

ISSN 2181-922X

LANGUAGE & CULTURE

UZBEKISTON O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER
LINGVISTIKASI

2024 Vol. 1 (6)

www.compling.tsuull.uz

UZBEKISTAN

ISSN 2181-922X

O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2024 Vol. 1 (6)

compling.tsuull.uz

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

Bosh muharrir:

Botir Elov

Bosh muharrir o‘rinbosari:

Shahlo Hamroyeva

Mas‘ul kotib:

Oqila Abdullayeva

Tahrir kengashi

Shuhrat Sirojiddinov (O‘zbekiston), Eshref Adali (Turkiya), [Viktor Zaxarov] (Rossiya), Vladimir Benko (Slovakiya), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat O‘rxun (Turkiya), Suyun Karimov (O‘zbekiston), Abduvali Qarshiyev (O‘zbekiston), Muxammadjon Musayev (O‘zbekiston), Kamoliddin Shukurov (O‘zbekiston), O‘tkir Hamdamov (O‘zbekiston), Tal‘at Zuparov (O‘zbekiston), Bahodir Mo‘minov (O‘zbekiston), Faxriddin Nurullayev (O‘zbekiston), Zulxumor Xolmanova (O‘zbekiston), Muqaddas Abdurahmonova (O‘zbekiston), Habibulla Madatov (O‘zbekiston), Azizaxon Raxmanova (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).

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 Navo'i Tashkent State University of the Uzbek Language and Literature

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

Editorial board

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

Information about the magazine

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

The magazine supplement is published four times a year.

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

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

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

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

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

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

MUNDARIJA

Firuza Nurova

Jahon kompyuter lingvistikasida bir necha soʻz-shakldan iborat leksemalarga ishlov berish tajribasi haqida6

Iqbola Xolmonova

Oʻzbek-turk parallel korpusi uchun matnlar tokenizatsiyasi masalasi.....21

Botir Elov, Shahlo Hamroyeva, Marjona Hamroqulova

Nlpda semantik teglash usullari.....32

Oqila Abdullayeva

Oʻzbek tili matnlarida sintaktik teg va teglash masalasi.....46

Aziza Raxmanova

Modern methods of teaching the linguistic basics of the uzbek and english languages.....58

Nargiza Shamiyeva

The main principals of creating a bilingual thesaurus for the uzbek language.....66

Zarnigor Khayatova

Uzbek paraphrasing software: how your words get a makeover (without losing their meaning!).....78

CONTENT

Firuza Nurova

About the experience of processing lexemes consisting of several word forms in world computer linguistics.....19

Iqbola Xolmonova

The issue of text tokenization for the uzbek-turkish parallel corpus.....31

Botir Elov, Shahlo Hamroyeva, Marjona Hamroqulova

Semantic tagging methods in nlp.....44

Oqila Abdullayeva

The issue of syntactic tags and tagging in uzbek language texts.....56

Azizaxon Raxmanova

O'zbek va ingliz tillarining lingvistik asoslarini o'qitishning zamonaviy metodlari.....64

Nargiza Shamiyeva

O'zbek tili uchun bilingval tezaurus yaratishning asosiy tamoyillari.....76

Zarnigor Xayatova

O'zbekcha parafrazlash dasturi: sizning so'zlaringiz qanday o'zgaradi? (ma'noni saqlagan holda).....85

O‘ZBEK TILI MATNLARIDA SINTAKTIK TEG VA TEGGLASH MASALASI

Oqila Abdullayeva¹

Annotatsiya. Tabiiy tilni qayta ishlash bosqichlaridan biri sintaktik tahlilning vazifasi gap qurilmalarini tahlil qilishdir. Korpus lingvistikasida parsing deb ataluvchi jarayon gap va gaplarning tuzilishini, gapdagi leksik birliklarning sintaktik munosabatini aniqlaydi. Mazkur maqolada o‘zbek tilidagi gaplarni sintaktik teglash, tegset tanlash masalasi o‘rganildi, o‘zbek tilidagi misollar orqali tahlil qilindi.

