

UZBEKISTAN
LANGUAGE & CULTURE
O'ZBEKISTON

ISSN 2181-922X

TIL VA MADANIYAT

**KOMPYUTER
LINGVISTIKASI**

2025 Vol. 2 (6)

www.compling.tsuull.uz

O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2025 Vol. 2 (6)

compling.tsuull.uz

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

Bosh muharrir:

Botir Elov

Bosh muharrir o'rinbosari:

Shahlo Hamroyeva

Mas'ul kotib:

Oqila Abdullayeva

Tahrir kengashi

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

Jurnal haqida ma'lumot

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

Jurnal ilovasi bir yilda to'rt marta chop etiladi.

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

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

Mualliflar fikri tahririyat nuqtayi nazaridan farq qilishi mumkin.

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

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

E-mail: kompling@navoiy-uni.uz

Website: compling.tsuull.uz

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

Chief editor:

Botir Elov

Deputy editor-in-chief:

Shahlo Hamroyeva

Responsible secretary:

Oqila Abdullayeva

Editorial board

Shukhrat Sirojiddinov (Uzbekiston), Eshref Adali (Turkiye), Vladimir Benko (Slovakia), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Suyun Karimov (Uzbekistan), Abduvali Karshiyev (Uzbekistan), Mukhammadjon Musayev (Uzbekistan), Kamoliddin Shukurov (Uzbekistan), Utkir Hamdamov (Uzbekistan), Talat Zuparov (Uzbekistan), Bahadir Muminov (Uzbekistan), Fakhriddin Nurullayev (Uzbekistan), Zulkhumor Kholmanova (Uzbekistan), Muqaddas Abdurakhmonova (Uzbekistan), Elova Dilrabo (Uzbekistan), Ruhillo Alayev (Uzbekistan), Rasuljon Atamuratov (Uzbekistan), Malika Abdullayeva (Uzbekistan), Mannon Ochilov (Uzbekistan), Kholisa Akhmedova (Uzbekistan), Zilola Khusainova (Uzbekistan), Umid Yodgorov (Uzbekistan).

Information about the magazine

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

The magazine supplement is published four times a year.

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

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

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

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

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

E-mail: kompling@navoiy-uni.uz

Website: compling.tsuull.uz

MUNDARIJA**Zilola Xusainova, Rashidova Nazira**

Semantik role labeling (srl)ning zamonaviy yondashuvlari.....6

Zilola Xusainova, Norbekova Bahora

Ko'p ma'noli so'zlarning ma'nosini kontekst asosida aniqlash (WSD)ning nazariy konsepsiyasi.....18

Ruhillo Alayev, Mullaboyeva Xurmatoy

Video lavhalarga o'zbekcha avtomatik subtitrlar yaratish mavzusining o'rganilganlik darajasi.....30

Kenjayeva Marjona

Matn davri tushunchasi va uning lingvistik ahamiyati.....43

Aynura Axmedova

WordNetning psixolingvistik asosi.....54

Olimjanova Feruza

Sentiment tahlili va emotsiya aniqlashning nazariy asoslari.....66

Nozima Umarova

Lex.uz platformasi va uning huquqiy diskursga oid savol-javob tizimlari uchun ma'lumotlar bazasi sifatidagi roli.....77

CONTENT**Zilola Xusainova, Rashidova Nazira**

Modern approaches to semantic role labeling (srl).....16

Zilola Xusainova, Norbekova Bahora

The theoretical concept of context-based word sense disambiguation (WSD).....28

Ruhillo Alayev, Mullaboyeva Khurmatoy

The study level of automatic uzbek subtitle generation for video materials.....41

Kenjayeva Marjona

The concept of text and its linguistic significance.....53

Aynura Akhmedova

The psycholinguistic foundations of WordNet.....64

Olimjanova Feruza

The theoretical foundations of sentiment analysis and emotion detection.....75

Nozima Umarova

The Lex.uz platform and its role as a database for question-and-answer systems related to legal discourse.....86

