

ISSN 2181-922X

LANGUAGE & CULTURE

UZBEKISTON O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER
LINGVISTIKASI

2023 Vol. 4 (6)

www.compling.tsuull.uz

ISSN 2181-922X

O‘ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2023 Vol. 4 (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 tarjimasini, 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), [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

Mastura Primova

Til korpuslarida matnlarni annotatsiyalash: afzallik va kamchiliklari.....6

Nilufar Muradova

Clarin tizimidagi og'zaki korpuslar xususida.....19

Noila Matyakubova

Iboralarni moslashtirish (phrase alignment)da otli va
fe'lli so'z birikmalar mosligi.....28

Ruxsora Muftillayeva

Dialektal korpuslarning umumiy tavsifi: tajriba va tahlil.....38

Sabura Xudayarova

Jahon tilshunosligida tabiiy tilni modellashtirish nazariyasi va
amaliyoti.....49

Jahongir Berdiyev

Tensorflow kutubxonasining imkoniyatlari.....63

CONTENT

Mastura Primova

Advantages and disadvantages of corpus annotation.....17

Nilufar Muradova

Specifically oral corpuses in the clarin system.....27

Noila Matyakubova

Aligning noun and verb phrases in phrase alignment36

Ruxsora Muftillayeva

General description of dialectal corpuses: experiment and analysis.....48

Sabura Xudayarova

Theory and practice of natural language modeling
in world linguistics.....62

Jahongir Berdiyev

Tensorflow library capabilities.....72

IBORALARNI MOSLASHTIRISH (PHRASE ALIGNMENT)DA OTLI VA FE'LLI SO'Z BIRIKMALAR MOSLIGI

Noila Matyakubova¹

Annotatsiya. Moslashtirish jarayoni tabiiy tilga ishlov berish (NLP)ning ko'plab vositalari samaradorligini oshirish uchun juda muhim omil hisoblanadi. Asosan turli tuzilishga ega bo'lgan tillarni moslashtirishda so'z va iboralarni moslashtirish muhim ahamiyat kasb etadi. Bu moslashtirish bosqichlari gaplarni va abzatslarni moslashtirish bosqichlarini tog'ri va samarali ishlashini ta'minlab beradi. Ushbu maqolada o'zbek va ingliz tillaridagi so'z birikmalarini moslashtirishda yuzaga keluvchi holatlar ko'rib chiqiladi.

Kalit so'zlar: *So'z birikmalari, otli birikma, fe'lli birikma, token, moslashtirish jarayoni, erkin birikma, turg'un birikma.*

Kirish

So'z birikmalari gapda malum ma'nolarni ifodalab keluvchi birikmalar bo'lib, ma'lum grammatik qurilmaga ega bo'lishi bilan oddiy so'zlardan ajralib turadi. NLPning ko'plab sohalarida, asosan, mashina tarjimasida, parallel matnlarni moslashtirish jarayonida manba til va maqsadli tilda qo'llaniladigan iboralar va ularning grammatik tuzilishi, semantikasi haqida to'liq tushunchaga ega bo'lish talab etiladi. Asosan turlicha grammatik tuzilishga ega bo'lgan tillarda ularni tuzilishida va ma'nosida katta farq bo'ladi. Ingliz va o'zbek tillari aynan shunday tillar qatoriga kiradi va ba'zi hollarda moslikni topishda bir nechta murakkabliklar keltirib chiqaradi [MacCartney, Galley, 2008]. O'zbek-ingliz so'z birikmalarining to'liq jamlanishi va mashinani o'qitish (Machine learning) jarayoni orqali moslashtirish vositalarini, mashina tarjimonlarini va parallel matnlar bilan ishlovchi turli vostalarning samaradorligini oshirish mumkin.

¹*Matyakubova Noila Shakirjanovna* – Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti tayanch doktoranti.