Kalit so‘zlar: *sintaktik teglash, teg, parsing, parser, treebank, tarkibiy qismlar tahlili, tobelikni tahlil qilish.*

Kirish

Bugungi kunda inson tilini texnologiya bilan o‘zaro ta’sirini tushinishda tabiiy tilni qayta ishlash uchun kerakli dasturlarni yaratish ehtiyoji mavjud. Buning uchun esa katta hajmli ma’lumotlar to‘plamini tayyorlash, matnlarni avtomatik tahrir va tahlil qiluvchi tabiiy tilni qayta ishlash (NLP)ning vazifalari uchun tilning lingvistik ta’minotini ishlab chiqish muhim hisoblanadi. Tabiiy tilni tahlil qilish bosqichlari sifatida grafematik, morfologik (leksik), sintaktik, semantik, pragmatik tahlil qilish kabilar ko‘rsatiladi. Ba’zi tadqiqotlarda diskurs tahlil va fonologik tahlil ham alohida kiritiladi. Sintaktik tahlil til korpuslarida gapning sintaktik tuzilishini aniqlash, gap bo‘laklarini va ularning vazifalarini belgilash, o‘zaro bog‘lanish sintaktik munosabatini tahlil qilishdan iborat. NLPda avtomatik sintaktik tahlil - parsing atamasi bilan ham yuritiladi. Tabiiy tilni qayta ishlash kontekstida parsing atamasi so‘zlar ketma-ketligidan tuzilgan jumla va uning sintaktik tuzilmasini aniqlovchi avtomatik tahlil qilish jarayonini anglatadi [Nederhof, Satta, 2010]. Sintaktik tahlil yoki parsing - gap yoki matn qismini tahlil qilish va uning grammatik tuzilishini aniqlash jarayonidir. Bunga jumla orasidagi tarkibiy so‘z

¹Abdullayeva Oqila Xolmo‘minovna – filologiya fanlari bo‘yicha falsafa doktori. Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti doktoranti (DSc)

E-pochta: abdullayeva.oqila@navoiy-uni.uz

ORCID: 0000-0002-2524-4832

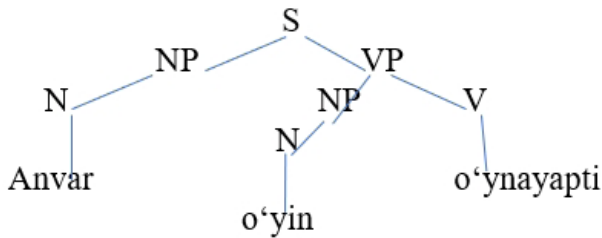
birikmalari va bog'liqliklarni aniqlash, shuningdek, har bir so'zning gapdagi rollarini (masalan, ega, kesim va obyekt (2-darajali bo'laklar nazarda tutilyapti) aniqlash kiradi [<https://github.com/GoktugOcal/turkish-syntactic-parser>]. Sintaktik parsing gapning grammatik tuzilishi bilan birga, gapning to'g'ri yoki noto'g'ri tuzilganligini, gap chegaralarini va so'zlarning o'zaro birikish munosabatlarini tahlil qiladi. Sintaktik tahlil bilan bog'liq tadqiqotlar o'rganilganda treebank atamasini uchratish mumkin. Treebanklar til korpuslarida matni sintaktik va semantik tahlil qilish usulidir. 80-yillarda Jefri Lich tomonidan kiritilgan bu termin matni ham sintaktik, ham semantik xususiyatlarini kompozitsion tahlil tasviri daraxt strukturasi-ga o'xshatilgan [Sampson, 2003].

Asosiy qism

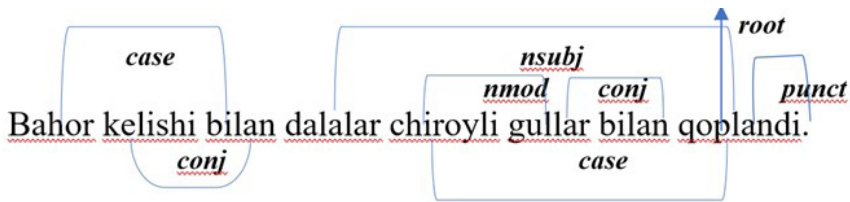
Sintaktik parsing boshqa tahlil usullari kabi lingvistik va kognitiv nazariyalarning natijalaridan mashinali o'qitishda murakkab hisoblanadi [Henderson, 2010]. Sintaktik parsingda 2 xil yondoshuv keng tarqalgan: **Constituency Parsing** (tarkibiy qismlar tahlili) va **Dependency Parsing** (tobelikni tahlil qilish).