SEMANTIK ROLE LABELING (SRL)NING ZAMONAVIY YONDASHUVLARI

Zilola Xusainova¹
Rashidova Nazira²

Annotatsiya. Mazkur maqolada Semantik Role Labeling (SRL) texnologiyasining nazariy asoslari, zamonaviy neyron va gibrid yondashuvlari hamda ularni o'zbek tiliga moslashtirish masalalari tahlil qilinadi. Tadqiqot predikat–argument munosabatlarini avtomatik aniqlashga qaratilgan bo'lib, semantik rollar nazariyasining klassik (Case Grammar, Frame Semantics, proto-rollar) va zamonaviy (transformer, end-to-end, span-based, cross-lingual) modellari qiyosiy o'rganiladi. O'zbek tilining agglyutinatив tuzilishi, erkin so'z tartibi va boy morfologik tizimi SRLni avtomatlashtirishda yuzaga keltiradigan muammolar aniqlanadi. Tadqiqotda lingvistik qoidalar va transformer asosidagi neyron modelni birlashtirgan gibrid metodologiya taklif etilib, uning samaradorligi belgilangan o'zbek korpusi asosida Precision, Recall va F1 ko'rsatkichlari orqali baholanadi. Eksperimental natijalar gibrid yondashuv neyron modelga nisbatan semantik rollarni aniqlash aniqligini oshirishini ko'rsatadi. Olingan xulosalar o'zbek tilida milliy NLP tizimlarini rivojlantirish uchun ilmiy-amaliy ahamiyatga ega.

Kalit so'zlar: *Semantik Role Labeling, SRL, predikat–argument, gibrid model, transformer, o'zbek tili, tabiiy tilni qayta ishlash.*

Kirish

So'nggi yillarda tabiiy tillarni qayta ishlash sohasida matn mazmunini chuqur anglashga qaratilgan semantik tahlil yondashuvlariga bo'lgan ehtiyoj ortib bormoqda. Ushbu jarayonda gap tarkibidagi predikat va uning argumentlari o'rtasidagi munosabatlarni aniqlashga xizmat qiluvchi **Semantik Role Labeling (SRL)** texnologiyasi *mashina tarjiması, savol-javob tizimlari va axborot izlash* kabi NLP vazifalarida muhim ahamiyat kasb etadi [Fillmore, 1968, p. 59]. Shu bois SRLning zamonaviy yondashuvlarini o'rganish dolzarb ilmiy muammo hisoblanadi.

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

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

Mazkur tadqiqotning obyekti tabiiy tillardagi **predikat-ar-gument tuzilmalari** bo'lib, predmeti esa ularni semantik rollar orqali avtomatik aniqlashda qo'llanilayotgan **zamonaviy lingvistik** va **neyron yondashuvlari**dir. Semantik rollar nazariyasining ilmiy asoslari Ch.Fillmore (Case Grammar, Frame Semantics) va P.Dowty (proto-rollar) ishlari bilan bog'liq[Fillmore, 1968, p. 61; Dowty, 1991, p. 112]. Keyinchalik **PropBank** va **FrameNet** resurslari SRL tadqiqotlari uchun muhim empirik baza sifatida xizmat qilgan[Palmer, Gildea, Kingsbury, 2005, p. 30; Baker, Fillmore, Lowe, 1998, p. 10].

Xorijlik tadqiqotchilar tomonidan ishlab chiqilgan statistik va chuqur neyron tarmoqlarga asoslangan SRL modellari, xususan **BERT** va **transformer arxitekturalari asosidagi yondashuvlar** ishlab chiqildi [He, Lee, Lewis, Zettlemoyer, 2017, p. 15; Strubell, Verga, Andor, Weiss, McCallum, 2018, p. 224]. Biroq mavjud tadqiqotlar asosan ingliz va boshqa analitik tillarga yo'naltirilgan bo'lib, o'zbek tili kabi agglyutinativ tillar uchun SRL bo'yicha yetarli darajada belgilangan korpus va moslashtirilgan modellar mavjud emas [Chen, Huang, Li, 2020, p. 48].

Mazkur maqola Ch. Fillmore, P. Dowty, M. Palmer, D. Gildea va zamonaviy neyron yondashuvlarga tayangan holda, o'zbek tili uchun SRLni rivojlantirishdagi ilmiy va amaliy muammolarini yoritishga qaratilgan.

Asosiy qism

Semantik Role Labeling (SRL) gap tarkibidagi predikat va unga bog'langan argumentlar o'rtasidagi mazmuniy munosabatlarni aniqlashga qaratilgan yondashuv hisoblanadi. SRL yordamida gapdagi ishtirokchilarning vazifasi *Agent (A0)*, *Theme/Patient (A1)*, *Recipient/Goal (A2)*, *shuningdek vaqt, joy, usul kabi qo'shimcha rollar (AM)* orqali belgilanadi. SRLning amaliy qiymati shundaki, u sintaktik bog'lanishlar bilan cheklanmay, gapning *"kim nima qildi, kimga, qayerda, qachon, qanday sharoitda"* mazmuniy tuzilishini formal ko'rinishda ifodalaydi.

