PARALLEL KORPUSLARNING TARJIMASHUNOSLIKDAGI AHAMIYATI

Akramdjanova Madina,
Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti tayanch doktoranti

DOI: https://doi.org/10.5281/zenodo.20427496

Annotatsiya. *Mazkur maqolada parallel matnlar korpusining tarjimashunoslik va zamonaviy lingvistika sohalaridagi nazariy hamda amaliy ahamiyati yoritilgan. Parallel korpuslarning tarjima jarayonida muhim manba sifatida xizmat qilishi, shuningdek, terminologik ma'lumotlar bazasini yaratishda tutgan o'rni tahlil etilgan. Shuningdek, parallel korpuslarning ikki tilli, uch tilli va ko'p tilli shakllari hamda ularni yaratish texnologiyalari haqida ma'lumot berilgan. Parallel korpuslar orqali esa tilning dialektal variantlari va turli ijtimoiy guruhlarga xos nutq xususiyatlarini o'rganish imkoniyatlari asoslab berilgan. Maqola natijalari parallel korpuslarning lingvistik tadqiqotlar, tarjimashunoslik va til ta'limidagi ahamiyatini ochib berishga xizmat qiladi.*

Kalit so'zlar: *parallel korpus, segment, birlik, valentlik, muvofiqlashtirish, teglash.*

Аннотация*. В данной статье рассматривается теоретическое и практическое значение параллельных корпусов в переводоведении и современной лингвистике. Анализируется роль параллельных корпусов как важного источника в процессе перевода, а также их роль в создании терминологической базы данных. Также представлена информация о двуязычных, трехъязычных и многоязычных формах параллельных корпусов и технологиях их создания. Обосновываются возможности изучения диалектных вариантов языка и речевых характеристик различных социальных групп с помощью параллельных корпусов. Результаты статьи позволяют выявить значимость параллельных корпусов в лингвистических исследованиях, переводоведении и языковом образовании.*

Ключевые слова*: параллельный корпус, сегмент, единство, валентность, координация, тегирование.*

Annotation. *This article explores the theoretical and practical significance of parallel text corpora in the fields of translation studies and modern linguistics. It analyzes the role of parallel corpora as an important resource in the translation process, as well as their contribution to the development of terminological databases. The paper also provides information on bilingual, trilingual, and multilingual parallel corpora and the technologies used to create them. Furthermore, it substantiates the potential of parallel corpora for studying dialectal variations of language and speech characteristics specific to different social groups. The findings of the article highlight the importance of parallel corpora in linguistic research, translation studies, and language education.*

Keywords: *parallel corpus, segment, unit, valency, alignment, tagging.*

Kirish. Parallel korpuslar tarjimashunoslikda muhim manba hisoblanadi. Shuningdek, ushbu sohada erishilgan natijalardan terminologik ma'lumotlar bazasini yaratishda keng foydalanmoqda. Parallel korpuslarning bugungi kunda ikki tilli, uch tilli va ko'p tilli ishlash texnologiyalari yaratilgan. Ushbu uch tilli ingliz, nemis va franzus tillaridagi parallel korpusga namuna keltirilgan.

Parallel korpuslar bir tilli, ikki tilli, uch tilli yoki ko'p tilli bo'lishi mumkin.

Bir tillardagi parallel korpus muayyan tildagi matnning dialektal variantlari yoki muayyam sotsiumga tegishli kishilar nutqida u yoki bu matnda qanday qo‘llanganligini ifodalanishi mumkin.

Adabiyotlar sharhi. Ikki tildagi parallel korpuslarni V.Zaharov ikki turga shartli ravishda ajratadi:

1) muayyan tildagi asliyat matnning ikkkinchi tildagi aynan tarjimasi;
2) biror sohaga aloqador matnlarning mavzulashtirilgan guruhlar bo‘yicha bir tasnifga kiritilishi.

Parallel matnlar asosida korpus yaratish u darajada oson emas. Chunki ikki tilning o‘zaro mos matn fragmentlariga ko‘ra juftlanadi. Ushbu masalani avtomatik tarzda hal qilish uchun turli metodlar ishlab chiqilgan, buning uchun ikki tildagi matndagi grammatik konstruksiyalar (gaplar va so‘z birikmalari), so‘zlar muvofiqlashtiriladi. A.V.Zubov va I.I.Zubovalarning tadqiqotida jumlalar darajasidagi moslashtirish orqali foydalanish mumkin. Ushbu moslashtirish olti xil usulda amalga oshirilishi mumkin[1,2013:20]:

1) asliyatdagi bir jumla ikkinchi tilga ham bir jumla tarzida tarjima qilinishi;
2) asliyatdagi ikkita jumla ikkinchi tilga bir jumla tarzida tarjima qilinishi;
3) asliyatdagi bir jumla ikkinchi tilga ikkita jumla qilib tarjima qilinishi;
4) asliyatdagi ikkita jumla ikkinchi tilga ham ikkita jumla tarzida tarjima qilinishi;
5) asliyatdagi jumla ikkinchi tilga tarjima qilinmasligi;
6) asliyatdagi matnning tarjima qilinadigan tilda mos ekvivalentini mavjud emasligi.