Tarkibiy qismlar tahlili (Constituency parsing) - berilgan gapni o'tli birikma (NP) va fe'lli birikma (VP) kabi tarkibiy qismlarga ajratadigan tahlil daraxtlarini quradi. U jumlaning iyerarxik tuzilishini ko'rsatadi va so'zlarning kattaroq grammatik birliklarga qanday joylashishini ko'rsatadi. Gapni tarkibiy qismlarga ajratib tahlil qilish (constituency parsing) CFGga asoslanadi va tahlil daraxti kichik birikmalarga bo'lingan gapni o'z ichiga oladi. Gapdagi har bir lingvistik birlik asosiy tuguni va POSiga ega terminal tugunga ega [<https://www.upgrad.com/blog/dependency-parsing-in-nlp/>]. Ingliz tilida gap tahlili umumiy *Sentence = S = Noun Phrase (NP) + Verb Phrase (VP) + Preposition Phrase (PP)* tuzilishida bo'ladi. Masalan, "John is playing a game" kabi.

Sintaktik tahlil tillarning tipologik xususiyatiga va grammatik tuzilishiga ko'ra farqli bo'ladi. O'zbek tilida gap qurilishi *Sentence = S = Noun Phrase (NP) + Preposition Phrase (PP) + Verb Phrase (VP)* tuzilishida bo'ladi. O'zbek tilida keltirilgan jumla misolida tahlil qilamiz. Masalan, "Anvar o'yin o'ynayapti" gapi quyidagi pars daraxtini hosil qiladi:



Dependency parsing - tobelikni tahlil qilish bo'lib, gapdagi har bir leksik birlik boshqasiga bog'liq bo'lgan daraxt tuzilmasini qurish orqali so'zlar orasidagi grammatik aloqalarni tasvirlaydi. U tez-tez ma'lumot olish va mashina tarjimasi kabi vazifalarda qo'llaniladi, chunki u ega-kesim-obyekt aloqalarining munosabatlariga qaratilgan [What is parsing in NLP: Its types and techniques. <https://intellipaat.com/>]. Dependency Parsing (DP) atamasi gapning grammatik tuzilishini aniqlash uchun birikmalar orasidagi bog'liqlikni tekshirish jarayonini bildiradi. Gap, asosan, shunga asoslanib ko'p kichik bog'liqlik bo'limlarga bo'linadi. Bu jarayon gapdagi har bir li-soniy birlik o'rtasida to'g'ridan-to'g'ri bog'liqlik bor degan taxminga asoslanadi. Quyidagi sxemada jumlaning bog'liqlik tuzilishi tushuntiriladi:



Sxemadagi teglar qaysi bog'liqlikni tahlil qilishini aniqlaymiz: *nsubj* – nominal ega (ega va kesim o'rtasidagi munosabat); *root* – o'zak (grammatikada gapning o'zagiga ishora qiladi); *obj* – obyektga yo'naltirish (kesimni obyektga yo'naltirish); *det* – determinator (aniqlovchi); *nmod* – nominal modifikator (ot yoki otli birikmani o'zgartiruvchisi; holat munosabati qaramligi); *case* – holatni belgilash (holatni munosabati, qo'shimchali holatlar).

Til korpuslarida matnlarni lingvistik annotatsiyalash uchun maxsus tegsetlar ishlab chiqilgan. Sintaktik parsing uchun tillarda maxsus teglar belgilangan. Ba'zi tadqiqotlarda esa maxsus ishlab chiqilgan universal tegsetlardan foydalaniladi. Tarkibiy qismlarni tahlil qilish (CP) gapni yaxshiroq tahlil qilish uchun mantiqiy birikmalarga ajratib, gap orasidagi obyektни aniqlaydi va semantik munosabatlarida yordam beradi. CP tahlilida Penn Treebank POS teglaridan foydalanib, gapning pars daraxtini quradi, gap bo'lak-

larini aniqlab, vazifalarini belgilaydi. CP Penn Treebank tomonidan ishlab chiqilgan eng ko'p uchraydigan 9 ta birikma toifani ichiga oluvchi universal birikmalar to'plamidan foydalanadi. Bular otli birikma (NP), fe'lli birikma (VP), sifatlovchi birikma (AJP), ergash gapdagi qo'shimcha so'z birikmasi (AVP), predlogli birikma (PP), Gap yoki ergashgan gap (S), bog'lovchili birikma (CONJP), muvofiqlashtirilgan birikma (COP) va boshqalar (X) [Han, Wong, Chao, 2014]. Quyidagi jadvalda Penn Treebank tomonidan taklif qilingan teglar to'plami keltirilgan [https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_p.os.html]