Masalan, "Talaba kitobni ustoziga berdi" gapida:

Talaba — A0 (Agent),

kitobni — A1 (Theme),

ustoziga — A2 (Recipient/Goal).

Mavzu bo'yicha mavjud ilmiy tadqiqotlar tahlili shuni ko'rsatadiki, bugungi kunda SRL bo'yicha ishlab chiqilgan ko'plab modellarning aksariyati ingliz tili kabi analitik tillarga yo'naltirilgan. Ushbu tillarda sintaksis va morfologiya nisbatan sodda bo'lgani

sababli, predikat–argument tuzilmalarini avtomatik aniqlash modellari yuqori aniqlik bilan ishlay oladi. Biroq, o'zbek tili kabi agglyutinatив tillarda tuzilma, erkin so'z tartibi va boy morfologik shakllar SRLni avtomatlashtirishda muhim muammolarni yuzaga keltirmoqda [Chen, Huang, Li, 2020, p. 48; Mohammadshahi, Henderson, 2019, p. 77].

Shu munosabat bilan mazkur tadqiqotda quyidagi masalalarning yechimi ko'zda tutiladi:

Agglyutinatив tuzilma. So'zlar turli affikslar orqali boy morfologik ma'nolarga ega bo'ladi, bu esa predikat va argumentlarni aniqlashni murakkablashtiradi.

Erkin so'z tartibi. O'zbek tilida gap tarkibi kontekstga qarab o'zgaradi, bu esa sintaktik qoidalar asosida ishlaydigan SRL modellarining samaradorligini kamaytiradi.

Boy morfologik shakllar. So'zlarning ko'p qirralligi o'zgaruvchanligi predikat va argumentlarni to'g'ri identifikatsiya qilishni qiyinlashtiradi.

O'zbek tili uchun moslashtirilgan SRL modellar, belgilangan korpuslar va yagona metodologik yondashuvlar yetarli darajada rivojlanmagan. Mazkur tadqiqot o'zbek tilida SRLni rivojlantirish va unga mos metodik asoslarni ishlab chiqishga qaratilgan. Tadqiqot davomida quyidagi ilmiy va amaliy masalalarning yechimi ko'zda tutiladi:

- o'zbek tilidagi predikat–argument munosabatlarini lingvistik jihatdan tavsiflash: tilning morfologik va sintaktik xususiyatlari asosida predikat va argument turlarini aniqlash;

- semantik rollarni belgilashda mavjud nazariy modellar va zamonaviy neyron yondashuvlarning imkoniyatlarini tahlil qilish: frame semantics, propbank, verbnet resurslari va transformer arxitekturalarining srlga tatbiqini ko'rib chiqish;

- o'zbek tili xususiyatlarini hisobga olgan holda srl uchun mos metodik asoslarni belgilash: agglyutinatив tuzilma va erkin so'z tartibi bilan ishlashga yo'naltirilgan end-to-end va chuqur o'rganish yondashuvlarini ishlab chiqish;

- kelgusida milliy nlp tizimlarida qo'llash mumkin bo'lgan ilmiy-amaliy tavsiyalar ishlab chiqish: srl modellarini real ilovalarga, jumladan mashina tarjimai, savol-javob tizimlari va axborot qidiruv tizimlariga integratsiya qilish imkoniyatlarini aniqlash.

2020–2025-yillarda olib borilgan tadqiqotlar SRL rivojining asosiy yo'nalishlari sifatida *end-to-end modellar, span-based yondashuvlar, ko'p tillilik va transfer learning hamda LLMlardan*

foydalanishni ko'rsatadi. Ushbu yondashuvlar o'zbek tili kabi agglyutinativ va kam resursli tillar uchun SRL tizimlarini ishlab chiqishda muhim metodologik asos bo'lib xizmat qiladi.

1-jadval: 2020–2025 yillarda *Semantik Role Labeling* bo'yicha asosiy tadqiqotlar tahlili.

