, -masin, -guncha, -r,

-sin, -ydigan, -maguncha, -moq

kabi fe'ning lug'aviy va sintaktik shakl yasovchi qo'shimchalari va ularning kombinatsiyalaridan iborat qo'shimchalar brikmasi qo'shiladi. Bu esa aniq qoida hisoblanadi. O'zbek tilida ba'zi omonimlar aynan shunday qoidalar to'plamiga bo'ysunadi. Omonim so'zlarning paydo bo'lishi sabablaridan ko'ra ularning qaysi so'z turkumlariga mansubligi, qanday qoidalar asosida ma'nolarini farqlanishi muhim. Endi esa so'z turkumlarini teglashning Stoxastik usulini ko'rib chiqamiz.

*ST teglashning Stoxastik usuli.* "Stochastik tagging" atamasi ST tegi muammosiga turli xil yondashuvlarni nazarda tutishi mumkin. Ba'zi manbalarda statistik metod deb ham qo'llaniladi. Har qanday holatda chastota yoki ehtimollikni o'z ichiga olgan har qanday model *to'g'ri stoxastik model* deb belgilanadi. Eng oddiy stoxastik teglash so'zlarni faqatgina so'zning ma'lum bir teg bilan paydo bo'lish ehtimoli asosida ajratib turadi. Boshqacha qilib aytadigan bo'lsak, ushbu so'z bilan uchrashi mumkin bo'lgan so'zlar to'plamida eng ko'p uchraydigan teg bu so'zning noma'lum misoliga berilgan tegdir. So'zlarning chastotali yondashuviga alternativ teglar ketma-ketligining yuzaga kelish ehtimolini hisoblashdir. Bu ba'zan **n-gramm** yondashuvi deb ataladi<sup>44</sup>, bunda ma'lum bir so'z uchun eng ma'qul teg avvalgi **n** teglar bilan yuzaga kelish ehtimoli bilan belgilanadi. Ushbu yondashuv ilgari aniqlanganidan ancha mantiqiydir, chunki u alohida so'zlar uchun teglarni kontekst asosida ko'rib chiqadi. Stoxastik teglashda kiritilishi mumkin bo'lgan keyingi murakkablik darajasi avvalgi ikkita yondashuvni ikkala teglar ketma-ketligi ehtimoli va so'z chastotasini o'lchovlari yordamida birlashtiradi. Bu esa Yashirin Markov Modeli deb ataladi. Ingliz tili grammatikasida bu modeldan foydalanish juda qulay va samarali. Chunki ingliz tilida so'zlarga qo'shiladigan lug'aviy va sintaktik shakl yasovchi qo'shimchalar o'mida predloglar uchraydi, o'zbek tilida esa aksincha. O'zbek tilida Stoxastik teglash usulidan bir so'z turkumi doirasidagi omonim so'zlarni yoki leksik, morfologik jihatdan bir biriga yaqin so'z turkumlari doirasidagi omonim so'zlarni farqlash jarayonida foydalanishni taklif qilamiz.

<sup>44</sup> Divya Godayal An introduction to part-of-speech tagging and the Hidden Markov Model.

<https://www.freecodecamp.org/news/an-introduction-to-part-of-speech-tagging-and-the-hidden-markov-model-953d45338f24/>

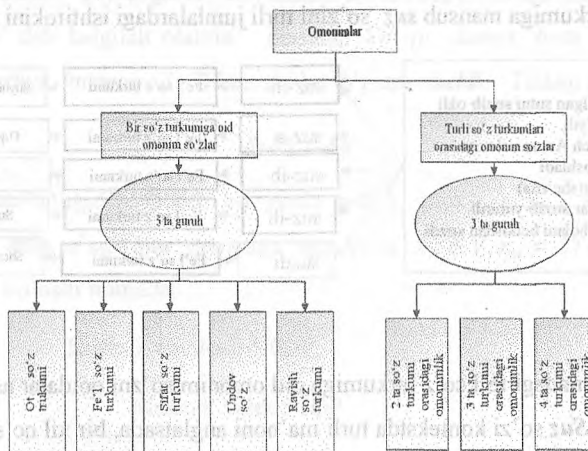
## II BOB. Ontosemantik qoidalarga asoslangan o'zbek tili semantik analizatori matematik modeli

### 2.1. O'zbek tilida omonimiyani bartaraf etish usullari

Milliy til korpus tizimida so'zlarni semantik tahlil qilish jarayonida qator lingvistik va matematik modellarga ehtiyoj seziladi. Ushbu bobda o'zbek tilida mavjud omonim (shakldosh) so'zlarni ma'noviy guruhlarga ajratish, ularni qurshovda aniqlay olish uchun qator matematik modellarni ko'rib chiqamiz. Ma'lumki, o'zbek tili semantik analizatori uchun omonim so'zlar asosiy lingvistik ta'minot hisoblanadi. Korpusda omonimlar muammosini yechish, omonimiyani bartaraf etish muhim masalalardan biridir. Yaratilajak Milliy korpusimiz uchun mana shu muammoni yechish kun tartibidagi muhim vazifa. Buning uchun omonim so'zlar, ularning izohi, filtri, lingvistik va matematik modellarini, ma'lumotlar bazasini shakllantirish masalasi dolzarblik kasb etadi.

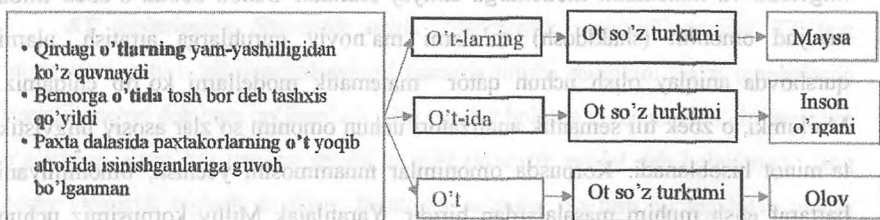
#### 2.1.1. Bir so'z turkumi doirasidagi omonimiyani bartaraf etish usullari

O'zbek tilida faqat bir so'z turkumi doirasida (ot, sifat, fe'l, ravish va taqlid so'z) omonimlik hosil qiluvchi so'zlar mavjud. Shakldosh so'zlarni so'z turkumlari doirasidagi tasnifini ushbu iyerarxiyada keltiriladi.

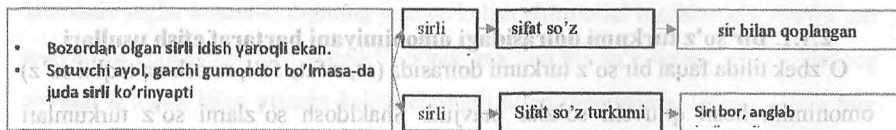


Rasm 1: Omonim so'zlarni so'z turkumlari doirasig tasnifi

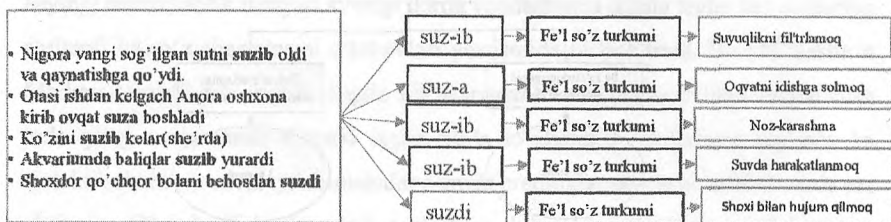
Ushbu iyerarxiyada keltirilgan bir soʻz turkumi doirasidagi omonim soʻzlarni farqlash modellarini koʻrib chiqamiz. Oʻzbek tilida eng koʻp uchraydigan omonimlar soʻzlar ot soʻz turkumiga mansub ekan. Ot soʻz turkumi doirasidagi omonim soʻzlar soni 1060 tasi aniqlangan. Shunday omonimlardan biri *oʻt* soʻzi matnda turli maʼnolarda kelishi mumkin.



Sifat soʻz turkumi doirasida 82 ta omonim soʻz toʻplandi. Misol tariqasida *sirli* soʻzini kontekstda qanday maʼnolarda uchrashini koʻramiz.

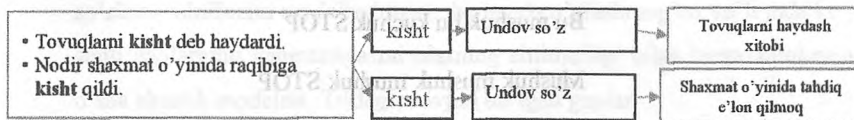


Feʼl soʻz turkumiga mansub *suz* soʻzini turli jumalalardagi ishtirokini qarasak.



Ushbu namunadagi feʼl soʻz turkumiga oid omonim soʻzni qoidalar asosida farqlab boʻlmaydi. *Suz* soʻzi kontekstda turli maʼnoni anglatmada, bir xil qoʻshimcha bilan brikib kelgan.

Undov so'zlar orasida ham omonimlik kuzatilganiga guvoh bo'ldik.



Undov so'zlar otlashish xususiyatiga ega, demak ikkala holatda ham egalik/kelishik/shaxs son qo'shimchalari bilan birikishi mumkin. Xuddi shunday ravish, olmosh so'z turkumlari orasidagi omonimlik uchraydi.

Yuqorida keltirilgan bir so'z turkumi doirasidagi omonim so'zlarni ma'nolarini farqlashda so'zlar teglashning Stoxastik usulidan foydalaniladi, chunki bir so'z turkumi doirasidagi omonimlikni farqlashda hech qanday fiksirlangan qoida mavjud emas. Agar bo'lsa ham umumiy emas.

Bir so'z turkumi doirasidagi omonim so'zlarni avtomatik aniqlash jarayonida tilni modellashtirish muammosiga duch kelamiz.

#### Tilni modellashtirish muammosi.

Aytaylik bizda o'zbek tildagi gaplar ketma-ketliklaridan iborat korpus bor. Korpusdagi til modellarini parametrlarini aniqlab olamiz.

Til modellari quyidagicha aniqlanadi. Birinchidan, tildagi mavjud so'zlar to'plamini  $V$  deb belgilab olamiz: masalan  $V = \{u, daraxt, bola, oila, ta'lim, ko'rmoq, mushuk, hurmoq \dots\}$ . Bilamizki bu to'plam **chekli**. Tildagi har bir gap

$$x_1 x_2 \dots x_n$$

so'zlar ketma-ketligidan iborat, bu yerda  $n \geq 1$ ,  $x_1, x_2, \dots, x_{n-1} \in V$  va faraz qilaylik  $x_n$  aniq o'zgarmas qiymatga ega ya'ni  $x_n \notin V, x_n = STOP$ . Gaplar quyidagicha bo'lishi mumkin.

Kuchuk vovilladi STOP

Mushuk kuldi STOP

Bu mushuk kuchukni ko'rdi STOP

Bu STOP

Bu mushuk bu kuchuk STOP

Mushuk mushuk mushuk STOP

STOP

$V^t - V$  to'plamdagi elementlar ketma ketligidan tashkil topgan gaplardan iborat cheksiz to'plam. Til modeli esa quyidagicha ta'riflanadi.

**1-Izoh: (Til modeli).** Til modeli chekli  $V$  to'plam elementlari va shunday  $p(x_1 x_2 \dots x_n)$  funksiyadan iboratki:

1. Har bir  $\langle x_1 x_2 \dots x_n \rangle \in V^t$  uchun  $p(x_1 x_2 \dots x_n) \geq 0$
2.  $\sum_{\langle x_1 x_2 \dots x_n \rangle} p(x_1 x_2 \dots x_n) = 1$

Demak  $p(x_1 x_2 \dots x_n) - V^t$  gaplarning ehtimollik taqsimotini aniqlovchi funksiya.

Ta'lim korpusidan til modelini o'rganishni bir misolda ko'rsatib beramiz. Ushbu misolda quyidagi metodda o'rganishni o'z ichiga oladi.  $c(x_1 \dots x_n)$  - ta'lim korpusidagi  $x_1 \dots x_n$  jumlaning nechta marta uchrashlari soni va  $N$ -korpusda mavjud bo'lgan jumlar soni aniqlagan holda  $p(x_1 \dots x_n)$  ni qiymatini aniqlanadi, ya'ni

$$p(x_1 \dots x_n) = \frac{c(x_1 \dots x_n)}{N}$$

Ammo bu juda eski model: agar jumla ta'lim korpusida mavjud bo'lmasa 0 ehtimollikni beradi.

Bir qarashda tilni modelashtirish muammosi juda g'alati tuyuladi, nima uchun buni o'ylashimiz kerak? Bir nechta sabablarga ko'ra:

1. Til modellar mashina tarjimai va nutqni aniqlash jarayonidagi eng aniq va juda foydalisidir. Ko'pchilik ilovalarda qaysi jumlar tilda mavjud bo'lishi mumkin yoki mumkin emasligining oldindan aniq bo'lishi  $p(x_1 \dots x_n)$



taqsimot uchun juda qulay. Masalan nutqda tilni modellashtirish bu turli so'zlarni talaffuzini modellashtirish: bu haqda o'ylashning bir yo'li juda ko'p sonli gaplarning generatsiyasini ularning ehtimolligi bilan birga aniqlovchi o'sha akustik modeldir. Tildagi mavjud bo'lgan gaplar

2. Biz  $p$  funksiyasini aniqlash va baholash uchun tasvirlab beradigan usullar o'quv misollaridan olingan modelning parametrlari foydali bo'ladi. Kurs davomida boshqa bir qancha kontekstlarda: masalan, yashirin Markovda biz ko'rib chiqadigan modellar va tabiiy tilni tahlil qilish modellarida.

Endi esa, ta'lim korpusini hisobga olgan holda  $p$  funksiyaning qiymatini qanday aniqlash mumkin degan savolga javob izlaymiz. Birinchi navbatda ehtimollar nazariyasi markazida bo'lgan Markov modellarini tasvirlaymiz. Modellashtirishni quyidagi gap misolida ko'rib chiqamiz.

### **Bugungi quyosh namlikdan asar qoldirmaydi**

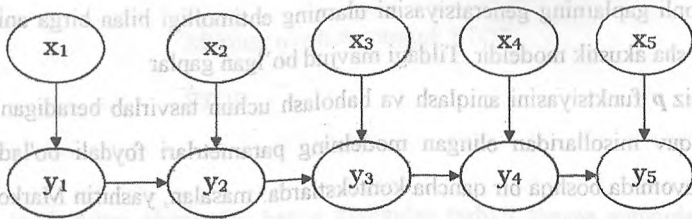
Teglar ketma-ketligi quyidagi ko'rinishni oladi

Adv N N N V

(Bu yerda  $N$ -ot,  $Adv$ -ravish,  $V$ -fe'l) Jumlani tarkibidagi har bir so'zga bitta teg mos keladi, ya'ni berilgan jumla uzunligi teglashtirilgan jumla uzunligiga teng bo'ladi (yuqoridagi misolda, **Bugungi** uchun  $Adv$ , **quyosh** uchun  $N$ , **namlik** uchun  $N$ , **asar** uchun  $N$  va **qoldirmoq** uchun  $V$  teglar mos qo'yilgan).

Teglash modellarida  $x_1, x_2, \dots, x_n$  belgilashlardan foydalanamiz va buni berilgan gap sifatida qaraymiz. Yuqoridagi misolda gap uzunlig  $n=5$ , ya'ni  $x_1=bugungi$ ,  $x_2=quyosh$ ,  $x_3=namlikdan$ ,  $x_4=asar$ ,  $x_5=qoldirmaydi$ . Gap tarkibidagi so'zlarning teglari ketma-ketligi yoki holati ketma-ketligini modellashtirishda  $y_1, y_2, \dots, y_n$  belgilashlardan foydalanamiz. Yuqoridagi misolni quyidagicha belgilaymiz:  $y_1=Adv$ ,  $y_2=N$ ,  $y_3=N$ ,  $y_4=N$ ,  $y_5=V$ . Gaplarning  $x_1, x_2, \dots, x_n$  ketma-ketligini  $y_1, y_2, \dots, y_n$  teglar ketma-ketligiga almashtirish muammosi **ketma-ketlikni belgilash** yoki **teglash muammosi** deb yuritiladi.<sup>45</sup>

<sup>45</sup> Michael Markov, "Tagging with Hidden Markov models" 2011



Faraz qilaylik bizda  $(x^{(i)}, y^{(i)})$  misollar to'plami mavjud, (buyurda,  $(x^{(i)}, y^{(i)}), i = 1..m$ , har bir  $x^{(i)}$  to'plam  $x_1^{(i)}, x_2^{(i)}, \dots, x_{n_i}^{(i)}$  gaplardan iborat va har bir  $y^{(i)}$  to'plam esa  $y_1^{(i)}, y_2^{(i)}, \dots, y_{n_i}^{(i)}$  teglar ketma-ketligidan iborat to'plam ( $n_i - i$  - misolning uzunligi). U holda  $x_j^{(i)}$  esa  $i$ - misoldagi  $j$ - so'z,  $y_j^{(i)}$  esa  $j$ - so'zning tegi.

Bizga misol sifatida  $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})$  juftliklar berilgan, bu yerda har bir  $x^{(i)}$  berilgan gap va  $y^{(i)}$  uning tegi. Mavjud berilgan jumlar to'plamini  $X$  va barcha teglar to'plamini  $Y$  belgilashdan foydalanamiz. Endi esa bizning vazifamiz  $f: X \rightarrow Y$  moslik, ya'ni har bir berilgan  $x$  uchun unga mos  $f(x)$  qonuniyatni o'rganishimiz lozim. Tabiiy tilni qayta ishlashdagi tizimidagi ko'p muammolar nazorat ostiga olingan. Masalan, har bir  $x^{(i)}$  berilgan jumla uchun  $x_1^{(i)}, x_2^{(i)}, \dots, x_{n_i}^{(i)}$  so'zlar ketma-ketligi va har bir  $y^{(i)}$  uchun  $y_1^{(i)}, y_2^{(i)}, \dots, y_{n_i}^{(i)}$  teglar ketma-ketligini teglashtirishi bu muammo ( $n_i - i$  - misolning uzunligi).  $X$  sifatida  $x_1, x_2, \dots, x_n$  ketma-ketliklarning to'plami va  $Y$  sifatida  $y_1, y_2, \dots, y_n$  teglar ketma-ketliklarning to'plami belgilanadi. Ayni paytda bizning vazifamiz gaplar ketma-ketligini teglar ketma-ketligiga mos qo'yuvchi  $f: X \rightarrow Y$  funksiyani o'rganishdan iborat. Mashina tarjimasida berilgan har bir  $x$  manbadagi gaplar (Xitoy a.g.) va har bir "label" (belgi)lar tilidagi gap sifatida qabul qilinadi. Bizning vazifamiz berilgan

barcha  $x$  va ularning teglari  $y$  belgilardan iborat barcha  $(x^{(i)}, y^{(i)})$ ,  $(i=1..n)$  juftliklarni aniqlaydigan funksiyani o'rganishdan iborat.

$f(x)$  - funksiyani aniqlashning yana bir yo'li - shartli model usulini qaraymiz. Bunda biz har qanday  $(x, y)$  juftlik shartli ehtimollikni aniqlaydigan modelni aniqlaymiz.

$$\sum_{y \in Y} p(y|x) = 1$$

Model parametrlari esa traininglarda baholanadi. Kiritilgan yangi  $x$  dan olinadigan natijaviy misol namunasi

$$f(x) = \arg \max_{y \in Y} p(y|x) \quad (1)$$

Ushbu modeldan olinadigan  $y$  belgi, belgilarning eng katta qiymati.  $p(y|x)$  modelimiz berilgan namunalarning haqiqiy shartli taqsimotiga yaqin bo'lsa  $f(x)$  funksiya optimal hisoblanadi. Mashina tili va tabiiy tilni qayta ishlashdagi eng tez foydalaniladigan va eng muqobil model bu **generative** modeldir. Shartli ehtimollikni to'g'ridan-to'g'ri baholaydigan  $p(y|x)$  model o'rimga,  $(x, y)$  juftliklari uchun qo'shma ehtimollik  $p(x, y)$  modellashtiramiz.

$(x^{(i)}, y^{(i)})$ ,  $(i=1..n)$  keltirilgan misollar uchun  $p(x, y)$  model parametrlari yana faraz qilinadi. Ko'p hollarda  $p(x, y)$  ehtimollikni quyidagicha taqsimot qilamiz.

$$p(x, y) = p(y)p(x|y) \quad (2)$$

va  $p(y)$  va  $p(x|y)$  modellar uchun ehtimolliklar alohida bo'ladi. Bu ikki model tarkibiy qismlari quyidagicha talqin qilinadi:

$p(y)$  -  $y$  teglari taqsimotining dastlabki ehtimolligi.

$p(x|y)$  -  $y$  asosiy teg ekanligini hisobga olib, berilgan  $x$  ni hosil qilish ehtimolligi.

Quyidagi usul bilan modellarni tarkibiy qismlarga ajratish ko'p hollarda qulay ekanligini ko'rib chiqamiz; masalan nutqni aniqlashda klassik yondashuv dekompozitsiyaning bu turiga asoslangan. Generativ model, ya'ni har qanday  $(x, y)$

juftlik uchun  $p(y|x)$  shartli ehtimollikni hisoblashda Bayes formulasidan foydalanamiz.

$$p(y|x) = \frac{p(y)p(x|y)}{p(x)}$$

Bu yerda,

$$p(x) = \sum_{y \in Y} p(x, y) = \sum_{y \in Y} p(y)p(x|y)$$

Ushbu qo'shma model juda ko'p qirrali bo'lib, bunda biz  $p(x)$  va  $p(y|x)$  ehtimolliklarni ham keltirib chiqarishimiz mumkin. Qo'shma modelni yangi misollarni test qilishda Bayes qoidasidan bevosita foydalanamiz.

Kiritilgan  $x$  qiymatlar uchun bizni modelimizning  $f(x)$  natijasi quyidagicha aniqlanadi:

$$f(x) = \arg \max_y p(y|x)$$

$$= \arg \max_y \frac{p(y)p(x|y)}{p(x)} \quad (3)$$

$$= \arg \max_y p(y)p(x|y) \quad (4)$$

(3) Bayes tenglamasida (4) tenglama kelib chiqadi, chunki  $p(x)$  ning qiymati  $y$  ning qiymatiga bog'liq emas shuning uchun  $\arg \max$  ning qiymati o'zgarmaydi. Bu bizga ancha qulaylik tug'diradi, ya'ni  $p(x)$  ni hisoblashimiz shart emas.

Qo'shma ehtimolliklarni  $p(x)$  va  $p(x|y)$  shartlarga ajratuvchi modellar noisychannel (shovqinli kanal) modellar deyiladi<sup>46</sup>. Intuitiv ravishda, misol sifatida berilgan  $x$  ni 2 xil bosqichda yaratilgan deb qaraymiz:

Birinchi, har bir  $y$  tegni tanlanganlik ehtimoli  $p(y)$ ;

Ikkinchi, berilgan  $x$  namunalari  $p(x|y)$  taqsimotdan hosil bo'lganligi.

<sup>46</sup> Michael Collins. Tagging with Hidden Markov Models

<http://www.cs.columbia.edu/~mcollins/courses/nlp2011/notes/hmms.pdf> -2011

$p(x|y)$  model sifatida,  $y$  tegni qabul qilib natijada  $x$  hosil qiladi deb oladigan bo'lsak, u holda bu model qo'l kelmaydi. Bizning vazifamiz esa,  $x$  qiymatni qabul qilgan holda  $y$  teglarni hosil qilish.

Xulosa qiladigan bo'lsak:

- Bizning vazifamiz  $x$  berilganlarni  $y = f(x)$  teglarga almashtiruvchi funksiyani o'rganish. Bizga berilgan misollar va ularni teglaridan iborat qiymatlar  $(x^{(i)}, y^{(i)})$ ,  $(i=1..n)$ ;
- Noisy channel (shovqinli kanal) yondashuvi asosan, berilgan  $x$  misollari yordamida  $p(x)$  va  $p(x|y)$  modellarni hosil qilamiz. Bu modellar esa qo'shma modellar quyidagicha

$$p(x, y) = p(y)p(x|y)$$

- Berilgan  $x$  qiymatni  $y$  tegini quyidagi formula yordamida aniqlaymiz

$$f(x) = \underset{y}{\operatorname{arg\,max}} p(y|x)$$

$X$  kiritilganlarni  $f(x)$  teglarini aniqlash jarayoni dekodlash deb ham ataladi.