E-pochta: matyakubovanoila@navoiy-uni.uz, nailya89mm@mail.ru

ORCID: 0009-0009-3154-723X

Asosiy qism

O'zbek tilida so'z birikmalari birikma qismlarining grammatik tabiatiga ko'ra va tuzilishiga ko'ra tasnif qilinadi, ya'ni grammatik tabiatiga ko'ra tasnif qilinganda hokim a'zoning qaysi so'z turkumi ekanligi va tobe a'zoning sintaktik vazifasi asosga olinadi. Ular turg'un va erkin so'z birikmalari ko'rinishida uchraydi, erkin birikmaning o'zi ham o'z navbatida teng bog'lanish va tobe bog'lanish usullarida yasaladi [Abdullayev, 1976]. Tarkibiy jihatiga ko'ra ingliz tilidagi so'z birikmalari bilan deyarli bir xil bo'lib **bosh so'z** (Head word) va **ergash so'z** (Auxiliary word)dan tashkil topadi [Brinton, 2000]. Har ikkala tilda bosh so'z qaysi so'z turkumiga kirsam, aynan o'sha so'z turkumli so'z birikmasi deb yuritiladi. Tuzilishiga ko'ra har ikkala tilda oddiy so'z birikmasi va murakkab so'z birikmasi deb yuritiladi, ammo ingliz tilida o'zbek tilidan farqli ravishda qo'shma so'z birikmasi ham mavjud. Ingliz tilida so'z birikmalarining bir nechta turlari mavjud bo'lib, biz ularni otli birikma (noun phrase), fe'lli birikma (verb phrase), sifatli birikma (adjective phrase), ravishli birikma (adverbial phrase), predlogli birikma (prepositional phrase), verbal phrase, absolute phrase va bundan tashqari frazema (phrasal verb), kollakatsiya (collocation) va ibora (idiom) ko'rinishida uchratamiz [Brinton, 2000]. Ammo o'zbek tilda ular ingliz tilidagidan birmuncha farq qiladi va otli, fe'lli, sifatli, ravishli va modal so'zli birikma ko'rinishida yasaladi. Parallel matnlarni tarjima qilinganda asos tilidagi gapda uchraydigan so'z birikmalari maqsadli tilda ham birikma holatida ham tarqoq holatda kelishi mumkin. Ushbu turdagi so'z birikmalari moslashtirish jarayonida bir nechta chalkashliklarga olib keladi. Maqolada o'zbek va ingliz tilidagi otli va fe'lli so'z birikmalarining yasaliishi va gapda qo'llanilishini batafsil ko'rib chiqamiz.

O'zbek tilida otli birikma (NP) bosh so'z vazifasida ot keladi, u bitta yoki ikkita so'zlarning bog'lanishidan hosil bo'ladi. NP qismlari bitishuv, muvofiqlashuv hamda boshqaruv yo'li bilan grammatik aloqaga kirishadi va quyidagicha yasaladi:

- Bosh so'z bo'lib ot keladi: bag'rikeng inson, mard yigit.
- Bosh so'z bo'lib sifat keladi: onaning o'zi, Noilaning o'zi.
- Bosh so'z bo'lib otlashgan sifat keladi: talabalardan eng aql-lisi, gullarning sarasi.
- Bosh so'z bo'lib otlashgan son keladi: talabalardan biri, ularning uchovi.
- Bosh so'z bo'lib otlashgan sifatdosh keladi: bizning aytayot-ganimiz, ularning o'qiyotganlari.
- Bosh so'z bo'lib otlashgan ravish keladi: talabalarining ko'pi, so'zlarining ozi.

• Bosh soʻz boʻlib otlashgan undov soʻz keladi: nochorlarning dod-voyi.

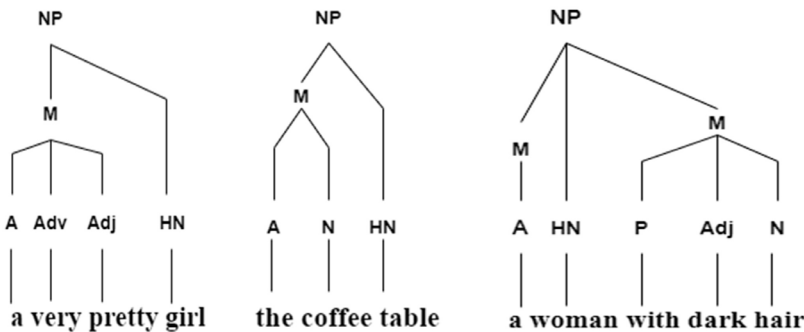
• Bosh soʻz boʻlib otlashgan taqlid soʻz keladi: kosalarining qasir-qusiri, otlarning dupur-dupuri.