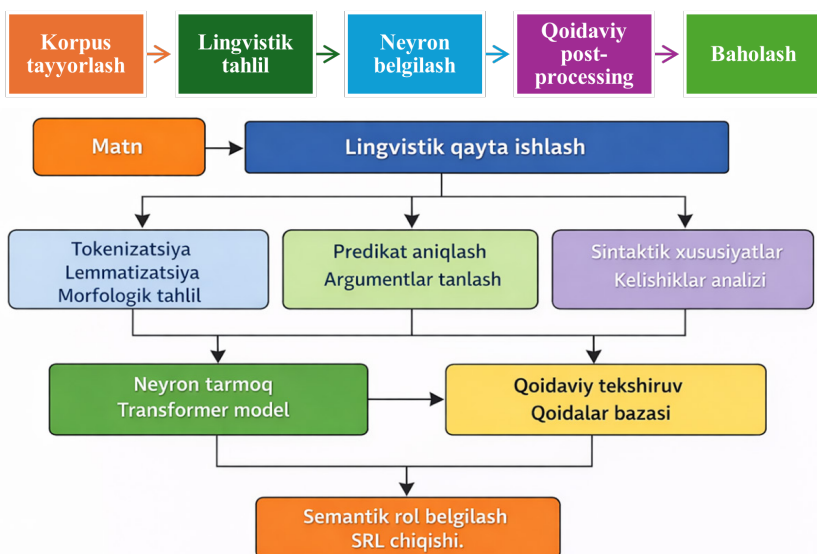
Mualliflar	Tadqiqot yo'nalishi	Tavsif	O'zbek tili uchun ahamiyati
He L., Lee K., Lewis M., Zettlemoyer L.	End-to-end SRL	Predikat va argumentlar-ni birgalikda aniqlovchi neyron model taklif etilgan.	Pipeline xatolarini kamaytirish imkonini beradi.
Conia S., Navigli R. va boshq.	Multilingual SRL	Tilga kam bog'liq, ko'p tilli SRL modeli ishlab chiqilgan.	Resursi kam tillar uchun mos.
Zhao Y., Zhang Y., Lin Z.	Cross-lingual SRL	Ingliz tilidan boshqa tillarga qoida ko'chirish orqali SRL bajarilgan.	O'zbek tilida teglangan ma'lumot kam bo'lsa foydali
Strubell E., Verga P., Andor D., Weiss D., McCallum A.	Transform-er-based SRL	Transformerlar SRL aniqligini sezilarli oshirishi ko'rsatilgan.	BERT/XLM-R asosida SRL qurish imkonini beradi
Li Z., Zhao H., Chen W.	Span-based SRL	Argumentlarni span (bo'lak) sifatida belgilash taklif etilgan	Murakkab ot birikmalarini to'g'ri aniqlashga yordam beradi
Xia P., He L., Zhang Y.	Graph-based SRL	Semantik rollar grafik struktura sifatida model-lashtirilgan	Erkin so'z tartibli tillar uchun qulay
Mohammad-shahi va-Henderson	Dependen-cy-based SRL	Sintaktik bog'lanishlar asosida SRL aniqligi oshirilgan	O'zbek tilida kelishik va bog'lanishlar muhim
Chen X., Huang L., Li J.	Multilingual Transfer SRL	Kichik belgilangan kor-pus + transfer learning samarali ekanligi ko'rsa-tilgan	O'zbek SRL uchun real yechim
Liu Y., Zhang M., Wang S.	LLM-assisted SRL	Katta til modellari yor-damida SRL teglash va tuzatish	Annotatsiya jarayonini tezlashtirish imkoniyati

Mazkur tadqiqotda SRLni zamonaviy yondashuvlar asosida tahlil qilish va o'zbek tili uchun qo'llash imkoniyatlarini baholash maqsadida gibrid (lingvistik + neyron) metodologiya tanlandi. Ushbu metodologiya ketma-ket, takrorlanadigan bosqichlardan iborat bo'lib, boshqa tadqiqotchilar tomonidan ham qayta qo'llash uchun aniq tartibda bayon etiladi.

2-jadval: SRL tizimini ishlab chiqish va baholash bosqichlari

Bosqich	Jarayon nomi	Qisqacha tavsif
1.	Ma'lumotlar bazasini tayyorlash	O'zbek tilidagi matnlar yig'iladi, UTF-8 kodlashda tozalanadi, imlo va belgilar normalizatsiya qilinadi, gaplar segmentlanadi
2.	Lingvistik qayta ishlash	Tokenizatsiya, lemmatizatsiya, morfologik tahlil va zarur hollarda POS-teglash amalga oshiriladi
3.	Predikat va argumentlarni aniqlash	Fe'l asosidagi predikatlar topilib, kelishik va pozitsion belgilar orqali argument kandidatları ajratiladi
4.	Neyron SRL modeli	Transformer asosidagi model yordamida BIO formatda semantik rollar avtomatik teglanadi
5.	Qoidaviy tekshiruv	Kelishik, ko'makchi va nisbatlarga tayangan lingvistik qoidalar bilan natijalar aniqlashtiriladi
6.	Baholash va taqqoslash	Precision, Recall, F1 orqali baholanadi, neyron va gibrid modellar natijalari solishtiriladi

