

PARALLEL KORPUSDA TARJIMANI AVTOMATLASHTIRISH (kollokatsiyalar misolida)

Guli ERGASHEVA

Filologiya fanlari doktori, dotsent
O'zbekiston davlat jahon tillari universiteti
Toshkent, O'zbekiston

Zahriddin XAITQULOV

Tayanch doktoranti
O'zbekiston davlat jahon tillari universiteti
Toshkent, O'zbekiston

Nilufar KUCHIMOVA

Falsafa doktori (PhD)
Samarqand davlat universiteti
Samarqand, O'zbekiston

ПАРАЛЛЕЛЬНЫЙ КОРПУС В АВТОМАТИЗИРОВАННОМ ПЕРЕВОДЕ (на примере коллокаций)

Гули ЭРГАШЕВА

Доктор филологических наук, доцент
Узбекский государственный университет мировых языков
Ташкент, Узбекистан

Захриддин ХАЙТКУЛОВ

Докторант, Узбекский государственный университет мировых языков
Ташкент, Узбекистан

Нилуфар КУЧИМОВА

Доктор философии (PhD)
Самаркандский государственный университет
Самарканд, Узбекистан

PARALLEL CORPUS IN AUTOMATED TRANSLATION (regarding to the collocations)

Guli ERGASHEVA

Doctor of Philology, Associate Professor
Uzbekistan State World Languages University
Tashkent, Uzbekistan ergashevaguli327@gmail.com

Zakhriddin KHAITKULOV

Doctoral student (PhD)
Uzbekistan State World Languages University
Tashkent, Uzbekistan haitkulovzakhriddin@gmail.com

Nilufar KUCHIMOVA

Doctor of Philosophy (PhD)
Samarkand State University

Samarkand, Uzbekistan Kuchimovanilufar7@gmail.com

UDC (UO'K, УДК): 81.276.6

For citation (iqtibos keltirish uchun, для цитирования):

Ergasheva G., Xaitqulov Z., Kuchimova N. Parallel korpusda tarjimani avtomatlashtirish (kollokatsiyalar misolida). // O'zbekistonda xorijiy tillar. — 2023. — № 5 (52). — B. 162-173.

<https://doi.org/10.36078/1697008599>

Received: September 21, 2023

Accepted: October 03, 2023

Published: October 05, 2023

Copyright © 2023 by author(s).

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Annotatsiya. Texnologiyaning rivoji ortidan ilm-fan sohasida bir qancha o'zgarishlar yuzaga kelmoqda. Asosan tarjima sohasida texnologik o'zgarishlarni alohida ta'kidlash o'rinli. Turli elektron tarjimon lug'at dasturlarining yaratilishi va undan foydalanish natijasida fan tarmoqlarida ham sezilarli statistiklarni misol keltirish mumkin. Tarjimani avtomatlashtirishdagi asosiy kutilayotgan natijalar samarasi o'laroq, uning rivoji uchun ham fan va texnologiyaning uzviy bog'liqlik jihatini kuzatish mumkin. Elektron lug'atlardan foydalanish ilm egalari va til o'rganuvchilar uchun asosiy dastur sifatida qo'llanilib kelinmoqda. Bunga sabab internetdagi kerakli ma'lumotlar turli tillarda mavjud bo'lib, mazkur ma'lumotlarni tarjima qilish uchun foydalanuvchilarning elektron lug'atga ehtiyoji ortadi. Shunday ekan, tarjimon dastrulardan foydalanish barcha soha vakillarining kundalik ehtiyoj manbaiga aylanib bo'lgan. Hozirgi vaqtga kelib bir nechta billingval va multilingval tarjimon dasturlar yaratilgan bo'lib, ularning mobil ilovalari ham foydalanuvchilar tomonidan keng qo'llanilib kelmoqda. Tarjimada bir vaqtning o'zida tarjima qilingan matnning ikki variantini tekshirish juda qiyin jarayon, shu bilan birga elektron tarjimon dasturlarining xotirasida yetarlicha ma'lumotlarni o'zida jamlanmaganligi tufayli sifatli tarjimadan foydalana olmaslik ehtimoli oshadi. Korpus lingvistikasi texnologiyasidan foydalangan holda, tarjima qilingan matndan keng foydalanish, matnni tahlil qilish va boshqa bo'limlardan unumli foydalanish mumkin. Parallel til korpusi, ya'ni asosiy tildagi matnning avtomatlashtirilgan to'plamini ko'zlangan tildagi tarjimalari bilan son-sanoqsiz xususiyatlariga avtomatik kirishni ta'minlaydi. Xususan, parallel korpus ixtirosidan ilgari elektron tarjima dasturlarida bunday imkoniyatlar bo'lmagan. Shu o'rinda parallel korpus orqali tarjimon dastur yaratish jarayonlarini alohida ta'kidlash o'rinli. Ushbu maqolada korpus va parallel korpusning maqsadi va vazifasi, tarjimon dasturlarini parallel korpus orqali yaratishning asosiy bosqichlari muammo va unga optimal yechimlar haqida so'z boradi.

