

O‘ZBEK TILIDA SEMANTIK ROLLARNING O‘ZIGA XOSLIGI

*Zilola Xusainova¹,
Rashidova Nazira²*

Annotatsiya: Maqolada semantik rollar belgilash (SRL) jarayoni va uning o‘zbek tilidagi kompyuter lingvistikasi tizimlaridagi ahamiyati tahlil qilindi. SRL jarayonining asosiy komponentlari - predikatlarini aniqlash va aktantlarning semantik rolini belgilash yoritildi. Tadqiqotda o‘zbek tilida morfologik va sintaktik teglash tizimi - Uzbek TAGSET ishlab chiqildi, lemmatizatsiya, POS-tagging va dependency parsing kabi vazifalarda samaradorlikni oshirishga qaratilgan tizimli yondashuvlar taqdim etildi. Maqola o‘zbek tilidagi NLP tizimlarini rivojlantirish va semantik tahlilni avtomatlashtirishda muhim ilmiy-amaliy asosni taqdim etdi. Shuningdek, rus va o‘zbek tillaridagi fe‘l valentligi va aktantlarning semantik-sintaktik vazifalari ham qiyosiy tahlil qilindi, bu esa tilshunoslik va kompyuter lingvistikasi sohasida yangi yondashuvlar va metodologiyalarni yaratishga xizmat qildi.

Kalit so‘zlar: semantik rollar, SRL, predikat-argument munosabatlari, Uzbek TAGSET, POS-tagging, lemmatizatsiya, dependency parsing, named entity recognition, o‘zbek tili, NLP, fe‘l valentligi, aktantlar, morfologik teglash, sintaktik teglash, agglutinatativ til.

Tilning mazmunini o‘rganishda semantik rollar juda muhim – ular gapning “yuragi” bo‘lgan fe‘l (predikat) atrofida kim nima qilayotganini, qanday ta’sir ko‘rsatayotganini aniq belgilab beradi. Bu rollar **aktantlar** deb ataladi va ularning fe‘l bilan bog‘lanishi gapning umumiy ma’nosini shakllantiradi. Turli tillarda bu jarayon o‘ziga xos kechadi, chunki u tilning tuzilishi, sintaksisi va semantikasiga bog‘liq.

Semantik rollar nazariyasining shakllanishi va rivojini chuqur anglash uchun, avvalo, xorijiy lingvistlar tomonidan ishlab chiqilgan konseptual yondashuvlar, predikat–argument modeli bo‘yicha yaratilgan ilmiy ishlanmalar hamda nazariy tasniflarning shakllanish bosqichlarini tahlil qilish zarur.

Bu borada rus tilshunosligida olib borilgan izlanishlar, xususan, O. A. Kukatovanning tadqiqotlari semantik rollarning mazmuniy mohiyatini anglashda muhim ilmiy asos vazifasini bajaradi. Kukatovanning 2021-yilda chop etilgan ishlari

¹ Xusainova Zilola Yuldashevna – filologiya fanlari falsafa doktori (PhD), dotsent v.b. Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti.

E-pochta: xusainovazilola@navoiy-uni.uz

ORCID: 0000-0003-4357-7515

²Rashidova Nazira – ToshDO‘TAU, Kompyuter lingvistikasi mutaxassisligi I bosqich magistranti.

semantik rollarni aniqlashda aktantlarning turli tasnif mezonlarini ko'rib chiqish bilan ajralib turadi. Muallif predikat–argument tizimini lingvistik nuqtai nazardan chuqur yoritib, aktantlarning gap tarkibidagi funksional yukini, ularning mazmuniy shakllanishida morfologik hamda sintaktik omillarning tutgan o'rnini batafsil izohlaydi. Bu yondashuv predikat markazidagi mazmuniy strukturaning shakllanish mexanizmlarini tushunishda konseptual nazariy asos yaratadi. [1]

Kukatovanning 2012-yilda chop etilgan maqolasi esa semantik rollar doirasida kauzativlik fenomenini talqin qilishga bag'ishlangan. Unda, ayniqsa, emotsional fe'llar ishtirokida kuzatiladigan semantik rol o'zgarishlari tahlil qilinadi. Muallif kauzatorning gap mazmunini shakllantirishdagi roli, sababchi sub'yektning semantik funksiyasi va uning predikat bilan o'zaro munosabatini izchil asoslaydi. Ushbu natijalar o'zbek tilida uchraydigan kauzativ fe'l shakllarini o'rganishda metodologik yo'nalish sifatida alohida ahamiyatga ega. [2]