$f(x)$  – funksiya yordamida jumladagi har bir so'zning so'z turkumini belgilaymiz. Bu usuldan turli turkum doirasidagi, lekin bir xil grammatik xususiyatga ega bo'lgan omonim so'zlarni farqlashda foydalaniladi. Shunga qaramasdan bu ma'lumotlarni bir xil so'z turkumidagi omonimlarni farqlash bo'limida keltirib o'tdik, chunki Markov modellaridan foydalanishda asqotadi.

Gap tuzilishini modellashtirishning 2 xil usulini ko'rib chiqamiz: *o'zgarmas uzunlikdagi gaplar uchun Markov ketma-ketligi* va *o'zgaruvchan uzunlikdagi gaplar uchun Markov ketma-ketligi*

### **O'zgarmas uzunlikdagi gaplar uchun Markov ketma-ketligi**

Tasodifiy  $X_1, X_2, \dots, X_n$  o'zgaruvchilar ketma-ketligi berilgan bo'lsin. O'zgaruvchilarning har bir qiymati chekli  $V$  to'plamga tegishli. Hozir biz ketma-ketlik uzunligini o'zgarmas  $n$  deb faraz qilamiz (mas.  $n=100$ ). Keyingi bo'limda biz

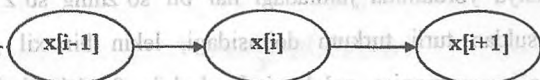
turli uzunlikdagi  $n$  ta turli ketma-ketliklarni umumlashtirish jarayonlarini ko'rib chiqamiz.

Bizning maqsadimiz quyidagicha: har qanday  $x_1 \dots x_n$  ketma-ketlikning ehtimolligini modellashtirish, bu yerda  $n \geq 1, x_i \in V, i = 1 \dots n$  lar uchun ehtimolliklarni qo'shish formulasi quyidagicha bo'ladi

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

$|V|^n$  — mumkin bo'lgan  $x_1 \dots x_n$  ketma-ketliklardan iborat.  $|V|$  ning o'rtacha qiymatlari va  $n$  ning barcha ehtimolliklari ro'yhati  $|V|^n$  uchun bo'lishi mumkin emas, bu aniq. Biz bunga qaraganda ancha ixcham modelni qurishimiz kerak.

Birinchi darajali Markov jarayoni bilan ushbu farazlarni va modellarni ancha soddalashtirish mumkin. Ingliz va rus tillarida omonimlarni farqlashda Markov modelidan foydalanish ancha qulay, ya'ni joriy so'zni o'zidan oldingi so'zlar bilan baholash. O'zbek tilidagi omonim so'zlarni farqlashda esa nafaqat o'zidan oldingi, balki o'zidan keyingi so'zlarni ham inobatga olish talab qilinadi.



Shuni hisobga olib, Markov modelini o'zbek tili grammatikasi uchun moslashtiramiz.

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

$$= P(X_1 = x_1) \prod_{i=2}^n P(X_i = x_i | X_1 = x_1 \dots X_{i-1} = x_{i-1}) \quad (5)$$

$$= P(X_1 = x_1) \prod_{i=2}^n P(X_i = x_i | X_{i-1} = x_{i-1}) \quad (6)$$

Birinchi qadam, (1) formula: ehtimolliklar zanjiri qoidasi bo'yicha har qanday taqsimotni  $P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$  shaklda yozish mumkin. Shunday qilib bu bosqichdagi farazlarni yo'q deb qaraymiz. Biroq 2-qadamda ya'ni (6)



tenglamada bu aniq emas, biz har qanday  $i=2..n$  va ixtiyoriy  $x_1 \dots x_n$  ketma-ketlik uchun

$$P(X_i = x_i | X_1 = x_1 \dots X_{i-1} = x_{i-1}) = P(X_i = x_i | X_{i-1} = x_{i-1})$$

Bu birinchi darajali Markov farazi. Faraz qilaylik  $i$ -so'z ketma-ketlikdagi o'zidan oldingi ya'ni  $x_{i-1}$  so'zning xususiyatlariga bog'liq bo'lsin.

Ikkinchi darajali Markov jarayoni, ko'pincha trigram modeli deb ataymiz, bu jarayonda murakkab bo'lmagan farazni ilgari suramiz, ya'ni har bir so'z ikkita o'zidan oldingi va keyingi so'zlarga bog'liq bo'lsin.

$$P(X_i = x_i | X_1 = x_1 \dots X_{i-1} = x_{i-1}) = P(X_{i+1} = x_{i+1} | X_{i-1} = x_{i-1}, X_i = x_i)$$

Bundan kelib chiqadiki, butun ketma-ketlikning ehtimoli quyidagicha yoziladi:

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_{i=1}^n P(X_{i+1} = x_{i+1} | X_{i-1} = x_{i-1}, X_i = x_i) \quad (7)$$

Qulaylik uchun,  $x_0 = x_{-1} = *$  deb faraz qilamiz, bu yerda  $*$  – “Start” (boshlash) gapni boshlanishini bildiruvchi simvol.

### O'zgaruvchan uzunlikdagi jumlar uchun Markov ketma-ketliklari

Jumlalarning uzunliklarini o'zgaruvchan deb fazar qilamiz.  $n$  tasodifiy o'zgaruvchi. Ushbu o'zgaruvchanlikni tilni modellashtirishning eng keng tarqalgan usulini ko'rib chiqamiz.

**Oddiy usul:**  $X_n - n$ -ta so'zdan iborat jumla. Uni uzunligini o'zgartiramiz. Ketma-ketlikdagi  $n$ -so'z maxsus belgi bo'lib, uning qiymati har doim *STOP* ga teng. Bu belgi ketma-ketlikning oxirini bildiradi.

$$p(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_{i=1}^n p(X_{i+1} = x_{i+1} | X_{i-1} = x_{i-1}, X_i = x_i)$$

Bu yerda  $n \geq 1$  hosil bo'lgan  $x_1, x_2, \dots, x_n$  ketma-ketlik uchun  $x_n = STOP$   $x_i \in V, i=1..(n-1)$ .

$$p(X_{i+1} = x_{i+1} | X_{i-1} = x_{i-1}, X_i = x_i)$$

Bu yerda  $x_i \in V$  yoki  $x_i = STOP$ . Agar  $x_i = STOP$  bo'lsa u holda omonim so'zning o'zidan oldingi birikuvchisi ya'ni,  $p(x_{i-2}|x_{i-1})$  hisoblanadi. Ehtimollikni hisoblashda quyidagilarni inobatga olish so'raladi:

1.  $i=1$  qadamda  $x_0 = x_{-1} = *$  qiymat beriladi;
2. har bir  $x_i$  - uchun  $p(X_{i+1} = x_{i+1} | X_{i-1} = x_{i-1}, X_i = x_i)$  aniqlanadi
3. Agar  $x_i = STOP$  bo'lsa  $x_1, x_2, \dots, x_{i-1}$  ketma-ketlik qaytariladi, aks holda  $i=i+1$  qiymat beriladi va 2-qadamga o'tiladi.

Har bir ketma-ketlik uchun  $n$  o'zgaradi. Ushbu usuldan korpus ma'lumotlari yordamida omonimlarni semantic tahlil qilishda foydalaniladi. Bir xil uzunlikdagi jumalardan omonimlarni farqlash vazifasi qo'yilganda o'zgarmas uzunlikdagi jumalar uchun Markov modelidan foydalaniladi.

Til modellarini aniqlashning muhim modeli Trigramm til modeli hisoblanadi. Yuqorida keltirilgan Markov modellarida bo'lgani kabi trigram til modelida ham, har bir jumlaning tasodifiy  $n$  uzunlikdagi  $X_1, X_2, \dots, X_n$  ketma-ketlik sifatida modellashtiramiz.  $X_n$  doimiy  $X_n = STOP$  qiymatga ega. 2-tartibli Markov modelida har qanday  $x_1, x_2, \dots, x_n$  ketma-ketlikni ehtimoli

$$P(X_i = x_i | X_1 = x_1 \dots X_{i-1} = x_{i-1}) = q(x_i | x_{i-2}, x_{i-1})$$

Bu yerda har bir  $(w, u, v)$  uchun  $q(w|u, v)$  modelning parametri. Bu parameter korpus ma'lumotlari asosida hisoblangan ehtimollik. Bizning model quyidagi ko'rinishni oladi

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n q(x_i | x_{i-2}, x_{i-1})$$

$q(w|u, v)$ -qiymat uchun quyidagi mulohazani keltiramiz:

Trigram til modelida har qanday  $(w, u, v)$  trigramma uchun  $V$  - cheksiz to'plam mavjud bo'lib,  $w \in V \cup STOP$  va  $u, v \in V \cup \{*\}$ .  $q(w|u, v) - (u, v)$

so'zlar birgalikda uchraganida ular bilan birgalikda  $w$  so'z uchrashining ehtimolini

bildiradi. Har bir  $x_1, x_2, \dots, x_n$  jumla  $x_i \in V, i = 1 \dots (n - 1)$  va  $x_n = STOP$  uchun ehtimollik

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n q(x_i | x_{i-2}, x_{i-1}) \quad (8)$$

formula yordamida aniqlanadi. Bu yerda  $x_{-1} = x_0 = *$

Misol tariqasida quyidagi

*Dalada paxta gulladi STOP.*

Jumlani ehtimolligini hisoblashni

$p(\text{Dalada paxta gulladi STOP})$

$$= q(\text{Dalada} | **, *) \times q(\text{paxta} | *, \text{dalada}) \times q(\text{gulladi} | \text{dalada}, \text{paxta}) \times q(\text{STOP} | \text{paxta}, \text{gulladi})$$

Ushbu namuna misolida jumalarni semantic tahlil qilish mumkin. Biz esa jumla tarkibidagi omonim so'zning birikuvchilari bilan birgalikda uchrash ehtimolligini hisoblaymiz.

$$q(w|u, v) = \frac{c(u, v, w)}{c(u, v)} \quad (9)$$

Bu yerda,  $c(w, u, v)$  –  $w, u, v$  trigrammaning korpusda uchrashlar soni,  $c(u, v)$  – esa  $u, v$  bigrammaning korpusda uchrashlar soni.

Yuqoridagi jumladagi trigrammaning ehtimoli

$$q(\text{gulladi} | \text{dalada}, \text{paxta}) = \frac{c(\text{dalada}, \text{paxta}, \text{gulladi})}{c(\text{dalada}, \text{paxta})}$$

Bu usul orqali aniq natija olish uchun katta hajmdagi korpus kerak bo'ladi. Bir so'z turkumi doirasidagi omonimlarni farqlash uchun korpus ma'lumotlari tarkibiga omonim so'zlar qatnashgan jumalarni kiritish talab qilinadi, aks holda yuqoridagi qiymat 0 ga teng bo'lishi mumkin.

Omonim so'zlar orasida shunday so'zlar borki, ular turli so'z turkumlariga mansub bo'lsada bir xil qo'shimcha bilan birikadi. Lekin qo'shimcha qo'shilgandan keyin bu omonimning brikuvchilari turli so'z turkumlariga mansub bo'lishi, yoki

birikuvchilari turli giponimlarga tegishli bo'lishi mumkin. Bunday vaziyatlarda omonimlarni farqlashda Trigramm Yashirin Markov modelidan (YMM) foydalaniladi. Trigramm YMMdan foydalanish uchun jumla tarkibidagi so'zlarning teglari ham aniqlanishi lozim. Trigramm YMM,  $V$ - mumkin bo'lgan so'zlarning chekli to'plami va  $K$  shu so'zlarning mumkin bo'lgan teglaridan iborat chekli to'plamdan iborat bo'lib, quyidagicha parametrlarda.

- $q(s|u, v)$  parametr har bir  $s|u, v$  trigram uchun  $s \in K \cup \{STOP\}$  va  $u, v \in V \cup \{*\}$ ,  $q(s|u, v)$  ehtimollik  $(u, v)$  teglarining bigrammidan keyin  $s$  tegining uchrash ehtimollikni anglatadi, \*-gapning boshini bildiradi.
- $e(x|s)$  parametr, har bir  $x \in V, s \in K$ ,  $e(x|s)$  qiymati,  $x$  so'zning  $s$  so'z turkumi bilan juftlashish ehtimolligini aniqlaydi.

$S(x_1, \dots, x_n, y_1, \dots, y_{n+1})$  so'zlar ketma-ketligi va teglar ketma-ketligi juftliklaridan iborat to'plam, bu yerda  $n \geq 0, x_i \in \vartheta, i = 1 \dots n, y_i \in K, i = 1 \dots n$  va  $y_{n+1} = STOP$ .

Biz har bir  $(x_1, \dots, x_n, y_1, \dots, y_{n+1}) \in S$  uchun quyidagi

$$p(x_1, \dots, x_n, y_1, \dots, y_{n+1}) = \prod_{i=1}^n q(y_i | y_{i-2}, y_{i-1}) \prod_{i=1}^n e(x_i | y_i) \quad (10)$$

ehtimollikni aniqlashimiz kerak. Bu yerda  $y_0 = y_{-1} = *$ .

Masalan, bizda  $n=5$ ,  $x_1 \dots x_5$  ga teng bo'lgan "Qirdagi o'tlarning yam-yashilligidan ko'z quv naydi" gap berilgan bo'lsin, ushbu gap uchun qo'shma ehtimollikni qo'llash uchun uni o'zak va qo'shimchalarga ajratib olinadi va qiymati  $N N Adj N V STOP$  lardan iborat  $y_1 \dots y_6$  to'plam ham aniqlanadi va qo'shma ehtimollik formulasiga quyidagicha hisoblanadi,

$$\begin{aligned} p(x_1, \dots, x_n, y_1, \dots, y_{n+1}) &= q(Adv | *, *) \times q(N | *, Adv) \times q(N | Adv, N) \times q(N | N, N) \\ &\times q(V | N, N) \times q(STOP | N, V) \times e(Bugungi | Adv) \times e(Quyosh | N) \\ &\times e(namlilik | N) \times e(asar | N) \times e(qoldirmoq | V) \end{aligned}$$

Bu model noise-channel modeli hisoblanadi:

$$q(Adv | *, *) \times q(N | *, Adv) \times q(N | Adv, N) \times q(N | N, N) \times q(V | N, N) \\ \times q(STOP | N, V)$$

ning qiymatini hisoblashda, *Adv N N N V STOP* teglar ketma-ketligining ehtimolligini aniqlash uchun ikkinchi darajali Markov modeli (trigram modeli)dan foydalanamiz.

$$e(\text{Bugungi} | Adv) \times e(\text{Quyosh} | N) \times e(\text{namlik} | N) \times e(\text{asar} | N) \times \\ e(\text{qoldirmoq} | V) - - \\ p(\text{Bugungi quyosh namlikdan asar qoldirmaydi} | Adv N N N V STOP)$$

shartli ehtimollikni bildiradi, bu yerda

$p(x|y)$  – "Bugungi quyosh namlikdan asar qoldirmaydi"

jumlasidan olingan  $x$  va  $Adv N N N V STOP$  teglaridan olingan  $y$  larning shartli ehtimolligini bildiradi. Ushbu ehtimollikni hisoblashimiz uchun esa ba'zi

parametrlar kerak bo'ladi. Endi shu parametrlarni baholaymiz. Analizator uchun namunalar to'plami  $X_1 \dots X_n$  berilgan. Har bir jumla uchun  $x_1 \dots x_n$  so'zlar va  $y_1 \dots y_n$  teglar ketma-ketligini aniqlanadi. Ushbu ma'lumotlarni hisobga olgan holda model parametrlarini qanday baholaymiz?

Bu savolga oddiy va juda intuitive javob borligini ko'ramiz.  $c(u, v, s)$  berilgan ma'lumotlardagi  $u, v, s$  so'z turkumlari ketma-ketligining soni aniqlaydi, masalan  $c(N, N, Adj)$  -berilgan jumladagi *o't* omonim so'zi ot so'z turkumi vazifasida kelganida bu so'zdan oldin va keyin brikib keluvchi so'zlarning  $N, N, Adj$  3 ta teg ketma-ketlikning sonini bildiradi. Xuddi shunday  $c(u, v)$  esa  $(u, v)$  bigramm belgilarini nechta marta uchrashini bildiradi.  $c(s)$  esa berilgan ma'lumotlar korpusida  $s$  ni nechta marta ko'rilganligini aniqlaydi. Va nihoyat  $c(s \rightsquigarrow x)$   $s$  so'z turkumiga mansub  $x$  so'zning korpusda uchrashlari soni: masalan  $c(N \rightsquigarrow O't)$  – o't so'zining korpusda  $N$  (ot) tegi ko'rinishida uchrashlari soni.

Ushbu izohlarni inobatga olgan holda maksimal-ehtimollik quyidagicha keltiriladi

$$q(s|u, v) = \frac{c(u, v, s)}{c(u, v)} \quad (11)$$

va

$$e(x|s) = \frac{c(s \sim x)}{c(s)} \quad (12)$$

Masalan, biz misolimiz uchun ushbu ehtimollik quyidagicha hisoblanadi

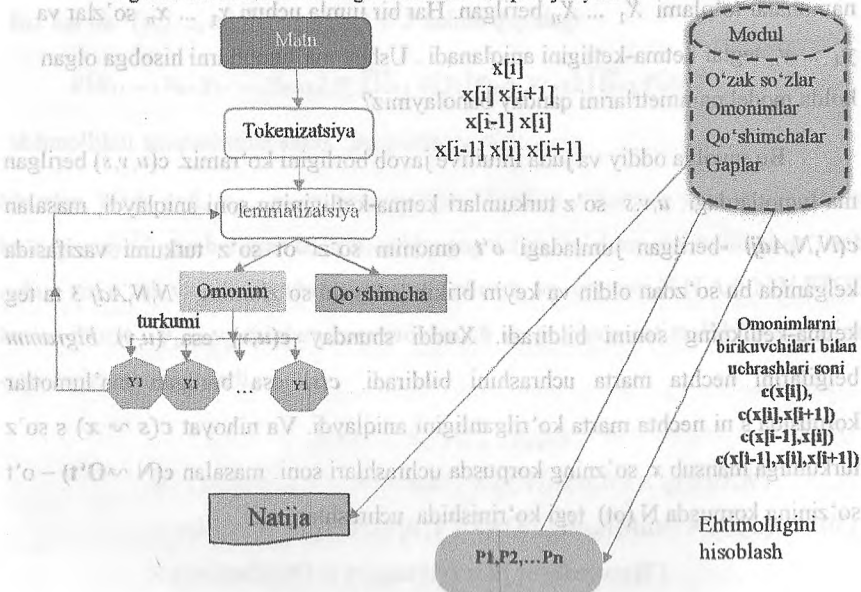
$$q(V|N, N) = \frac{c(N, N, V)}{c(N, N)}$$

Va

$$e(o't|N) = \frac{c(N \sim asar)}{c(N)}$$

Shunday qilib model parametrlarini baholash uchun til korpusidan sonlarni sanash va maksimal ehtimollikni formulalar orqali hisoblash kifoya. Turli so'z turkumlari orasida bo'lsada grammatik xususiyatlari yaqin bo'lgan omonim so'zlari farqlashda

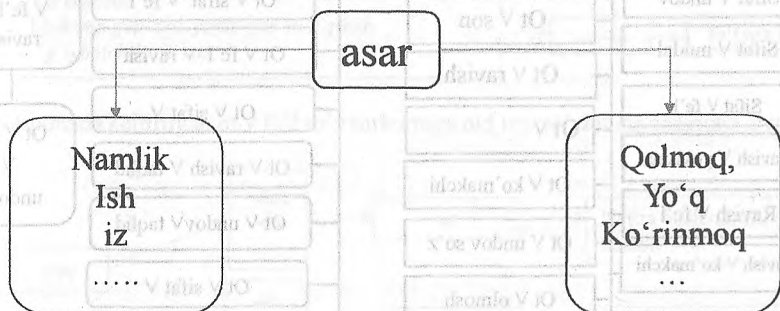
1-grafik. Bir turkum doirasidagi omonimlarni farqlash jarayoni





Trigram Yashirin Markov modelidan foydalanish mumkin. Yuqorida keltirilgan nazariy ma'lumotlarni grafik ko'rinishda ifodalanishi quyidagicha:

Grafikda keltirilgan statistik hisoblashlar har doim ham noldan farqli bo'lmasligi mumkin. Ya'ni gap tarkibidagi omonim so'z va uning birikuvchisi korpus ma'lumotlari ichida uchramasligi ham mumkin, natijada  $p=0$  qiymatga ega bo'ladi. Bunday hollarda Semantik konkardanse metodi qo'llaniladi. Semantik konkardans shunday metodki, omonim so'zning birikuvchilari (valentliklari) dan iborat to'plamlar hosil qilinadi.

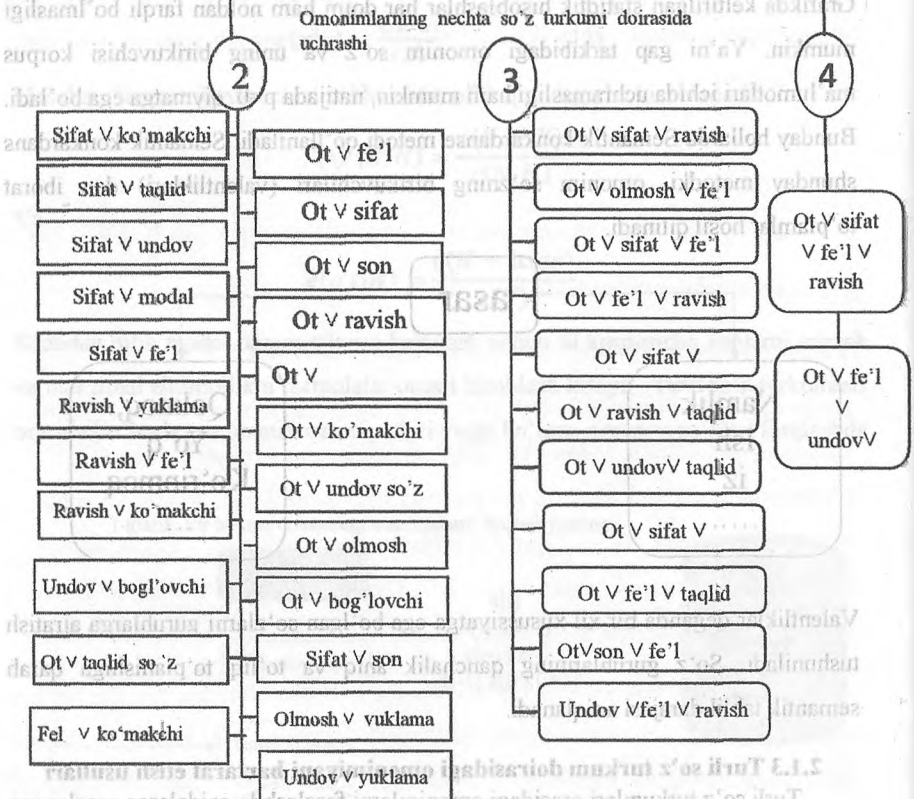


Valentliklar deganda bir xil xususiyatga ega bo'lgan so'zlarni guruhlarga ajratish tushuniladi. So'z guruhlarning qanchalik aniq va to'liq to'planishiga qarab semantik tahlil darajasi aniqlanadi.

### 2.1.3 Turli so'z turkum doirasidagi omonimiyani bartaraf etish usullari

Turli so'z turkumlari orasidagi omonimlarni farqlashda qoidalarga asoslangan teglash usullaridan foydalanish tavsiya qilinadi. Ba'zi so'z turkumlari orasidagi omonimlikni aniqlashda Trigramm YMM dan foydalanish mumkin. O'zbek tilidagi omonimlar so'zlarning turli turkumlar (2, 3 va 4 ta)ga mansub bo'lish ierarxiyasida quyidagi rasmda keltirilgan turkumlar tahlil qilinadi:

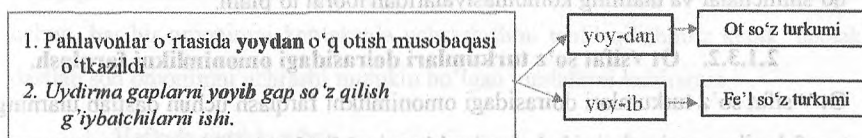
Turli so'z turkumiga oid omonim so'zlar



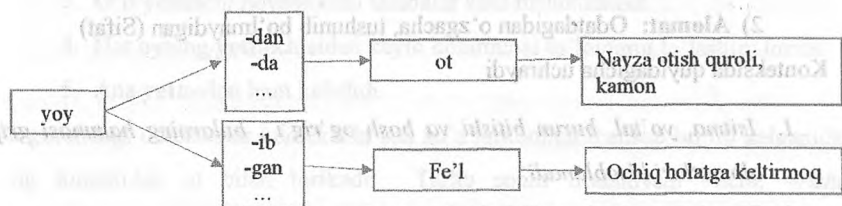
Terarxiyada keltirilgan bir soʻz turkumi doirasidagi omonimlikni farqlash uchun matematik modellar ishlab chiqamiz.