Tadqiqot metodologiyasi. Parallel matnlarning mos juftliklarini tuzish uchun inson tomonidan tahrir qilingan matnlar kiritilib, lug‘at orqali matnda uchragan segment birliklarga qarab avtomatik tarzida tarjima birliklari aniqlanadi. Buning uchun ikki tilga mos leksik va grammatik jihatdan ishlab chiqilgan mos variantlar lug‘atga kiritilgan bo‘lishi kerak.

Parallel korpuslar quyidagi maqsadlarda qo‘llanilishi mumkin[2,2013:20]:

1) ikki tilli yoki ko‘p tilli tarjima lug‘atlarni tuzish;
2) mashina tarjima tizimlari uchun lug‘at yaratish va uni doimiy tarzida to‘ldirib borish;
3) kontekstda uchraydigan ko‘p ma’noli so‘zlarni kompyuter tahlili orqali leksik birliklarni polisemiyaga oidligini aniqlash;
4) matndagi terminologik va frazeologik birliklarni tarjima qilish;
5) korpusdan foydalangan holda ularning mos variantlaridan foydalanib, kompyuter xotirasi usuli orqali mashina tarjimasi tizimi uchun to‘liq avtomatik tarjimani amalga oshirish.

Leksikani avtomatik tasniflash matnni avtomatik tushunish uchun asosiy jarayonlardan biridir[3,2004:55-64].

Tahlillar va natijalar. Parallel korpuslar yordamida so‘z va so‘z birikmalarining o‘zaro valentliklarini aniqlash, bu muayyan guruhga kiruvchi leksik-semantik guruhdagi so‘zlarning semantik maydonini aniqlash, qolaversa, mashina tarjimasi uchun tarjima sifatini oshirishga xizmat qiladi.

Parallel matnlarni tekislash(moslashtirish) – ikki til (tarjima va asl) matnidagi bir-biriga to‘g‘ri keluvchi gap, abzaslarni identifikatsiyalash jarayondir. Matn qismlarini bunday moslashtirish turlicha bo‘lishi mumkin. Qismlar muallif maqsadiga qarab har xil bo‘ladi: bob, bo‘lim, abzas, gap va h.k. Foydalanuvchi uchun eng qulay moslashtirilgan shakl bu asl va tarjima matnlarining gaplar bo‘yicha muvofiqlashtirilishi hisoblanadi. Bu juda murakkab, ko‘p kuch talab etiladigan jarayonni talab etadi. Bugungi kunga qadar dunyoda to‘g‘ri muvofiqlashtirilgan, namunali sanaladigan parallel korpuslar sanoqlidir. Chunki asl va tarjimadagi matn lavha(qism)larini muvofiqlashtirishning avtomatik usuli yaxshi natija bermaydi. Eng mukammal ishlangan avtomatik tarjima dasturi ham gaplarning chegarasini aniqlashda aniq yechim berolmaydi. Asl matndagi fikr bir gap bilan, ikki-uch gap bilan ifodalanishi mumkin yoki aksincha. Shuning uchun bitta buyruq bilan asl va tarjima matn qismlarini tenglashtirishning imkoni bo‘lmaydi. To‘g‘ri, shunga mo‘ljallangan texnik vositalar mavjud, ammo ularning hech biri inson aralashuvisiz muammoni hal qilmaydi. Parallel korpus tarkibidagi matnni gaplar bo‘yicha mosligiga erishish turli lingvistik amallar bajarishga yo‘l ochib beradi. Ma’limki, tarjima jarayonida asl matndagi bitta jumla tarjima matnda ikki jumlaga bo‘linishi, aksincha, asl matndagi bir necha gap tarjimada bitta gap holida berilishi, gaplarning izchillik o‘rni almashtirilishi mumkin. Shuning uchun parallel korpus tuzishda eng murakkab jarayon asl va tarjima matndagi moslashtirish jarayoni hisoblanadi[4]

“Parallel korpuslar matnni qanday tarjima qilishni o‘rgatmaydi, balki tarjimadagi o‘zaro mosliklarni ko‘rsatadi, ya’ni muayyan leksema, so‘z shakl yoki grammatik formaga boshqa tilda unga muvofiqlashtirilgan birlik konkordansda taqdim etiladi”. [5,2023:40]