Ingliz tilida NP ot soʻz turkumiga sifat, olmosh yoki artikl kabi aniqllovchilarning birikishi orqali hosil qilinadi va ular erkin birikma hisoblanadi. Misol uchun, **my family, a tall building, a very pretty girl**. Oʻz oʻrnida NP uch turga boʻlinadi: **simple NP (sodda), compound NP (qoʻshma), complex NP (murakkab)**.

1-Jadval. Ingliz tilida NPning yasalishi

NP	Yasalishi	Misol
Simple NP	Aniqllovchi + Ot	The black car, a friendly dog
Compound NP	Aniqllovchi + ot + ot	The swimming pool, a cotton shirt
Complex NP	Aniqllovchi + ot + aniqllovchi (ergash gap bilan aniqlanishi ham mumkin)	The lady waiting outside (The lady who is waiting outside)

Moslashtirish jarayonidagi tokenizatsiyadan foydalanilganda hosil boʻladigan tokenlardagi tafovut ham aynan birikmalardagi qatnashgan soʻzlarning grammatik jihatdan har ikkala tilda mavjudligi va ularning gapdagi oʻrniga ham bogʻliq boʻladi. Misol uchun baʼzi hollarda otga birikib kelgan qism otning oldida emas otdan keyin ham joylashishi mumkin. Bunday hollarda otdan oldin va keyin kelgan soʻzlar aniqllovchi vazifasini bajarib, ingliz tilida otdan oldin keladiganlari premodifier va otdan keyin kelganlari modifier deb ataladi. Bir nechta modifierlarning qatnashishi asosan murakkab NPda kuzatiladi va otga predlogli birikmaning birikishi orqali sodir boʻladi. Buni quyidagi misollarda koʻrib chiqishimiz mumkin:



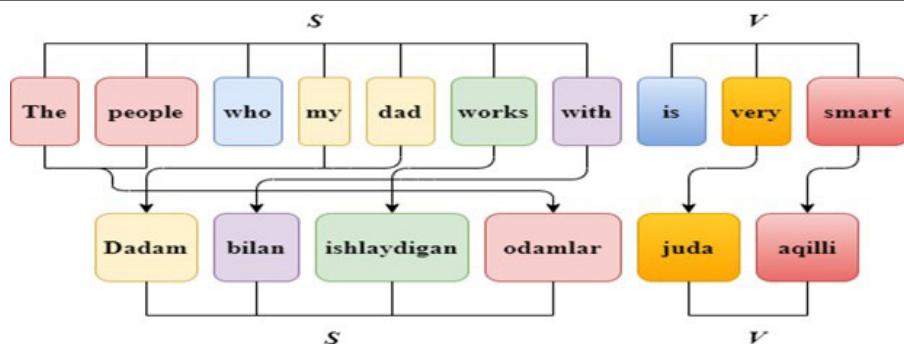
1-rasm. Sodda NP, qoʻshma NP, murakkab NP