### 2.1.3.1. OtVfe'l soʻz turkumidagi omonimlik.

Ot Vfe'l soʻz turkumidagi omonimlikni farqlashda ularning morfologik xususiyatlari asosida matematik mulohazalar ishlab chiqish lozim bo'ladi. Ot Vfe'l soʻz turkumi doirasidagi omonim soʻzlarni farqlash qoidasi asosida matematik model yoki algoritim tavsiya qilamiz:



Namunada keltirilgan ot V fe'l soʻz turkumiga oid yoy soʻzini quyidagicha ajratamiz.



Bu guruhlashlar asosida quyidagi matematik formulani keltiramiz.

$$H_{N,V} = \begin{cases} N + x, & x \in \{aff_i^N\}, i = 1..n \\ V + x, & x \in \{aff_j^{(1,3)}\}, j = 1..m \end{cases}$$

$H_{N,V}$ -ot va fe'l orasidagi omonimlik,  $N$ -ot,  $x$ - lug'aviy va sintaktik shakl yasovchi qo'shimchalar va ularning kombinatsiyalaridan iborat to'plam,  $V$ -fe'l,  $aff_i^N$ <sup>47</sup>- ot soʻz turkumi doirasidagi omonim soʻzga qo'shiluvchi sintaktik va lug'aviy shakl yasovchi qo'shimcha va qo'shimchalar kombinatsiyalari to'plami,  $aff_j^{(1,3)}$ <sup>48</sup>- fe'lga qo'shiluvchi sintaktik va lug'aviy shakl yasovchi qo'shimcha va ularning kombinatsiyalar to'plami.

<sup>47</sup> Ot suz turkumiga qushiluvchi qushimchalar kombinatsiyalarining hozircha 1155 tasi aniqlandi

<sup>48</sup> Fe'l so'z turkumiga qo'shiluvchi qo'shimchalar kombinatsiyalarining hozircha 4132 tasi aniqlandi

Namunadagi jumladagi omonim soʻzlar uchun ushbu keltirilgan matematik formulani tadbiiq yetamiz.  $H_{N,V} = \{yoy\}$  – ot va feʼl soʻz turkumlari doirasidagi omonim soʻzlar toʻplamining bir elementi, qoʻshimchalar va ularning kombinatsiyalaridan iborat toʻplamlarning koʻrinishi quyidagicha:

$aff_i^N = \{-ni, -ning, -lar, -da, -ga, -larga, -larni, -larning, -larda, -cha, -lardagi, -chada, -chaga, -chani, -chani, -chalar, -chalariga, -chalarining, -chalarda, -larniki, -niki, -dir, -nikidir, \dots\}$

$aff_i^V = \{-di, -maydi, -gan, -maydigan, -masdi, -masdan, -magan, -masin, -sin, -ydi, \dots\}$

$x = \{aff_i^V, aff_i^N, aff_i^{Adj}, aff_i^{Adv}, \dots\}$  – lugʻaviy va sintaktik shakl yasovchi qoʻshimchalar va ularning kombinatsiyalaridan iborat toʻplam.

### 2.1.3.2. Ot vsifat soʻz turkumlari doirasidagi omonimlikni farqlash.

Ot v sifat soʻz turkumlari doirasidagi omonimlikni farqlash uchun dastlab ularning morfologik xususiyatlarini koʻzdan kechiramiz. Misol tariqasida

1) **Alomat:** Belgi, koʻrsatkich (ot)

2) **Alomat:** Odatdagidan oʻzgacha, tushunib boʻlmaydigan (Sifat)

Kontekstda quyidagicha uchraydi

1. *Isitma, yoʻtal, burun bitishi va bosh ogʻrigʻi, bularning hammasi gripp alomatlari hisoblanadi.*

2. *Bu alomat qiz ekanmi? Yoki Bu qiz alomat ekanmi?*

1-gapdagi alomat soʻz ot soʻz turkumi doirasida boʻlib otga xos qoʻshimchalar bilan brikishi, u oʻng va soʻl tomonidan ot, soʻl tomonidan feʼl bilan brikishi mumkin.

2-gapda esa sifat soʻz turkumiga doir boʻlib u oʻng va soʻl tomonidan ot bilan brikishi mumkin, yoki boʻlmasa otlashishi yaʼni keyin kelgan ot tushib qolib sifat otga xos xususiyatlarni qabul qilishi ham mumkin. Bundan kelib chiqadiki ot va sifat

soʻz turkumi doirasi omonimlarni farqlashda ham stoxastik teglash usulining trigramm Yashirin Markov modeliga murojaat qilinib brikuvchilari bilan farqlanadi. Korpus maʼlumotlari orasidan sifat va ot soʻz turkumlariga doir omonim soʻzlari mavjud boʻlgan jumlar yigʻiladi. Har bir jumladagi omonim soʻzning brikuvchilari va izohi mutaxassislar tomonidan belgilab qoʻyiladi. Maʼlumotlar, omonim soʻzlarning brikuvchilari toʻliq va toʻgʻri yigʻilgandagina semantik

tahlilning aniqligi darajasi yuqori bo'ladi. Buning uchun korpusga omonim so'zlar ishtirokidagi ko'plab jumlar kiritilishi lozim.

### 2.1.3.3. Ot Vson so'z turkumlari orasidagi omonimlikni farqlash.

Ot V son so'z turkumlari doirasidagi omonimlikni farqlovchi matematik modellarni yaratishdan oldin ularning morfologik xususiyatlarini ko'rib chiqamiz.

1) *Yetti: Olti bilan sakkiz orasidagi butun sonning nomi (son)*

2) *Yetti: Kishi vafotining o'rtinchi kuni o'tkaziladigan marosim(ot)*

Tadqiqot ishimizning asl maqsadi omonimlarni kontekstda farqlash bo'lagani uchun, har bir omonimni kontekstda uchrashishini tahlil qilishimiz kerak, demak dastlab son omonimni uchrashi mumkin bo'lgan jummalarni keltiramiz.

1. Haftada yetti kun bor
2. Mehmonlar yettinchi kuni kelishadi.
3. O'n yettinchi noyabr kuni talabalar kuni nishonlanadi.
4. Har oyning yettinchisidan keyin kommunal to'lovlarni to'lashim lozim.
5. Ana yettiovlon ham kelishdi.

Bu guruhdagi omonimlar kontekstda son so'z turkumiga mansub bo'lib kelganida o'ng tomonidan ot bilan birikadi. Tartib sonni ifodalovchi *-nchi, -inchi* qo'shimchalar bilan birikib kelganda ham so'l tomonidan ot bilan birikadi. Tartib son qo'shimchalarini qabul qilganida omonim so'z o'ng tomonidan ot so'z turkumidagi so'z yoki o'nlik, yuzlik, minglik kabi sonlarni ifodalovchi so'zlar bilan birikadi. Undan tashqari egalik va kelishik qo'shimchalari va ularning kombinatsiyalaridan iborat qo'shimchalar ham qo'shilishi mumkin. Nafaqat tartib sonni ifodalovchi qo'shimchalar, balki jamlovchi *-ov, -ovlon* qo'shimchalar bilan ham birikib keladi. Jamlovchi qo'shimchalar bilan birikib kelgani uchun sondan keyin fe'l so'z turkumiga doir so'z bilan birikadi.

- **Yetti:**(son)+kishi/so'm/metr/kun/mahalla/xufton/iqlim/og'ayni/qaroqchi/pusht (*ot*)
- **Yetti:**(son)+o'lcha/bukil (*fe'l*)
- **Yetti:**(ot) +ni/ -ga/ -da/ -dan+o'tkaz/bo'l/ko'rish/gaplash/xayrlash/bor (*fe'l*)

2.1.3.3. - **Yetti**:+ -ning+kelishi/bo'lishi (*harakat nomi*)

Yuqoridagilarni hisobga olgan holda ot va son omonimlarni farqlash uchun quyidagi modelni tavsiya qilamiz.

$$H_{num,N} = \begin{cases} \downarrow W_i^{dec} \oplus H^{num} + \downarrow aff_i^{ord} \oplus \downarrow (W_k^N + \downarrow aff_i^{ord}) \oplus N_j^c \\ H^{num} \oplus V^{num} \\ H^N + aff^{CS} \oplus V \end{cases} \quad (14)$$

Bu yerda  $H^{num}$ - son so'z turkumiga oid omonim so'z,  $aff_i^{ord}$  (ordinal number)- tartib sonlarni bildiruvchi qo'shimchalar

$aff_i^{ord} = \{-nchi, -inchi, -ta, -tacha, \dots\}$

$W_i^{dec}$  (decimal words)-o'nliklar, yuz, ming va boshqa sonlarning nomlaridan iborat chekli to'plam,  $W_i^{dec} =$

$\{o'n, yigirma, o'ttiz, \dots, yuz, ming, million, milliard, trillion, \dots, dodekalon\}$ ,

$i = 1..23$

$W_k^N$  (number)-raqam so'zlar,  $W_k^N = \{bir, ikki, \dots, to'qqiz\}$ ,  $k = 1..9$ ,

$N_j^c$  (countable)-sanaladigan otlar to'plami,  $V^{num}$ -son omonimga binkuvchi fe'llar

to'plami,  $H^N$ -son omonimning ot so'z turkumi vazifasida kelgan ko'rinishi,

$aff^{CS}$  - kelishik qo'shimchalari to'plami.

Bu guruhga mansub bo'lgan omonim so'zlardan 3 tasini jamladik.

#### 2.1.3.4. Ot V ravish so'z turkumlari orasidagi omonimlikni farqlash

Ot/ ravish so'z turkumlari orasidagi omonimlikni ifodalovchi so'zlarga misol

sifatida oqibat so'zini keltiramiz

- Oqibat:**
1. natija, yakun, samara; iz, asar; ergashish, ta'qib qilish.
  2. Qadrdonlik, do'stlik tuyg'usi, mehr, muhabbat va shundan kelib chiqadigan yaxshi munosabat; muruvvat.
  3. Oxirgi yakun, natija. (*ot*)

**Oqibat:** Hamma narsa (voqea yoki harakatlar) bo'lib o'tgandan keyin, oxiri, natijada, oxir-pirovardida, nihoyat. (*ravish*)

Oqibat so'zini kontekstlarda kuzatilganida quyidagicha uchradi:



1. Bu uchrashuvning oqibati qandoq bo'larkin? (ravish) (Ufq r. 612 b)
2. Uzoq o'ylash oqibati shu bo'ldiku, u ertangi kundan birob xiralik ko'rmadi (ot) (Ufq r. 651 b)
3. Bu bevafoqlikning oqibati nimalarga sabab bo'lishini o'ylamadi (Ufq r. 604 b)
4. Buning oqibatida ular qariyb yarim asr vatanidan judo bo'lib yashadi (korpus m-t.)
5. Uning salbiy oqibatlarini tasavvur etish (korpus m-t.) (ot)
6. Oxir-oqibatda ruhiy qashshoq bo'lib qolishiga olib kelishi mumkin (korpus m-t.) (ot)
7. Uning Xirosima va Nagasaki degan shaharlariga tashlangan atom bombasining fojiali oqibatlari hozirgacha bu yurt odamlarini bezovta qilib kelmoqda (korpus m-t.) (ot)

Keltirilgan misollardan ko'rinadiki ot yoki ravish orasidagi omonimlikni farqlashda ularning valentliklaridan foydalanish maqsadga muvofiq.

### 2.1.3.5. Ot ∨ predikative so'z orasidagi omonimlikni farqlash

Ot ∨ predikative so'z orasidagi omonim so'zlardan 4 tasi yig'ilgan bo'lib unga misol keltiramiz

**Lozim:** Albatta zarur, albatta kerak (predikativ so'z)

**Lozim:** Sharq ayollari kiyadigan uzun ishton (ot)

Bu tipdagi omonim so'zlarni farqlovchi matematik modellar tavsiya qilish uchun uni kontekstda kuzatish kerak bo'ladi.

1. Qayerni lozim topsangiz, o'shaqda ishlaganim bo'lsin. (Ufq r. 59 b) (predikativ so'z)
2. Azizaxonni tasodifan uchratib, shuncha odam oldida uni yupatishni lozim ko'rmadi (Ufq r. 225b) (predikativ so'z)
3. Hokimiyatga uzun lozim kiygan juvon kirib keldi (og'zaki nutq) (ot)

1,2-jumlalarda lozim predikativ so'z quyidagi xususiyatlarga ega,

– Lozim ⊕ topmoq/ko'rmoq/deb bilmoq/...

(d) – *Lozim* + *–ligi/–ligini* ⊕ *aytmoq/bilmoq/aytmoq/... /V*  
 – *Bilishi/ko'rishi/qilishi/yupatishi/.../harakat nomi* ⊕ *lozim*  
 3-jumladagi *lozim* so'zi ot so'z turkumida bo'lib, ot so'z turkumiga xos xususiyatlarga ega. Undan tashqari kiyim-kechaklar guruhiga mansub so'z.

$$H_{pv} = \begin{cases} V^H \oplus H^{pv} \oplus V^{pv} \\ H^{pv} + \begin{matrix} ligi \\ ligini \end{matrix} \oplus V \end{cases} \quad (15)$$

Bu yerda  $V^{pv}$ -predikative so'z bilan o'ng tomonidan birikuvchi fe'llardan iborat chekli to'plam,  $V^{pv} = \{topmoq, ko'rmoq, bilmoq, deb bilmoq, \dots\}$ ;

$V^H$ - fe'lining harakat nomi shaklidan iborat chekli to'plam,

$V^H = \{borishi, kelishi, olishi, topishini, \dots\}$

### 2.1.3.6. Ot V ko'makchi so'z orasidagi omonimlarni farqlash

Ot V ko'makchi orasidagi omonim so'zlarga misol

- 1) **Ko'ra:** Kabob pishirishda cho'g' solinadigan cho'ziq temir manqal. (*Ot*)
- 2) **Ko'ra:** eskirgan. Vaza (*ot*)
- 3) **Ko'ra:** Binoan, qaraganda (*ko'makchi*)

Bu omonimlarni kontekstda quyidagicha ishtirok etadi.

– *Azixxonni shaloladagi ishlardan ko'ra Esh polvonning qando ishlayotgani qiziqtirardi. (Said Ahmad Ufq r. 201-b)*

– *Har dam olish kuni Anvar aka ko'rada kabob pishirar edi.*

Bu jumalardan ko'rinib turibdiki asosiy maqsad ot va ko'makchi orasidagi omonimlikni farqlash. Ko'makchi hech qanday qo'shimcha qabul qilmaydi, ot esa aksincha. Ammo ot ham mustaqil holda uchraydi. Demak boshqa xususiyatlarini inobatga olishimiz kerak.

$$H_{N,Aux} = \begin{cases} H^N + \downarrow affi^N \\ W + \begin{matrix} ga \\ dan \end{matrix} \oplus H^{Aux} \oplus W_j^{Aux} \end{cases} \quad (16)$$

Bu yerda W-ixtiyoriy so'z bo'lishi mumkin, ot, fe'l yoki boshqa so'z turkumidagi so'zlar to'plami, bu so'zlar *–ga, –dan* qo'shimchalarni qabul qilgan holda ko'makchi so'z bilan birikadi.  $W_j^{Aux}$  – ko'makchi omonim so'zning so'ng tomonidan birikishi mumkin bo'lgan so'zlar to'plami. Ot so'z turkumiga oid omonim so'zni esa

yuqorida keltirilgan matematik model yordamida farqlash mumkin. Bu to'plamlarning to'g'ri va to'liq to'planishi uchun mutaxassislar jalb qilinishi lozim.

### 2.1.3.7. Ot V undov so'z orasidagi omonimlikni aniqlash

Ikkita so'z turkumlari orasidagi omonimlikni hosil qiluvchi Ot va undov so'z orasidagi omonim so'zlarga:

**BAS:** Boshqa kerak emas, shu bilan yetar; to'xtat (*undov so'z*)

**BAS:** Garov bog'lashmoq, garov o'ynamoq (*Ot*)

**BAS:** 1. Erkaklarning eng past, yo'g'on ovozi va shunday ovoza ega ashulachi.

2. Ko'p ovozli musiqa asarlaridagi eng pastki ovoz, partiya (*ot*)

Ot so'z turkumining grammatik xususiyatlari bizga ma'lum. Undov so'zning grammatik xususiyatlarini aniqlash uchun bir nechta jumlarlar keltiramiz.

- Xullas, o'zimizni aldashni **bas** qilaylik (*korpus ma-t*)
- qo'rquvlariga **bas** kelishlari (*korpus ma-t*)
- bunda "sotuvchi — kombinat" o'rtasidagi muammo ochib berilsa **bas** edi. (*korpus ma-t*)
- Ularga **bas** kelolmasligini anglagan bola o'zini himoya qilish uchun yerdan tosh olib, jon-jahdi bilan raqiblariga uloqtiradi. (*korpus ma-t*)
- *Alisher o'rtog'lar bilan bugungi futbol o'yini haqida bas bog'ladi* (*korpus ma-t*)

Kuzatuvlar shuni ko'rsatadiki, undov so'zi kontekst hech qanday lug'aviy va sintaktik shakl yasovchi qo'shimcha bilan birikmagan holda, mustaqil uchraydi.

*Bas* ⊕ *qilmoq/edi/kelmoq/demoq/yetar*

*Bas* ⊕ *bog'lamoq/o'ynamoq/kirishmoq/...*

$$H_E = \begin{cases} W + \begin{matrix} -ga \\ -da \end{matrix} \oplus H^E \oplus V^e \\ H^N + aff^N \oplus V^{N_E} \end{cases} \quad (17)$$

Bu yerda  $V^E$  – undov omonim bn birikuvchi fe'llardan iborat chekli to'plam,

$V^E = \{qilmoq, demoq, edi, \dots\}$ ,  $V^{NE}$  – ushbu guruhga mansub ot soʻz turkumidagi omonim soʻz bilan birika oladigan feʼllar toʻplami.

### 2.1.3.8. Ot v olmosh soʻz turkumlari orasidagi omonimlikni aniqlash

Ot olmosh soʻz turkumlari orasidagi omonim soʻzga misol keltiramiz

- 1) UL: Oʻgʻil. (Ot)
- 2) UL: esk. ayn. u II. (olmosh)

Bu soʻzlarni kontekstda uchrashiga namuna keltiramiz.

- *Bu ust-boshi rasvo yigit choponning uli boʻladi*
- *Ul sarvi gulroʻ kelmadi.*

Kuzatuvlar shuni koʻrsatadiki bu guruhdagi omonim soʻzlar kontekstda mustaqil holda ham, qoʻshimchalar bilan birikkan holda ham uchrashi mumkin. Demak ular morfologik xususiyatiga koʻra oʻxshash, shuning uchun bu turdagi omonimlarni ham Yashirin Markov modeli yordamida farqlashni tavsiya qilamiz.

### 2.1.3.9. Ot v bogʻlovchi orasidagi omonimlikni aniqlash

Ot v bogʻlovchi orasidagi omonim soʻzlarga misol sifatida goʻyo soʻzini keltiramiz.

*GOʻYO: Aytuvchi, gapiruvchi. (ot)*

*GOʻYO: Qiyoslash, oʻxshatish maʼnolarini ifodalash uchun xizmat qiladi. (bogʻlovchi)*

Bu turdagi omonim soʻzlarni farqlashda bogʻlovchini morfologik tahlili koʻrib chiqish kifoya. Bogʻlovchilar hech qanday qoʻshimcha bilan birikmaydi, gapda oʻzidan oldin vergul qoʻyiladi. Lekin ot soʻz turkumi doirasida boʻlganda ham xuddi shunday boʻlishi mumkin.

- *Maktabga har kuni goʻyo bayramga borayotgandek oshiqadi. (korpus ma-t.) (bogʻlovchi)*
- *Goʻyo ular men anglagan hikmatni allaqachon bilib boʻlgandek. (korpus ma-t.) (bogʻlovchi)*

- *Qozi mashvarat qatnashchilari hofiz, go'yo, faylasuflarga yuzlandi.*  
(*koprus ma-t.*) (ot)

Ushbu jumalardan ko'rinadiki, bu turdagi omonim so'zlar kontekstda qo'shimcha bilan birikkan holda uchrasa demak bu omonim so'z ot so'z turkumiga mansub, aks holda bu so'zlarni brikuvchilari ajratib olinadi va ular orqali farqlanadi.

### 2.1.3.10. Sifat v son orasidagi omonimlikni aniqlash

Son va sifat so'z turkumlari orasidagi omonimlikni farqlash jarayonida ham qoidalarga asoslangan teglash usulidan foydalanamiz. Yuqorida sifat so'z turkumining morfologik xususiyatlari haqida ma'lumot bergan edik. Quyida son va sifat so'z turkumlari orasidagi omonimlikni farqlovchi modellar ko'rib chiqamiz.

**Ellik:** Beshinchi o'nlikning nomi(son)

**Ellik:** Barmoqning kengligiga teng, serbar(sifat)

Ushbu omonimlarni qurshovda uchrashiga namunalar keltiramiz.

Son so'z turkumi:

1. *Yaktagining bariga to'rt-beshta handalak solib olgan ellik yoshlar chamasidagi bir kishi shoshilmay boradi. (Said Ahmad, Ufq, 7 b)*
2. *Yoz bo'yi ellik-olimish mirob bu yerds ko'rpa to'shak qilib yotib oladi (Said Ahmad, Ufq, 30b)*
3. *Ellikta gul sotib oldik.*
4. *Ikki yuz ellik ikkinchi maktabda ishlayman.*

- **Ellik:(son)** ⊕ {bosh, nafar, dona, marta, karra, bir, ikki, ...} (ot)

- {yuz, ming, million, ...} ⊕ **Ellik (son)**+ta/-tacha/-inchi/... ⊕ {N} (ot)

Sifat so'z turkumi:

- *Ellik gaz gazlama sotib olindi*

- *Ellik yog'di.*

**Ellik:(sifat)** ⊕ {yog'} (fe')

**Ellik: (sifat)** ⊕ {chang} (ot)

Yuqoridagi morfologik xususiyatlarni inobatga olgan holda quyidagi matematik modelni keltiramiz.

$$H_{num,Adj} = \begin{cases} H^{num} + \downarrow affl^{ord} \oplus \downarrow (W_k^N + \downarrow affl^{ord}) \oplus N_f^f \\ H^{Adj} \oplus V \\ H^{Adj} \oplus N \end{cases} \quad (18)$$

Bu model yordamida sifat va son orasidagi omonimlarni farqlashda foydalanishni tavsiya qilamiz.

### 2.1.3.11. Olmosh $\vee$ yuklama orasidagi omonimlikni aniqlash

Badiiy asarlar, omonimlarning izohli lug'ati va korpus ma'lumotlari ichidan Olmosh  $\vee$  yuklama orasidagi omonimlik hosil qiluvchi u so'zi topildi.

U: ko'rsatish olmoshi sifat kabi ishlatiladi. (*olmosh*)

U: III shaxs birlik kishilik olmoshi-ot kabi ishlatiladi. (*olmosh*)

U: Ta'kid, ajablanish kabi ma'nolarni ifodalaydi (*yuklama*)

U so'zi ishtirok etgan jumalarni keltiramiz

- *Men bugun senga u bu ishlarni buyurib ketuvdim. (ko'rsatish olmoshi)*
- *Chunki u doim darsga mas'uliyat bilan yondashar, o'z metodlaridan ko'ngli to'lardi. (kishilik olmoshi) (korpus ma-t.)*
- *Nurin sochgay bog '-u cho'lga, diyori O'zbekiston. (A.Xishraviy)*

Olmoshlar otlashish xususiyatiga ega, yuklama esa juft so'zlarning tarkibida uchrashi kuzatildi:

*bog '-u rog', do'st-u dushman, yosh-u qari, oq-u qora, sen-u men, oz-u ko'p, ...*

-u shaklda uchrasa yuklama deb xulosa chiqariladi. U omonim so'zi yuklama bo'lib barcha mustaqil so'z turkumlariga doir juft so'zlar tarkibida uchraydi.