Kalit so'zlar: parallel korpus; mashina tarjimasi; MT mexanizmi; korpus; kontekst; kollokatsiya; ilmiy; publitsistik matnlar; elektron lug'atlar; tarjima xotirasi; billingval; multilingval korpus.

Аннотация. Вслед за развитием технологий в сфере науки происходит ряд изменений, в частности, в сфере перевода. В результате создания и использования различных программ-переводчиков можно привести значительную статистику в исследованиях. Электронные словари использовались в качестве базовой программы для ученых и изучающих языки. Это связано с тем, что необходимая информация в Интернете доступна на других языках, а потребность пользователей в электронном словаре для перевода этой информации растет. Поэтому использование программ-переводчиков стало источником повседневных потребностей представителей всей отрасли. К настоящему времени создано несколько двуязычных и многоязычных программ-переводчиков, а их мобильные приложения широко используются пользователями. При переводе сложно проверять одновременно два варианта переведенного текста, при этом увеличивается вероятность не иметь возможности воспользоваться качественным переводом из-за недостаточности данных в памяти электронных программ-переводчиков. Используя технологию корпусной лингвистики, можно широко использовать переведенный текст, анализ текста и др. Корпус параллельного языка обеспечивает автоматический

доступ к бесчисленным функциям автоматического сбора текста на исходном языке с переводами на целевые языки, здесь уместно подчеркнуть процессы создания программы-интерпретатора посредством параллельного корпуса. В статье рассматриваются назначение и функции корпуса и параллельного корпуса, основные этапы создания программ-трансляторов посредством параллельного корпуса.

Ключевые слова: параллельный корпус; машинный перевод; механизм МГ; корпус; контекст; коллокация; научный; публицистические тексты; электронные словари; память переводов; выставление счетов; многоязычный корпус.

Abstract. Since the development of technology, a number of changes occurred in the field of science. It is worth noting technological changes mainly in the field of translation. The development and application of numerous translator-dictionary programs have made it possible to gather important statistics in the scientific domains. The primary anticipated outcomes in the area of automated translation allow one to see the dependency of science and technology for its advancement. For scientists and language learners, using computerized dictionaries has been employed as a fundamental program. This is because consumers increasingly need an electronic dictionary to interpret the necessary information they can get on the Internet, which is available in different languages. As a consequence, the use of translation software has evolved into a necessity for professionals across all levels. Numerous bilingual and multilingual translator programs have been developed to date, and users frequently use their mobile applications. When translating, it is exceedingly challenging to compare two translations of the same text at once, and the risk that an inferior translation will be used rises due to a lack of data in the memory of electronic translator systems. Text that has been translated can be extensively used, and text analysis and other fields can benefit greatly from the usage of corpus linguistics technology. Innumerable functions for automatically compiling source language material with translations into target languages are accessible through the Parallel Language Corpus. Particularly, such capabilities were not included in electronic translation tools prior to the advent of the parallel corpus. Here, it is appropriate to focus on the steps involved in developing an interpreter program using a parallel corpus. This article covers the key stages of the problem of developing translation programs using a parallel corpus and the best approaches to solving it. It also addresses the objective and function of the corpus and parallel corpus.

Keywords: parallel corpus; machine translation; MT mechanism; corpus; context; collocation; scientific; journalistic texts; electronic dictionaries; translation memory; billing a multilingual corpus.

Kirish. Bugungi kunda dunyoda yuz berayotgan madaniy, siyosiy-ijtimoiy, iqtisodiy o'zgarishlar tilda o'z ifodasini topmoqda. XX asr boshlarida kompyuter texnikasining rivojlanishi va internet tarmog'ining keng tarqalishi kompyuter tilshunosligi va korpus lingvistikasi sohalarida zamonaviy tilshunoslikning paydo bo'lishiga zamin yaratdi. Axborotlashtirilgan jamiyatda har qanday til sun'iy tilga aylanishi kerak deb hisoblash — to'g'ri tanlov bo'lishi mumkin. Shu maqsadda, har qanday tilning sun'iy shaklini yaratish korpus lingvistikasi uchun muhim va dolzarblik kasb etadi.