Berilgan misollarda A- artikl, Adv- ravish, Adj- sifat, N- ot, HN- asosiy ot, M- aniqlovchi, NP- o'tli birikmani ifodalaydi. Ot oldidan kelgan artikl, ravish va sifat otga birikib NPni hosil qilyapti. Qo'shma va murakkab NP bir nechta otlar qatnashishi orqali yasaliib, ulardan bittasi asosiy ot (head noun) hisoblanadi va gapning kesimi aynan o'sha HNga moslashishi talab etiladi. Misol uchun, "**A child holding a lot of toys is my son**" gapini tahlil qiladigan bo'lsak, gapning egasi **A child holding a lot of toys**, kesim bo'lib kelgan **is** murakkab NP ko'rinishida kelgan yaxlit egaga emas HNga moslashishi talab etiladi. Bunda HN va kesim oralig'idagi barcha so'zlar modifier hisoblanaadi. Bundan tashqari murakkab NP **relative clause (RC)** yordamida ham yasaliib, NPda kelgan HNdan keyin bog'lovchi va ega kesimli gap keladi, ya'ni qo'shma gap ko'rinishida. Ammo bunday shaklda yasalgan qo'shma gaplar o'zbek tiliga tarjima qilinganda, sodda gap ko'rinishiga keladi, chunki ingliz tilidagi bog'lovchiga birikib kelgan ega kesimli bo'lakning vazifasi HN aniqlash bo'lib modifier vazifasini bajaradi va gapning istalgan qismida kelishi mumkin.

S [RC]+V+O [RC]

NP gapda ega va to'ldiruvchi vazifasida kelishi mumkin. Bunday holatda RC yordamida yasalgan murakkab NP bog'lovchi yordamida yasaliishi yoki sifatdosh yordamida yasaliishi ham mumkin. Har ikkala holatda ham o'zbek tiliga bir xil tarjima qilinadi, ammo tokenlar sonida va gap turlarida farq qiladi. Ingliz tilida berilgan gap qo'shma gap shaklida bo'lsa, o'zbek tilida sodda gap ko'rinishida keladi.

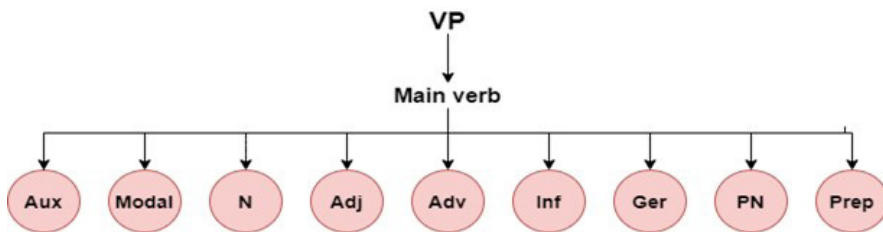
<i>The man who is working in the garden is my dad.</i>	<i>Bog'da ishlayotgan kishi mening dadam.</i>
<i>The man working in the garden is my dad.</i>	



2-rasm. Murakkab NP qatnashgan gapning moslashuvi

O'zbek tilida fe'lli birikma (VP) bosh so'z sifatida fe'lning kelishi bilan ifodalanadi va fe'lning boshqa boshqa so'zlar bilan birikib kelishi orqali hosil bo'ladi. Misol uchun: *tez gapirmoq, kulib yubormoq, diqqat bilan tinglamoq*.

Ingliz tilida VP gaplarni grammatik jihatdan to'g'ri tuzish, gapning ma'nosini to'liq yetkazish uchun zarurdir. VP odatda asosiy fe'l va bir yoki bir nechta yordamchi fe'llarni birikib kelishidan hosil bo'ladi. Bundan tashqari fe'l, sifat, ravish, ot, fe'lning otlashgan shakli (infinitive/gerund) va inkor yuklamalari bilan ham birikib kelishi mumkin. Ingliz tilida hozirgi va o'tgan oddiy zamondagi gaplarning darak shakli asosiy fe'lning o'zi bilan yasaladi, ammo ularning inkor va so'roq shakllari va qolgan barcha zamonlar VP yordamida yasaladi.



3-rasm. Ingliz tilida VPning yasalishi

VP yasalishi jihatdan *essential* va *non-essential VP* ga ajraladi. Essential VP gapning grammatik jihatdan to'g'ri shakllantirish uchun juda muhim bo'lib, asosan ko'makchi va modal fe'llar yordamida yasaladi. Misol uchun, "**will have gone**" fe'l birikmasi Future perfect tense da yasalishi uchun muhim bo'lsa, "**will probably have gone**" non-essential fe'l birikmasi hisoblanadi, chunki bunda probably gapda ehtimollik holatini ifodalash uchun ishlatilib, gapning ma'nosini ifodalashda ahamiyatli bo'lsa gapning grammatik jihatdan to'g'ri shakllantirish uchun muhim emas. Phrase alignment (birikmalarini moslashtirish) jarayonida o'zbek va ingliz tilidagi birikmalarini moslashtirishda yuzaga keluvchi ko'plab tafovutlar aynan VPlarning bir-biridan tubdan farq qilishida hisoblanadi.