### 2.1.3.12. Undov $\vee$ yuklama orasidagi omonimlikni aniqlash

Undov  $\vee$  yuklama orasidagi omonim so'zlardan faqat *ha* so'zi topildi. Bu so'zni kontekstda undov yoki yuklama ekanligini aniqlash modelini keltirsak. Matematik model uchun bu turdagi omonim so'zlarning morfologik, sintaktik xususiyatlarini ko'rib chiqamiz.

HA: 1 Javobda biror fikrni tasdiqlashni, rozilikni bildiradi.



2 Biror fikrni, fakti alohida qayd etadi, ta'kidlaydi.

3 Kutilmaganda biror yangi narsa, hodisa, fikr yodga tushganda, biror kimsa yoki narsani eslashga urinishda, suhbat mavzusini o'zgartirish kabi hollarda «aytganday», «aytmoqchi» degan ma'noni bildiradi yoki shu so'zlariga xos ma'noni ta'kidlab, unga e'tiborni tortadi. (yuklama)

**HA:** 1 So'roq va taxmin ma'nolarini ta'kidlash uchun ishlatiladi; so'roqqa e'tiborni jalb etadi.

2 Undov gaplar tarkibida (a tovushi turlicha — kuchli, cho'ziq va boshqacha aytilib) ma'qullash, norozilik, taajjub kabi his-tuyg'ularni ifodalaydi.

3 Biror ish-harakatga undashni, shu ma'noda qo'llab-quvvatlashni bildiradi.

4 Biror ish-harakatga undash, qistash ma'nolarini anglatadi.

5 Taajjub, norozilik, rad etish kabi ma'nolarni anglatadi.

6 Birovga murojaat qilganda, «hoy» yoki «ho» so'zlari o'rni ishlatiladi. (undov so'z)

**Namuna:**

1. *Jonivor ha deganda bo'y sunavermaydi.*

2. *Ha, xuddi shunday.*

o *Ha demay xorijdan qaytadilar*

3. *Ah-ha... men ko'ray*

Namunada berilgan jumladan ko'rinadiki, **ha** so'zi kontekstdagi birikuvchilari yordamida farqlanadi.

### 2.1.3.13. Fe'l V ko'makchi orasidagi omonimlikni aniqlash

Fe'l va ko'makchi orasidagi omonim so'zlari quyidagilar

1. **Osha:** eskirgan. 1. Yemoq. 2. Palovni biror kishining panjasidan og'ziga olmoq. (fe'l)

2. **Osha:** Ma'lum vaqtni o'tkazib; ma'lum manzil orqali (ko'makchi)

Bu so'zni kontekstlarda quyidagicha uchrashlari kuzatildi.

1. Asrlar **osha** yashab kelayotgan qadriyatlaridan. (korpus ma-ti.)

2. π sonni yillar **osha** olimlar e'zozlashadi. (korpus ma-ti.)

3. Qariya oshini **oshab**, yoshini yashab bo'lgan edi. (korpus ma-ti.)

4. Oshini **oshatib** bo'ldi. (korpus ma-ti.)

1 va 2 jummalarda keltirilgan *osha* soʻzi koʻmakchi vazifasida kelgan. Bilamizki koʻmakchilar qoʻshimcha bilan birikmaydi va chap tomonidan zamon va makonni ifodalovchi soʻzlar bilan birikadi.

*zamonlar/ yillar/asrlar/kamlar/soatlar/... ⊕ osha*

3 va 4 jummalarda *osha* soʻzi feʼl soʻz turkumiga oid. U eskirgan soʻz boʻlib, manbalarda *osh* soʻzi bilan birikkan holda uchrashi kuzatildi.

*Oshini ⊕ osha + b/-t/-tib/-ttirib/-ttir...*

Ushbu maʼlumotlarni inobatga olib quyidagi matematik modelni shakllantiramiz

$$H_{V,Aux} = \begin{cases} W^T \oplus H^{Aux} \\ osh + \downarrow aff^N \oplus (H^V + aff_k^V) \end{cases} \quad (19)$$

Bu yerda  $W^T$ —makon va zamoni ifodalovchi soʻzlardan iborat chekli toʻplam,  $W^T = \{zamon, asr, yil, kun, oy, kun, makon, tog', \dots\}$ ;  $aff_k^V$  — feʼl omonimga qoʻshiluvchi baʼzi qoʻshimchalar,

$$aff_k^V = \{-t, -b, -tib, -tir, -tirib, -gan, -tgan, \dots\}$$

$aff_k^V$  va  $W^T$  toʻplamlarning qiymatlari qancha toʻliq va toʻgʻri boʻlsa feʼl va koʻmakchi orasidagi omonimlikni farqlash shuncha sifatli boʻladi.

#### 2.1.3.14. Sifat V koʻmakchi orasidagi omonimlikni aniqlash

Sifat V koʻmakchi orasidagi omonim soʻzlarga orqali soʻzini namuna qilib keltiramiz

**Orqali:** 1 Biror narsaning ichidan, orasidan, oʻrtasidan yoki ustidan, biror narsa oralab, bir tomondan ikkinchi tomonga. 2 Biror kimsa yoki narsa vositasida, yor-damida, yordami bilan. 3 Tufayli, sababli.

(koʻmakchi)

**Orqasi:** Orqasi, suyanchigʻi, homiysi bor. (sifat)

Sifat soʻz turkumining otlashishi xususiyatini hisobga olgan holda, bu turdagi omonimlikni (16) formula orqali farqlashni tavsiya qilamiz.

### 2.1.3.15. Sifat $\vee$ taqlid so'z orasidagi omonimlikni aniqlash

Sifat va taqlid so'z turkumi doirasidagi omonimlikni farqlash modellarini yaratishda shunga alohida e'tibor berish kerakki, bu turdagi omonim so'z kontekstda sifat bo'lib kelgan bo'lsa bilingki bu so'z tojik tilidan kirib kelgan so'z. Bundan kelib chiqadiki ushbu sifat o'zbek tilidagi sifatga qo'shiladigan qo'shimchalar qo'shilmaydi. Ammo o'ng va so'l tomonidan o'zbek tilidagi sifat bilan bir xil bo'lishi mumkin. Bunday hollarda birikuvchilari orqali farqlash madsadga muvofiq. Taqlid so'zning morfo-semantik filteri yuqorida keltirgan edik. Yuqorida keltirilgan taqlid va sifat omonimlarni farqlovchi modellarni birlashtirsak ushbu turdagi omonimlarni farqlovchi matematik model hosil bo'ladi.

### 2.1.3.16. Sifat $\vee$ undov so'z orasidagi omonimlikni aniqlash

Sifat  $\vee$  undov so'z orasidagi omonim so'zlarga misol sifatida **balli** so'zini keltiramiz.

- Balli:** 1 Ma'qullash, tasdiqlash, maqto'v, tahsin ma'nolarini ifodalaydi.  
2 Cho'zib talaffuz qilinganida, norozilik, kesatiq ma'nolarini ifodalaydi. (*undov so'z*)

**Balli:** 1 Muayyan ball kuchiga ega bo'lgan.

2 Raqamlar sistemasi bilan ifodalangan. (*Sifat*)

Sifat  $\vee$  undov so'zlar orasidagi omonim so'zlarni farqlashda ot  $\vee$  undov so'z orasidagi omonimlikni farqlovchi modeldan foydalanish maqsadga muvofiq

### 2.1.3.17. Sifat $\vee$ modal so'z orasidagi omonimlikni aniqlash

Sifat va modal so'z orasidagi omonimlarni farqlashda ham qoidalarga asoslangan teglash usulidan foydalanish maqsadga muvofiq. Garchi bunday so'zlar sanoqli bo'lsa-da farqlovchi omillarni keltirib o'tishni joiz deb topdik. Misol tariqasida quyidagi omonim so'zni keltiramiz.

**Essiz:** 1 Es-hushi o'zida emas.

2 Xotirasi sust, hech narsani eslab qololmaydigan. (*sifat*)

**Essiz:** Eh Essiz, afsus, achinish hissini ifodalovchi so'z. (*Modal so'z*)

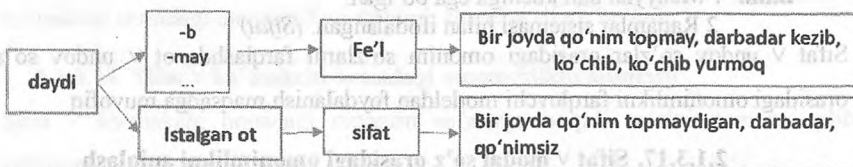
Yuqorida sifat soʻz turkumidagi omonim soʻzni farqlovchi omillarni koʻrib chiqqan edik. Endi esa modal soʻzlarning morfologik xususiyatlarini oʻrganib chiqamiz. Oʻzbek tili grammatikasidan maʼlumki, modal soʻz gap boshi boʻlsa undan keyin, gap oʻrtasida boʻlsa uning ikkala tomonidan, gap oxirida kelsa undan oldin vergul qoʻyiladi. Faqat shu xususiyati bilangina farqlashimiz mumkin.

$$H_{adj,m} = \begin{cases} \downarrow (N_i + x) \oplus Adj \oplus \downarrow (N_i + x), & x \in \{aff_{s_{ij}}^N\}, i = 1..n \\ Adj + x, & x \in \{aff_{s_{ij}}^N\}, i = 1..n \end{cases} \quad (20)$$

Ushbu model yordamida modal soʻz va sifat orasidagi omonimlikni farqlash mumkin.

### 2.1.3.18. Sifat va feʼl orasidagi omonimlikni aniqlash

Sifat va feʼl soʻz turkumlari orasidagi omonimlikni farqlashda qisman ot va feʼl orasidagi omonimlikni farqlash qoidalaridan ham foydalanamiz, chunki sifat oʻtlashish xususiyatiga ega. Shunday qilib sifat va feʼl soʻz turkumi doirasidagi omonimlikni farqlash jarayonini koʻrib chiqamiz.



Undan tashqari gapda sifat oʻtlashib kelishini inobatga olishimiz lozim, yaʼni

*Bu odam daydiga oʻxshar ekan.*

gapda *daydi* soʻziga otga qoʻshiluvchi sintaktik shakl yasovchi qoʻshimcha qoʻshilgan va undan keyingi ot tushib qolgan. Bu misollar asosida esa sifat va feʼl soʻz turkumi doirasidagi omonimlikni farqlash uchun quyidagi matematik modelni keltiramiz:

$$H_{Adj,V} = \begin{cases} \downarrow (N+x) \oplus H^{adj} + \downarrow aff^{Ad,d} \oplus \downarrow (N+x), & x \in \{aff^N\} \\ H^{adj} + x, & x \in \{aff^N \cup aff^{Ad,d}\} \\ H^V + x, & x \in \{aff_{s,j}^V\}, j = 1..m \end{cases} \quad (21)$$

$H_{Adj,V}$ -Sifat va fe'l orasidagi omonim so'zlar,  $H^{adj}$ -sifat omonim,  $\oplus$ -belgidan keyin so'z kelishini mumkinligini anglatadi,  $N_i$ - Ot so'z turkumidagi so'zlar to'plami,  $aff_{s,i}^N$ - OT so'z turkumi doirasidagi omonim so'zga qo'shiluvchi sintaktik va lug'aviy shakl yasovchi qo'shimcha va qo'shimchalar kombinasiyalari to'plami,  $H^V$ -fe'l omonim,  $x$ -lug'aviy va sintaktik shakl yasovchi qo'shimchalardan iborat,  $aff_{s,j}^V$ - Fe'lga qo'shiluvchi sintaktik va lug'aviy shakl yasovchi qo'shimcha va ularning kombinasiyalari to'plami, ya'ni  $aff_{s,j}^V = \{-ma, -magan, -maganman, -magansan, -magandir, -magandi, -magansiz, -may, \dots\}$

Yuqoridagi matematik model orqali ot va sifat orasidagi omonimlarni farqlashga samarali natija beradi.

### 2.1.3.19. Ravish $\vee$ yuklama orasidagi omonimlikni aniqlash

Ravish  $\vee$  yuklama orasidagi omonim so'zga axir so'zini keltiramiz

**Axir:** So'ng, pirovard. Oxiri, pirovardida, nihoyat. (*ravish*)

**Axir:** Ma'noni kuchaytiradi, ta'kidlaydi. (*yuklama*)

Bu so'zning kontekstda uchrashi:

1. Ataylab tosh yemagandir, axir! (*yuklama*)
2. Va buyuk hikmatga axir yuz burdim. (A. Oripov) (*ravish*)
3. Axir, u butun respublikadagi o'quvchilarni baholay olmasligi aniq-ku! (korpus m-ti) (*yuklama*)
4. Ko'pdan beri kutilayotgan tekshiruvchi axiri keldi. (korpus m-ti) (*ravish*)

Kontekstda axir yuklamasi gap boshida yoki oxirida uchrashi kuzatildi. Gap boshida uchraganida undan keyin vergul, gap oxirida uchraganida undan keyin

undov(!) va undan oldin vergul belgisi qo'yiladi. Bu guruhga mansub omonim yuklama bo'lsa u hech qanday qo'shimcha bilan birikmaydi. Ravish esa otlashish xususiyatiga ega. Mana shularni hisobga olib quyidagi matematik modelni tavsiya qilamiz.

$$H_{Adv,P} = \left\{ \begin{array}{l} W+, \oplus H^P+, \\ H^{Adv} + \downarrow aff^N + W^{Adv} \end{array} \right. \quad (22)$$

Bu yerda W-yuklamadan oldin kelgan ixtiyoriy so'z,  $W^{Adv}$ -ravish so'z turkumidagi omonim so'zlar bilan birikadigan so'zlar to'plami.

### 2.1.3.20. Ravish v fe'l orasidagi omonim so'zlar

Navbatda ravish va fe'l orasidagi omonimlikni farqlovchi omillarni ko'zdan kechirib o'z tavsiyamizni keltiramiz.

**Oz:** Kam, miqdori nisbatan ko'p bo'lmagan. (*ravish*)

**Oz:** Oriqlamoq, etidan yo'qotmoq. (*fe'l*)

**Oz:** Noxush bo'lmoq, kuchsizlanmoq, holsizlanmoq. (*fe'l*)

**Oz:** Adashmoq, to'g'ri yo'ldan chetga chiqmoq (*fe'l*)

3 ta o'rinda fe'l va bitta ravish vazifasida keluvchi omonim so'z keltirilgan. Ushbu tipdagi omonim so'zlarni farqlovchi modellarni ko'rib chiqamiz.

Bu tipdagi omonim so'zlardan qaysi biri fe'l so'z turkumiga oid ekanligini ajratib olishni ko'rgan edik. Endi esa ravish so'z turkumiga oid ekanligini aniqlashimiz qoldi xalos. Buning uchun ravishning morfologik xususiyalarini ko'rib chiqamiz.

1) oz+ (-gina) +odam/mehnat/mashaqqat/bola/vaqt

2) oz+gapir/so'zla/kul/o'yna/ishla/tishla/qol/kut/sabr qil/sarfla

3) oz ravishi otlashish xususiyatiga ega bo'lib, oz fe'lida bunday hodisa kuzatilmaydi. Ravishga -i egalik qo'shimchasi qo'shilsa undan oldin ot so'z turkumiga oid so'z keladi va u qaratqich kelishigini oladi, ya'ni



gapning ozi yaxshi, ishning ko'pi

mevaning ozini olib kelmoq

Bu qoidalarni inobatga olgan holda quyidagi matematik modelni keltiramiz

$$H^{adv} = \begin{cases} H+\downarrow aff_i^{adv} \oplus N, i = 1..n, aff_i^{adv} \in aff^N \\ H+\downarrow aff_i^{adv} \oplus V, i = 1..n, aff_i^{adv} \in aff^N \\ \downarrow (N + ning) \oplus H+\downarrow i+\downarrow ni \oplus V \end{cases} \quad (23)$$

Bu yerda  $aff_i^{adv}$  –ravishga qo'shilishi mumkin bo'lgan qo'shimchalar, yani kelishik, egalik, otning kichraytirish erkalash shaklini ifodalovchi qo'shimcha, sifat darajasini ifodalovchi qo'shimchalar va ularning kombinatsiyalaridan iborat qo'shimchalardan iborat chekli to'plam:

$$aff_i^{adv} = \{-gina, -roq, -roqqina, -si, -ginasi, -ning, -ginasini, \dots\},$$

$N$  –ot(noun),  $V$  –fe'l(verb)

Bu to'plamlarni hosil qilib ma'lumotlar bazasiga joylansa bu turdagi omonimlarni avtomatik farqlash 100% amalga oshiriladi.

#### 2.1.3.21. Ravish V ko'makchi orasidagi omonimlikni aniqlash

Bu guruhga mansub omonim so'zga misol tariqasida *baravariga* so'zini misol qilib olish mumkin.

**baravariga:** O'miga, eva-ziga. (*ko'makchi*)

**baravariga:** Bir vaqtda, paytda, baravar, bir-biridan qolishmay, birday. (*ravish*)

Bu guruhga mansub omonim so'zlarning birikuvchilarini korpus ma'lumotlari ichidan yig'ish maqsadga muvofiq. To'plangan birikuvchilar asosida Trigramm Yashirin Markov modelidan foydalaniladi va omonimlar semantik tahlil qilinadi.

#### 2.1.3.22. Undov va bog'lovchi orasidagi omonimlikni aniqlash

Undov va bog'lovchi orasidagi omonimlik farqlovchi matematik mulohazani tavsifiya qilishdan bu turdagi so'zlarga namuna keltiriladi.

undov(!) va undan oldin vergul belgisi qo'yiladi. Bu guruhga mansub omonim yuqum bo'lsa u hech qanday qo'shimcha bilan birikmaydi. Ravish esa otlashish xususiyatiga ega. Mana shularni hisobga olib quyidagi matematik modelni tavsiya qilamiz.

$$H_{Adv,P} = \left\{ \begin{array}{l} W+, \oplus H^P+, \\ H^{Adv} + \downarrow aff^N + W^{Adv} \end{array} \right. \quad (22)$$

Bu yerda W-yuklamadan oldin kelgan ixtiyoriy so'z,  $W^{Adv}$ -ravish so'z turkumidagi omonim so'zlar bilan birikadigan so'zlar to'plami.

### 2.1.3.20. Ravish $\vee$ fe'l orasidagi omonim so'zlar

Navbatda ravish va fe'l orasidagi omonimlikni farqlovchi omillarni ko'zdan kechirib o'z tavsiyamizni keltiramiz.

**Oz:** Kam, miqdori nisbatan ko'p bo'lmagan. (*ravish*)

**Oz:** Oriqlamoq, etidan yo'qotmoq. (*fe'l*)

**Oz:** Noxush bo'lmoq, kuchsizlanmoq, holsizlanmoq. (*fe'l*)

**Oz:** Adashmoq, to'g'ri yo'ldan chetga chiqmoq (*fe'l*)

3 ta o'rinda fe'l va bitta ravish vazifasida keluvchi omonim so'z keltirilgan. Ushbu tipdagi omonim so'zlarni farqlovchi modellarni ko'rib chiqamiz.

Bu tipdagi omonim so'zlardan qaysi biri fe'l so'z turkumiga oid ekanligini ajratib olishni ko'rgan edik. Endi esa ravish so'z turkumiga oid ekanligini aniqlashimiz qoldi xalos. Buning uchun ravishning morfologik xususiyatlarini ko'rib chiqamiz.

1) oz + (-gina) + odam/mehnat/mashaqqat/bola/vaqt

2) oz + gapir/so'zla/kul/o'yna/ishla/tishla/qol/kut/sabr qil/sarfla

3) oz ravishi otlashish xususiyatiga ega bo'lib, oz fe'lida bunday hodisa kuzatilmaydi. Ravishga  $-i$  egalik qo'shimchasi qo'shilsa undan oldin ot so'z turkumiga oid so'z keladi va u qaratqich kelishigini oladi, ya'ni

*gapning ozi yaxshi, ishning ko'pi*

*mevaning ozini olib kelmoq*

Bu qoidalarni inobatga olgan holda quyidagi matematik modelni keltiramiz

$$H^{adv} = \begin{cases} H + \downarrow aff_i^{adv} \oplus N, & i = 1..n, aff_i^{adv} \in aff^N \\ H + \downarrow aff_i^{adv} \oplus V, & i = 1..n, aff_i^{adv} \in aff^N \\ \downarrow (N + ning) \oplus H + \downarrow i + \downarrow ni \oplus V \end{cases} \quad (23)$$

Bu yerda  $aff_i^{adv}$  –ravishga qo'shilishi mumkin bo'lgan qo'shimchalar, yani kelishik, egalik, otning kichraytirish erkalash shaklini ifodalovchi qo'shimcha, sifat darajasini ifodalovchi qo'shimchalar va ularning kombinatsiyalaridan iborat qo'shimchalardan iborat chekli to'plam:

$aff_i^{adv} = \{-gina, -roq, -roqqina, -si, -ginasi, -ning, -ginasni, \dots\}$ ,

$N$  –ot(noun),  $V$  –fe'l(verb)

Bu to'plamlarni hosil qilib ma'lumotlar bazasiga joylansa bu turdagi omonimlarni avtomatik farqlash 100% amalga oshiriladi.

### 2.1.3.21. Ravish V ko'makchi orasidagi omonimlikni aniqlash

Bu guruhga mansub omonim so'zga misol tariqasida *baravariga* so'zini misol qilib olish mumkin.

**baravariga:** O'miga, eva-ziga. (*ko'makchi*)

**baravariga:** Bir vaqtda, paytda, baravar, bir-biridan qolishmay, birday. (*ravish*)

Bu guruhga mansub omonim so'zlarning birikuvchilarini korpus ma'lumotlari ichidan yig'ish maqsadga muvofiq. To'plangan birikuvchilar asosida Trigramm Yashirin Markov modelidan foydalaniladi va omonimlar semantik tahlil qilinadi.

### 2.1.3.22. Undov va bog'lovchi orasidagi omonimlikni aniqlash

Undov va bog'lovchi orasidagi omonimlik farqlovchi matematik mulohazani tavsiya qilishdan bu turdagi so'zlarga namuna keltiriladi.

**Va:** Teng huquqli gan bo'laklari va gaplarni bog'lash uchun xizmat qiladi. (*bog'lovchi*)

**Va:** Narsa, kimsa, harakat kabilarning belgisi me'yoriy darajadan pastligini, maqtaydigan darajada emasligini ifodalaydi (*undov so'z*)

*Va* so'zida badiiy asarlardan namunalar keltiramiz.

- 1) *Qor va yomg'ir/ Kecha va kunduz/ oq va qora/...*
- 2) *Ko'ngil xazinasining qulfi tildir va ul xazinaning kaliti so'zdir. (Alisher Navoiy)*
- 3) *Va-a, bor qilgan ishing shumi?*
- 4) *Ba'zilar ko'zlarini o'ynatib, "o'xo'"! "o'xo'"! desa, ba'zilar labini burib, "va-a'"! "va-a'"! derdi.*

Kuzatishlar va undov so'z va-a ko'rinishda uchrashini ko'rsatdi. Boshqa hollarda bog'lovchi deb xulosa qilish maqsadga muvofiq.

#### **2.1.3.23. Undov v taqlid so'zlar orasidagi omonimlikni aniqlash**

Undov v taqlid so'zlar orasidagi omonim so'zlarga misol sifatida *ba* so'zini keltiramiz.

**Ba:** Ba'zi uy hayvonlarini chaqirish uchun ishlatiladi. (*undov so'z*)

**Ba:** Qo'ylar ma'ragandagi tovushni ifodalaydi. (*taqlid so'z*)

Taqlid so'z otlashish xususiyatiga ham ega. Taqlid so'zi gapda asos ko'rinishida ishtirok etganida esa undov so'zda birikuvchilari yordamida farqlanadi

#### **2.1.3.24. Ot v taqlid so'z orasidagi omonimlikni aniqlash**

Ot va taqlid so'z turkumlari orasidagi omonimlikni farqlashda ham qoidalarga asoslangan teglash usulidan foydalanish tavsiya qilinadi.