Korpus — til birliklarining imkoniyatlarini aniqlash borasida qidiruv dasturiga bevosita bo‘ysundirilgan matnlar majmui maqsadida yaratilgan. Til korpuslari til bo‘yicha tadqiqot va amaliy topshiriqlar yechimi uchun inkor etib bo‘lmas ish quroli sanaladi(16). Korpus bir til doirasini yoki ko‘p tillarni qamrab olishi mumkin. Korpuslardan foydalanishning qator afzalliklari mavjud. Masalan, korpuslar vaqt o‘tishi bilan tildagi o‘zgarishlarni tadrijiy o‘rganish imkoniyatini beradi shu bilan birga foydalanuvchilar auditoriyasiga muvofiq bo‘ladi.

Korpusning ixtisosligiga qarab bir nechta turlari tasniflangan: informatsion, monolingval, billingval, multilingval, parallel, qiyosiy (comparable), diaxronik (diachronic) va shu kabi boshqa bir nechta tilni qayta ishlash uchun, tilni tahrir qilish uchun, tarjimoni avtomatlashtirish hamda elektron lug‘atga xos dasturlarni yaratish. Parallelizm mezoniga ko‘ra, korpuslar bir tilli, ikki tilli va ko‘p tilli toifalarga bo‘linadi. Ikki tilli va ko‘p tilli korpus bir xil tematik sohada ikki yoki undan ortiq tilda mustaqil ravishda yozilgan matnlarni birlashtiradi. Masalan, turli tillardagi konferensiyalar, ilmiy anjumanlar, siyosiy konferensiyalar va boshqa materiallar to‘plami va boshqalar. Mazkur qiyosiy tadqiqotlar, tarjima nazariyasi va amaliyoti bo‘yicha tadqiqotlar olib borish uchun muhim manba bo‘lib xizmat qiladi.

Tarjimada elektron lug‘atlar mashina tarjimasidan foydalanish jarayonining tez va samarador bo‘lishiga xizmat qilish bilan birga, yangi ilmiy muammolarni o‘rganishga bo‘lgan ehtiyojni ham keltirib chiqarmoqda. Mashina tarjima tizimlari, avtomatlashtirilgan lug‘atlar, terminologik ma‘lumot banklarini yaratish parallel matnlarni tadqiq va tahlil qilishni talab qiladi. Buning samarasi sifatida dunyo bo‘ylab parallel matndan foydalanish parallel korpuslarni tadqiq qilish va natijalarni mukammallashtirish imkoniyatini yaratadi. Shu o‘rinda parallel korpus bo‘yicha quyidagi ta‘riflarni keltirish o‘rinli.

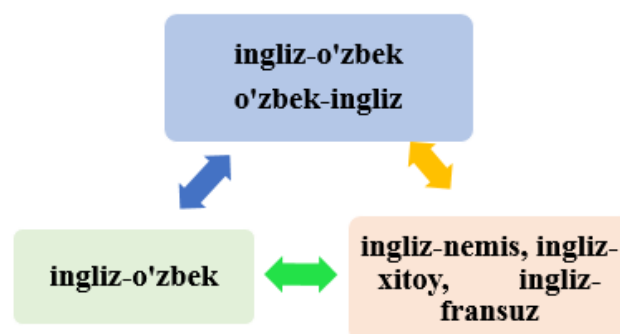
Parallel korpus — tarjima matnlarining parallel elektron analogi hisoblanib, asosan ikki tilli tarjima matnlardan iborat. Korpusdagi elektron matnlar asl matnning o‘zi yoki uning bir qismi ham bo‘lishi mumkin (12). Bir qancha olimlar, jumladan, D. Dobrovolskiy, Yu. Tao, V. Zaxarov, A. Kokoreva (11, 53) hamda E. Sosninalar (13) parallel korpusning tuzilishi, tarkibi va imkoniyatlari haqida o‘zlarining tadqiqot ishlarida alohida e‘tirof etishgan. Parallel korpusda bajariladigan amallarni takomillashtirish yuzasidan ham ko‘plab olimlar ilmiy izlanishlar olib borishgan. Ulardan M. Kay, M. Roscheisen, W. Gale, K. Church, P. Resnik, P. Brawn, J. Lai, R. Merserlar parallel korpus yaratish bosqichlari, yo‘llari xususida o‘z fikr-mulohazalarini bildirishgan(10). Ushbu olimlar orasidan Key va Roscheisenning fikricha, tarajimada gaplar bir-biriga mos kelishi uchun, avvalo, so‘zlar muqobilligiga erishish muhim o‘rin tutadi, lekin ikki tilda matn so‘zlarining muqobilligiga erishish oson emas va bu muqobillik adekvat bo‘lmasligi mumkin(4, 125). Shunday bo‘lsa, parallel korpusdagi izlangan so‘zning bir qancha matndagi ma‘nolarining berilishi orqali tadqiqotning o‘rganish jarayoni osonlashadi.