Number	Tag	Description	
1.	CC	Coordinating conjunction	Bog'lovchi birikmalar
2.	CD	Cardinal number	Son
3.	DT	Determiner	Aniqlovchi (artikl, olmosh)
4.	EX	Existential there	Mavjudlikni?
5.	FW	Foreign word	Xorijiy so'zlar
6.	IN	Preposition or subordinating conjunction	Ko'makchi va bog'lovchi
7.	JJ	Adjective	Sifat
8.	JJR	Adjective, comparative	Sifat, Qiyosiy Daraja
9.	JJS	Adjective, superlative	Sifat, ortirma Daraja
10.	LS	List item marker	Elementlar ro'yxati
11.	MD	Modal	Modal
12.	NN	Noun, singular or mass	Ot, birlik, jamlovchi ot
13.	NNS	Noun, plural	Ot, ko'plik
14.	NNP	Proper noun, singular	Nomlar (atoqli ot) birlikda
15.	NNPS	Proper noun, plural	Nomlar (atoqli ot) ko'plikda
16.	PDT	Predeterminer	Ot turkumidagi so'zlardan oldidan keladigan aniqlovchi so'zlar
17.	POS	Possessive ending	Egalik qo'shimchasi
18.	PRP	Personal pronoun	Kishilik olmoshi
19.	PRP\$	Possessive pronoun	Egalik olmoshi (his, her, their)
20.	RB	Adverb	Ravish
21.	RBR	Adverb, comparative	Ravish, qiyosiy
22.	RBS	Adverb, superlative	Ravish, ortirma
23.	RP	Particle	Fe'llarga, so'zlarga bog'lanib keladigan so'zlar (up, off...) Yuklama

24.	SYM	Symbol	Ramzlar
25.	TO	to	To predlogi
26.	UH	Interjection	Undov so'zlar
27.	VB	Verb, base form	Fe'l, asosiy fe'l
28.	VBD	Verb, past tense	Fe'l, o'tgan zamon
29.	VBG	Verb, gerund or present participle	Fe'l, sifatdosh
30.	VBN	Verb, past participle	Fe'l, o'tgan zamon sifatdoshi
31.	VBP	Verb, non-3rd person singular present	3-shaxs hozirgi zamon birlik fe'li
32.	VBZ	Verb, 3rd person singular present	
33.	WDT	Wh-determiner	Wh bilan boshlanuvchi aniq
34.	WP	Wh-pronoun	Wh bilan boshlanuvchi olmoshlar
35.	WP\$	Possessive wh-pronoun	Wh bilan boshlanuvchi egalik olmoshlari
36.	WRB	Wh-adverb	Wh bilan boshlanuvchi ravishlar

Tobelikni tahlil qilish (DP) gapdagi tarkibiy qismlarning bosh so'zlari bilan yo'naltirilgan bog'liqlikni kodlaydi. DP tahlili Universal Stanford dependencyning 37 ta sintaktik munosabatlarni belgilovchi teglaridan foydalaniladi. 2005-yilda ishlab chiqilgan Stanford Dependency 2014-yilga 30 ta til uchun universal teglarni taklif qildi [<https://universaldependencies.org/u/dep/>].

	Nominals	Clauses	Modifier words	Function Words
Core arguments	nsubj	csubj		
	obj	ccomp		
	iobj	xcomp		
Non-core dependents	obl vocative expl dislocated	advcl	advmod discourse	aux cop mark
Nominal dependents	nmod appos nummod	acl	amod	det clf case
Coordination	Headless	Loose	Special	Other
conj cc	fixed flat	list parataxis	compound orphan goeswith reparandum	punct root dep

Individual tillar uchun bu yerda belgilangan universal tiplar-ning maxsus subtiplari ishlab chiqilgan, bu orqali aniqroq munosabatlarni belgilashi mumkin. Subtipli munosabat har doim asosiy turdan boshlanadi. Umuman olganda, subtiplar tilga xos va ixtiyoriydir. Biroq, ba'zi kichik tiplar ko'p tillarga taalluqli deb taxmin qilinadi va ular yarim majburiy deb hisoblanishi kerak. Quyidagi jadvalda tobelikni tahlil qilish uchun ishlab chiqilgan tegset bilan tanishishingiz mumkin, ammo bu barcha tillar uchun majburiy emas:

T/r	Dependency Tag	Description		Dependency Tag	Description
1.	acl	clausal modifier of a noun (adnominal clause)	2.	expl:move	reflexive pronoun utilized in reflexive passive
3.	acl:relcl	relative clause modifier	4.	expl:pvt	reflexive clitic with an inherently reflexive verb
5.	advcl	adverbial clause modifier	6.	mounted	mounted multiword expression
7.	advmod	adverbial modifier	8.	flat	flat multiword expression
9.	advmod:emph	emphasizing phrase, intensifier	10.	flat:overseas	overseas phrases
11.	advmod:lmod	locative adverbial modifier	12.	flat:title	names
13.	amod	adjectival modifier	14.	goeswith	goes with
15.	appos	appositional modifier	16.	iobj	oblique object
17.	aux	auxiliary	18.	checklist	checklist
19.	aux:move	passive auxiliary	20.	mark	marker
21.	case	case-marking	22.	nmod	nominal modifier
23.	cc	coordinating conjunction	24.	nmod:poss	possessive nominal modifier
25.	cc:preconj	preconjunct	26.	nmod:tmod	temporal modifier
27.	ccomp	clausal complement	28.	nsubj	nominal topic
29.	clf	classifier	30.	nsubj:move	passive nominal topic
31.	compound	compound	32.	nummod	numeric modifier

33.	compound: lvc	gentle verb building	34.	nummod: gov	numeric modifier governing the case of the noun
35.	compound: prt	phrasal verb particle	36.	obj	object
37.	compound: redup	reduplicated compounds	38.	obl	indirect nominal
39.	compound: svc	serial verb compounds	40.	obl:agent	agent modifier
41.	conj	conjunct	42.	obl:arg	indirect argument
43.	cop	copula	44.	obl:lmod	locative modifier
45.	csubj	clausal topic	46.	obl:tmod	temporal modifier
47.	csubj:move	clausal passive topic	48.	orphan	orphan
49.	dep	unspecified dependency	50.	parataxis	parataxis
51.	det	determiner	52.	punct	punctuation
53.	det: numgov	pronominal quantifier governing the case of the noun	54.	reparandum	overridden disfluency
55.	det: nummod	pronominal quantifier agreeing with the case of the noun	56.	root	root
57.	det:poss	possessive determiner	58.	vocative	vocative
59.	discourse	discourse ingredient	60.	xcomp	open clausal complement
61.	dislocated	dislocated parts			
62.	expl	expletive			
63.	expl: impers	impersonal expletive			

Tillarning grammatik strukturasi turli xil bo'lsa-da, umumiy universal lingvistik qoidalari mavjud degan qarash asosida ishlab chiqilgan mazkur tegsetlar kelajakda o'zbek tilida ham qo'llanilishi mumkin. Oxirgi izlanishlarda CoNLLU va UD teglari orqali o'zbek tilidagi matnlarni tahlil qilish ishlarini kuzatishimiz mumkin [Abduraxmanova, 2020].

O'zbek tilida sintaktik parsing, sintaktik analizator yaratish mavzusidagi tadqiqotlar sanoqlidir. Sh.Hamroyeva, O.Xidirov tadqiqotlarida sintaktik razmetkaning lingvistik asoslari, sintaktik parsing yaratish masalalari yuzasidan nazariy fikrlar berilib, teglarning

berilishi xususida takliflar berilgan. Sh.Hamroyeva tadqiqotida 34 ta sintaktik teg taklif qilingan. Tadqiqotda berilgan sintaktik razmetka teglari gap va ularning turlarini ifodalovchi teglar hisoblanadi [Hamroyeva, 2018]:

T/r	Teg	Tavsifi
1.	<SG>, </SG>	Sodda gap
2.	<UG>, </UG>	Uyushgan gap
3.	<QG>, </QG>	Qo'shma gap
4.	<dg>	Darak gap
5.	<sg>	So'roq gap
6.	<bg>	Buyruq gap
7.	<E+>	Egali gap
8.	<E->	Egasiz gap
9.	<sh.n.g>	Shaxsi noma'lum gap
10.	<a.g>	Atov gap
11.	<s.fsh.g>	Semantik-funksional shakllangan gap
12.	<yg>	Yig'iq gap
13.	<yog>	Yoyiq gap
14.	<u>, </u>	Undalma
15.	<k>, </k>	Kiritma
16.	<QG>, </QG>	Qo'shma gap
17.	<BQG>, </BQG>	Bog'langan qo'shma gap
18.	<EQG>, </EQG>	Ergashgan qo'shma gap
19.	<B-siz QG>, </B-siz QG>	Bog'lovchisiz qo'shma gap
20.	<b.BQG>	Biriktiruv munosabatli bog'langan qo'shma gap
21.	<z.BQG>	Zidlov munosabatli bog'langan qo'shma gap
22.	<a.BQG>	Ayiruv munosabatli bog'langan qo'shma gap
23.	<i.BQG>	Inkor munosabatli bog'langan qo'shma gap
24.	<KQG>	Kesim tobe gapli qo'shma gap
25.	<EQG>	Ega tobe gapli qo'shma gap
26.	<HQG>	Hol tobe gapli qo'shma gap
27.	<TQG>	To'ldiruvchi tobe gapli qo'shma gap
28.	<AQG>	Aniqllovchi tobe gapli qo'shma gap
29.	<e.MQG>, </e.MQG>	Bir necha ergash gapli murakkab qo'shma gap
30.	<b.MQG>, </b.MQG>	Bir necha bosh gapli murakkab qo'shma gap
31.	<a.MQG>, </a.MQG>	Aralash murakkab qo'shma gap
32.	<u.MQG>, </u.MQG>	Kesimlari uyushgan murakkab qo'shma gap
33.	<KG>, </KG>	Ko'chirma gap
34.	<MG>, </MG>	Muallif gapi

O.Xidirov tadqiqot ishida keyinchalik ham ikki xillik yuzaga kelmasligi uchun Sh.Hamroyeva taklif etgan sintaktik teglardan foydalaniganini yozadi [Xidirov, 2021].

Xulosa qilib shuni aytish mumkinki, o'zbek tili matnlarida to'liq sintaktik teglash masalasi tilning xususiyatidan kelib chiqib hal etiladi, ya'ni gapning tuzilishi, qurilishi, so'z birikmasi tuzilishi qoidalari batafsil ishlab chiqilib, nafaqat otli va fe'lli birikma, balki boshqa so'z turkumlarining gapda birikish imkoniyatlari, gapda gap bo'laklarining kelish o'rinlari aniqlanadi. Shuningdek, sintaktik teglar yuzasidan amalga oshirilgan tajribalar o'rganilib, o'zbek tilida mustaqil ravishda maxsus sintaktik teglar ro'yxati ishlab chiqilishi mumkin.

Foydalanilgan adabiyotlar

Aaron Li-Feng Han, Derek F. Wong, Lidia S. Chao, Yi Lu, Liangye He, Liang Tian. A Universal Phrase Tagset for Multilingual Treebanks. 2014.

https://www.researchgate.net/publication/267544606_A_Universal_Phrase_Tagset_for_Multilingual_Treebanks

Abduraxmanova N. Dependency Parsing Based On Uzbek Corpus. 2020.

<https://www.researchgate.net/publication/338065749>
<https://github.com/GoktugOcal/turkish-syntactic-parser>
https://link.springer.com/chapter/10.1007/978-3-030-99329-0_16;

<https://www.studocu.com/in/document/university-of-mumbai/natural-language-processing/uddesh-nlp2022-falkfasdf/31623646>;

<https://www.guru99.com/nlp-tutorial.html>;

<https://shashank-prasad11.medium.com/the-5-phases-in-natural-language-processing-c67a72ec742f>

<https://www.analyticsvidhya.com/blog/2021/12/dependency-parsing-in-natural-language-processing-with-examples/>

https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_p.os.html

<https://www.upgrad.com/blog/dependency-parsing-in-nlp/>

James B. Henderson. Artificial Neural Networks. 2010. The Handbook of Computational Linguistics and Natural Language Processing. Edited by Alexander Clark, Chris Fox, and Shalom Lappin.

MARK-JAN NEDERHOF AND GIORGIO SATTA. Theory of Parsing.