2016-yilda e'lon qilingan tadqiqotida esa Kukatova semantik rollarni predikat–argument tuzilmasi negizida talqin qiladi. Muallif aktantlarning predikat bilan bog'lanish mexanizmini, semantik yukning shakllanish jarayonini va rol turlarining izchil tasnifini ilmiy asosda yoritadi. Mazkur yondashuv semantik rollarni aniqlash jarayonida nazariy aniqlik va izchillikni ta'minlashga xizmat qilib, predikat markazidagi komponentlararo munosabatni chuqurroq anglashga yordam beradi. [3] Umuman olganda, Kukatovanning semantik rollar bo'yicha olib borgan izlanishlari ushbu nazariyaning konseptual asoslarini mustahkamlovchi, aktantlarning struktura va mazmuniy vazifalarini aniqlashtiruvchi hamda kauzativ konstruksiyalarni izohlash imkonini beruvchi muhim ilmiy platformani yaratadi. Uning ishlari fe'l valentligi, uch aktantli tuzilmalarning semantik xususiyatlari va predikat–argument tizimida rol belgilash jarayonining nazariy asoslarini shakllantirishda metodik yo'nalish sifatida katta ahamiyat kasb etadi. Shuningdek, bu yondashuvlar o'zbek tilida semantik rol belgilash (SRL), fe'l semantikasi va sintaktik bog'lanishlar bo'yicha keyingi tadqiqotlar uchun mustahkam ilmiy bazani taqdim etadi.

Bundan tashqari, Shamsieva rus va o'zbek tillaridagi semantik hamda sintaktik valentlik xususiyatlarini qiyosiy tahlil qilgan bo'lib, fe'llarning mazmuniy imkoniyatlari va ular talab etadigan aktantlar soni hamda turini izchil o'rganilgan. Muallif har ikki tilda predikatning markaziy roli, aktantlarning gap tarkibida qanday shakllanishi va semantik yukning qanday taqsimlanishini tahlil qiladi. Tadqiqotda o'zbek tilining agglutinatив tuzilishi valentlikning morfologik ko'rsatkichlar orqali ifodalanishini ta'minlashi, rus tilida esa sintaktik vositalarning nisbatan faolroq o'rin tutishi ko'rsatib beriladi. Uch aktantli konstruksiyalar, kauzativlik hodisasi hamda fe'llarning semantik roli bilan bog'liq ayrim farqlar alohida

yoritiladi. Shamsievaning bu yondashuvi o'zbek tilidagi aktantlar tizimini aniqlash, ularning predikatga bog'lanish mexanizmlarini tushunish va semantik rollarni tizimlashtirish uchun qiymatli nazariy asos yaratadi. [4]

Hasanova (2025) maqolasida rus va o'zbek tillaridagi grammatik kategoriyalar, xususan, case (padej) tizimining semantik va grammatik o'zaro bog'liqligi tahlil qilinadi. Tadqiqotda har ikki tilning case kategoriyalari qanday semantik rollarni ifodalashi, sintaktik funksiyalari va ularning o'zaro moslik/nomoslik jihatlari izchil yoritilgan. Muallif grammatik kategoriya nutq tuzilishidagi semantik vazifani qanday kodlayotgani, argumentlar o'rtasidagi munosabatlar va til strukturalaridagi farqlar bo'yicha qiyosiy yondashuvni taklif etadi. Ushbu ish semantik va sintaktik valentlik, predikat–argument munosabatlari hamda semantik rollar tuzilishi masalalarini o'rganayotgan tadqiqot uchun muhim qo'shimcha ilmiy manba hisoblanadi. Sotsial omillar (jamiyatdagi ijtimoiy o'zgarishlar, madaniy ta'sirlar va kommunikativ kontekstlar) so'zlarning semantikasi (ma'no tuzilishi va evolyutsiyasi)ga ta'sirini o'rganishga bag'ishlangan bo'lib, rus va o'zbek tillarining leksik semantikasini qiyosiy tahlil qiladi. [5]

