

Qodirova Madinabonu Murodjon qizi

Toshkent davlat o'zbek tili va adabiyoti universiteti tayanch doktoranti

Tel: +998943839909 e-mail: begonammmm@gmail.com

YORDAMCHI TIL BIRLIKLARI SINONIMIYASI LINGVISTIK BAZASINI YARATISH ASOSLARI.

Annotatsiya. Sinonimiya hodisasi deyarli barcha dunyo tillarida uchraydi. Mashina tarjimasida, umuman, NLPning barcha yo'nalishlarida bu hodisa asosiy muammolardan sanaladi. Chunki bir-biridan ma'no nozikliklariga ko'ra farqlanuvchi sinonimlar qatoridan uslub va davr nuqtayi nazaridan mos variantni qo'llashni mashinaga o'rgatish talab etiladi. Ushbu tadqiqotimizda o'zbek tilidagi yordamchi til birliklarida uchrovchi sinonimlik hodisasi va uni ingliz-o'zbek mashina tarjimasida uchun hal etishning bir necha usullarini taklif etamiz va lingvistik ta'minotini yaratish asoslarini ko'rib chiqamiz.

Kalit so'zlar: mashina tarjima, NLP, modellashtirish, grammatik shakl, grammatik sinonimiya, morfologik sinonimiya, sintaktik sinonimiya, lingvistik bo'yoq, periodizatsiya.

Кадирова Мадинабону

*Докторант Ташкентского государственного
университета узбекского языка и литературы*

Тел: +998943839909 почта: begonammmm@gmail.com

ОСНОВЫ СОЗДАНИЯ ЛИНГВИСТИЧЕСКОЙ БАЗЫ СИНОНИМИИ ВСПОМОГАТЕЛЬНЫХ ЯЗЫКОВЫХ ЕДИНИЦ.

Аннотация: Явление синонимии встречается практически во всех языках мира. Это явление является одной из основных проблем машинного перевода, в целом, во всех областях НЛП. Потому что необходимо научить машину использовать подходящий с точки зрения стиля и периода вариант из синонимов, отличающихся друг от друга тонкостями значения. В данном исследовании мы предлагаем несколько путей решения явления синонимии, встречающегося во вспомогательных языковых единицах узбекского языка и для англо-узбекского

машинного перевода, и рассматриваем основы создания лингвистического обеспечения.

Ключевые слова: машинный перевод, НЛП, моделирование, грамматическая форма, грамматическая синонимия, морфологическая синонимия, синтаксическая синонимия, языковая окраска, периодизация.

Kadirova Madinabonu

*Doctoral student of Tashkent State University of
Uzbek Language and Literature*

Tel : +998943839909 e-mail: begonammmm@gmail.com

FUNDAMENTALS OF CREATION OF THE LINGUISTIC BASE OF SYNONYMY OF AUXILIARY LANGUAGE UNITS.

Abstract. The phenomenon of synonymy is found in almost all languages of the world. This phenomenon is one of the main problems in machine translation, in general, in all areas of NLP. Because it is necessary to teach the machine to use the appropriate variant from the point of view of style and period from the synonyms that differ from each other according to the subtleties of meaning. In this study, we propose several ways to solve the phenomenon of synonymy found in auxiliary language units of the Uzbek language and for English-Uzbek machine translation, and consider the basics of creating linguistic support.

Key words: machine translation, NLP, modeling, grammatical form, grammatical synonymy, morphological synonymy, syntactic synonymy, linguistic color, periodization.

KIRISH.

Grammatik sinonimiya tushunchasini muomalaga A. M. Peshkovskiy kiritib, sinonimlarni morfologik va sintaktikga ajratgan. U grammatik sinonimiyani “grammatik ma’nosiga ko’ra bir-biriga yaqin bo’lgan so’z va iboralarning ma’nolari” sifatida nihoyatda keng tushungan.[1] Leksik sinonimlardan farqli ravishda grammatik sinonimlar matn mazmuniga ko’ra turli grammatik ma’nolarni ifodalashi mumkin.