2-Jadval. Ingliz tilidagi turg'un VP larning yasalishi

Tenses	Auxiliaries	Modals
Present Simple	-	Modal+V1
Present Continuous	To be (am/is/are) + Ving	Modal be + Ving
Present Perfect	To have been + V3	Modal have been + V3
Present Perfect Continuous	To have been + Ving	Modal have been + Ving
Past simple	-	Modal have been + V3

Iboralarni moslashtirish (phrase alignment)da o'tli va fe'lli so'z birikmalar mosligi

Past Continuous	To be (was/were) + Ving	Modal have been + Ving
Past Perfect	Had been + V3	Modal have been + V3
Past Perfect Continuous	Had been + Ving	Modal have been + Ving
Future Simple	Will + V1	Modal+V1
Future Continuous	Will be + Ving	Modal be + Ving
Future Perfect	Will have been + V3	Modal have been + V3
Future Perfect Continuous	Will have been + Ving	Modal have been + Ving

Ushbu jadvalda berilgan VP da qatnashgan ko'makchi va modal fe'llar gapda zamonlarning to'g'ri shakllantirish uchun muhim hisoblanadi, ammo o'zbek tiliga tarjima qilganimizda ko'makchi fe'llar tarjimaga ega bo'lmaydi va tokenlardagi asosiy farqlar aynan shundan kelib chiqadi. Misol uchun, "*He **has been working** in the office since 7 p.m*" ushbu gap Present Perfect Continuous zamonida yasali gapda murakkab kesim bo'lib kelgan VP, **has been working**, uchta tokenidan tashkil topgan bo'lsa, "*U soat yettidan beri **ishlayapti***", o'zbek tilidagi tarjimasida, **ishlayapti**, ya'ni bitta tokeniga to'g'ri keladi.

3-Jadval. Qoidaga asosan so'z birikmalarini moslashtirish

So'z birikmasining turi	O'zbek tilida	Ingliz tilida	Tokenlardagi farq	Moslashtirishdagi muammolar va sababi
O'tli birikma	Talabalardan biri	One of the students	2	O'zbek tilida mavjud bo'lmagan grammatik shakllarning qo'llanilishi
	Bag'ri keng inson	a generous person	Yo'q	Ayni ma'noni ifodalovchi lekemaning tarjimasi yo'q, "bag'rikeng" so'zi ko'chma ma'noda kelgan; O'zbek tilida mavjud bo'lmagan grammatik shakllarning qo'llanilishi

Fe'li birikma	Yo'qolgan bo'lsa kerak	Must have been lost	1	O'zbek tilida mavjud bo'lmagan grammatik shakllarning qo'llanilishi
	Diqqat bilan tingladi	Listened carefully	1	Ingliz tilida mavjud bo'lmagan grammatik shakllarning qo'llanilishi

Xulosa

So'z birikmalarini moslashtirish birmuncha murakkab jarayon bo'lib, umumiy moslashtirish jarayoniga qaraganda ancha chalkashliklarga olib keladi. Turg'un so'z birikmalarining jamlanishi va erkin birikmalarning qoidalarini mashinaga to'liq o'rgatish orqali bu muammolarni hal qilish mumkin hisoblanadi. Olib borilayotgan ilmiy ishimiz doirasida o'zbek va ingliz tilidagi turg'un so'z birimalari bazasi va erkin so'z birikmalarining qoidalar to'plami tuzilib, inson aralashuvi orqali kompyuterga o'rgatilmoqda va bu "Aligner" dasturiy vositasining samaradorligini oshiruvchi birlamchi omillardan deb hisoblanmoqda.