Hasanova Parizod Jalil qizi ham rus va o'zbek tillaridagi kelishik kategoriyalarining semantik va grammatik mosliklarini qiyosiy nuqtai nazardan tahlil qilgan. Tadqiqotning asosiy maqsadi -ikki tilning morfologik kelishik tizimi o'rtasidagi struktur, funksional va ma'no farqlarini aniqlash hamda kelishiklarning sintaktik vazifalari va semantik yuklamalarini o'rganishdan iborat. Muallif rus tilining flektiv tuzilishi va o'zbek tilining agglutinatib tabiatini taqqoslab, kelishik qo'shimchalari orqali ifodalanadigan subyekt–obyekt munosabatlari, lokativ, instrumental, direktiv, qiyosiy va boshqa ma'no turlari qanday farqlanishini ko'rsatadi. Shuningdek, kelishik shaklining faqat grammatik vazifa emas, balki semantik rol ham tashishi, ya'ni kontekstdagi ma'no o'zgarishlariga ta'sir qilishi alohida ta'kidlanadi. O'zbek tilida kelishik ma'nolarining ko'p hollarda **morfologik shakl + postpozitsiya + kontekst** orqali ifodalanishini, rus tilida esa kelishik affiksatsiyasi yetakchi rol o'ynashini ko'rsatadi. Bu farqlar, ayniqsa, rus tilini o'rganuvchi o'zbek talabalari uchun boshqaruv, predikat–argument aloqalari va gap qurilishida muayyan qiyinchiliklarni yuzaga keltirishi aytiladi. Muallif kelishik kategoriyalarining semantik va grammatik uyg'unligini chuqur o'rganish ikki til o'rtasidagi tipologik tafovutlarni anglash, shuningdek til o'qitish metodikasini takomillashtirish uchun muhim ilmiy asos bo'lib xizmat qilishini ta'kidlaydi. [6]

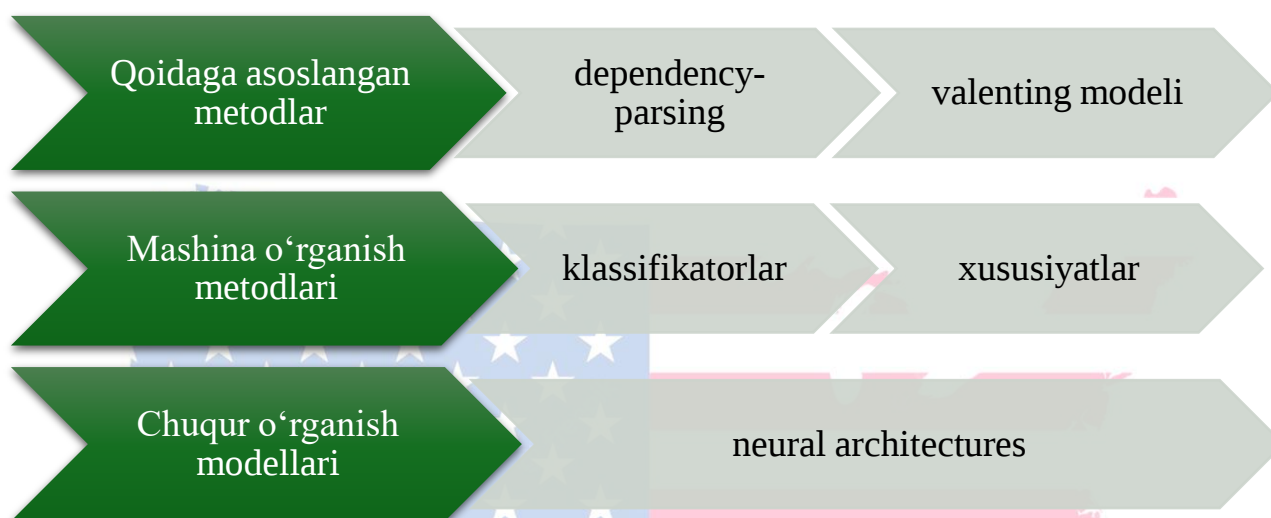
Ignatyev va Abdullayev hujjatlarni semantik rollar bo'yicha avtomatik va yarim-avtomatik belgilash (Semantic Role Labeling – SRL) jarayonini nazariy va amaliy nuqtai nazardan yoritadilar. Taqqidiotning maqsadi matndagi predikat–argument

tuzilmasini aniqlash, fe'ning talab qiladigan semantik rollarini (agent, patient, instrument, location, time va boshqalar) aniqlash va ularni ruxsat etilgan ontologik model asosida markirlashdan iborat. [7]