Qaysidir grammatik ma'nosiga ko'ra bir o'rinda sinonim bo'lgan so'z boshqa ma'nosiga ko'ra boshqa sinonimik qatorda turadi. Yordamchi so'z turkumiga oid sinonimik qatordagi so'zlar quyidagi hodisalar nuqtai nazaridan o'zaro farqlanadilar:

1) *qo'llanish uslubiga ko'ra* so'zlashuv, badiiy, publitsistik, ilmiy, rasmiy uslublardan biriga tegishli yoki neytral bo'ladi. Masalan: *qadar*, *dovur*, *-gacha*. Ushbu sinonimik qatorda *qadar* – badiiy uslub, *dovur* – so'zlashuv uslubiga tegishli. *-gacha* esa uslub jihatidan neytral.

2) *adabiy tilga munosabatiga ko'ra* adabiy yoki dialektal: *avval* – adabiy, *burun* – dialektal.

Ko'p ma'noli so'z har bir ma'nosi bilan boshqa-boshqa sinonimik qatorda bo'lishi mumkin. Masalan: ko'p ma'noli «*bilan*» so'zi turli xil grammatik ma'nosi bilan turli sinonimik qatorga birlashadi:

1. Biriktiruv bog'lovchisi sifatida *va*, *ham*, *hamda*, *-u*, *-yu*, *-da* bog'lovchilari bilan sinonim bo'ladi:

- Olma **bilan** anor
- Olma **va** anor
- Olma **hamda** anor
- Olma-**yu** anor

2. To'siqsizlik ma'nosini ifodalab, *ham* yuklamasiga sinonim bo'ladi:

*Erta kelgani **bilan** bir ishni do'ndirmadi.*

*Erta kelsa **ham** bir ishni do'ndirmadi.*

3. Sabab ma'nosini ifodalab, *sababli*, *tufayli*, *uchun* so'zlariga sinonim bo'ladi:

*Sening gaping **bilan** uni ketkazib yubordim.*

*Sening gaping **sababli** uni ketkazib yubordim*

Grammatik sinonimlar quyidagi yo'llar bilan paydo bo'ladi:

- bir tildan ikkinchi tilga soʻz oʻtishi natijasida: *haqida* – arabcha, *toʻgʻrisida* – turkiycha.
- mahalliy shevalardan soʻz olish natijasida: *qadar* – adabiy tilga xos, *dovur* – dialektal.

Leksik birliklarda soʻz yasalishi yoʻli bilan sinonim hosil qilishvasoʻzning yangi maʼno kasb etishi natijasida sinonimlik hosil qilish usullari ham mavjud. Lekin grammatik birliklarda bunday usullar deyarli kuzatilmaydi.Chunki tilning grammatik tizimi nisbatan barqarorlik xususiyatiga ega boʻladi.

ADABIYOTLAR TAHLILI

Grammatik sinonimlar kompyuter lingvistikasida oʻrganilishi muhim boʻlgan masalalardan biridir. Matnni avtomatik tahlil qilish jarayonida grammatik sinonimlarni aniqlash muhim ahamiyatga egadir.Shu sababli dunyo va oʻzbek tilshunosligida mashina tarjimasi uchun sinonimlik masalasini hal etish uchun bir qancha mexanizm, gʻoyalar ishlab chiqilgan. Masalan, Fransiyadagi Tuluza universitetining bir guruh olimlari E.Navarro, F.Sajous, B.Gaume, L.Prevot, tayvanlik olimlar H.ShuKai, Tzu-Yi, H.Chu-Renlar matn tarkibidan sinonimlarni aratib olishning lugʻatlarga asoslangan usulini taklif etadilar.Ular Vikipediya asoslangna Wikilugʻatni tabiiy tillarni qayta ishlash uchun potensial resurs sifatida baholaydilar. Tadqiqotchilar matn orasidan sinonimlarni aniqlash va ularni baholashning ikki usulini taklif etadilar: semantik yaqinlik oʻlchovidan foydalanish va Wikilugʻatning tarjima manbaalaridan foydalanish. [5]