Masalan, 4 ta ot va taqlid so'z orasidagi omonim so'zni farqlashda faqat ot orasidagi omonimlarni farqlashda bir turkum doirasida omonimlikni farqlovchi matematik modelidan foydalaniladi, endi esa taqlid so'zni aniqlovchi modelni keltiramiz:

- 1) to'p: Koptok, futbol koptogi. (*ot*)
- 2) to'p: Snaryad otuvchi uzun stvulli harbiy qurol, zambarak (*ot*)
- 3) to'p: so'zlash. Peshin (*ot*)

4) to'p: O'ram, bog'lam, guruh (*ot*)

5) to'p: Yumshoq jismning yerga urilishidan chiqadigan tovushni bildiradi  
(*taqlidiy so'z*)

To'p so'zi taqlid so'z bo'lib kelganida unga qo'shilishi mumkin bo'lgan qo'shimchalarni ham, brikuvchi so'zlarni ham aniqlab olish talab qilinadi.

Yaratilgan lingvistik modellarga tayanib quyidagi matematik formulani keltiramiz.

$$H^{imit} = \begin{cases} H + aff_i^{CS} \oplus V_j^{imit} \\ H \oplus \downarrow N_l^{imit} + aff_i^N \oplus V_j^{imit} \\ H \oplus V_k^{q2} \end{cases} \quad (24)$$

Bu yerda  $aff_i^{CS}$  - kelishik qo'shimchalar to'plamining qism to'plami, ya'ni

- Tushum kelishigi *-ni*: to'pni bas qil
- Jo'nalish kelishigi *-ga*: to'pga uyg'on/kel
- Chiqish kelishigi *-dan*: to'pdan qo'rq/uyg'on, demak taqlid so'zlar otlashish xususiyatiga ega ekan.

$V_j^{imit}$ -taqlid so'zlarga brikuvchi fe'l so'z turkumidagi so'zlardan iborat chekli to'plam, masalan:

$$V_j^{imit} = \{bas qilmoq, uyg'onmoq, kelmoq, \dots\}, \quad j = 1 \dots n$$

Taqlid so'z otlashganda otga xos grammatik shaklni qabul qilishi so'zni otomonimga o'xshab qolishiga va ayrim brikuvchilar bir xil bo'lishiga olib keladi. Shuni hisobga olgan holda

$N_l^{imit}$ - taqlid so'z bilan brika oladigan ot(noun) so'z turkumiga oid so'zlardan iborat chekli to'plam hosil qilinadi,  $aff_i^N$ - ot so'z turkumiga qo'shilishi mumkin bo'lgan qo'shimchalar:

Tushum kelishigi *-ni*: to'p degan ovozni bas qil

Jo'nalish kelishigi *-ga*: to'p ovoz+i/lari/miz/... +ga uyg'on/kel

Chiqish kelishigi *-dan*: to'p ovoz+i/lari/miz/... +dan qo'rq,

$V_k^{q2}$  — qo'shma fe'ning 2-komponenti, ya'ni taqlid so'zga qo'shilib qo'shma fe'l hosil qiladi:  $V_k^{q2} = \{etmoq, demoq, qilmoq, \dots\}$

*To'p qilmoq, to'p demoq, ...*

Yuqorida keltirilgan matematik model orqali ot va taqlid so'z turkumlari orasidagi omonimlikni farqlash mumkin.

### 2.1.3.25. Ot va sifat va ravish orasidagi omonimlikni aniqlash

Sifat, ot va ravish so'z turkumlari orasidagi omonimlikni farqlashni ko'rib chiqamiz:

Masalan:

**Sira:** Ochiq, aniq, to'g'ri (*sifat*)

**Sira:** Hech, aslo (*ravish*)

**Sira:** Olmosh, da'vat ifodalovchi so'z orasidagi omonimlik (*ot*)

Sifat, ravish va ot so'z turkumlari orasidagi omonim so'zlarni kontekstda farqlash uchun ularning xususiyatlari tahlil qilinadi. Sifat va ot so'z turkumlari doirasidagi so'zlar grammatik jihatdan o'xshash. Demak ot va ravish yoki sifat va ravish so'z turkumlari orasidagi omonim so'zlarni farqlovchi matematik modelni keltirish kifoya. Sifat va ot so'z turkumlari orasidagi omonim so'zlar Trigramm YMM orqali farqlanadi. Sifat va ravish so'z turkumlari orasidagi omonim so'zlarni qo'shimchalar orqali farqlash aniq natija bermaydi. Quyidagi gaplarda ularni ma'nolari yaqqol namoyon bo'ladi.

1. *Ammo mulla nikoh o'qishga siru unamas edi. (Said Ahmad, Ufq trilogiyasi, 26-b)*
2. *Endi ko'nglingga kelmasinu, gapning sirasini aytib qo'rganim durust (Said Ahmad, Ufq trilogiyasi, 327-b)*

Bu so'zni badiiy asarlarda kuzatganimizda ravish bo'lib kelganida undan keyin ham, oldin ham turli so'z turkumidagi so'zlar bilan birikishi, gapning boshida ham uchrashi mumkinligi kuzatildi. Ko'p hollarda o'zak holda uchraydi. Sifat so'z turkumi bo'lib kelganida esa egalik va kelishik qo'shimchalari bilan birikishi kuzatildi va undan oldin qaratqich kelishigini qabul qilgan ot bilan birikishini



uchratdik. Ushbu statistik ma'lumotlarni inobatga olib quyidagi matematik formulani tavsiya qilamiz.

$$H_{adv,adv} = \left\{ \begin{array}{l} W_i^{adv,b} \oplus H^{adv} + \downarrow aff^N \oplus W_i^{adv,a} \\ \downarrow (N_k + aff_j^N) \oplus H^{adj} \oplus \downarrow (N_k + aff_j^N) \\ H^{adj} + aff_i^N \end{array} \right. \quad (25)$$

Bu yerda  $W_i^{adv,b}$ - ravish omonimga chap tomonidan birikuvchi so'zlardan iborat chekli to'plam,  $W_i^{adv,a}$ - esa o'ng (omonimdan keyin) tomonidan birikishi mumkin bo'lgan so'zlardan iborat chekli to'plam. Agar ravish otga xos qo'shimchalarni qabul qilsa, u holda  $W_i^{adv,a}$  - so'z fe'l so'z turkumiga oid bo'ladi Bu to'plamlarning qiymatlari badiiy asarlardan filologlar nazorati ostida qo'l mehnati bilan shakllantiriladi va ma'lumotlar bazasiga kiritiladi. Bu turdagi omonimlarni birikuvchilarining aniqlik darajasi yuqori bo'lsa, omonimlarning semantik tahlili 100% natija beradi.

### 2.1.3.26. Ot v olmosh v fe'l so'z turkumlari orasidagi omonimlikni aniqlash

Ot, olmosh va fe'l so'z turkumlari orasida omonimlikni ifodalovchi so'zga misol keltiramiz.

**Un:** Donni tegirmonda tortib hosil qilinadigan kukun holatidagi

mahsulot. (Ot)

**Un:** eskirgan. Ovoz. (ot)

**Un:** 21 shaxs birlik kishilik olmoshining ba'zi affikslar qo'shiladigan asos i. (olmosh)

**Un:** O'sib, yuzaga chiqmoq. (fe'l)

Tadqiqotimizning maqsadi omonimlarni kontekstda farqlash bo'lganligi sababli, dastlab jumalarni keltiramiz.

1) *Hamma yoq qabriston jimligiday unsiz, vahimali.*  
(A.Hakimov.Ilonizidan)

2) *Dildor unga yugurgan edi (Ufq r. 582-b)*

3) *Ziyodaxon supra yozib, un elab o'tirar edi. (A.Qahhor, qo'shchinor chiroqlari)*

4) *Erta bahor, moychechaklar unib chiqqan payt edi.*

Ushbu jumalarda keltirilgan un so'zi olmosh bo'lib kelgan holini tahlil qiladigan bo'lsak, u olmoshiga *-ga, -da, -dan* qo'shimchalari qo'shilganida *n* orttiriladi. Demak bu turdagi omonimlarni farqlash uchun quyidagi modelni tavsiya qilamiz.

$$H_{N,V,Pr} = \begin{cases} H^N + aff_i^N \\ H^V + aff_j^V \\ \text{unga} \\ H^{Pr} = \text{unda} \oplus W_k^{Pr} \\ \text{undan} \end{cases} \quad (26)$$

$W_k^{Pr}$  —olmosh omonim bilan so'ng tomonidan qo'shilishi mumkin bo'lgan so'zlar to'plami. Bu so'zlar to'plami to'g'ri va to'liq bo'lsa semantik tahlilining aniqlik darajasi shuncha yuqori bo'ladi. Ot va fe'l so'z turkumlaridagi omonim so'zlarni farqlashni yuqorida keltirgan edik.

### 2.1.3.27. Ot ∨ sifat ∨ fe'l so'z turkumlari orasidagi omonimlikni aniqlash

Ot ∨ sifat ∨ fe'l so'z turkumlari orasidagi omonim so'zlardan biri **Oshiq** so'zini misol qilib olamiz.

**Oshiq:** Hayvonning orqa oyog'ida son bilan boldirni tutashtirib turuvchi so'ngakcha (bolalar o'yinida ishlatiladi) (ot)

**Oshiq:** Ortiq, keragidan ko'p (*sifat*)

**Oshiq:** Shoshilmoq (*fe'l*)

Oshiq so'zini korpus ma'lumotlari va badiiy asarlarda kuzatganimizda, quyidagi shakllarda uchratdik.

- 1) Bobom bunga parvo ham qilmasdan, haligi odamni kutib olishga **oshiqdi**. (korpus m-t)
- 2) Umri davomida 40 yildan **oshiq** faoliyati Sharqshunoslik institutida o'tdi. (korpus m-t)
- 3) Mening yoshligimda birga **oshiq** o'ynab, yaxmalak otib o'sgan bir og'aynim bor. (Oybek. Tanlangan asarlar)
- 4) Endi **oshiqning** ikki dunyosi islomiy, oxirati ulug'. (korpus m-t)

Kuzatishlar natijasiga ko'ra quyidagi matematik mulohazani keltiramiz

$$H_{Adj,V} = \begin{cases} \downarrow (N_i + x) \oplus H^{adj} \oplus \downarrow (N_i + x), & x \in \{aff_{s_{ij}}^N\}, i = 1..n \\ \downarrow W + dan \oplus H^{adj} + x, & x \in \{aff_{s_{ij}}^N\}, i = 1..n \\ H^V + x, & x \in \{aff_{s_{ij}}^V\}, j = 1..m \\ H^N + \downarrow x, & x \in \{aff_{s_{ij}}^N\}, i = 1..n \end{cases} \quad (27)$$

Bu yerda  $\downarrow(W+dan)$ - sifat omonimdan oldingi so'zning *-dan* chiqish kelishigi qo'shimchasi bilan birikishi mumkinligini anglatadi.

### 2.1.3.28. Ot $\vee$ fe'l $\vee$ ravish so'z turkumlari orasidagi omonimlikni aniqlash

Ot  $\vee$  fe'l  $\vee$  ravish so'z turkumlari orasidagi omonim so'zlarga *kech* so'zini misol qilamiz

**Kech:** Kechki payt, kechqurun (*ot*)

**Kech:** Belgilangan vaqtdan ancha keyin (*ravish*)

**Kech:** 1. Birin-ketin o'tmoq.

2. Suyuqlik, massa bo'lib harakatlanib o'tmoq.

3. Bahridan o'tmoq, da'vo qilmaslikka ahd qilmoq (*fe'l*)

Bu so'zni jumalalarda quyidagi shakllarda uchratdik.

1) Ular orasida *kech* qolgan abituriyentlar ham ko'pchilikni tashkil qilgan. (korpus m-ti)

2) *Kechga* borib jala quyib tong otguncha momaqaldiroq guldirab chiqdi. (Said Ahmad, Ufq.r.)

3) Bunda sinfdagi doskadan butunlay voz *kechiladi*. (korpus m-ti)

Kuzatuvlar natijasi bizni quyidagi mulohazaga olib keldi.

$$H_{N,Adv,V} = \begin{cases} H^{Adv} + \downarrow aff_i^{adv} \oplus N, & i = 1..n, aff_i^{adv} \in aff^N \\ H^{Adv} + \downarrow aff_i^{adv} \oplus V, & i = 1..n, aff_i^{adv} \in aff^N \\ \downarrow (N + ning) \oplus H^{Adv} + \downarrow i + \downarrow ni \oplus V & \\ H^N + \downarrow aff_j^N & \\ H^V + \downarrow aff_j^V & \end{cases} \quad (28)$$

Bu mulohaza *kech* so'zining morfologik xususiyatlaridan kelib chiqqan holda tavsiya qilindi. Agar bu kabi olmosh so'zlar asos shaklida uchrasa u holda, omonimning birikuvchilaridan foydalanishga to'g'ri keladi.

- *Kech (ot) + kirdi/bo'ldi/usshdi/...*

- *Kech(ravish)+edi/*

### 2.1.3.29. Ot ∨ sifat ∨ predikativ soʻzlar orasidagi omonimlikni aniqlash

Ot ∨ sifat ∨ predikativ soʻzlar orasidagi omonim soʻzlarga misol sifatida *chogʻ* soʻzini keltiramiz.

**Chogʻ:** Vaqt, payt, on (*ot*)

**Chogʻ:** Cama, taxmin (*ot*)

**Chogʻ:** Hajmi kichik (*sifat*)

**Chogʻ:** Yaxshi, koʻtarinki holatda (*predikativ soʻz*)

*Chogʻ* soʻzining jumalalarda uchrashiga namunalar keltiramiz:

- 1) *Beshikdagi qulogʻ ingga kirgan mungli va ayni chogʻ da nurli allami?*
- 2) *Shu chogʻ masjid darchasidan pir etib bir juft qush uchib kirdi-yu, Husanning xayoli boʻlindi.*
- 3) *Hammaning dimogʻ i chogʻ .*
- 4) *Darhaqiqat, birinchi qoʻngʻ iroq — koʻngillarni chogʻ aylaguvchi sado.*

Ot/sifat/predikativ soʻzlar orasidagi omonimlikni ularning oʻng va chap birikuvchilari yordamida farqlashni tavsiya qilamiz. Sifat va ot soʻz turkumidagi soʻzlarga bir xil lugʻaviy va sintaktik shakl yasovchi qoʻshimcha bilan birikishi mumkin. Badiiy asarlarda oʻtkazilgan kuzatuvlar natijasi shuni koʻrsatadiki, bu uchala soʻz turkumiga mansub soʻzlar jumalalarda asos shaklida ham koʻp uchraydi. Xulosa qilib aytganda, bu turdagi omonim soʻzlarni Oʻzbek milliy korpusi maʼlumotlari yordamida semantik tahlil qilinadi.

### 2.1.3.30. Ot ∨ ravish ∨ taqlid soʻzlar orasidagi omonimlikni aniqlash

Ot ∨ ravish ∨ taqlid soʻzlar orasidagi omonim soʻzlarga *vara-q-vaq* soʻzini keltiramiz. Bu soʻz quyidagi maʼnolarga ega:

**VARAQ-VARAQ:** Juda koʻp varaqlardan iborat boʻlgan, varagʻi koʻp (*ot*)

**VARAQ-VARAQ:** ayn. jaraq-jaraq (*ravish*)

**VARAQ-VARAQ:** Qaynayotgan yoki oqayotgan suyuqlik tovushi haqida (*taqlid soʻz*)

Ushbu soʻzni kontekstda uchrashiga namunalar keltiramiz:

- 1) *Aziza maktab yillarida vara-q-vaq insho yozar, kitob oʻqir edi.*
- 2) *Kimyogar qoʻshnimiz bor. Laboratoriyasidagi probirkalardan doim vara-q-vaq tovush eshitiladi.*

Varaq-vaq soʻzi ravish soʻz turkumiga mansub shaklida hozirgi paytda jaraq-jaraq soʻzi bilan almashgan. Taqlid va ravish soʻz turkumlari otlashishini inobatga olib bu

turdagi omonim so'zlarni semantik tahlil qilishni trigramm HMM yordamida amalga oshiramiz.

### 2.1.3.31. Ot V undov V taqlid so'z orasidagi omonimlikni aniqlash

Bu guruhdagi omonimlarga misol keltiramiz.

- 1) **Chuv:** Arava o'qining g'ildirakni tutib turuvchi qulog'i (*ot*)
- 2) **Chuv:** Qimor o'yinida (to'rt oshiq tanlashda) yutqiziqli holat (*ot*)
- 3) **Chuv:** chu, chuh undov. Otni harakatlanishga chaqiruvchi xitob (*undov*)
- 4) **Chuv:** Shovqin-suron tovushlarini bildiradi (*taqlid so'z*)

Yuqorida ot va taqlid so'z orasidagi omonimlikni farqlashni ko'rgan edik. Bu modelga esa undov so'zlarning xususiyatini ham qo'shamiz.

*Cho'pon otlarni chuv-chuv deya daraga haydab ketardi.*

Izlanishlar shuni ko'rsatadiki **chuv** so'zi odatda takror so'zlar ko'rinishida keladi. Morfologik tahlil qiladigan bo'lsak, undov so'zlar hech qanday qo'shimcha qabul qilmaydi. Bunday hollarda so'zning birikuvchilariga tayanamiz. Ya'ni undov omonim so'z bilan birikishi mumkin bo'lgan so'zlar to'plamini shakllantirishga ehtiyoj seziladi. Undan kelib chiqadi bu guruhga mansub omonim so'zlarni faqatgina birikuvchilari bilan farqlashimiz mumkin.

### 2.1.3.32. Ot V sifat Vko'makchi orasidagi omonimlikni aniqlash

Ot V sifat Vko'makchi orasidagi omonim so'zga misol sifatida **sari** so'zini keltiramiz.

- Sari:** 1 Ma'lum tomonga, tomon qarab.  
2 (-gan affiksli sfdsh. bilan) ayn. sayin (*ko'makchi*)
- Sari:** Eng yaxshi, sara (*sifat*)
- Sari:** Sariq, sariqrangli (*sifat*)
- Sari:** Hind ayollarining yubka shaklida belga bir necha bor o'ralib, bir uchi ro'mol qilib boshga tashlab qo'yiladigan kiyimi (*ot*)

Sifat otlashish xususiyatiga ega. Shuning uchun ot V sifat so'z turkumiga doir ekanligini ularning birikuvchilari orqali, qaysi biri ko'makchi so'z ekanligini (16) model orqali aniqlanadi.

**2.1.3.33. Ot V son V fe'l so'z turkumlari orasidagi omonimlikni aniqlash**

O'zbek tilida 3 ta so'z turkumiga doir omonim so'z ham mavjud. Bunday omonimlarni farqlashda har bir turkumdagi so'zning xususiyatlarini inobatga olish lozim.

**Uch:** Biror narsaning boshlanish yoki tugash qismi.

**Uch:** Ikki bilan to'rt orasidagi butun son. (Son)

**Uch:** 1. Havo bo'ylab harakatlanmoq. (fe'l)

2. Biror qirrasini cho'rt uzilib tushmoq.

3. Beixtiyor berilmoq, laqillamoq

Kuzatuvlar natijasida quyidagi qoidalarga ega bo'ldik

*uch+-i/-iga/-ni/-da* ⊕ *bor/ulamoq/boshlamoq/turmoq/...* (ot)

*uch+-inchi/-nchi/-ta/-tacha/-tadan/-larcha/* ⊕ *oy nomlari/dars/uy/...* (son)

*uch+-ov/-ovlon/..* ⊕ *gaplashmoq/ko'rishmoq/kelmoq/ketmoq/...* (son)

*uch+dan* ⊕ *bir/ikki/uch/to'rt/...* (son)

*Balandga/osmonga...* ⊕ *uch* (fe'l)

*Toshkentdan/shahardan...* ⊕ *uch* (fe'l)

*Uch+-ib* ⊕ *kel/ket/...*

{*Uch+-ar* ⊕ *qush/gilam/...*

Yuqoridagi morfologik xususiyatlarni inobatga olib son so'z turkumidagi omonim farqlovchi modelni quyidagicha keltirish mumkin

$$H_{num,N,V} = \begin{cases} H^{num} + \downarrow aff_i^{ord} \oplus \downarrow (W_k^N + \downarrow aff_i^{ord}) \oplus N_j^c \\ \downarrow W_i^{dec} \oplus H^{num} + \downarrow aff_i^{ord} \oplus N_j^c \\ H^{num} + aff_j^{agg} \oplus V \\ H^{num} + dan \oplus W_n^{num} \\ N + x, x \in \{aff_{s_i}^N\}, i = 1..n \\ V + x, x \in \{aff_{s_j}^V\}, j = 1..m \end{cases} \quad (29)$$

Bu yerda  $aff_j^{agg}$  —jamlovchi qo' shimchalar to'plami,  $aff_j^{agg} = \{-ov, -ovlon, -ala, -ovi, -oviga, -ovloni, \dots\}$ ,  $W_n^{num}$  —sonlarning nomlaridan iborat so'zlar to'plami,  $W_n^{num} = \{bir, ikki, uch, \dots\}$



Bu turdagi omonimlarni avtomatik farqlash uchun ularning har biri bilan birikuvchi soʻzlar toʻplamini shakllantirish talab qilinadi. Yaʼni son omonimga tartibni ifodalovchi qoʻshimchalar qoʻshilgan u bilan birikuvchi otlar toʻplami, sonlarning nomlaridan iborat toʻplam, songa qoʻshiluvchi qoʻshimchalar toʻplami kabi maʼlumotlarini toʻplab maʼlumotlar bazasiga kiritish talab qilinadi.

### 2.1.3.34. Otʻ feʼl V taqlid soʻz orasidagi omonimlikni aniqlash

Ot, feʼl va taqlid soʻz orasidagi omonimlikni farqlashda ham ularning morfologik va lugʻaviy xususiyatlariga nazar solamiz.

**Qiy:** Qoʻyning uzoq tepkilanib, qatlanib yotgan goʻngi. (Ot)

**Qiy:** Biror tomonga yotiq holda uzunasiga kesmoq. (feʼl)

**Qiy:** Odatda boʻlishsiz formada keladi. (feʼl)

**Qiy:** Qiy-chuv (tartibsiz baqiriq-chaqiriqlar) (taqlidiy soʻz).

Qiy soʻzi 2 ta feʼl, ot va taqlid soʻz koʻrinishlarida uchraydi. Bunday soʻzlarni kontekstda farqlash uchun badiiy asarlarda naʼmunalar keltirsak.

*Koʻchada qizlarning qiy-chuvi eshitildi (Said Axmad, Ufq, 417-b)*

*Keyin yeyishga koʻzi qiymagandek avaylab ogʻziga soldi (Ufq r. 443-b)*

*Uchi yirtilsa, qiyib chatib oladi-da. (Ufq r. 581-b)*

*Issiqxonadagi egatlarga qiy toʻkib chiqishdi.*

Qiy soʻzi taqlid soʻz boʻlib kelganida asosan qiy-chuv brikma koʻrinishida uchraydi. Undan tashqari tasdiq soʻzlarning morfologik xususiyatiga koʻra otlashadi, demak ot bilan taqlid soʻzlarni bir biridan birikuvchilari orqali farqlaymiz. Bilamizki taqlid soʻzlarning miqdori otga nisbatan kamroq, shu sababli taqlid soʻzlar bilan birikishi mumkin boʻlgan soʻzlarni toʻplash nisbatan osonroq. Feʼl soʻz turkumidagi omonimlarni esa qoʻshimchalar bilan farqlaymiz. Ushbu mulohazalardan kelib chiqib quyidagi matematik modelni tavsiya qilamiz.

$$H_{V,N,imit} = \begin{cases} H^{imit} - \oplus chuv + aff_i^N \oplus W_j^{imit} \\ H^N + aff_i^N \oplus W_k^N \\ H^V + aff_i^V \end{cases} \quad (30)$$

Bu yerda  $W_j^{imit}$  — mana shu turdagi omonim taqlid soʻz boʻlib kelganida u bilan birikishi mumkin boʻlgan soʻzlar toʻplami. Xuddi shunday  $W_k^N$  —shu turdagi omonim ot soʻz turkumiga oid boʻlganida u bilan birikishi mumkin boʻlgan soʻzlar toʻplami. Ot, feʼl va taqlid soʻzlari orasidagi omonimlikni farqlashni mana shu model yordamida amalga oshirishni tavsiya qilamiz.

### 2.1.3.35. Undov $\vee$ feʼl $\vee$ ravish soʻz turkumlari orasidagi omonimlikni aniqlash

Bu guruhga mansub omonim soʻzlarga ura soʻzini misol qilamiz.