Mualliflar SRL jarayonining ikki asosiy komponentini ko'rsatadilar:

1. Predikatlarni aniqlash, ya'ni matnda harakatni bildiruvchi markaziy birliklarni topish;
2. Aktantlarni aniqlash, ya'ni predikat bilan bog'liq bo'lgan sintaktik bo'laklarning semantik rolini belgilash.

Ignatyev va Abdullayev turli usullar taqqoslanadi:



Mualliflar semantik rol belgilashning axborot izlash, matnni avtomatik tahlil qilish, konseptual xaritalar yaratish, hujjatlarni indekslash, sun'iy intellektga asoslangan bilimlar bazalarini boyitish kabi amaliy sohalaridagi ahamiyatini ham ko'rsatadilar. SRL texnologiyalari tabiiy tilni "yuzaki sintaksis"dan "chuqur semantik struktura"ga o'tkazuvchi asosiy mexanizm ekanini ta'kidlab, bu yondashuv zamonaviy NLP tizimlarining ajralmas qismi ekanini ta'kidlaydi.

Maqola o'zbek tilida ishlaydigan mashina o'rganish (ML) va tabiiy tilni qayta ishlash (NLP) tizimlari uchun zarur bo'lgan morfologik va sintaktik teglardan iborat yagona standart -**Uzbek TAGSET**ni yaratishga bag'ishlangan. Mualliflar o'zbek tili uchun umumiy qabul qilingan, barqaror va tizimli taglar majmuasi yo'qligi tufayli NLP modellari qurishda beqarorlik, mos kelmaslik va aniqlikning pastligi kabi muammolar yuzaga kelayotganini asoslab beradilar.

Tadqiqotda Uzbek TAGSET quyidagi asosiy mezonlar asosida ishlab chiqilgan:

- 1) Agglutinatив til xususiyatlarini to'liq qamrab olish;
- 2) Morfologik kategoriyalar (ot, fe'l, sifat, son, olmosh, ravish, yordamchi so'zlar, bog'lovchilar, yuklamalar, affikslar va b.) uchun aniq teglash tizimi yaratish;

- 3) Kelishiklar, shaxs-son, zamon, mayl, nisbat, egalik, ko'plik kabi grammatik kategoriyalarni formalizatsiya qilish;
- 4) Sintaktik rollar va gap bo'laklarining funksional tasnifini belgilash (subyekt, obyekt, hol, kesim va b.).

Mualliflar tagsetni yaratishda xalqaro tajriba -*Universal Dependencies (UD)*, *Penn Treebank*, *Russian National Corpus (RNC)* -kabi tizimlarni solishtirgan holda o'zbek tili uchun moslashtirilgan yechimlarni ishlab chiqadilar. TAGSETning asosiy ustunligi u ***o'zbek tilining murakkab morfologik tuzilishini kompyuter uchun standartlashtirish, shu orqali lemmatizatsiya, POS-tagging, dependency parsing, named entity recognition*** kabi vazifalarda ML modellarining samaradorligini sezilarli oshirishdir.

Xulosa sifatida, Zaynidinov va Sharipovlar tomonidan taklif etilgan Uzbek TAGSET o'zbek tili uchun chuqur morfologik annotatsiya tizimini shakllantirish, til texnologiyalarini rivojlantirish va korpus lingvistikasini standartlashtirish yo'lida muhim ilmiy-amaliy qadamlardan biri ekani ta'kidlanadi.

Maqolada o'zbek tili korpusi matnlarini qayta ishlashning zamonaviy kompyuter lingvistikasi asosidagi metod va texnologiyalari tizimli ravishda yoritilgan. Mualliflar o'zbek tilining agglutinativ tuzilishi, morfologik murakkabligi va variantlilik xususiyatlari tufayli korpusni qayta ishlash jarayonlari maxsus moslashtirilgan algoritmlarni talab qilishini asoslab beradilar.