NLP uchun sinonimlar masalasini hal etish uchun oʻzbek tilshunosligida ham bir qancha olimlar tadqiqot olib bordilar. Masalan, filologiya fanlari doktori, professor N.Abdurahmonova “lingvistik taʼminotning semantik qatlami va til va nutq birliklarining alohida yaratilgan bazalari mashina tarjimasi oldida turgan omonimiya va sinonimiya masalasini hal qilishga yordam beradi” deb hisoblaydi. Shuningdek, olima mashina tarjimasi muammolarini hal etishning birdan-bir yoʻli modellashtirish usuli ekanligiga alohida urgʻu beradi.Modellashtirish mashina tarjimasida eng samarali metod

sanaladi. Bunda lisoniy hodisalarni muayyan darajada tahlil qilish, tekshirish va solishtirish imkoniyati mavjud. Nutqiy hodisalarning barchasini o'zida to'liq aks ettirmasa-da ular haqida muayyan xulosa berishda ahamiyatga ega.[6]

Sh.Orifjonova esa mashina tarjimasida sinonimlar roli haqida fikr yuritib, shunday deydi: "Albatta bir tilda mavjud so'zning ikkinchi tildagi aynan ma'nosini beruvchi muqobili bo'lmasligi mumkin. Bunday vaziyatlarda, eng to'g'ri yo'l sinonimik qatordan ularning dominantasini tanlab olish maqsadga muvofiqdir. [7]"

METODOLOGIYA.

O'zbek tilida holat, sabab, payt, o'rin, modallik kabi ko'plab grammatik ma'nolar turli til vositalari yordamida ifodalanadi. Ma'lum grammatik ma'noni ifodalashga qaratilgan bir necha til birliklari grammatik sinonimlar qatoriga birlashadi. Biroq bu birliklarni bir grammatik ma'noni ifodalaganligi uchun har doim ham biri o'rnida ikkinchisi qo'llanavermaydi. Nutq taqozosiga ko'ra, ya'ni uslub, davr, so'zlovchining fikrga munosabatiga ko'ra sinonimlar qatoridan munosib birliklarni topib qo'llash muhim. Inson tarjimon tomonidan tarjima qilinganda bu hech qanday muammo tug'dirmaydi, mashina tarjimada esa bunday emas. Chunki mashina tilning ma'noviy o'ziga xosliklarini 'tushunmaydi'. Shu sababli mashina tarjimasi uchin o'zaro qiyoslanayotgan ikki tilning grammatik sinonimlarini, ularning ma'noviy o'ziga xosliklarini aniqlash va o'zaro taqqoslash imkonini beruvchi lingvistik ta'minot vositasini yaratish zarurati paydo bo'ladi. Biz ushbu maqolada ta'minot vositasining umumiy ko'rinishini tavsiflashga harakat qilamiz. Bunda modellashtirish metodiga alohida e'tibor qaratamiz: matnni kompyuter tiliga o'tkazamiz.

Grammatik birlikning har bir ko'rinishi va har bir ma'nosiga alohida id-raqam beriladi. Shuningdek, o'zbek tilidagi yordamchi so'zlar ifodalagan grammatik ma'nolar uchun maxsus kodlar beriladi. Keyingi bosqich shu kodlar yordamida ma'lumotlarni qayta ishlash bo'ladi.

TAHLILLAR VA NATIJALAR

Ma'lumki sinonimlar bir-biri bilan funksional va emotsional-ekspressivligiga ko'ra ham farqlanadi. Leksik birliklardan farqli ravishda grammatik shakllar, xususan, yordamchi so'zlar emotsional-ekspressivlikdan qisman holi bo'ladi. Ayrim yordamchi so'zlar, xususan, ba'zi yuklamalarda ijobiylik-salbiylik bo'yog'i qisman mavjud bo'lishi mumkin, masalan, *xuddi* – ijobiy bo'yoqdor, *g'irt* – salbiy bo'yoq ottenkasiga ega. Biroq yordamchi so'zlardagi sinonimik qatorlarning hammasida ham bunday holat kuzatilavermaydi. Shu sababli emotsional-ekspressivlikni yordamchi so'zlar sinonimiyasidagi xarakterli belgi deb qarolmaymiz. Funksionallik esa tildagi barcha birliklar uchun umumiy bo'ladi va sinonimlarning ma'noviy farqlanishida asosiy rol o'ynaydi. Yordamchi til birliklarini sinonimik jihatdan modellashtirishda so'zning uslublarda qo'llanishiga alohida e'tibor qaratish zarur. O'zbek tilida 5 ta nutq uslubi bor. Bir so'z bir necha uslub doirasida ishlatilishi mumkin. Barcha uslublarda birdek qo'llana oluvchi so'z uslubiy jihatdan neytral sanaladi. Grammatik shakllarni lingvistik bazada berishda uslubiy xoslanishiga e'tibor qaratamiz. 1-jadvalga qarang.