Ura: Hu-jum qilayotgan jangchilarning jangovar nidosi; koʻgarinki ruhbilan maʼqullash yoki xursandlikni ifodalovchi nido, xitob (*undov soʻz*)

Ura: Urmoq fl. rvdsh. (*feʼl*)

Ura: Epchilik bilan, tezda (*ravish*)

Ura soʻzini farqlovchi model tuzish uchun morfologik tahlil qilish zarur boʻladi.

- 1) Qayerdadir pulemyot tarilladi: yakka multiq tovushlari, ura sadolari yangradi. (*A. Qahhor, Oltin yulduz*) (*undov soʻz*)
- 2) Demak, ishlarni shunchaki “qogʻozda”, “ura-ura”chilik asosida bajarish bilan ish bitmaydi. (*korpuz m-ti*) (*undov soʻz*)
- 3) Koʻchalardan birining nomini eshitganida yigitning tomiri yana tez ura boshlabdi. (*korpuz m-ti*) (*feʼl*)
- 4) Bola bilan men qopqon tomon ura qochdik. (*N. Safarov, Olovli izlar*) (*ravish*)
- 5) Nigora maosh olishi bilan ura solib uyiga joʻnaydi. (*P. Tursun, Oʻqituvchi*) (*ravish*)

Keltirilgan jumalardan kelib chiqib quyidagi matematik mulohazani keltiramiz

$$H_{E,V,Adv} = \begin{cases} H^{Adv} \oplus \begin{cases} solmoq \\ qochmoq \end{cases} \\ H^V \oplus V \setminus \{qochmoq, solmoq\} \\ H^E + \downarrow aff^N \\ H^E, \oplus H^E \end{cases} \quad (31)$$

Ushbu matematik mulohaza yordamida undov soʻz, feʼl va ravish soʻz turkumlari orasidagi omonimlikni farqlashni tavsiya qilamiz.

### 2.1.3.36. Ot $\vee$ sifat $\vee$ feʼl $\vee$ ravish soʻz turkumlari orasidagi omonimlikni aniqlash

Ot  $\vee$  sifat  $\vee$  feʼl  $\vee$  ravish soʻz turkumlari orasidagi omonim soʻzga *quv* soʻzini keltiramiz.

**Quy:** eskirgan. Oqqush. (*Ot*)

**Quy:** Mug'ombir, ayyor (*sifat*)

**Quy:** Butunlay (*ravish*)

**Quy:** 1. Haydamoq.

2. Qochgan kishini ketma-ket ta'qib etib bormoq (*fe'l*)

*Quy* so'zini korpus ma'lumotlari, badiiy asarlardan izlaganimizda ot va ravish so'z

turkumlari sifatida deyarli kuzatilmadi. Garchi shunday bo'lsada ot, sifat, fe'l,

ravish so'z turkumlarining morfologik xususiyatlaridan kelib chiqib quyidagi

matematik modelni tavsiya qilamiz.

$$H_{N,Adj,V,Adv} = \begin{cases} H^{Adv} + \downarrow aff_i^{adv} \oplus N, & i = 1..n, aff_i^{adv} \in aff^N \\ H^{Adv} + \downarrow aff_i^{adv} \oplus V, & i = 1..n, aff_i^{adv} \in aff^N \\ \downarrow (N + ning) \oplus H^{Adv} + \downarrow i + \downarrow ni \oplus V \\ H^N + \downarrow aff_j^N \\ H^V + \downarrow aff_i^V \\ \downarrow (N + x) \oplus H^{Adj} + \downarrow aff^{Adj} \oplus \downarrow (N + x), x \in \{aff^N\} \end{cases} \quad (32)$$

### 2.1.3.37. Ot $\vee$ undov $\vee$ fe'l $\vee$ taqlid so'z orasidagi omonimlikni aniqlash

Ot  $\vee$  undov  $\vee$  fe'l  $\vee$  taqlid so'z orasidagi omonim so'zga misol sifatida quyidagi qur so'zini keltiramiz.

**Qur:** Payt, vaqt (*ot*)

**Qur:** Tovuqqa o'xshash katta qora (yoki ola) patli qush, qoraqur (*ot*)

**Qur:** Yasamoq (*fe'l*)

**Qur:** Qo'yarni haydashda ishlatiladi (*undov so'z*)

**Qur:** Baqa chiqaradigan tovushni bildiradi (*taqlidiy so'z*)

Ushbu turdagi omonim so'zlarni farqlashda (2.1.3.32) da keltirilgan mulohaza va (13) matematik modeldan foydalanish maqsadga muvofiq. O'zbek tilidagi omonim

so'zlarni semantik tahlil qilish uchun ma'lumotlar bazasini shakllantirib yuqorida

keltirilgan matematik modellar asosida tuzilgan algoritmlar asosida amalga

oshiriladi.

## 2.2. Kontekstda polisemiyanı bartaraf etishning matematik modellar

“Polisemiya har qanday tilda ham salmoqli o‘rin tutadi. Til boyligi faqat so‘zlar, iboralar bilangina emas, so‘zlarning leksik ma‘nolari bilan ham o‘lchanadi<sup>49</sup>” deya ta‘riflaydi M.Mirtojiyev. Ko‘p ma‘noli so‘zlarda ma‘no ko‘chish usullari kuzatiladi. Polisemantik so‘zlarni semantik tahlil qiluvchi tizim yaratish uchun tahlil qilish omillari, matematik modellari zarur. Dunyoning qator tillarida polisemantik so‘zlarni farqlovchi omillar ishlab chiqilgan. Rus olimlari T.Кикинеджи, П. А Соболева, Н.М.Меркурьева, А.А. Кожанов И.К. Архипов, И. Горбушина kabi ko‘plab olimlar ilmiy izlanishlar olib bormoqda.<sup>50</sup> Hozirgacha omonimiya va polisemiyanı farqlash masalasi munozarali bo‘lib qolmoqda. Polisemiyanı tizimli semantik tahlil qilish masalasi ustida taqiqotlar davom etmoqda [Beleeva, 2005].

Turkiy tillar doirasida ko‘p ma‘nolilik muammosi ko‘rib chiqilmagan desak mubolag‘a bo‘ladi. Turkiy tillardagi leksik polisemiya hodisasi haqida ba‘zi hollarda tadqiqotchilar qarama-qarshi fikrlar bildiradilar. Shunday qilib, turkiyshunoslikda ko‘p ma‘nolilik masalalari asarlarda o‘rganilgan: o‘zbek tilida – M.Mirtojiyev [1984], qorachay-bolgar tilida S.R.Bidjjeva [2001], F.K.Appaeva [2002], Z.X.Tramova [2006], qozoq tilida - K. Axanov [1958], R. Sodiqbekov [1973], T. Aliqulov [1966], qirg‘iz R. Bekjanova [1967], boshqird- M.X.Axtyamov [1966], tatarcha - I.B.Bashirova [2006], umumiy turkiy tillar miqyosida esa - A.A.Yo‘ldoshev [1972] va J.M.Guzeev [1985]. Bir qator turkiy tillarda so‘zning leksik o‘zgarishi muammosiga oid tadqiqotlar paydo bo‘ldi, masalan, qozoq tilida I.K.Uyuqboyev

<sup>49</sup> Miraziz Mirtojiyev. Ozbek tilida polisemiya // Toshkent fan nashriyoti. 1984;

<sup>50</sup> Кикинеджи Т. Г. Особенности формирования полисемии сложных слов // Словосочетание, словосложение и аффиксация как способы словообразования: сб. ст. Владивосток: Изд-во ДВНЦ АН СССР, 1990. С. 82-94; Соболева П. А. Словообразовательная полисемия и омонимия. М.: Наука, 1980. 294; Меркурьева Наталия Михайловна. Полисемия и омонимия композитов. Лексикографический аспект // Филологические науки. Вопросы теории и практики Тамбов: Грамота, 2017. № 7(73); в 3-х ч. Ч. 3. С. 146-148. ISSN 1997-2911.; А.А. Кожанов. Неоднозначность: полисемия и омонимия в юридической терминологии (на материале немецкого языка) // аспирант Брянского государственного университета имени академика И.Г. Петровского; Архипов, И.К. Новые тенденции анализа лексической полисемии // Вестник Московского государственного лингвистического университета. – 2008. – №544. – С.23-32.; Горбушина И. Торить ‘прокладывать дорогу’ и торить ‘учить’ — омонимия или полисемия?

[1976]; Qirg'iz tilida Q. Seydakmatov [1982]; Turkman tilida A. Geldimurodov [1974]. So'zning semantik tuzilishi muammosiga, xususan, polisemiya bag'ishlangan turkologik adabiyotlar tahlili shuni ko'rsatadiki, uning ko'p jihatlari yetarlicha o'rganilmagan va bahsli bo'lib qolmoqda.

Jahon tajribasi shuni ko'rsatadiki, ko'p ma'nolilikni o'rganishga oid nazariy qoidalarni tizimlashtirish va umumlashtirish, har bir tabiiy tillardagi so'zlarning ko'p ma'noliligining o'ziga xos xususiyatlarini aniqlash maqsadga muvofiq. Ko'plab tadqiqodlar tildagi ko'p ma'noli so'zlarni turkumlar doirasida farqlashning o'ziga xos xususiyatlarini o'rganishga bag'ishlangan. Masalan, tildagi fe'l so'z turkumi oid polisemantik so'zlar [Султанова Алина Петровна, 2008// Полисемия глаголов деструкции в русском и английском языках.], ma'lum bir sohadagi ko'p ma'noli so'zlar [И.А. Аграновская, Полисемия термина «экология»], huquqiy sohada [А.А. Кожанов, Неоднозначность: полисемия и омонимия в юридической терминологии (на материале немецкого языка)].

O'zbek tilida so'zlar orasidagi, qo'shimchalar orasidagi ko'p ma'nolilik haqida M. Mirtojdiyevning "O'zbek tilida polisemiya" kitobida batafsil ma'lumot berilgan. M.Mirtojdiyev ko'p ma'noli so'zlarning vujudga kelishini quyidagicha xarakterlaydi:

1. O'zbek tilidagi polisemantik so'zlar yo ko'chma ma'no, yo evfemistik, disfemistik ma'no, yo yasama ma'no hisobiga sof holatda tarkib topishi ham mumkin
2. Ular ham hosil bo'lgan, ham yasalgan ma'nolardan aralash holatda ham tarkib topaveradi.
3. Polisemantik so'zlar vujudga kelishida ko'chma ma'nolar asosiy o'rinda turadi. Ko'pincha evfemistik vadisfemistik ma'nolar ham ularga teng bo'ladi.
4. Yasama ma'nolar hamma vaqt leksik bo'lib, polisemantik so'zlarni so'zsiz vujudga keltiradi. Ko'chma, evfemistik va disfemistik ma'nolarda bunday deb bo'lmaydi.

Polisemantik va omonim so'zlar bir biriga yaqin so'zlar. Shu bois ham omonimiya va polisemiya hodisalarini o'rganish masalasi hozirgacha dolzarbligini

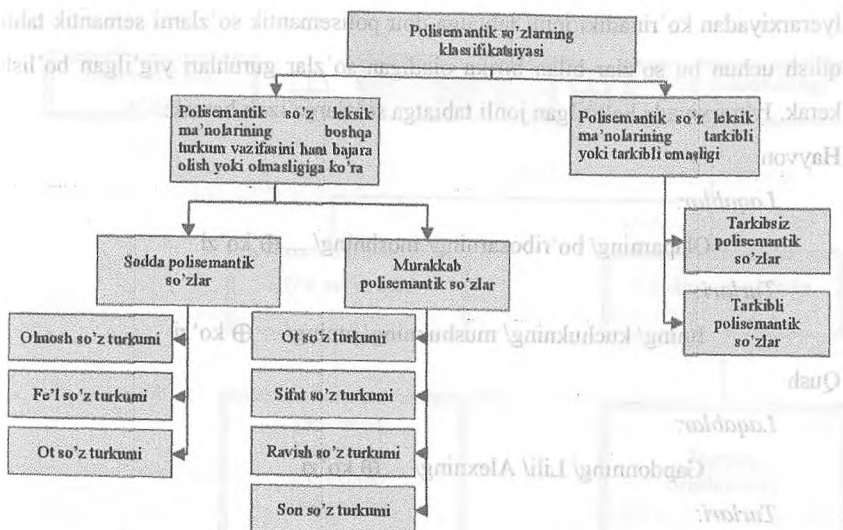
yo'qotmagan. Ularning kichik farqli xususiyati sifatida shuni aytishimiz mumkinki omonim tor ma'nodagi tushuncha, polisemiya esa keng ma'nodagi tushuncha ekanligini keltirish mumkin. Tilshunoslar tomonidan o'zbek tilida omonim va polisemantik so'zlarni farqlashning 8 ta jihati borligi aytib o'tilgan.<sup>51</sup> Ularni farqlovchi jihatlardan kelib chiqib quyidagi jadvalni keltiramiz.

Xususiyat	Polisemantik so'zlar	Omonim so'zlar
Semantik farqi	Kontekstda	Ma'no va etimologiyasiga ko'ra
Kontekstdagi roli	Muhim	Muhim bo'lmasligi mumkin

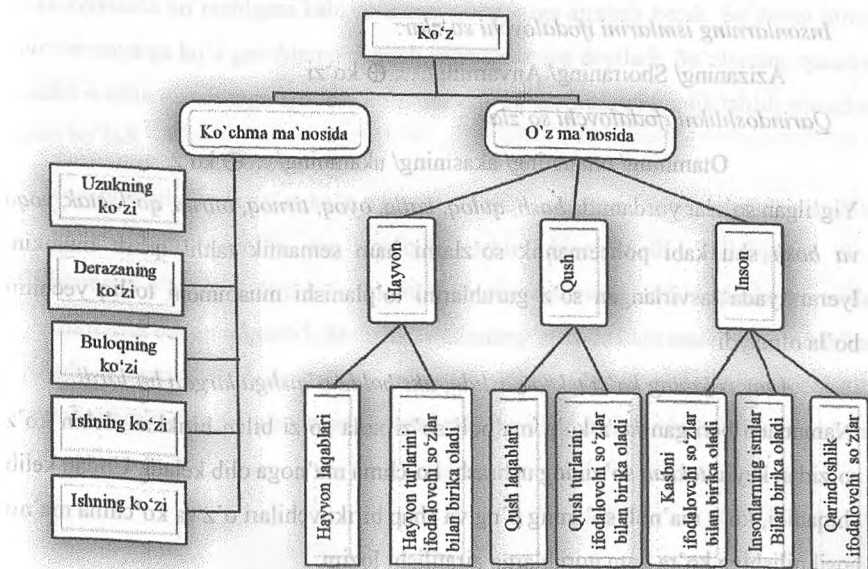
Demak, polisemantik so'zlarni farqlashni kontekst asosida amalga oshirish mumkin. Dunyo bo'yicha eng ko'p muloqot qilinadigan til bu xitoy tili hisoblanadi. Bu til polisemantik so'zlardan holi emas. Xitoy lug'atining 14% ko'p ma'noli so'zlar bo'lib, bu so'zlar tez-tez ishlatiladigan so'zlardir. Xitoy tili korpusidagi ma'lumotlarning 42% ini polisemantik so'zlar tashkil qiladi. Ko'p ma'noli so'zlar dunyoning barcha tabiiy tillarida o'zining ko'rsatkichlariga ega. Jahon tajribasiga ko'ra, ko'p ma'noli so'zlar ehtimollar nazariyasiga asoslangan holda semantik tahlil qilinadi. Kontekst asosida farqlash uchun ham albatta mutaxassislarning qo'l mehnati zarur. Ko'p ma'noli so'zlarni kontekstda farqlovchi modellarni ishlab chiqish uchun ularni guruhlariga ajratiladi. Ajratilgan guruhlarining o'ziga xos xususiyatlari o'rganiladi. O'rganilgan xususiyatlar asosida qoidalar ishlab chiqiladi. Polisemantik so'zlar quyidagicha klassifikatsiya qilinadi

<sup>51</sup> Hamroyeva M. Omonimlar va polisemantik so'zlarning fonetik mutanosibligi va semantik tafovutlari// 2021





Tadqiqot ishida polisemantik so'zlarni so'z turkumlariga ajratgan holda farqlashni tavsiya qilmoqchimiz. Masalan, *ko'z* so'zining quyidagi iyerarxiyasini keltiramiz.



Iyerarxiyadan ko'rinadiki jonli tabiatga doir polisemantik so'zlarni semantik tahlil qilish uchun bu so'zlar bilan birika oladigan so'zlar guruhlarini yig'ilgan bo'lishi kerak. Iyerarxiyada keltirilgan jonli tabiatga so'zlarga izoh bersak:

Hayvon:

*Laqablar:*

Olaparning/ bo'ribosarning/ moshning/ ... ⊕ ko'zi

*Turlari:*

Itning/ kuchukning/ mushukning/ otning/... ⊕ ko'zi

Qush:

*Laqablar:*

Gapdonning/ Lili/ Alexning/ ... ⊕ ko'zi

*Turlari:*

lochinning/ burgutning/ to'tining/ ... ⊕ ko'zi,

Inson:

*Kasbni ifodalovchi so'zlar:*

o'qituvchining/ shifokorning/ tadbirkorning/ ... ⊕ ko'zi

*Insonlarning ismlarini ifodalovchi so'zlar:*

Azizaning/ Shoiraning/ Anvarning/.....⊕ ko'zi

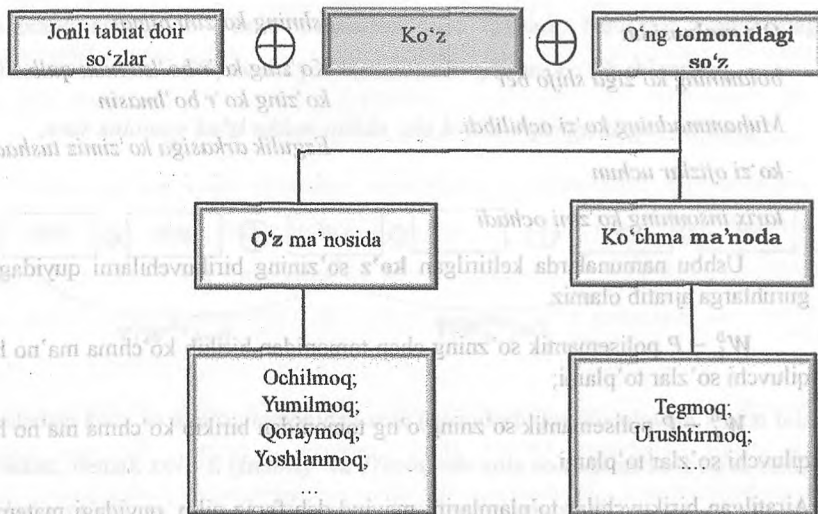
*Qarindoshlikni ifodalovchi so'zlar:*

Otamning/ onamning/ akasining/ ukamning/ ... ⊕ ko'z

Yig'ilgan so'zlar yordamida *bosh, quloq, kalla, oyoq, tirnoq, burun, qo'l, etak, yoqa va bosh* shu kabi polisemantik so'zlarni ham semantik tahlil qilish mumkin. Iyerarxiyada tasvirlangan so'z guruhlarini to'planishi muammoni to'liq yechimi bo'la olmaydi.

*Anor xolaning ko'zi tekkan aslida, aks holda o'qishga kirgan bo'lardi.*

Namunada berilgan *ko'z ko'p ma'noli so'zi xola so'zi bilan birikkan, lekin ko'z so'zidan keyin tekkan so'zining uchrashi ko'chma ma'noga olib keladi. Undan kelib chiqadiki, ko'p ma'noli so'zning o'ng va chap birikuvchilari o'z va ko'chma ma'no hosil qilishiga ko'ra ham guruhlariga ajratilishi lozim.*



Ko'z so'zining o'z va ko'chma ma'nolarini hosil qiluvchi birikuvchilari tasvirlangan ierarxiyalarda ko'rsatilgani kabi so'zlarni guruhlariga ajratish kerak. So'zlarni biror bir xususiyatga ko'z guruhlariga ajratish *segmentatsiya* deyiladi. So'zlarning qancha kichik-kichik guruhlariga ajratilsa ko'p ma'noli so'zlarning semantik tahlili shuncha aniq bo'ladi.

Xorijiy tajribalarga ko'ra ko'p ma'noli so'zlarni semantik tahlil qilishda statistik metodlardan foydalanish eng ma'qul yechim hisoblanadi. Tadqiqotchi Sh. Gulyamova o'zbek tilidagi ko'p ma'noli so'zlarni semantik tahlil qilishi borasida o'z fikrlarini bayon qilgan<sup>52</sup>. Statistik usullarning asosida ham matematik modellar yotadi. Yuqorida keltirilgan ko'z so'zi misolidagi klassifikatsiyalarga asoslanib mulohaza qilaylik. Ko'z so'zining kontekstda quyidagi birikuvchilar bilan uchraydi.

<sup>52</sup> Sh.K.Gulyamova. Polisemiya lingvistik xodisa sifatida// "O'zbekistondagi amaliy tadqiqotlar" mavzusidagi online konferensiya

**Ko'chma ma'nosida:**

**O'z ma'nosida:**

*bolamning ko'ziga shifo ber*

*Muhammadning ko'zi ochilibdi.*

*ko'zi ojizlar uchun*

*tarix insonning ko'zini ochadi*

**Ishning ko'zini biladi**

*Ko'zing ko'r bo'lsa ham qalb*

*ko'zing ko'r bo'lmasin*

*Ezgulik arkasiga ko'zimiz tushadi*

Ushbu namunalarda keltirilgan ko'z so'zining birikuvchilarni quyidagicha guruhlarga ajratib olamiz.

$W_i^b - P$  polisemantik so'zning chap tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami;

$W_j^a - P$  polisemantik so'zning o'ng tomonidan birikib ko'chma ma'no hosil qiluvchi so'zlar to'plami.

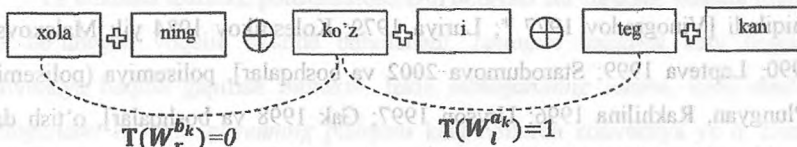
Ajratilgan birikuvchilar to'plamlarini mavjud deb faraz qilib, quyidagi matematik mulohazani keltiramiz

$$W_r^{bk} + \downarrow aff^{sl} \oplus P_k + \downarrow aff^{sl} \oplus W_l^{ak} + \downarrow aff^{sl} = \begin{cases} P_k^s, T(W_r^{bk}) \vee T(W_l^{ak}) = 0 \\ P_k^p, T(W_r^{bk}) \vee T(W_l^{ak}) = 1 \end{cases} \quad (33)$$

Bu yerda  $P_k$ - polisemantik so'z;  $aff^{sl}$  - sintaktik va lug'aviy shakl yasovchi qo'shimchalar to'plami;  $W_r^{bk}, W_l^{ak} - P_k$  so'zning o'ng va chap (oldin (*before*) va keyin (*after*)) birikuvchilari;  $P_k^s - P_k$  polisemantik so'zning o'z (self meaning),  $P_k^p$  - ko'chma (portable) ma'nosida ekanligini bildiradi;  $T(W_r^{bk}), T(W_l^{ak}) - W_r^{bk}, W_l^{ak}$  so'zlarni mos ravishda  $W_i^b, W_j^a$  to'plamlar tarkibida mavjud yoki mavjud emasligini ifodalovchi Boolean tipidagi qiymatlar;  $\vee$  - dizyunksiya (mantiqiy qo'shish) amali, dizyunksiya *yoki* degan ma'noni anglatadi.  $\vee$ - dizyunksiya belgisini qo'llanilishining sababi shundaki, gapdagi polisemantik so'z o'z (asl) ma'no yoki ko'chma ma'no kasb etadi. Bir vaqtda ikkalasi ham bo'lishi mumkin emas.

Ushbu mulohazaning aniqlik darajasi yuqori bo'lishi uchun juda ko'p mutaxassislarning qo'li, ko'z mehnatlari talab qilinadi. Navbatda kontekstdagi polisemantik so'zni (33) model yordamida tahlil qilishni ko'rib chiqamiz.