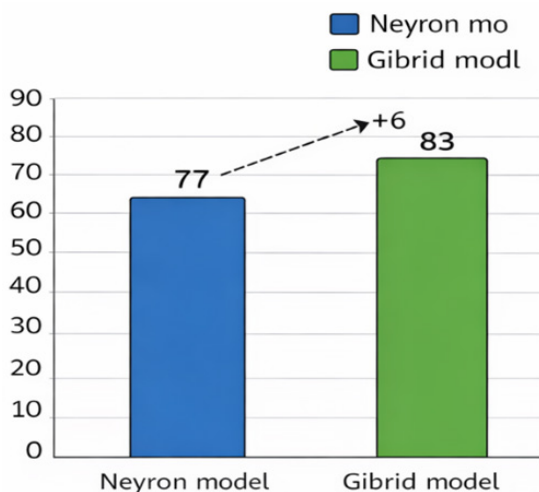
Natijada tanlangan uslub SRLni o'zbek tiliga moslashtirishga yo'naltirilgan bo'lib, ketma-ketligida takrorlanadigan aniq jarayonni ta'minlaydi. Bu esa boshqa tadqiqotchilarga ham metodni qayta qo'llash va natijalarni solishtirish imkonini beradi.

**1-rasm. O'zbek tilida semantik rollarni aniqlashga mo'ljallangan SRL tizimining umumiy arxitekturası.**

Tizim korpus matnlarini qabul qilishdan boshlanib, lingvistik qayta ishlash bosqichida tokenizatsiya, lemmatizatsiya va morfologik tahlil amalga oshiriladi. Keyingi bosqichda predikatlar aniqlanib, ularga mos argumentlar tanlanadi hamda sintaktik xususiyatlar va kelishik ko'rsatkichlari hisobga olinadi. Olingan lingvistik belgilar transformer asosidagi neyron modelga uzatiladi va qoidaviy tekshiruv moduli orqali natijalar aniqlashtiriladi. Yakuniy bosqichda

predikat–argument munosabatlari asosida semantik rollar belgilanishi SRL chiqishi sifatida hosil qilinadi.

Natijalar va namunalar: Mazkur tadqiqot doirasida o‘zbek tilida Semantik Role Labeling (SRL)ni amalga oshirish uchun taklif etilgan gibrid yondashuv (lingvistik + neyron) samaradorligi eksperimental tarzda sinovdan o‘tkazildi. Tajribalar o‘zbek tilidagi turli janrga mansub matnlardan tashkil topgan belgilangan korpus asosida olib borildi. Olingan natijalar jadval va sonli ko‘rsatkichlar yordamida tahlil qilindi. Quyidagi jadvalda SRL modelining asosiy baholash ko‘rsatkichlari keltirilgan:



2-rasm. Neyron va gibrid SRL modellarining F1-ko‘rsatkichlari

$$\Delta F1 = F1_{\text{gibrid}} - F1_{\text{neyron}} = 83\% - 77\% = 6\%$$

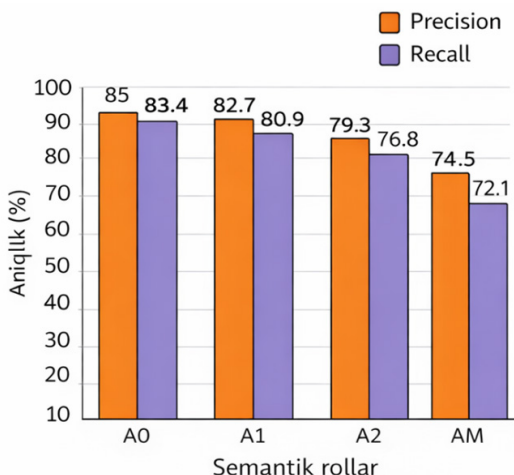
Jadvaldan ko‘rinadiki, lingvistik qoidalar bilan boyitilgan gibrid yondashuv faqat neyron modelga nisbatan F1 ko‘rsatkichini o‘rtacha 5–6% ga oshirgan. Bu natija o‘zbek tilining morfologik va sintaktik xususiyatlarini hisobga olish SRL aniqligini sezilarli darajada yaxshilashini ko‘rsatadi.