2010. The handbook of computational linguistics and natural language processing / edited by Alexander Clark, Chris Fox, and Shalom Lappin 105-130
- Sampson, G. (2003) 'Reflections of a dendrographer.' In A. Wilson, P. Rayson and T. McEnery (eds.) *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*, Frankfurt am Main: Peter Lang, pp. 157-184
- What is parsing in NLP: Its types and techniques. <https://intellipaat.com/>
- What is parsing in NLP: Its types and techniques. <https://intellipaat.com/>
- Xidirov O. Milliy korpus uchun parsing dasturi yaratishning lingvistik asoslari: Filol. fan. bo'yicha fa;safo doktori (PhD)... diss. – Jizzax, 2021. –136 b.
- Ҳамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD)...дисс. – Бухоро, 2018. –253 б.

THE ISSUE OF SYNTACTIC TAGS AND TAGGING IN UZBEK LANGUAGE TEXTS

Oqila Abdullayeva¹

Abstract. One of the stages of natural language processing, the task of syntactic analysis is the analysis of sentence devices. In corpus linguistics, the process called parsing determines the structure of sentences and sentences, the syntactic relationship of lexical units in a sentence. In this article, the issue of syntactic tagging of sentences in the Uzbek language and the lists of tags was studied and analyzed through examples in the Uzbek language.

Key words: *syntactic tagging, tag, parsing, parser, treebank, constituency parsing, dependency parsing.*

References

Aaron Li-Feng Han, Derek F. Wong, Lidia S. Chao, Yi Lu, Liangye He, Liang Tian. A Universal Phrase Tagset for Multilingual Treebanks. 2014.

https://www.researchgate.net/publication/267544606_A_Universal_Phrase_Tagset_for_Multilingual_Treebanks

Abduraxmanova N. Dependency Parsing Based On Uzbek Corpus. 2020. <https://www.researchgate.net/publication/338065749>

<https://github.com/GoktugOcal/turkish-syntactic-parser>

[https://link.springer.com/chapter/10.1007/978-3-030-99329-0_16;](https://link.springer.com/chapter/10.1007/978-3-030-99329-0_16)

[https://www.studocu.com/in/document/university-of-mumbai/natural-language-processing/ud-desh-nlp2022-falkfasdf/31623646;](https://www.studocu.com/in/document/university-of-mumbai/natural-language-processing/ud-desh-nlp2022-falkfasdf/31623646)

[https://www.guru99.com/nlp-tutorial.html;](https://www.guru99.com/nlp-tutorial.html)

<https://shashank-prasad11.medium.com/the-5-phases-in-natural-language-processing-c67a72ec742f>

<https://www.analyticsvidhya.com/blog/2021/12/dependency-parsing-in-natural-language-processing-with-examples/>

https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_p.os.html

<https://www.upgrad.com/blog/dependency-parsing-in-nlp/>

James B.Henderson. Artificial Neural Networks. 2010. The Hand-

¹Abdullayeva Oqila Xolmo'minovna – PhD, post-doctorate student at Tashkent State University of Uzbek Language and Literature named after Alisher Navo'i.

E-pochta: abdullayeva.oqila@navoiy-uni.uz

ORCID: 0000-0002-2524-4832

- book of Computational Linguistics and Natural Language Processing. Edited by Alexander Clark, Chris Fox, and Shalom Lappin.
- MARK-JAN NEDERHOF AND GIORGIO SATTA. Theory of Parsing. 2010. The handbook of computational linguistics and natural language processing / edited by Alexander Clark, Chris Fox, and Shalom Lappin 105-130
- Sampson, G. (2003) 'Reflections of a dendrographer.' In A. Wilson, P. Rayson and T. McEnery (eds.) *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*, Frankfurt am Main: Peter Lang, pp. 157-184
- What is parsing in NLP: Its types and techniques. <https://intellipaat.com/>
- What is parsing in NLP: Its types and techniques. <https://intellipaat.com/>
- Xidirov O. Milliy korpus uchun parsing dasturi yaratishning lingvistik asoslari: Filol. fan. bo'yicha falsafa doktori (PhD)... diss. – Jizzax, 2021. –136 b.
- Hamroyeva Sh. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari: Filol. fan. bo'yicha falsafa doktori (PhD)...diss. – Buxoro, 2018. –253 b.

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 29.02.2024-yilda ruxsat etildi.
Bichimi 70x100 1/16, Ofset bosma. "Cambria" garniturasida.
Shartli b.t. 7,51. Nashr b.t. 7,62.

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