D. Dobrovolskiy parallel korpusga quyidagicha ta'rif beradi: "Parallel korpus — asliyat va ko'zlangan tildagi tarjimalarining elektron matnlar yig'indisidan iborat. Parallel korpusda maqsadli va tarjima matnlar yonma-yon joylashmaydi, balki bu ikki matndagi gaplar, jumlar mazmunan mos kelish darajasida bir nechta qismga bo'linadi. So'ngra mazkur birliklar yonma-yon, bir-biriga havola qilinadi. Asliyat tilidagi matn fragmentiga mos keluvchi tarjimadagi fragment belgilab qo'yiladi. Aynan mana shu holat bunday korpuslar yordamida turli lingvistik amallar bajarishga imkon yaratadi" (9).

Korpus lingvistikasida parallel korpusning bir qancha ko'rinishi mavjud. Iste'molchilar tomonidan ko'p foydalanilib kelinayotganlar asosan a) ko'p tilli korpus (comparable/ multilingual corpora) va b) tarjima korpusi (translation corpora). Bunday xususiyatga ega korpus tarkibi uning maqsadidan kelib chiqib turlicha bo'lishi mumkin:

- 1) tarjimaga havola qiluvchi odatiy matn tarzida;
- 2) qiyoslash uchun qulay bo'lgan interfeysdagi matnlar shaklida;
- 3) ma'lumotlar bazasi ko'rinishida.

Tuzilishi, matnning joylashuvi, birliklarning moslashtirishiga ko'ra, parallel korpusning bir necha tildagi ko'rinishi mavjud va ular quyidagi tarzda bo'ladi (15).



1-rasm: Parallel korpusning bir necha tildagi ko'rinishi.

Parallel korpus tarjima tilining struktur tuzilishidan tashqari so'z, ibora kollokatsiya, idiom va so'z birikmalarining muayyan kontekstdagi imkoniyatlarini kuzatishga ham xizmat qiladi. Ushbu korpus turlari zamonaviy tilshunoslikda, qiyosiy tilshunoslikda (ikki til strukturasi tahlil qilish maqsadida); tarjima sohasida (asliyat tildagi matnning kontekstual tarjimasini boshqa tillardan qidirish maqsadida); avtomatik tarjima (elektron lug'atda); leksikografiya sohaslarida keng foydalanish mazkur soha tarmoqlarini kengayishiga imkon yaratadi. Bu esa parallel korpus yaratishning tizimlarining ilmiy va amaliy jihatlarini chuqurroq o'rganishni talab etadi.

Ba'zi olimlar parallel korpuslarni yaratishda yagona til va qiyosiy matnlardan farqli o'laroq, madaniyatlararo munosabatlar omilini hisobga olish kerak, deb hisoblaydi. Asliyat tildagi manba materiallarini ko'zlangan tilga tarjima qilishda asosan statistikadagi tarjima qilingan materiallardan foydalanish parallel korpusga bo'lgan ishonchni yanada oshiradi. Parallel korpus yaratish bo'yicha

madaniyatlararo munosabatlar omilini inobatga olish xususida bir nechta o'zbek tadqiqotchilari ham o'z fikr-mulohazalarini bildirishgan. Ulardan V. Tangriyevning mulohazasiga ko'ra, "parallel korpusni yaratishda ikki tilda madaniyatlararo aloqa mavjud bo'lmasa, parallel korpus yaratish mumkin emas. Madaniyatlararo bog'liqlik qanchalik zaif bo'lsa, ikki tildagi tarjimaning sifati ham shunchalik zaif bo'ladi. Tarjimalarni kam amalga oshirish ortidan korpus va parallel korpus yaratish shunchalik qiyinlashib boraveradi. Masalan, ikki davlat o'rtasida siyosiy va madaniy aloqalarning mavjudligi turli hujjatlar, yo'riqnomalar, qo'llanmalar, risolalar va hokazolarni bir tildan ikkinchi tilga tarjima qilishni osonlashtiradi va manbalar sonini oshiradi" (9, 1005). Yuqoridagi fikrdan ko'rinadiki, tadqiqotchi V. Tangriyev parallel korpusni yaratishda asosan duch keladigan muammo — madaniyatlararo aloqalarda ekanligiga urg'u bergan.