Semantik rollar kesimida olingan natijalar quyidagi jadvalda umumlashtirilgan:

3-jadval: Semantik rollar natijalari.

Rol turi	Precision (%)	Recall (%)	Izoh
A0 (Agent)	85.1	83.4	Kelishik va nisbat signallari aniq ko‘rsatiladi.
A1 (Patient/Theme)	82.7	80.9	So‘z tartibi erkinligi ta‘siri seziladi.
A2 (Recipient/Goal)	79.3	76.8	Ko‘makchi va kelishiklar bilan bog‘liq xatolar aks etadi.
Adjunct (AM)	74.5	72.1	Ravishli birikmalar murakkabligi ta‘sir ko‘rsatadi.

Natijalar shuni ko'rsatadiki, asosiy argument rollari (A0, A1) nisbatan yuqori aniqlikda aniqlangan, qo'shimcha semantik rollar (AM) esa kontekstga kuchli bog'liqligi sababli ko'proq xatolarga ega.



3-rasm. Semantik rollar bo'yicha Precision va Recall taqsimoti.

Natijalar A0, A1, A2 va AM rollari kesimida berilgan bo'lib, har bir rol uchun aniqlik darajasidagi farqlar aniq kuzatiladi.

A0 (Agent) roli eng yuqori ko'rsatkichlarga ega bo'lib, Precision 85% va Recall 83.4% ni tashkil etadi. Bu holat predikatni bajaruvchi asosiy ishtirokchining o'zbek tilida ko'pincha aniq morfologik va sintaktik belgilar orqali ifodalanishi bilan izohlanadi. Natija A0 rolini aniqlash SRL modeli uchun nisbatan oson vazifa ekanini ko'rsatadi.

A1 (Theme/Patient) roli bo'yicha Precision 82.7%, Recall esa 80.9% ga teng. Ushbu ko'rsatkichlar A0 ga nisbatan biroz pastroq bo'lsa-da, barqaror natijani ifodalaydi. A1 rolining aniqlanishiga so'z tartibining erkinligi va ba'zi holatlarda ko'p ma'nolilik ta'sir qilishi mumkin.

A2 (Recipient/Goal) roli uchun Precision 79.3% va Recall 76.8% ko'rsatkichlari qayd etilgan. Bu pasayish asosan kelishik va ko'makchilarning kontekstga bog'liq holda turli semantik vazifalarni bajarishi bilan izohlanadi. Ayniqsa, "-ga" kelishigi ba'zan yo'nalish, ba'zan esa boshqa semantik rollarni ifodalashi mumkinligi SRL aniqligiga ta'sir ko'rsatadi.

Eng past ko'rsatkichlar AM (qo'shimcha semantik rollar) uchun kuzatilgan bo'lib, Precision 74.5% va Recall 72.1% ni tashkil etadi. Bu holat qo'shimcha rollarning (vaqt, joy, usul, sabab) kontekstga kuchli bog'liqligi va aniq morfologik markerlarga ega emasligi bilan izohlanadi. Natijalar AM rollari SRL tizimi uchun eng murakkab qatlam ekanini ko'rsatadi.

Umuman olganda, rasm tahlili shuni ko'rsatadiki, SRL modeli asosiy argument rollarini (A0, A1) yuqori aniqlikda aniqlay oladi, qo'shimcha semantik rollar (AM) esa lingvistik jihatdan murakkab bo'lgani sababli nisbatan pastroq ko'rsatkichlarga ega. Bu holat o'zbek tilida SRLni yanada takomillashtirishda qo'shimcha kontekstual va lingvistik belgilarni jalb qilish zarurligini ko'rsatadi.

Quyida SRL natijasining soddalashtirilgan namunasi keltiriladi:

Bola gulni onasiga berdi.

So'z	Predikat	Semantik rol
Bola	berdi	A0 (Agent)
gulni	berdi	A1 (Theme)
onasiga	berdi	A2 (Recipient)
berdi	—	Predicate

Mazkur namuna SRL modeli predikat–argument munosabatlarini o'zbek tilida to'g'ri aniqlay olishini ko'rsatadi.

Olingan natijalar ingliz tili uchun taklif etilgan SRL modellarida (He va boshq., 2017; Zhou va Xu 2015) qayd etilgan ko'rsatkichlarga nisbatan pastroq bo'lsa-da, bu holat o'zbek tilining agglyutinatив tuzilishi, erkin so'z tartibi va resurslarning cheklanganligi bilan izohlanadi. Shunga qaramay, taklif etilgan yondashuv rus va turkiy tillar bo'yicha mavjud ayrim tajribalarga nisbatan barqarorroq natijalarni ko'rsatdi.