*Anor xolaning ko'zi tekkan aslida, aks holda o'qishga kirgan bo'lardi*



Jumladagi *ko'z* so'zi o'ng tomonidan *xola* (qarindoshlikni ifodalovchi) so'zi bilan birikkan, demak  $xola \in \{inson\}$  va  $T(xola)=0$ - *xola* so'zi bilan *ko'z* so'zi birikib bu ko'p ma'noli so'z o'z (ko'rish organi) ma'nosini anglatadi;

Chap tomondagi *teg* so'zi ko'zning funksiyasiga tegishli emas va  $T(teg)=1$ . Xulosa qiladigan bo'lsak

$$T(xola) \vee T(teg) = 1$$

$$0 \vee 1 = 1$$

Keltirilgan gapdagi *ko'z* so'zining (33) matematik model yordamida qilingan semantik analizi 100% aniq natija berdi. To'g'ri, bu model ixtiyoriy kontekstdagi ko'p ma'noli so'zlarni to'liq semantik tahlil qilmasligi mumkin, ammo bu model o'zbek tilidagi polisemantik so'zlarni morfologik va sintaktik xususiyatlariga asoslangan desak mubolag'a bo'lmaydi.

### 2.3. Kontekstda polifunksional so'zlarni farqlovchi matematik modeli.

Leksik birliklarning polifunksionalli muammosi zamonaviy tilshunoslikning dolzarb muammolaridan biridir. Polifunksional so'z masalasi jahon tilshunosligida rus, ingliz, tatar, xitoy tillarida ko'p uchraydi. Polifunksionallik tushunchasi ba'zi manbalarda ko'p funksiyalilik ko'rinishida uchraydi. Ko'pgina xorijiy lingvistik asarlarda ko'p funksiyalilikni omonimiya hodisalari bilan bog'liq holda ko'rib chiqiladi [Vinogradov 1977 \*; Luriya 1979; Kolesnikov 1984 yil; Malaxovskiy 1990; Lapteva 1999; Starodumova 2002 va boshqalar], polisemiya (polisemiya) [Plungyan, Rakhilina 1996; Uryson 1997; Gak 1998 va boshqalar], o'tish davri [Peshkovskiy 1956; Migirin 1971; Tixomirova 1973 yil; Prokopovich 1974 yil; Rogojnikova 1984 yil; Vinogradov 1986; Babaytseva 2000; Sergeeva 2000 va boshqalar]. Tilshunos Gorina Irina Ivanovna o'zining "Лексема словна полифункциональное слова современного русского языка" maqolasida "Ko'p funksionallik gapning bir o'zgarma qismining boshqasiga o'tish qobiliyatidir" deya ta'rif beradi. I.I.Gorina o'zining izlanishlarida rus tilidagi *словна* so'zining boshqa polifunksionallik ususiyatlari haqida batafsil ma'lumotlar keltirgan<sup>53</sup>. Xitoy tilshunoslaridan Syun Xun o'zining nomzodlik dissertatsiyalarida rus tili va rus-xitoy lug'atidagi polifunksional so'zlarni xitoycha ekvivalenti topilib ularning ma'nolari tushuntirilgan<sup>54</sup>. Ingliz tilida polifunksional so'zlarga olmoshlarni misol qilish mumkin. О.Д.Вишнякова *-able* modal so'zining polifunksional so'z sifatida qarash mumkinligini izohlar yordamida dalillagan<sup>55</sup>. Е.Н. Виноградова rus tilida mavjud predloglar orasidan ko'p funksionalliklarini ajratgan<sup>56</sup> va ularni guruhlariga bo'lib chiqqan.

Turkiy tillar oilasiga mansub manbalarni kuzatganimizda tarar tili korpusida Polifunksional so'z masalasi jahon tilshunosligida to'liq yechim topmagan bo'lib,

<sup>53</sup> Горина Ирина Ивановна. Лексема словна полифункциональное слова современного русского языка// Инновации и инвестиции №-4, 2014

<sup>54</sup> Syun Xun Полифункциональные слова в русском языке как проблема русско-китайских словарей : Дис. ... канд. филол. наук: 10.02.20: Москва, 2003 196 с. РГБ ОД, 61:04-10/225-9

<sup>55</sup> О.Д.Вишнякова. СУЩЕСТВУЕТ ЛИ "МОДАЛЬНОЕ ПРИЧАСТИЕ" В АНГЛИЙСКОМ ЯЗЫКЕ?// Вестн. Моск. ун-та. Сер. 19. Лингвистика и межкультурная коммуникация. 2013. № 3

<sup>56</sup> Е.Н. Виноградова. Вплотную к наречным предлогам (квопросу о полифункциональных единицах)// РУССКИЙ ЯЗЫК ЗА РУБЕЖОМ № 1/2021



uning avtomatik aniqlanishi, korpusda bartaraf etilishi, lingvistik filtri va modeli borasida biror tadqiqot uchratmadik. Mazkur faslda o'zbek tilida mavjud nazariy qarashlarni umumlashtirib, ulardan foydalanib, semantik analizator uchun polifunksional so'zlarni modellashtirish borasidagi dastlabki ishlarni bajarishga harakat qilindi.

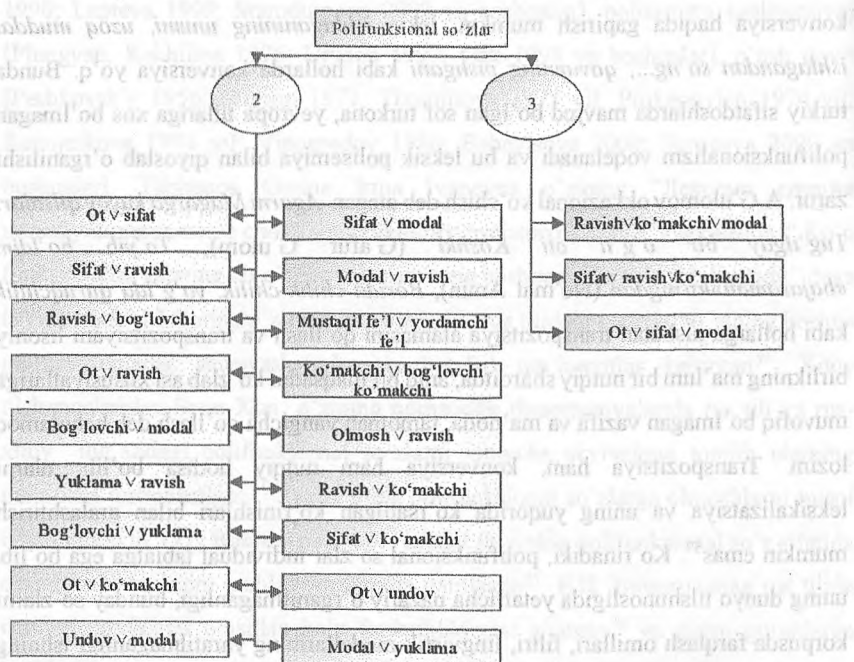
Ta'kidlash lozimki, polifunksionalizm hodisasi sof turkona, boshqa tillarga xos bo'lmagan voqelik sifatida baholanadi. *Ishlagan tishlaydi* kabi hollarda konversiya haqida gapirish mumkin, lekin *ishlaganining unumi, uzoq muddat ishlagandan so'ng...*, *qovunning pishgani* kabi hollarda konversiya yo'q. Bunda turkiy sifatdoshlarda mavjud bo'lgan sof turkona, yevropa tillariga xos bo'lmagan polifunksionalizm voqelanadi va bu leksik polisemiya bilan qiyoslab o'rganilishi zarur. A. G'ulomov okkazional ko'chish deb atagan *Agarni Magarga kuyov qilsalar, Tug'ilgay bir o'g'il oti Koshki* (G'afur G'ulom), *To'yib bo'ldim «bajaraman»laringdan* (Ne'mat Amin), *Borida chilik-chilik, yo'g'ida quruqchilik* kabi hollarga nisbatan transpozitsiya atamasini qo'llash va transpozitsiyani lisoniy birlikning ma'lum bir nutqiy sharoitda, aniq bir maqsadni ko'zlab asl xususiyatlariga muvofiq bo'lmagan vazifa va ma'noda, tamoman yangicha qo'llash deb baholamoq lozim. Transpozitsiya ham, konversiya ham nutqiy hodisa bo'lib, ularni leksikalizatsiya va uning yuqorida ko'rsatilgan ko'rinishlari bilan aralashtirish mumkin emas<sup>57</sup>. Ko'rinadiki, polifunksional so'zlar individual tabiatga ega bo'lib, uning dunyo tilshunosligida yetarlicha nazariy o'rganilmaganligi, bunday so'zlarni korpusda farqlash omillari, filtri, lingvistik modellarining yaratilmaganligi ishning dolzarbligini namoyon etadi.

Polifunksional so'zlarni modellashtirish uchun filtrlar ish beradi<sup>58</sup>. Yaratilgan filtrlar va lingvistik modellarni asoslanib matematik modellar yaratish muhim. Tadqiqotchi Sh. K. Gulyamova "Polifunksional so'zlarni lingvistik modellashtirishni

<sup>57</sup> Sayfullayeva R. Hozirgi o'zbek adabiy tili. – Toshkent, 2010. – B. 153.

<sup>58</sup> Гулямова Ш. Полифункционал ва кўп маъноли сўزلарни фарқлашда локализация ҳодисаси // "O'zbek amaliy tilshunosligi va lingvodidaktika masalalari" mavzusidagi respublika ilmiy-amaliy konferensiya materiallari to'plami. – Toshkent, 2021.

turkumlar kesimida aniqlab olish kerak”,–deydi<sup>59</sup>. Tadqiqot ishining mazkur qismida o‘zbek tilida mavjud nazariy qarashlarni umumlashtirib, ulardan foydalanib, semantik analizator uchun polifunksional so‘zlarni modellashtirish borasidagi dastlabki ishlarni bajarishga harakat qilindi. Tadqiqotchi Sh. Gulyamova polifunksional so‘zlarni so‘z turkumlar kesimida ja‘mi 21 ta guruhini aniqlagan<sup>60</sup>. Mazkur tadqiqot ishida Sh. Gulyamovaning yondashuvlariga asoslaniladi.



<sup>59</sup> Гулямова Ш. К. The problem of polyfunctionality in the traditional and computational linguistics // Middle European scientific bulletin. ISSN 2694-9970. Journal Impact Factor = 5.985. – 104-107 p

<sup>60</sup> Гулямова Ш. Полифункционал ва кўп маъноли сўзларни фарқлашда локализация ҳодисаси // “O‘zbek amaliy tilshunosligi va lingvodidaktika masalalari” mavzusidagi respublika ilmiy-amaliy konferensiya materiallari to‘plami. – Toshkent, 2021.

Har bir turkum bo'yicha polifunksional so'zni farqlashning matematik modelini ishlab chiqish kerak. Polifunksional so'zlar ham omonimlarga o'xshab birikuvchilar asosida farqlanishi mumkin. Ammo faqat bu metodning o'zi barcha polifunksional so'zlarni farqlab, model tuzish uchun yetarli emas. Chunki ba'zi so'zlar turkumi kontekstda aniqlanadi. Garchi shunday bo'lsa-da, polifunksionallikni farqlovchi tizim uchun aniq qonuniyatlar, modellar kerak bo'ladi. Sh.Gulyamova turkumlar kesimida klassifikatsiya qilganida, har bir guruh polifunksional so'zlari uchun lingvistik model ishlab chiqqan. Yaratilgan lingvistik modellar asosida tizim uchun qonuniyat, matematik modellar ishlab chiqamiz. Umimiy polifunksional so'zlar to'plamini  $Pf$  kabi belgilash kiritamiz.

### 2.3.1. Ot-sifat so'z turkumlari orasidagi polifunksionallikni aniqlash

ot-sifat orasidagi polifunksionallikni hosil qiluvchi so'zga misol sifatida *kasal* so'zini olamiz. *Kasal* so'zini kontesktida qay holatda ot va qay holatlarda sifat so'z turkumida doir bo'lishiga na'munalar keltiramiz

1. «O'g'lingizning *kasali* ma'lum bo'ldi. – debdi Ibn Sino

2. *Bahonani oshirishmi yoki kasalni yashirish?!*

3. *Qo'shim kasal sigirini yetaklab qassobxonaga yo'l oldi*

4. *Palataga og'ir kasal bemor joylashtirildi*

1,2- gaplarda keltirilgan *kasal* so'zi kasallik ma'nosi bo'lib ot so'z turkumiga mansub. 3 va 4 esa *kasal* so'zi o'zidan keying otni sifati ifodalaydi.

– *Kasal*+-i/ni/ga/lar/da/ini/... ⊕ aniqlandi/ko'rmoq/keldi/...

– *Kasal* ⊕ odam/bemor/hayvon/qush/...

Ushbu tahlillarga asoslanib xulosa qilamiz

$$Pf_{adj,N} = \begin{cases} Pf^N + \downarrow aff^N \oplus V \\ Pf_{adj} + \downarrow aff^{deg} \oplus N \end{cases} \quad (34)$$

Bu yerda  $Pf_{adj,N}$  –sifat- ot orasidagi polifunksionallik hodisasi,  $Pf^N$  –ot so'z turkumidagi polifunksional so'z,  $Pf^{adj}$  –sifat so'z turkumidagi polifunksional so'z.

(34) yordamida ot-sifat orasidagi polifunksionallikni aniqlashni taklif qilamiz.

### 2.3.2. Sifat – modal soʻz turkumlari orasidagi polifunksionallikni aniqlash

Sifat – modal soʻz turkumlari orasidagi polifunksionallikni aniqlovchi matematik modelni *aniq* soʻzi yordamida aniqlab xulosa qilamiz.

*Aniq – sifat:*

1. Saodatxon orziqib javob kutar, Tojiboy aka bo'lsa, *aniq* javob bermay, gapni aylantirar edi. (S.Zunnunova, Olov).
2. Ko'pchilik tomonidan ma'qullangan amaliy, *aniq* takliflar kiritildi. (Gazetadan)

*Modal soʻz:*

1. Xolmirza aka ichidan «shu g'ayrating bo'lsa, ikki yilda Xirmontepadan *aniq* ayirilamiz», deb o'ylaydi. (A.Qodiriy, G'irvonlik Mallavoy).
2. Endi kimga og'iz solsang, g'iring demay tegishi *aniq*. (S. Ahmad, Yulduz)

*Aniq – modal soʻz shubhasiz, shaksiz soʻzlari bilan ma'nodosh.*

*Jumlalardan ko'rinadiki, polifunksional soʻzdan keyin ikkala holda ham fe'1, oldin ot soʻz turkumidagi soʻz uchrashi mumkin. Bundan kelib chiqadiki, bu turdagi soʻzlar uchun birikuvchilar ro'yhatini shakllantirish orqali sifat- modal soʻzlar orasidagi polifunksionallikni farqlash mumkin.*

$$Pf_{M,adj} = \begin{cases} Pf^{adj} \downarrow aff^{adj} \oplus W^{Pf-adj} \\ W^M \oplus Pf^M \oplus W^M \end{cases} \quad (35)$$

$Pf_{M,adj}$ -modal soʻz va sifat soʻz turkumi orasidagi polifunksionallik,  $Pf^{adj}$ -sifat polifunksional soʻz,  $W^{Pf-adj}$ -sifat polifunksional soʻzning birikuvchilaridan iborat chekli to'plam,  $W^M$  –modal polifunksional soʻzlarning birikuvchilari,  $Pf^M$  –modal polifunksional soʻz.

### 2.3.3. Sifat –ravish soʻz turkumlari orasidagi polifunksionallikni aniqlash

Sifat–ravish soʻz turkumlari orasidagi polifunksional soʻzlarga *boq* soʻzini keltiramiz.

*Boq – fe'1.1.Nazar solmoq, qaramoq. 2.Uzoq vaqt oziq-ovqat berib parvarish qilmoq*

Boq – ko'makchi fe'l. Mustaqil fe'l anglatgan leksik ma'noga "harakatning sinash, tekshirish maqsadida bajarilishi" Grammatik ma'nosini qo'shadi<sup>61</sup>.

Bolani	Boq
menga, unga, bizga	

O'qib	Boq
Yeb	
Sakrab	
Kulib	

Yuqoridagi tahlilni inobatga olib quyidagi matematik modelni keltiramiz

$$Pf_{i,L} = \begin{cases} W^x + aff^{CS} \oplus Pf^I \\ V^L + aff^L \oplus Pf^L \end{cases} \quad (36)$$

$Pf_{i,L}$  – yetakchi (leader) va mustaqil (independent) fe'l orasidagi polifunksional so'z,

$W^x$  – ot, ravish, ravish xarakteridagi sifat, son, olmosh so'z turkumidagi so'zlardan iborat chekli to'plam,  $V^L$  – yetakchi fe'llardan iborat chekli to'plam,  $aff^{LI}$  – fe'lga qo'shilib yetakchi fe'l hosil qiluvchi qo'shimchalar to'plami,  $aff^L =$

$\{-a, -ay, -y, -b, -ib\}$ ,  $Pf^L$  – yetakchi fe'l so'z turkumidagi polifunksional

so'z. Ushbu model orqali yetakchi va ko'makchi fe'l bo'lib kela oladigan

polifunksional so'zlarni farqlash mumkin.

### 2.3.4. Sifat – ravish so'z turkumlari orasidagi polifunksionallikni aniqlash

Sifat – ravish so'z turkumlari orasidagi polifunksionallikni hosil qiluvchi so'zlardan biri sifatida *betartib* so'zlarning tahlilini ko'rib chiqamiz.

– **betartib+** – roq/-gina  $\oplus$  odam/xona/uy/ko'cha/

– **betartib**  $\oplus$  kiyinmoq/yumoq/kelmoq/...

$$Pf_{adj,adv} = \begin{cases} Pf^{adj} + aff^{deg} \oplus N^{Pf_{adj}} \\ Pf^{adv} \oplus V \end{cases} \quad (37)$$

$Pf_{adj,adv}$  – sifat  $\vee$  ravish so'z turkumlari orasida polifunksionallik,  $Pf^{adj}$  – sifat so'z

turkumidagi polifunksional so'z,  $Pf^{adv}$  – ravish so'z turkumidagi polifunksional

<sup>61</sup> Раҳматуллаев Ш. Ўзбек тили омонимларининг изоҳли луғати. – Тошкент: Ўқитувчи, 1984. – Б. 36.

so'z,  $N^{Pf-adj} - Pf^{adj}$  so'z bilan birikuvchi ot so'z turkumidagi so'zlar to'plami.

(37) model orqali –sifat– ravish so'z turkumidagi polifunksionallikni aniqlash

mumkin.	Bog	bolani
---------	-----	--------

### 2.3.5. Bog'lovchi yuklama orasidagi polifunksionallikni aniqlash

Bog'lovchi va ko'makchi orasidagi polifunksionallik quyidagicha modellashtiriladi. (agar bilan bog'lovchi bo'lib kelsa):

- *Ota bilan bola/ gul bilan lola/oq bilan qora/kelish bilan ketish/oy bilan quyosh/....*
- *Qunt bilan o'qi/ aql bilan o'ylamoq/tuni bilan ishlamoq/...*

$$Pf_{conj,aux} = \begin{cases} W^{conj} \oplus Pf^{conj} \oplus W^{conj} \\ W^{aux} \oplus Pf^{aux} \oplus V \end{cases} \quad (38)$$

$Pf_{conj,aux}$  –bog'lovchi va ko'makchi orasidagi polifunksionallik hodisasi,  $W^{conj}$  - polifunksional so'z bo'ladigan bog'lovchilar bilan birikuvchi so'zlar to'plami,  $Pf^{conj}$  – bog'lovchi polifunksional so'z,  $W^{aux}$  –ko'makchi polifunksional so'z bilan birikuchi so'zlar to'plami,  $Pf^{aux}$  –ko'makchi polifunksional so'z.

Bu matematik model yordamida bog'lovchi va ko'makchi so'z orasida polifunksionallik hosil qiluvchi so'zlarni aniqlash mumkin. Yuqoridagi polifunksional so'zlarning klassifikatsiyasi keltirilgan ierarxiyadagi boshqa guruh polifunksional so'zlarini aniqlash uchun ham xuddi shunday matematik modellar ishlab chiqish mumkin. Polifunksionallikni to'g'ri aniqlash uchun bu so'zlar bilan birika oladigan so'zlar to'plamini ajratish talab qilinadi.



## XULOSA

O'zbek tilidagi omonim, polifunksional va polisemantik so'zlarni semantik analiz qiluvchi tizim albatta matematik model va algoritmlarga asoslanadi.

Tizimga foydalanuvchi tomonidan kiritilgan matn ustida dastlabki qayta ishlash jarayoni amalga oshiriladi: *tokenizatsiya – matnni gaplarga ajratish, matndagi nomuhim so'zlarni<sup>62</sup> olib tashlash, gapdagi tinish belgilarini olib tashlash va gaplarni teglashtirish* kabi jarayonlar olib boriladi. Gapdagi nomuhim so'zlarni olib tashlash orqali tadqiqot obyektlarining mustaqil so'zlardan iborat birikuvchilari to'plami hosil qilinadi.

Semantik analizator uchun katta hajmdagi ma'lumotlar sifatida o'zbek tili milliy korpusining ma'lumotlaridan foydalaniladi. Birikuvchi (valentlik)larni to'plashda birlashgan yondoshuv deb nomlangan *probabilistik tegger* yondashuvlaridan foydalaniladi. Ushbu turdagi mashhur yondoshuvlar sifatida TBL (Transformation Based Learning) va Brillning morfologik analizatori (Brill tegger) qaraladi. Bu ikki yondoshuv qoidalarga asoslangan va deterministik algoritmlar asosida qayta ishlangan matn, ya'ni oldindan belgilangan so'zlarning Grammatik sinflari korpusdan avtomatik aniqlanadi. Keyin barcha turdagi teg o'zgarishlari tekshiriladi (ishlab chiqarish qoidalari) va tahlil qilish sifatini maksimal darajada oshiradiganlari qoldiriladi. Korpus ma'lumotlari yangilanganda undagi so'zlar qayta teglanadi.

### **Mazkur vazifaning markazida bir qancha g'oyalar mavjud**

1. Yuqori aniqlik bilan manba sifatida omonim so'zlar lug'atini yaratish lozim
2. Kontekstlarda omonimlarni joylashuviga qarab ularning ma'nolarini kuchli yoki kuchsizlarga ajratish talab qilinadi. Bu oddiy ehtimollik modellari asosida son ko'rinishida ifodalangani.
3. Ushbu parametrdagi algoritmninng maqsadi lemmani olishdir. Flective tillari normalizatsiya almashinuvi juda ko'p qirrali grammatik xususiyatlarini

<sup>62</sup> Madatov Xabibulla Axmedovich. O'ZBEK TILI MATNLARIDAGI NOMUHIM SO'ZLAR.// Toshkent 2021//<http://compling.navoiy-uni.uz/index.php/conferences/article/download/5/2>

ifodalash vositasi, shuning uchun omoformalarni kontekst elementlari sifatida ishlatishga harakat qildik.

#### 4. Ishni qurishda o'zbek tilining omonimlari chastotasi tartibini belgilash "lemma tanlashning qiyinchiliklari" g'oyasi qo'llaniladi

Algoritm ishlashi uchun berilgan ma'lumot sifatida matnlarni morfologik tahlil qiluvchi dasturi natijalari kerak bo'ladi, o'zbek tilida esa morfologik analizator yaratish ustida tadqiqot ishlari olib borilmoqda. Morfologik analizator yordamida gapdagi har bir so'zning quyidagicha tahlili olinadi:

$$\text{Kulibdi}\{kul=H, -ibdi=aff^V\}$$

Demak *kul*-omonim so'z, *-ibdi*- fe'lga qo'shiluvchi sintaktik yoki lug'aviy shakl yasovchi qo'shimcha yoki qo'shimchalar kombinatsiyasi, agar bu qo'shimcha tarkibida so'z yasovchi qo'shimcha mavjud bo'lsa, u holda bu so'z omonim emas deb qaraladi.

O'zbek tilidagi polisemantik so'zlarni ko'nteksda farqlash uchun albatta so'zlar orasidagi bog'lanishlar muhim ahamiyat kasb etadi. Bu bog'lanish turkiy tillarda Grammatik aloqa deb ataladi<sup>63</sup>. Grammatik aloqa gaplarni sintaktik, semantik tahlil qilishda muhim ahamiyatga ega. Gapni semantik tahlil qilish uchun gap tarkibidagi so'zlarning o'zaro aloqasi yordamida analizatorlar yaratiladi. Link Grammar Parser sintaktik analizator bo'lib, u Daniel Slitor va Devi Temperllar tomonidan tilning Grammatik aloqalari asosida ishlab chiqilgan.<sup>64</sup>

Ye.L.Boyarskayaning "Концептуальное моделирование процесса разрешения полисемии с использованием корпуса" maqolasida konseptual modellashtirish usuli korpus materialida polisemantik so'z ma'nolarini aniqlashning yangi istiqbollarni ochib berishi haqida qimmatli ma'lumotlar keltirilgan<sup>65</sup>.