Bizningcha, parallel korpusni yaratishda madaniyatlararo aloqalarga e'tibor berishdan avval PK ixtirosida asosan so'z, ibora, kollokatsiya va idiomalarning tadqiqotchilar va mutaxassis tarjimonlar tomonidan tarjima qilingan materiallardan foydalanish o'rinli, deb hisoblaymiz. Bunga sabab, gazeta va elektron veb-saytalardan ilmiy, publitsistik, siyosiy, ijtimoiy va boshqa shu kabi sohalardagi ikki tilda tarjima qilingan materiallardan foydalanish fikrimizning isboti hisoblanadi. Shuni ham ta'kidlash joizki, mazkur veb-saytdagi va gazeta-jurnallardagi tarjima qilingan materiallar unchalik ishonchli bo'lmasa-da, iste'molchi bir so'zning bir nechta tarjima turlaridan foydalanish imkoniyatiga ega bo'ladi.

PK yaratish jarayonida yana bir duch keladigan muammo — bir tushuncha bir tilda bitta so'z bilan ifodalansa, ikkinchi tilda ikki yoki undan ortiq so'z bilan ifodalanishi mumkin. Maqsadli tildagi matnda fikr qisqa jumla bilan ifodalangan bo'lsa, tarjima matnida uzun jumla vositasida ifodalanishi mumkin. Ushbu muammoning yechimi sifatida parallel korpus yaratmoqchi bo'lgan mutaxassis tarjimaning qaysi usulidan foydalanish kerakligini bilishi joiz. Chunki tarjimon dasturlarini parallel korpusda yaratishda tarjima usuli va tanlanadigan matnlar uslubi bir-biriga mos kelishi kerak. Masalan, ko'p tadqiqotchilar hozirda parallel korpusni yaratishda asosan statistik mashina tarjimasi tizimidan foydalanishadi, bunga sabab korpusda kiritilgan ikki va ko'p tilli matnlar tekshirilgan, bazasi shakllantirilgan va manbalarga asoslangan bo'ladi.

Mazkur tadqiqotda kontesktual tarjimon dasturini parallel korpus orqali yaratishda ilmiy matnlardagi kollokatsiyalardan foydalangan holda quyidagicha tahlil qilindi. Kollokatsiyalar tilshunoslik va tarjimashunoslik sohalarida muhim o'rin tutadi. Bu borada S. Bader hamda P. Neymark kabi tilshunos olimlar alohida e'tirof etishgan. P. Neymark kollokatsiyani "*matn va leksikaning nervlari*" deb hisoblaydi, boshqacha aytganda, "kollokatsiya — transport vositasining dvigateli" deya olim e'tirof etadi (5, 184). Kollokatsiyaning yana bir muhim jihatini tilshunos olim S.N. Alfari quyidagicha ta'kidlaydi: "Ingliz tili barcha tillardan foydalanish uchun asosiy bo'lgan qo'shma juftliklar va bo'laklarga to'la

bo'lganligi sababli, ikki tilli birikmalar lug'atni yaratishda ham muhim ahamiyat kasb etadi" (5, 186).

Ingliz tilida kollokatsiyaning ko'p turlari mavjud bo'lib, tilshunos olimlardan M. Benson hamda E. Bensonlar (1, 3) uni asosan ikki kategoriyaga ajratishgan. Ular *leksik* va *grammatik* kategoriyalar. Tarjima jarayonida eng muhim o'rin tutuvchi turlari ham leksik va grammatik kollokatsiyalar hisoblanadi. Leksik va grammatik kollokatsiyalarni bir-biridan farqli jihatini ajratish ba'zan muammo tug'dirishi ham mumkin. Shuning uchun *leksik* va *grammatik* kollokatsiyalarni quyidagi usullar orqali bir-birdan ajratish o'rinli. Leksik birikmalar *faqat fe'llar sifatlar, otlar va qo'shimchalardan (verbs, adjectives, nouns, and adverbs)* va turli birikmalar asosida tuziladi. Grammatik kollokatsiyalarning tuzilishi *fe'l, sifat yoki ot (nouns verbs or adjectives combined with a preposition or a grammatical structure)* kabi so'zlarni bosh gap yoki grammatik tuzilishi bilan birlashtiradi.