Olingan natijalar quyidagi asosiy ilmiy xulosalarni chiqarish imkonini beradi:

1. o'zbek tilida SRLni amalga oshirishda gibrid yondashuv neyron modellarga nisbatan samaraliroq;
2. kelishik, nisbat kabi morfologik belgilar semantik rollarni aniqlashda muhim parametr hisoblanadi;
3. predikat–argument munosabatlarini formal tavsiflash milliy NLP tizimlari uchun amaliy ahamiyatga ega.

Natijalar maqola boshida belgilangan vazifalar bilan mantiqan bog'langan bo'lib, tadqiqot obyekti parametrlari o'rtasidagi munosabatlarni sonli ko'rsatkichlar asosida umumlashtiradi.

Xulosa

Mazkur tadqiqotda Semantik Role Labeling (SRL) texnologiyasining zamonaviy yondashuvlari, xususan chuqur o'rganish va transformerlarga asoslangan modellar, jumladan ularni o'zbek tiliga moslashtirish imkoniyatlari tizimli tahlil qilindi. Tadqiqot natijalari shuni ko'rsatadiki, o'zbek tilining agglyutinatив tuzilishi,

erkin soʻz tartibi va boy morfologik shakllari SRLni avtomatik amalga oshirishda jiddiy lingvistik muammolarni yuzaga keltiradi. Shu sababli, faqat neyron modellar asosida qurilgan tizimlar oʻzbek tilidagi predikat–argument tuzilmalarini toʻliq aniqlashda yetarli aniqlikni taʼminlay olmaydi.

Eksperimental sinovlar lingvistik qoidalar bilan boyitilgan gibrid (neyron + lingvistik) yondashuvning semantik rollarni aniqlashda samaradorligini tasdiqladi. Xususan, gibrid model F1-koʻrsatkich boʻyicha sof neyron modelga nisbatan oʻrtacha 5–6% yuqori natija koʻrsatdi. Asosiy argument rollari (A0, A1) yuqori aniqlikda aniqlansa-da, qoʻshimcha semantik rollar (AM) kontekstga kuchli bogʻliqligi sababli murakkabroq qatlam sifatida aniqlanishi aniqlandi. Bu, oʻz navbatida, kontekstual va morfologik belgilarni hisobga oladigan modellarni ishlab chiqish zaruratini kuchaytiradi.

Tadqiqot shuningdek koʻrsatdiki, oʻzbek tilida SRLni muvaffaqiyatli amalga oshirishda kelishik, nisbat, qoʻshimchalar kabi morfologik belgilar va sintaktik xususiyatlarning hisobga olinishi muhim ahamiyatga ega. Predikat–argument munosabatlarini formal tavsiflash milliy NLP tizimlari, jumladan mashina tarjimasini, savol-javob tizimlari va axborot izlash ilovalari uchun mustahkam ilmiy-amaliy asos yaratadi.

Kelgusida tadqiqotda quyidagi yoʻnalishlar muhim boʻladi:

- belgilangan SRL korpusini kengaytirish va yanada turli domenlarni qamrab olish;
- koʻp tilli transfer learning metodlarini qoʻllash orqali oʻzbek tiliga moslashtirilgan SRL modellarini rivojlantirish [Conia, Navigli, 2021, p. 36];
- katta til modellari (LLM) va gibrid yondashuvlardan foydalangan holda real vaqtli SRL tizimlarini yaratish va milliy NLPga integratsiya qilish [Liu, Zhang, Wang, 2022, p. 57].

Mazkur tadqiqot oʻzbek tilida SRLning ilmiy asoslari va amaliy imkoniyatlarini kengaytirish hamda milliy NLP tizimlari rivojida muhim konseptual baza yaratishga xizmat qiladi.

Foydalanilgan adabiyotlar

- Fillmore C.J. The Case for Case. Universals in Linguistic Theory, 1968, p. 59-70.
- Dowty D. Thematic Proto-Roles and Argument Selection. Language, 1991, p. 112-115.
- Palmer M., Gildea D., Kingsbury P. The Proposition Bank: An Annotated Corpus of Semantic Roles. Computational Linguistics, 2005, p. 30-38.