1-JADVAL. SINONIMLARNING USLUBGA KO'RA FARQLANISHI

ID	Badiiy uslub	Ilmiy uslub	Publitsistik uslub	Rasmiy uslub	So'zlashuv uslubi
B01065	1	1	1	1	1
B18065	1	0	1	0	0
B31065	1	0	0	0	1
B04065	1	0	1	0	0

Ushbu jadvalda shart ma'nosini ifodalagan agar, gar, mabodo, basharti sinonimlar qatorining uslubiy farqlanishlari tasvirlangan.

Birinchi ustun yordamchi til birligining maxsus kodini ifodalaydi. Bunda birinchi belgi so'zning o'zbek tilida mansub bo'lgan turkumi: B – bog'lovchi, K – ko'makchi, Y – yuklama. Keyingi 2 belgi so'zning maxsus kodi. Oxirgi 3 ta belgi esa so'z ifodalagan grammatik ma'no sanaladi. Mashina tarjima dasturida ingliz tilida ham yordamchi birliklar shu tarzda kodlanadi va so'zlar orasidan shu kodga mos raqamli so'z olinadi. Keyingi bosqichda so'zlar uslubiy tomondan solishtiriladi. Agar so'z muayyan uslubda

qo'llanish imkoniyatiga ega bo'lsa 1, ega bo'lmasa 0 bilan belgilanadi. Har besh uslubda ham qo'llana olsa uslubiy jihatdan neytral hisoblanadi. Shu tariqa sinonimlar qatoridagi so'zlar bir-biridan ajratiladi.

XULOSA

Tadqiqotdan quyidagi xulosalar chiqarildi:

1. Modellashtirish orqali NLPning, xususan, mashina tarjimasining ko'plab muammolariga yechim topish mumkin. Shu sababli dunyo olimlari til muammolarini hal etishda modellashtirishga katta e'tibor qaratishadi.
2. Modellashtirish tadqiqotchidan ko'proq vaqt va mehnat talab etsa-da, sifatli tarjimani ta'minlaydi.
3. Grammatik shakllar tizimi leksik til birliklaridan farqli o'laroq barqaror va lisoniy qoliplarga mos keladi.
4. To'g'ri va mukammal tarjimaga erishish uchun grammatik shakllarning aniq tarjimalariga erishish zarur.
5. Grammatik sinonimiya leksik sinonimiyadan farq qiladi. Ularda emotsional-ekspressivlik bo'yog'i kam.

FOYDALANILGAN ADABIYOTLAR RO'YXATI

1. Пешковский А.М. Принципы и приемы стилистического анализа и оценки художественной прозы. М., 1964. С. 57.
2. Asqarova M., Yunusov R. O'zbek tili praktikumi. Toshkent, 2006. p. 159.
3. Майоров В.Д. Основы обработки текстов. Москва, 2018. p. 77-78.
4. <https://robotdreams.cc/blog/233-kak-iskat-sinonimy-s-pomoshchyu-word2vec>.
5. Navarro E. Wiktionary and NLP: Improving synonymy networks // Proceedings of the 2009 Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources. Toulouse 2009. P. 478-488
6. Abdurahmonova N. Mashina tarjimasining lingvistik ta'minoti Akademnashr, Toshkent – 2018, 131-bet

7. Orifjonova Sh.O'zbek tilida lug'aviy graduonimiya: Filol. fan.nomz. diss.avto-ref. – Toshkent, 1996.
8. Jamolxonov H.Hozirgi o'zbek adabiy tili. -Toshkent -Talqin. 2005.
9. Sayfullaeva R., Mengliev B., Boqiyeva, Qurbonova M., Yunisova Z., Abdullayeva M. Hozirgi o'zbek adabiy tili. -Toshkent 2009.