1-jadval.

Ilmiy matnlarda uchraydigan "the article investigate, the author speaks about, the paper considers" kabi kollokatsiyalarning kontekstual tarjimalari

<i>The paper investigates</i> the problem of equivalency and adequateness in literary translation.	<i>Maqolada</i> badiiy tarjima jarayonidagi ekvivalentlik va adekvatlik <i>masalasi ochib berilgan</i>
<i>The author investigates</i> theoretical works of foreign and native authors and comes to the conclusion that conversation and interview are different	<i>Muallif</i> suhbat va intervyu janrlarini o'zbek hamda xorijlik olimlarning qarashlariga tayanib <i>tadqiq etgan.</i>
<i>This article investigates</i> thematical principles of selection of phraseological for up—to-date dictionary, which will be based on three languages (English, Russian, and Uzbek). Different aspects of a new dictionary are being developed.	<i>Mazkur ilmiy ishda</i> uch tilli (ingliz, rus, o'zbek) o'quv lug'ati tipidagi frazeologik lug'at tuzishning parametrlarini o'rganishga <i>e'tibor qaratgan.</i>
<i>The author speaks about</i> the role of independent study, including home assignment in the organization of educational process and of	<i>Muallif</i> ta'lim sohasida bilim olish va uni mustahkamlashda, shuningdek, o'rganilgan materiallarni amalda qo'llay bilish ko'nikmalarini

their importance in obtaining linguistic knowledge by students.	shakllantirishda talabalarning mustaqil ishlari katta ahamiyat kasb etishini <i>ilmiy jihatdan dalillagan</i> .
<i>The work speaks about</i> virtual reality in the internet, accounting to its basic features, as well as separates the notion of virtual reality of identical one.	<i>Ishda</i> internetda virtual reallik mavjudligi uning muhim belgilari orqali <i>asoslangan</i> hamda virtual reallik unga o'xshash tushunchalardan chegaralab <i>ajratilgan</i> .
<i>This article speaks about</i> the semantic and stylistic features of abstract nouns according to their syntactic level. Additionally it was told that with the help of the features one can create laconic and short text units by their content	<i>Maqolada</i> o'zbek va rus tillaridagi mavhum otlarning semantik va stilistik imkoniyatlari hajman ixcham va mazmunan lo'nda matn birliklarini yaratishda muhim ekanligi haqidagi <i>mulohazalar bayon qilingan</i> .
<i>The author speaks about</i> the place and importance of political, historical, and pedagogical factors in the formation of healthy lifestyle of young students	<i>Muallif talaba-yoshlarning sog'lom</i> tafakkurini shakllantirishda ijtimoiy, iqtisodiy, siyosiy, tarixiy, pedagogik, psixologik, milliy-mental va boshqa omillarning o'rni va ahamiyatini <i>belgilab bergan</i> .
<i>The study considers</i> the main principles of journalism such as truthfulness, objectivity, freedom of word.	<i>Ilmiy ishda</i> jurnalistik faoliyatning asosiy prinsiplari — haqqoniylik, xolislik va fikr erkinligi borasida <i>mulohaza yuritiladi</i> .
<i>This work considers</i> the actual tendencies of the newspaper headline and its Font presentations.	<i>Maqolada</i> gazeta sarlavhalarida qo'llaniladigan shriftli bezashning dolzarb tendensiyalari haqida <i>gap borgan</i> .
The actual tendencies of the newspaper headline complex are <i>considered in this work</i> .	<i>Ushbu tadqiqotda</i> gazeta dizaynini yaratishda rangli tasvirlarning o'rni haqida <i>munozara olib borilgan</i> .
<i>The paper considers</i> love in the lyrics of N. Gumilyov and	<i>Ilmiy ishda</i> N. Gumilyov hamda Sharqda turli davrlarda ijod

poets of the East of different ages.	qilgan shoirlar lirikasida muhabbat konseptining mushtarak tomonlari <i>tahlil etilgan.</i>
--------------------------------------	---

Yuqorida keltirilgan “*the article investigate, the author speaks about, the paper considers*” kabi kollokatsiyalarning bir nechta ma’nodagi tarjimalarini ko’rish mumkin. Ushbu kontekstual tarjimalar grammatik kollokatsiya usulidan foydalanilgan. Mutaxassis tarjimon, o’z ilmiy ishlarida ushbu kollokatsiyalarning kontekstual tarjimalari orqali o’z fikrini asliyat tilidan ko’zlangan tarjima tiliga tarjima qilgan. Bu esa o’z navbatida tarjima jarayonida kollokatsiyalardan foydalanish maqolaning ilmiy saviyasini oshiradi. Shuning bilan bir qatorda muallif tomonidan yozilgan maqolani o’quvchi undan foydalanganda adekvat tarjima qilingan maqoladan foydalanishi mumkin.