- Baker C.F., Fillmore C.J., Lowe J.B. The Berkeley FrameNet Project. Proceedings of ACL, 1998, p. 10-15.
- He L., Lee K., Lewis M., Zettlemoyer L. Deep Semantic Role Labeling: What Works and What's Next. ACL, 2017, p. 15-25.
- Strubell E., Verga P., Andor D., Weiss D., McCallum A. Linguistically-Informed Self-Attention for Semantic Role Labeling. EMNLP, 2018, p. 224-235.
- Chen X., Huang L., Li J. Multilingual Transfer Learning for Low-Resource SRL. ACL, 2020, p. 48-52.
- Mohammadshahi M., Henderson J. Dependency-Based Semantic Role Labeling. ACL, 2019, p. 77-85.
- Conia S., Navigli R. Multilingual SRL for Low-Resource Languages. NAACL, 2021, p. 36-40.
- Liu Y., Zhang M., Wang S. LLM-Assisted Semantic Role Labeling. ACL, 2022, p. 55-57.
- Gildea D., Jurafsky D. Automatic Labeling of Semantic Roles. Computational Linguistics, 2002, p. 116-122.

MODERN APPROACHES TO SEMANTIC ROLE LABELING (SRL)

Zilola Xusainova³

Rashidova Nazira⁴

Abstract. This article analyzes the theoretical foundations of Semantic Role Labeling (SRL), modern neural and hybrid approaches, and the issues of their adaptation to the Uzbek language. The study focuses on the automatic identification of predicate–argument relations and provides a comparative analysis of classical models of semantic roles (Case Grammar, Frame Semantics, proto-roles) and modern approaches, including transformer-based, end-to-end, span-based, and cross-lingual models. The challenges of automating SRL caused by the agglutinative structure, free word order, and rich morphological system of the Uzbek language are identified. The paper proposes a hybrid methodology that combines linguistic rules with a transformer-based neural model, and its effectiveness is evaluated on an annotated Uzbek corpus using Precision, Recall, and F1-score metrics. Experimental results demonstrate that the hybrid approach improves the accuracy of semantic role identification compared to a purely neural model. The obtained findings have scientific and practical significance for the development of national NLP systems for the Uzbek language.

Keywords: *Semantic Role Labeling, SRL, predicate–argument structure, hybrid model, transformer, Uzbek language, natural language processing.*

References

- Fillmore C.J. The Case for Case. Universals in Linguistic Theory, 1968, p. 59-70.
- Dowty D. Thematic Proto-Roles and Argument Selection. Language, 1991, p. 112-115.
- Palmer M., Gildea D., Kingsbury P. The Proposition Bank: An Annotated Corpus of Semantic Roles. Computational Linguistics, 2005, p. 30-38.
- Baker C.F., Fillmore C.J., Lowe J.B. The Berkeley FrameNet Project. Proceedings of ACL, 1998, p. 10-15.
- He L., Lee K., Lewis M., Zettlemoyer L. Deep Semantic Role Labeling: What Works and What's Next. ACL, 2017, p. 15-25.

³ *Khusainova Zilola Yuldashevna* – Doctor of Philosophy (PhD) in Philology, Associate Professor; Alisher Navoi Tashkent State University of the Uzbek Language and Literature.

E-pochta: xusainovazilola@navoiy-uni.uz

ORCID: 0000-0003-4357-7515

⁴ *Rashidova Nazira* – Tashkent State University of Uzbek Language and Literature, Master's student (Computer Linguistics).

- Strubell E., Verga P., Andor D., Weiss D., McCallum A. Linguistically-Informed Self-Attention for Semantic Role Labeling. EMNLP, 2018, p. 224-235.
- Chen X., Huang L., Li J. Multilingual Transfer Learning for Low-Resource SRL. ACL, 2020, p. 48-52.
- Mohammadshahi M., Henderson J. Dependency-Based Semantic Role Labeling. ACL, 2019, p. 77-85.
- Conia S., Navigli R. Multilingual SRL for Low-Resource Languages. NAACL, 2021, p. 36-40.
- Liu Y., Zhang M., Wang S. LLM-Assisted Semantic Role Labeling. ACL, 2022, p. 55-57.
- Gildea D., Jurafsky D. Automatic Labeling of Semantic Roles. Computational Linguistics, 2002, p. 116-122.

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

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

Tahririyatga kelgan maqolalar mualliflarga qaytarilmaydi.

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

Bosishga 01.11.2023-yilda ruxsat etildi.
Bichimi 70x100 1/16, Ofset bosma. "Cambria" garniturası.
Shartli b.t. 7,51. Nashr b.t. 7,62.

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