Foydalanilgan adabiyotlar

- B. MacCartney, M. Galley, Ch. Manning, "A Phrase-Based Alignment Model for Natural Language Inference". Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, 2008.
- Abdullayev F.A. O'zbek tili grammatikasi. 2-jild sintaksis.18-39-betlar. 1976.
- Laurel J. Brinton, "The Structure of Modern English" John Benjamins Publishing Company Amsterdam /Philadelphia, 2000.
- Y. Arase, J. Tsujii, "Compositional Phrase Alignment and Beyond". Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 1611–1623, November 16–20, 2020.
- Abdullayeva O., Xudayarova S. O'zbek tilshunosligida so'z birikmasiga ta'rif, tavsif va tasnif masalasi. // International scientific-theoretical conference on the topic: «Problems of research and education of the Uzbek language»
www.myscience.uz.

Iboralarni moslashtirish (phrase alignment)da otli va fe'lli so'z birikmalar mosligi

Bas Aarts, Liliane Haegeman, "English Word Classes and Phrases",

Bas Aarts and Liliane Haegeman, January 2008.

Guelailia Ahmed, "Lists of general phrases that can be very helpful to write a good research paper." <https://www.researchgate.net/publication/340666349>

Khin Thandar Nwet, "Developing Word to Phrase Alignment for Myanmar-English Machine Translation", 13th International Conference on Computer Applications, 2015.

R. Sennrich, Volk, Martin (2011). "Iterative, MT-based sentence alignment of parallel texts." In: NODALIDA 2011, Nordic Conference of Computational Linguistics, Riga, 11 May 2011 - 13 May 2011.

<https://www.masterclass.com/articles/what-is-a-verb-phrase>

ALIGNING NOUN AND VERB PHRASES IN PHRASE ALIGNMENT

Noila Matyakubova¹

Abstract. The alignment process is a critical factor in maintaining much of the performance of natural language processing (NLP). Aligning words and phrases is important in alignment of the languages that have different grammatical sentence structures. These alignment steps ensure that sentence and paragraph alignment steps work correctly and efficiently. In this article alignment process of Uzbek and English phrases have been analyzed.

Key words: *Phrasal verbs, verb phrase, noun phrase, token, alignment, free phrase, fixed phrase.*

References

- B. MacCartney, M. Galley, Ch. Manning, "A Phrase-Based Alignment Model for Natural Language Inference". Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, 2008.
- Abdullayev F.A. O'zbek tili grammatikasi. 2-jild sintaksis.18-39-betlar. 1976.
- Laurel J. Brinton, "The Structure of Modern English" John Benjamins Publishing Company Amsterdam /Philadelphia, 2000.
- Y. Arase, J. Tsujii, "Compositional Phrase Alignment and Beyond". Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 1611–1623, November 16–20, 2020.
- Abdullayeva O., Xudayarova S. O'zbek tilshunosligida so'z birikmasiga ta'rif, tavsif va tasnif masalasi. // International scientific-theoretical conference on the topic: «Problems of research and education of the Uzbek language»
www.myscience.uz.
- Bas Aarts, Liliane Haegeman, "English Word Classes and Phrases", Bas Aarts and Liliane Haegeman, January 2008.
- Guelailia Ahmed, "Lists of general phrases that can be very helpful to write a good research paper." <https://www.researchgate.net/publication/340666349>
- Khin Thandar Nwet, "Developing Word to Phrase Alignment for

¹Matyakubova Noila Shakirjanovna – PhD student of Tashkent State University of Uzbek Language and Literature named after Alisher Navoi.

E-mail: matyakubovanoila@navoiy-uni.uz, nailya89mm@mail.ru

ORCID: 0009-0009-3154-723X

Myanmar-English Machine Translation”, 13th International Conference on Computer Applications, 2015.

- R. Sennrich, Volk, Martin (2011). “Iterative, MT-based sentence alignment of parallel texts.” In: NODALIDA 2011, Nordic Conference of Computational Linguistics, Riga, 11 May 2011 - 13 May 2011.

<https://www.masterclass.com/articles/what-is-a-verb-phrase>

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 25.12.2023-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.