Tadqiqotda quyidagi asosiy jarayonlar batafsil tahlil qilingan:

1.	<i>Matnlarni tozalash (Text Cleaning)</i>	Matnlarni qayta ishlash jarayonida ortiqcha belgilar chiqarib tashlanadi, kodlash bilan bog'liq xatolar tuzatiladi hamda noto'g'ri bo'linmalar normallashtiriladi.
2.	<i>Tokenizatsiya va segmentatsiya</i>	O'zbek tilida tokenlarni ajratish jarayoni murakkab bo'lib, bunda qo'shimchalar bilan almashinuvchi shakllar, qo'shma so'zlar, chiziqli birliklar hamda o'zgaruvchan fonetik variantlarni normallashtirish kabi jihatlar muhim ahamiyat kasb etadi.
3.	<i>Lemmatizatsiya va morfologik tahlil</i>	O'zbek tilida lemmatizatsiya jarayonini morfologik qoidalar, affikslar ketma-ketligi va paradigmatic o'zgarishlarga asoslangan modellar orqali amalga oshirishni ta'riflaydilar.
4.	<i>POS-tagging va grammatik teglash</i>	O'zbek tili uchun teglash jarayonida kontekstga asoslangan qoidalar, statistik modellar va neyron tarmoq yondashuvlari qiyosiy tahlil qilinadi.

(11th international scientific and practical conference)

5.	<i>Korpus arxitekturalari va annotatsiya</i>	Korpuslarni morfologik, sintaktik va semantik sathlarga ajratib teglash, ularni metama'lumotlar bilan boyitish hamda qidiruv tizimlari uchun optimallashtirish masalalari ham ko'tarilgan.
----	--	--

Maqola yakunida mualliflar o'zbek tili korpusini izchil qayta ishlash sun'iy intellekt modellarini yaratish, til resurslarini standartlashtirish, va o'zbek NLP ekotizimini rivojlantirishda hal qiluvchi ahamiyatga ega ekanini ta'kidlaydilar.

Tahlil qilingan tadqiqotlar semantik rollar nazariyasining predikat–argument munosabatlarini aniqlashda asosiy nazariy model ekanini ko'rsatadi. Rus va o'zbek tillarida fe'l valentligi hamda aktantlarning semantik-sintaktik vazifalari tilning tipologik xususiyatlari bilan bevosita bog'liq. Kelishik kategoriyalari, kauzativ konstruksiyalar va fe'l shakllarining mazmuniy imkoniyatlari semantik rollarning shakllanishida muhim rol o'ynaydi. Zamonaviy kompyuter lingvistikasi doirasida esa SRL, morfologik teglash va korpusga asoslangan modellar o'zbek tilida semantik tahlilni avtomatlashtirish va NLP tizimlarini rivojlantirish uchun muhim ilmiy-amaliy asosni ta'minlaydi.

FOYDALANILGAN ADABIYOTLAR:

1. **Кукатова О. А.** Типология семантических ролей актантов предикатов в современной лингвистической науке [Электронный ресурс]. – CyberLeninka, 2021.
2. **Кукатова О. А.** Особенности семантической роли каузатора в высказываниях с возвратными эмотивными глаголами // *Вестник лингвистики*. – 2012.
3. **Кукатова О. А.** Семантические роли и их интерпретация в рамках теории предикатно-аргументной структуры высказывания // *Foreign Philology Journal*. – 2016.
4. **Шамсиева Г.** Семантика и синтаксическая валентность в русском и узбекском языках // *AcademicsBook Conference Proceedings*. – 2023.
5. **Hasanova P.** Semantic and grammatical correlation of case categories in Russian and Uzbek // *Journal of Social Sciences*. – 2025.
6. **Игнатъев Н. А., Абдуллаев К. Д.** Разметка документов по семантическим ролям // *AIRI Journal*. – 2024.
7. **Ignatev N. A., Abdullaev K. D.** Semantic role labeling dataset for Uzbek // *AIRI Journal*. – 2024.