Tahlil qilingan kontekstual tarjimani parallel korpusda yaratish uchun ma’lumotlarni tabular formatiga keltirib olish lozim. Ya’ni *Excel, TMX, XML, XLIFF* kabi oddiy formatdagi fayllarga keltirib olish kerak. Parallel korpuslarni yaratish jarayonida dastlab mashina tarjimasi bosqichlaridan foydalaniladi. Matnlarning bir-biriga tog’ri tushishi va tarjimada adekvatlik hosil qilishi uchun biror-bir so’zning bir nechta ma’nosini ichidan keraklisini tanlashda muhim o’rin tutadi. Shuningdek, ma’lumotlarni yuqorida keltirilgan ushbu formatlardan biriga keltirib olgach, (*Excel, TMX, XML, XLIFF*) dasturni yaratish uchun yirik korpuslar platformasiga rasmiy ravishda murojaat qilib, unga a’zo bo’lgach, o’sha korpus platformasidagi yo’riqnomalarni amalga oshirgan holda parallel korpusni yaratish mumkin.

Xulosa. Xulosa sifatida qayd etish joizki, parallel matnlarni tahlil qilishning statistik usullari turli tillarda axborot qidirish tizimlarining bir qismiga aylanib ulgurdi. Shuningdek, hozirda tarjimani avtomatlashtirishda parallel korpusdan foydalanish va uni amaliyotga tatbiq qilish natijasida bir qancha yutuqlarga erishilmoqda. Shunday ekan, qanday korpus yaratishdan qat’i nazar dastlab korpus yaratish uchun duch keladigan muammolarni chuqurroq o’rganish muhim va unga yechim topish uchun korpus turlarining tizimlarini o’zlashtirish lozim. Masalan, parallel korpusni yaratishda duch keladigan muammolar asosan matnlarning qaysi formatda tanlanishi, matnlar va manbalar bazasining to’g’ri shakllantirilishi hamda korpus mutaxassisi uchun tanlangan tarjima matnlarining qaysi sohada, usul va uslubda ekanligi bilan bog’liq. Xususan, ilmiy matnlarning kontekstual tarjimalarini ingliz va o’zbek tillaridagi bazalarini shakllantirish jarayonida asosan e’tibor uning *Excel* formatida bir-biriga mos joylashtirishiga bog’liq. Parallel korpuslar tuzilishi, tarkibi, tuzilish tamoyillari bilan mushtaraklikka, shu xususiyatlari bilan bir qancha faqrlarga ham ega bo’ladi. Dunyo kompyuter lingvistikasida konkordans dasturi, korpus menejeri, ular asosida maxsus maqsadlarni ko’zlovchi parallel yoki boshqa turdagi korpuslarni yaratish imkoni katta. O’zbek tili uchun ham mana shu

kabi korpus turlarini yaratish va o'zbek tilining kontentini kengaytirish amaliy tilshunoslik va kompyuter lingvistikasi sohasida hamda ushbu soha mutaxassislari oldida turgan muhim vazifa hisoblanadi.