Tarjimashunoslikda parallel korpuslardan foydalanish ikki va undan ortiq tillarning kontrastiv tahlilini samarali amalga oshirish imkonini beradi. Ular tarjimonlarga manba va maqsadli til o‘rtasidagi ekvivalentlarini aniqlashda yordam beradi, xususan, to‘g‘ridan to‘g‘ri muqobili mavjud bo‘lmagan birliklar uchun mos kontekstual yechimlar topishga xizmat qiladi. Bundan tashqari, parallel korpuslar so‘zlar, birikmalar, frazeologik birliklar, metaforik ifodalar va stilistik ko‘chimlarning har ikki til sistemasi doirasidagi ishlatilish chastotasi haqida aniqlovchi ma’lumotlar beradi. Ushbu empirik asos tarjimonlarga kontekstga mos tarjima strategiyalarini ishlab chiqishda real vosita bo‘lib xizmat qiladi.[6,2026:24]

Dobrovolskiyning ta’kidlashicha ikki tildagi matn bo‘laklari – gaplarni bir-biriga moslashtirish bilan birga matn birliklari teglanishi ham zarur. Bunday teglar grammatik,

semantik va boshqa xil bo‘lishi mumkin. Teglar ma’lum bir so‘zga yoki butun matnning o‘ziga biriktirilgan bo‘lishi mumkin. Agar matn birliklari teglanmagan bo‘lsa, unda korpusdan ma’lum bir grammatik shaklni qidirish imkoni bo‘lmaydi. Demak, matn grammatik teglarga ega bo‘lganda, bir tildagi so‘zshaklning ikkinchi tildagi ekvivalentini qidirish, kuzatish mumkin. Agar so‘zshakllar grammatik teglanmagan bo‘lsa, ikki til ekvivalentlarini tekshirish faqat sintaktik butunliklar orasidagina amal qiladi.[7]

Unga ko‘ra, parallel korpuslar faqat tarjima nazariyasi, ikki tilli leksikografiya, tillarni qiyoslash bilan bog‘liq tilshunoslikning boshqa sohalari uchun foydali ekanligi haqidagi xulosa bilan chegaralanish noo‘rin[8]. Qiyoslanayotgan tillar leksikasidagi so‘zlarning semantik tarkibini yanada chuqurroq tavsiflashda ham muhim ahamiyat kasb etadi.

Xulosa qilib aytganda, dunyo tilshunosligidagi parallel korpuslarga oid tadqiqotlar e’tiborlidir. Ular o‘z tilining milliy korpuslarini yaratish asnosida butun dunyo bilan bog‘lanmoqda hamda ushbu tilni o‘rganuvchilar uchun qator yangiliklar yaratilmoqda.

Parallel matnlar korpusi qiyosiy-chog‘ishtirma aspektdagi tadqiqotlarda o‘ziga xos o‘rin tutadi. Mukammal tarzda qilingan tarjimalar parallel korpuslarni tuzishga asos bo‘ladi. Parallel korpuslar tarjimashunoslikda muhim manba sifatida qaraladi va terminologik ma’lumotlar bazasini yaratishda foydalaniladi.

Foydalanilgan adabiyotlar ro‘yxati:

1. Захаров В.П., Богданова С.Ю. Корпусная лингвистика. Учебник Санкт-Петербург. 2013. – С.20
2. Захаров В.П. Богданова С.Ю. Корпусная лингвистика. Учебник. Санкт-Петербург. 2013. – С.21
3. Беляева Л.Н. Лексикографический потенциал параллельного корпуса текстов // Труды международной конференции «Корпусная лингвистика – 2004». – СПб., 2004. – С. 55-64
4. Добровольский Д. Какие задачи решают параллельные корпусы. (Электрон ресурс): https://postnauka.ru/video/54851
5. Abjalova M. Parallel korpuslarni yaratish asoslari // Kompyuter lingvistikasining zamonaviy texnologiyalari –CTCL. –Toshkent, 2023. –B.40
6. Muhammadiyeva D. Boburnoma matni o‘zbekcha-turkcha-qozoqcha parallel korpusining lingvistik va tarjimashunoslik asoslari. Filol.fan.dok.(DsC)…diss. –Toshkent, 2026. –B.24
7. Добровольский Д. Какие задачи решают параллельные корпусы. (Электрон ресурс): https://postnauka.ru/video/54851
8. Добровольский Д.О. Корпус параллельных текстов в исследовании лексической семантики. http://www.dialog-21.ru/media/2503/dobrovolskij.pdf