<sup>63</sup> Tukeyev U. A., Raximova D.R. Семантические связи в автоматической обработке текста казахского языка // Вестн. КазНТУ. Серiya: Matematika, mexanika, informatika. 2012. № 2. С. 320–325.

<sup>64</sup> Temperley D. An Introduction to the Link Grammar Parser. 2014. URL: <http://www.abisource.com/projects/link-grammar/dict/introduction.html#1>

<sup>65</sup> Боярская Е.Л. Концептуальное моделирование процесса разрешения полисемии с использованием корпуса // <https://cyberleninka.ru>

**Birinchi**dan, lug'atlar bilan bog'liq muammo. Ko'pincha, so'zning ma'nolarini bir-biridan farqlash vazifasi qiyinchilik tug'dirmaydi, ammo ba'zi hollarda so'zning turli ma'nolari semantik jihatdan bir-biriga juda yaqin bo'lishi mumkin va bunday holatlarda turli lug'at va tezaruslarda ma'nolarga bo'linish sezilarli darajada farq qilishi mumkin.

**Ikkinchi**dan, ayrim tillarda qisman razmetkalash (belgilash) ko'p ma'nolilikni bartaraf etish muammosi bilan juda yaqin bog'liq bo'lishi mumkin, natijada bu ikki vazifa bir-biriga xalaqit berishi mumkin. Olimlar ularni ikki avtonom tarkibiy qismga ajratishga arziydimi-yo'qmi, degan fikrga kelishmagan<sup>66</sup>.

**Uchinchi**dan, leksik ko'pma'nolilikni hal qilish tizimlari har doim inson mehnati natijalari bilan taqqoslash orqali baholangan. Odamlar uchun esa bu vazifa POS-tagging kabi oddiy bo'lmasligi mumkin<sup>67</sup>.

Odamlarning natijalari har doim ham bir-biriga mos kelmaydi, ular ko'pincha so'zning muayyan kontekstda qanday ma'no anglatishi haqida bir xil qarorga kelishmaydi. Shunga qaramay, olimlar inson natijasini standart, kompyuter natijalari bilan taqqoslash uchun etalon sifatida qabul qiladilar. Mazkur monografiyada quyidagi natijalarga erishildi:

1. O'zbek tilidagi omonim so'zlarni bartaraf etishda ularni ikkita guruhga 1). Bir so'z turkumi doirasidagi omonimlar. 2). Turli so'z turkumidagi omonimlarga ajratish muhim ahamiyat kasb etadi.
2. Bir so'z turkumidagi omonimlikni hosil qiluvchi omonim so'zlar 5 (faqat ot, fe'l, sifat, olmosh, ravish) guruhga ajratilib, ular statistik usullar, Markov modellari yordamida aniqlanadi.
3. Turli so'z turkumidagi omonimlikni hosil qiluvchi omonimlarni 36 ta guruhga ajratib, jami 27 ta matematik model ishlab chiqildi.
4. O'zbek tilidagi polisemantik so'zlarning iyerarxiyasi ishlab chiqildi, ularni ma'nosini farqlash uchun matematik mantiqqa asoslangan matematik model

<sup>66</sup> Lynette Hirschmann, The evolution of evaluation. Computer Speech and Knowledge, 1998.

<sup>67</sup> Fellbaum C. 1997. Analysis of a handtagging task. In Proc. of ANLP-97 Workshop on Tagging Text with Lexical Semantics: Why, What, and How? Washington D.C., USA.

tuzildi. Kontekstda polisemantik so'zlarni to'g'ri farqlash uchun bu so'zlarning birikuvchilarini alohida guruhlariga ajratish maqsadga muvofiq.

5. O'zbek tilidagi polifunksional so'zlarni aniqlashda ularni so'z turkumlari kesimida guruhlariga ajratib, ot-sifat orasida, sifat-modal orasida, sifat-ravish orasida va ko'makchi-bog'lovchi orasida polifunksionallik hosil qiluvchi so'zlarni farqlovchi 5 ta matematik model ishlab chiqildi.

Ushbu erishilgan natijalar o'zbek tili semantik analizatori yaratish uchun dastlabki qadam sifatida xizmat qilsa ajab emas.

## Foydalanilgan adabiyotlar

1. Bar-Hillel Y. Language and information (Reading, Mass.: Addison-Wesley, 1964). – С.174- 179.
2. Bird S., Klein E., Loper E. Natural Language Processing with Python: Analyzing
3. *Brizgumova E.A.* Phones and intonation of the Russian speech. – Moscow, 1997. – P.12.
4. Cristopher D. Manning, Hinrich Schutze. Foundation of Statistical Natural Language Processing, 1999.
5. De Smet W., Moens M.F. Cross-language linking of news stories on the web using interlingual topic modelling // Proceedings of the 2Nd ACM Workshop on Social Web Search and Mining. – SWSM '09. – New York, NY, USA: ACM, 2009. – Pp. 57–64.
6. Fellbaum C. 1997. Analysis of a handtagging task. In Proc. of ANLP-97 Workshop on Tagging Text with Lexical Semantics: Why, What, and How? Washington D.C., USA.
7. Tony J. What is Semantic Search? Techulator. – March 15. 2012.
8. Hirschmann L. The evolution of evaluation. Computer Speech and Knowledge. – 1998.
9. Yoqubov T., Karimbekov S. Matematik mantiq elementlari. – Toshkent: O'qituvchi, 1996.
10. Younger D.H. Recognition and parsing of context-free languages in time n3. Information and Control 10 (2). – P.189-208.
11. Апресян Ю. Д., Богуславский И. М., Иомдин Л. Л. Лингвистический процессор для сложных информационных систем. – М.: Наука, 1992, – 256 с.
12. Большакова Б.И. Компьютерная лингвистика: методы, ресурсы, приложения // Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие. – М.: МИЭМ, 2011. С. 93—94.

13. 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.
14. Булаховский Л.А. Исторический комментарий к русскому литературному языку ... Учебник. – М.: Академический Проект: Парадигма, 2005. – 256 с.
15. Виноградов В.В. Об омонимии и смежных явлениях. Избранные труды. Исследования по русской грамматике. – Москва, 1975. – С. 295.
16. Гак В.Г. К проблеме семантической синтагматики. Проблемы структурной лингвистики. – М., 1971. – С. 367.
17. G'ulomov A. O'zbek tili morfologiyasiga kirish. O'zbek tili kursidan materiallar. – Toshkent, 1953. – 28 b.
18. Иорданская Л.Н., Мельчук И.А. Смысл и сочетаемость в словаре. – Москва, 2007. – 672 с.
19. Захаров В.П., Богданова С.Ю. Корпусная лингвистика. – Иркутск, ИГЛУ, 2011.
20. Mirtojyev.M. O'zbek tili leksik omonim manbalari. – Toshkent: Fan, 2014. – 126 b.
21. Mirtojyev M. O'zbek tilida polisemiya. – Toshkent: Fan, 1984. – 135 b.
22. Новиков Л.А. Семантика русского языка. Учебное пособие. – Москва, 1982. – С. 280.
24. Po'latov A. Kompyuter lingvistikasi. – Toshkent: Akademnashr, 2011. – 344 b.
25. Мочалова А.В., Захаров В.П., Кузнецов В.А., Мочалов В.А. Семантический анализ. Некоторые проблемы автоматической обработки текста / – Петрозаводск: ПИН, 2016. – 60 с.
26. Стернина М.А. Лексико-грамматическая полисемия в системе языка / М.А. Стернина. – Воронеж: Изд-во «Истоки», 1999. – 160 с.
27. Сулейманов Д.Ш., Гатиатуллин А.Р. Структурно-функциональная компьютерная модель татарских морфем. – Казань, 2003. – 220 с.
28. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. – М.: Наука, 1986.



29. Тузов В.А. Компьютерная семантика русского языка. – С.-Петербургского университета, 2003. – 391 с.
30. Тузов В.А. Компьютерная семантика русского языка. – СПб: Изд-во СПбГУ, 2003. – 391 с.
31. Фесенко Т.А. Концептуальное моделирование как метод познания ментальной реальности человека. Язык, сознание, коммуникация. – М., 2000.
32. Гатауллин Р.Р., Гатиатуллин А.Р., Неврозова О.А., Мухамедшин Д.Р., Сулейманов Д.Ш., Хакимов Б.Э., Хусайнов А.Ф. Формальные модели и программные инструменты компьютерной обработки татарского языка // – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Изд-во Академии наук Рт, 2019. – 260 с.
33. Hakimova M. Semasiologiya. – Toshkent, 2008. – 100 b.
34. Щипицина Л.Ю. Информационные технологии в лингвистике: учеб. пособие / Л.Ю.Щипицина. – М.: ФЛИНТА: Наука, 2013. – 128 с. – С. 27.
35. Б. А. Бобнев. Национально-культурные особенности в структуре многозначных слов ‘дом’ – ‘casa’ в русском и итальянском языках // Вестник Челябинского государственного университета. 2010. № 22 (203)
36. Abdurahmonova M., Raxmonova A. Korpus lingvistikasida polisemiya // “Kompyuter lingvistikasi: muammolar, yechim, istiqbollar” Respublika I ilmiy texnikaviy-konferensiya. – Vol. 1 №. 01 (2021).
37. Abdurhmonova 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.
38. Aliqulov T. Polisemiyalarning hosil bo‘lishi haqida // O‘zbek tili va adabiyoti. – 1963, №3 – В. 51-56.
39. Axmeova D. B. O‘zbek tilidagi omonimlarni teglash muammolari // “Zamonaviy ta’limda raqamli tizimlarni qo‘llash: filologiya va pedagogika sohasida

- zamonaviy tendensiyalar va rivojlanish omillari” mavzusidagi xalqaro ilmiy-amaliy masofaviy konferensiya. – Toshkent, 1 may 2020 yil. – B. 344-346.
40. Xabibulla Madatov. O‘zbek tili matnlaridagi nomuhim so‘zlar// “kompyuter lingvistikasi: muammolar, yechim, istiqbollar” Respublika ilmiy-texnikaviy konferensiya –Toshkent, 23-aprel 2021 yil.
41. Барабаш О.В. Разграничение омонимии и полисемии юридических терминов // <https://cyberleninka.ru/article/n/razgranichenie-omonimii-i-polisemii-yuridicheskikh-terminov>
42. Батура Т.В. Семантический анализ и способы представления смысла текста в компьютерной лингвистике // Международный журнал Программные продукты и системы. – December, 2016.
43. Гатауллин Р.Р. Аналитический обзор методов разрешения морфологической многозначности // Russian Digital Libraries Journal. 2016. V. 19. No 2.
44. Gulyamova Sh. Polisemiya lingvistik hodisa sifatida// “O‘zbekistonda ilmiy-amaliy tadqiqotlar” konf.materiallari. – Toshkent, 2020., №18. – B. 26-28.
45. Гурычева М.С., Серебренников Б.А. Задачи изучения основного словарного фонда языка // Вопросы языкознания. – 1953., № 6. – С. 3-20.
46. Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков // Филология и культура. Philology and culture. 2014. №2(36). – С. 20-26.
47. Zohidov R. “Sabot ul-ojizin”dagi ko‘p ma’noli so‘zlarning qiyosiy tahlili // O‘zbek tili va adabiyoti. – 2003., №2. – B. 49.
48. Исраилова Н.А., Бакасова П.С. Морфологический анализатор кыргызского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Казань, 2017. – 327 с.
49. Каневский Е.А., Боярский К.К. Семантико-синтаксический анализатор SemSin // Компьютерная лингвистика и интеллектуальные технологии. 2012.
50. Боярский К.К., Каневский Е.А. Предсинтаксический модуль в анализаторе

- SemSin // Интернет и современное общество. Санкт-Петербург, 2013. С. 280–286.
51. Кобозева И.М. Полисемия дискурсивных слов и возможности ее разрешения в контексте предложения (на примере слова ВОТ) // Труды международной конференции “Диалог 2007”. – С. 250-255.
52. Кобрицов Б.П., Ляшевская О.Н., Шеманаева О.Ю. Поверхностные фильтры для разрешения семантической омонимии в текстовом корпусе //
53. Qurbonmuratov A. O‘zbek tili milliy korpusidagi shakldosh so‘zlarini morfologik farqlash // So‘z san’ati xalqaro jurnali. 4-jild, 3-son. – Toshkent, 2021. – В. 284-287. – Doi 10.26739/2181-9297.
54. Ладанова Е.О., Ямашкин С.А. Семантический анализатор для выделения фактов из текстовых сообщений // Международный научно-исследовательский журнал. № 12 (66).
55. Ляшевская О.Н. Топологические классы имен в семантической разметке Национального корпуса русского языка // Труды международной конференции «Корпусная лингвистика-2008» (Санкт-Петербург, 6-10 октября 2008 г.). СПб.: С.-Петербургский гос. университет, Факультет филологии и искусств, 2008. – С. 276-284.
56. Мальковский М.Г., Старостин А.С. Модель синтаксиса в системе морфосинтаксического анализа «TREETON» // Тр. междунар. Конф. Диалог 2006. – М., 2006. – С. 481-492.
57. Mengliyev B. O‘zbek tili taraqqiyoti va rivojlanish omillari // O‘zbek tilini dunyo miqyosida keng targ‘ib qilish bo‘yicha hamkorlik istiqbollari // Xalqaro ilmiy-amaliy anjuman materiallari. – Toshkent, 2020, oktabr. – В. 358-359.
58. Mengliyev B. Lug‘aviy ko‘p ma’noqlik lison va nutq bo‘linishida // O‘zbek tili va adabiyoti. – 2000., №5. – В. 51-54.
59. Мозговой М.В. Семантический анализатор и задача информационного поиска // Вестник СПбГУ. Сер. 10, 2005. – С. 54-59.
60. Москвина А.Д., Орлова Д., Паничева П.В., Митрофанова О.А. Разработка ядра синтаксического анализатора для русского языка на основе

библиотек NLTK // Компьютерная лингвистика и вычислительные онтологии. Труды XIX Международной объединенной научной конференции «Интернет и современное общество», Санкт-Петербург, 22–24 июня 2016 г. СПб, 2016. С. 44–54.

61. Мочалова А.В. Алгоритм семантического анализа текста, основанный на базовых семантических шаблонах с удалением / А.В. Мочалова // Научно-технический вестник информационных технологий, механики и оптики. – 2014. – № 5. – С. 126–132.

62. Порохнин А. Анализ статистических методов снятия омонимии в текстах на русском языке // <http://www.mathnet.ru/vagtu286>

63. Рахилина Е.В., Ляшевская О.Н., Кобрицов Б.П., Кустова Г.И., Шеманаева О.Ю. Многозначность как прикладная проблема: Лексико-семантическая разметка в Национальном корпусе русского языка //

64. Лауфер Н.И., Нариньяни А.С., Селегей В.П. (ред.). Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции “Диалог 2006”, – 2006. – С. 445–450.

65. Селегей В. Лингвистические проблемы автоматического создания интернет-корпуса русского языка / В. П. Селегей // Инновации и высокие технологии: тр. 55-й науч. конф. МФТИ. – М.: Изд-во Московского физико-техн. ин-та, 2012. – С. 53–54.

66. Толдова С.Ю., Кустова Г.И., Ляшевская О.Н. Семантические фильтры для разрешения многозначности в Национальном корпусе русского языка: глаголы // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 4–8 июня 2008 г.). – Вып. 7 (14). – М.: РГГУ, 2008.

67. Тузов В.А. Компьютерная семантика русского языка / В.А. Тузов. – СПб.: Изд-во С.-Петерб. ун-та, 2004. – 400 с.;

68. Мозговой, М.В. Простая вопросно-ответная система на основе семантического анализатора русского языка / М.В. Мозговой // Вестник СПб университета. – 2005. – сер. 10. – вып. 1. – С. 116–122.

69. Федосин С.А. Технологический процесс решения задачи моделирования структуры землепользования на базе данных ДЗЗ / С.А.Федосин, С.А.Ямашкин // Науч.-техн. вестн. Поволжья. – 2014. – № 6. – С. 356–358.
70. Mengliyev B., Hamroyeva Sh. Turkiy tillardagi morfologik analizatorlar xususida // O'zbek tilini dunyo miqyosida keng targ'ib qilish bo'yicha hamkorlik istiqbollari. Xalqaro ilmiy-amaliy anjuman materiallari. – Toshkent, 2020, oktabr. – B. 360-364.
71. Шеманаева О.Ю., Кустова Г.И., Ляшевская О.Н., Рахилина Е.В. Семантические фильтры для разрешения многозначности в Национальном корпусе русского языка: прилагательные // Иомдин Л.Л., Лауфер Н.И., Нариньяни А.С., Селегей В.П. (ред.). Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог 2007». – 2007. – С. 582-587.
72. Ямашкин С.А. Гибридная система анализа данных дистанционного зондирования Земли / С.А.Ямашкин // Научно-технический вестник Поволжья. 2015. № 4. – С. 173–175.
73. Korobov M. Morphological Analyzer and Generator for Russian and Ukrainian Languages // Analysis of Images, Social Networks and Texts: 4th International Conference, AIST 2015, Yekaterinburg, Russia, April 9–11, 2015. Communications in Computer and Information Science, Springer, 2015. P. 320–332.
74. Mathematical model of an ontological-semantic analyzer using basic ontological-semantic patterns / A.V.Mochalova, V.A.Mochalov // Lecture Notes in Artificial Intelligence, Proceedings of 15th Mexican International Conference on Artificial Intelligence. – 2016. – P. 53–66.
75. Вдовин С.М. Получение, хранение и распространение геоданных как единый информационный процесс / С.М.Вдовин, С.А.Федосин, А.А.Ямашкин, С.А.Ямашкин // Природные опасности: связь науки и практики: материалы II Международной науч.-практ. конф. / отв. ред. С.М.Вдовин. – Саранск, 2015. – С. 82–90.

76. Каневский Е.А., Колпакова Н.В. К вопросу построения морфологического анализатора // Компьютерная лингвистика и интеллектуальные технологии. – Москва, 1999. Т. 2. С. 98–106.
77. Потылицына Н.В. Явления омонимии как одно из функциональных свойств языка (на материале современного английского языка) [Текст] / Н.В. Потылицына // Семантика, функция и грамматические категории лексических единиц: межвуз. сб. науч. тр. – Пятигорск: Изд-во ПГЛУ, 2004. – С. 99-103.
78. Usmonov S. Hozirgi zamon o'zbek adabiy tilida omonimlar va sinonimlar // Ilimiy asarlar. № 12. – Toshkent: Fan, 1959. – В. 40-44.
79. Хамроева Ш. О'zbek tili mualliflik korpusini tuzishning lingvistik asoslari: Filol. fan. bo'yicha falsafa doktori (PhD)...diss. – Qarshi, 2018. – 250 b.
80. Abjalova M.A. O'zbek tilidagi matnlarni tahrir va tahlil qiluvchi dasturning lingvistik modullari (Rasmiy va ilmiy uslubdagi matnlar tahriri dasturi uchun): Filol.fan.bo'yicha falsafa doktori (PhD)...diss. – Farg'ona, 2019. – 164 b.;
81. Abjalova M. Matnlarga avtolingvistik ishlov berish tizimlari // Shestaya Mejdunarodnaya konferensiya po kompyuternoy obrabotke tyurkskix yazykov «TurkLang-2018». (Trudy konferensii) – Tashkent, 2018. – 320 s.
82. Xoliyorov O'. O'zbek tili ta'limiy korpusini tuzishning lingvistik asoslari. Filol. fan. bo'yicha fal. dok-ri. diss. avtoref. – Termiz, 2021. – В. 16. – 52 b
83. Кириллов А.В. Математическое моделирование и программная реализация семантического преобразования поисковых запросов тема диссертации и автореферата по ВАК РФ 05.13.18, кандидат технических наук, 2012.
84. Ерискина Е.В., Курушин Д.С., Яруллин Д.В. Объектно-ориентированная модель морфологического анализатора русскоязычного текста //Корюкин Пермский национальный исследовательский политехнический университет, Пермь-2018
85. Кочетова Н. Национальная специфика семантем русской и английской адвербиальной лексики. Дисс. на соискание ученой степени кандидата филологических наук. – Воронеж, 2017. – 276 с.



86. Мозговой М.В. Машинный семантический анализ русского языка и его применения: дисс. канд. физико-матем. наук. – Санкт-Петербург, 2006.

87. Xoliyorov O'. O'zbek tili ta'limiy korpusini tuzishning lingvistik asoslari, Filol. fan. bo'yicha fal. dok-ri. diss avtoref. – Termiz, 2021. – B. 16. – 52 b.

88. Hamroyeva Sh. M. O'zbek tili morfologik analizatorining lingvistik ta'minoti. Filologiya fan. d-ri. diss... av-ti. – Farg'ona, 2021. – 76 b.

89. Султанова Алина Петровна, // Полисемия глаголов деструкции в русском и английском языках тема диссертации и автореферата по ВАК РФ

10.02.20, кандидат филологических наук Султанова, Алина Петровна. 2008

## Mundarija

<b>KIRISH</b> .....	5
<b>I BOB. Semantik munosabatlar va ularning matn tahlilida qoʻllanilishi</b> .....	9
1.1. Tabiiy tillardagi matnlarni tahlil qilish tizimlari.....	9
1.1.1. Morfologik analizatorlar.....	9
1.1.2. Sintaktik analizatorlar.....	13
1.1.3. Semantik analizatorlar.....	15
1.2. Matnlarni semantik analiz qilishning model va metodlari.....	18
1.3. Oʻzbek tili semantik analizatorida matnlarni tglash usullari.....	26
<b>II BOB. Ontosemantik qoidalarga asoslangan oʻzbek tili semantik analizatori matematik modeli</b> .....	31
2.1. Oʻzbek tilida omonimiyani bartaraf etish usullari.....	31
2.1.1. Bir soʻz turkumi doirasidagi omonimiyani bartaraf etish usullari.....	31
2.1.3 Turli soʻz turkum doirasidagi omonimiyani bartaraf etish usullari.....	47
2.2. Kontekstda polisemiyani bartaraf etishning matematik modellar.....	78
2.3. Kontekstda polifunksional soʻzlarni farqlovchi matematik modeli.....	86
<b>XULOSA</b> .....	93
<b>Foydalanilgan adabiyotlar</b> .....	97

ELOV BOTIR, AXMEDOVA XOLISXON

**O‘ZBEK TILI SEMANTIK  
ANALIZATORINIG MATEMATIK  
MODELLARI**

*Monografiya*

“Farovon yurt” nashriyoti.  
Toshkent sh., Shayxontohur t. Navoiy 30.

“Farovon yurt” XJB boshqaruvida bosildi.  
Toshkent sh., Shayxontohur t. Navoiy 30.

ELOY BOTIR AXMEDOVA XOLISSON

O'ZBEK TILI SEMANTIK  
ANALIZATORINING MATEMATIK  
MODELLARI

Monografiya

“Farovon yurt” nashriyoti.

Toshkent sh., Shayxontohur t., Navoiy 30.

“Farovon yurt” XJB bosmaxonasida bosildi.

Toshkent sh., Shayxontohur t., Navoiy 30.



Elov Botir Boltayevich Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti, Kompyuter lingvistikasi va raqamli texnologiyalar kafedrasi mudiri, texnika fanlari bo'yicha PhD.

**Yutuqlari:** 1) SMART LMS axborot tizimi asoschisi (<http://e.navoiy-uni.uz>).

2) ONLINE REPETITOR axborot tizimi muallifi (<http://test.navoiy-uni.uz>)

3) Ozbek tili talimiy korpusi axborot tizimi hammuallifi (<http://uzschoolcorpara.uz/>)

4) Ozbek poetik lugat dasturi hammuallifi (<http://narrativedictionary.navoiy-uni.uz/>)

**Qiziqishlari:** axborot tizimlarini loyihalashtirish, axborot xavfsizlig, biz-jarayonlarni modellashtirish, Python, C#, SQL, NLP, AI.



Axmedova Xolisxon Ilxomovna Alisher Navoiy nomidagi Toshkent davlat O'zbek tili va adabiyoti universiteti tayanch doktoranti

**Qiziqishlari:** Matematik modellashtirish, NLP, SQL, Python, Java.