Foydalanilgan adabiyotlar

1. Benson M., Benson E., Ilson R. The BBI combinatory dictionary of English: A guide to word combinations. — Amsterdam.: «John Benjamins press». 1986, — 462 p.
2. Brown P.F., Lai J.C., Mercer R.L. Aligning sentences in parallel corpora, proceedings of the 29th Annual Meeting of the Association for Computational Linguistics. Berkeley, 1991 — P. 169-176. — URL: <https://aclanthology.org/P91-1022>
3. Gale W.A., Church K.W. A program for aligning sentences in bilingual corpora. Computational Linguistics, 1993. — P. 75–102. — URL: <https://aclanthology.org/J93-1004>
4. Kay M., Roscheisen M. Text-translation alignment. Computational Linguistics: 1993. — P. 121–142. — URL: <https://aclanthology.org/J93-1006>
5. Newmark P.A. Textbook of translation. — New York.: «Prentice Hall International press». 1988. — 186 p.
6. Resnik P. Parallel Strands: A Preliminary Investigation into Mining the Web for Bilingual Text, Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA-98), — Langhorne, PA, 1998. — URL: https://link.springer.com/chapter/10.1007/3-540-49478-2_7
7. Riabi M. *Translating collocations in specific text from English into Arabic*. 2019. — 21 p. DOI:[10.5430/elr.v7n3p51](https://doi.org/10.5430/elr.v7n3p51)
8. V.Tangriyev. Problems in creating parallel corpora. Academic research in Educational sciences. Multidisciplinary Scientific journal. — Vol.3. Issue 3. —2022. — 1005 p.
9. Добровольский Д. Какие задачи решают параллельные корпусы. — URL: <https://postnauka.ru/video/54851>
10. Добровольский Д. Какие задачи решают параллельные корпусы? (Электрон ресурс). — URL: <https://postnauka.ru/video/>
11. Кокорева А.А. Корпус параллельных текстов в обучении иностранному языку. — Вестник ТГУ (Тамбовский государственный университет), выпуск 2 (118), 2013. — С. 57–62.
12. Соснина Е.П. Параллельные корпусы в обучении языку и переводу. (Электрон ресурс): — URL: http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_education_translation/
13. Соснина Е.П. Параллельные корпусы в обучении языку и переводу. (Электрон ресурс): — URL: http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_education_translation/
14. Тао Ю., Захаров В.П. Разработка и использование параллельного корпусарусского и китайского языков.

- Автоматизация обработки текста. (Электрон ресурс): — URL: <https://www.researchgate.net/publication/306911853>
15. <https://translationrating.ru/glossary/parallelnyj-korpus/>
16. zamin.uz/uz/jamiyat/79743-

References

1. Benson M., Benson E., Ilson R. *The BBI combinatory dictionary of English: A guide to word combinations*, Amsterdam: John Benjamins press. 1986, 462 p.
2. Brown P.F., Lai J.C., Mercer R.L. *Aligning sentences in parallel corpora, proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*. Berkeley, 1991, pp. 169–176. Available at: <https://aclanthology.org/P91-1022>
3. Gale W.A., Church K.W. *A program for aligning sentences in bilingual corpora*. Computational Linguistics, 1993, pp. 75-102. Available at: <https://aclanthology.org/J93-1004>
4. Kay M., Roscheisen M. *Text-translation alignment*. *Computational Linguistics*: 1993, pp. 121–142. Available at: <https://aclanthology.org/J93-1006>
5. Newmark P.A. *Textbook of translation*, New York.: Prentice Hall International press, 1988, 186 p.
6. Resnik P. *Parallel Strands: A Preliminary Investigation into Mining the Web for Bilingual Text, Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA-98)*, Langhorne, PA, 1998. Available at: https://link.springer.com/chapter/10.1007/3-540-49478-2_7
7. Riabi M. Translating collocations in specific text from English into Arabic. 2019, 21 p. Available at: DOI:[10.5430/elr.v7n3p51](https://doi.org/10.5430/elr.v7n3p51)
8. Tangriyev V. Multidisciplinary Scientific journal. Vol.3. Issue 3. 2022,
9. Dobrovolskii D. *Kakie zadachi reshayut parallel'nye korpusy?* (What tasks are solved by parallel buildings?), available at: <https://postnauka.ru/video/54851>
10. Dobrovolskii D. *Kakie zadachi reshayut parallel'nye korpusy?* (What tasks are solved by parallel buildings?), available at: <https://postnauka.ru/video/>
11. Kokoreva A.A. *Korpus parallel'nykh tekstov v obuchenii inostrannomu yazyku. Vestnik TGU*, vypusk 2 (118) (The corpus of parallel texts in foreign language teaching. Bulletin of TSU (Tambov State University), issue), 2013, pp. 57-62.
12. Sosnina E.P. *Parallel'nye korpusy v obuchenii yazyku i perevodu* (Parallel corpora in language teaching and translation). Available at: http://ling.ulstu.ru/linguistics/resourses/literature/articles/corpus_education_translation/
13. Sosnina E.P. *Parallel'nye korpusy v obuchenii yazyku i perevodu* (Parallel corpora in language teaching and translation). Available at: http://ling.ulstu.ru/linguistics/resourses/literature/articles/corpus_education_translation/
14. Tao Yu., Zakharov V.P. *Razrabotka i ispol'zovanie parallel'nogo korpusarusskogo i kitaiskogo yazykov. Avtomatizatsiya obrabotki*

teksta (Development and use of parallel corpus of the Russian and Chinese languages. Automation of text processing), available at: <https://www.researchgate.net/publication/306911853>

15. <https://translationrating.ru/glossary/parallelnyj-korpus/>
16. zamin.uz/uz/jamiyat/79743