

KORPUSGA ASOSLANGAN MASHINA TARJIMASI UCHUN IKKI TILLI FRAZEOLOGIK LUG‘AT YARATISHNING LINGVISTIK AHAMIYATI (BADIY ASAR MISOLIDA)

Sirojiddinova Muharram

O‘zbekiston Milliy universiteti Kompyuter
lingvistika mutaxassisligi magistranti.

Annotatsiya: Ushbu tezisdagi kompyuter texnologiyasi va tarjimashunoslik sohalarining integratsiyalashuvi natijasida mashina tarjimasi sohasining paydo bo‘lishi, ushbu sohaning yo‘nalishlari va korpusga asoslangan mashina tarjimasining muammolari, ya’ni badiiy asarlardagi frazeologik birliklarni boshqa tilga tarjima qilishdagi nomuvofiqliklar hamda bu muammolarni hal qilishda badiiy asarlarning ikki tilli yoki ko‘p tilli parallel korpuslarini yaratish va ular asosidagi frazeologik birliklar lug‘atini yaratishning lingvistik ahamiyati yoritilgan.

Kalit so‘zlar: frazeologik lug‘at, qoidaga asoslangan mashina tarjimasi (MT), korpus asosidagi mashina tarjimasi (Corpus-based MT), gibrid mashina tarjimasi (Hybrid machine translation (HMT)).

Abstract: In this thesis, the emergence of the field of machine translation as a result of the integration of the fields of computer technology and translation studies, the directions of this field and the problems of corpus-based machine translation, that is, inconsistencies in the translation of phraseological units in works of art into another language, and the use of works of art in solving these problems the linguistic importance of creating bilingual or multilingual parallel corpora and creating a dictionary of phraseological units based on them is highlighted.

Keywords: phraseological dictionary, rule-based machine translation (MT), corpus-based machine translation (Corpus-based MT), hybrid machine translation (Hybrid machine translation (HMT)).

Аннотация: В данной диссертации рассматривается возникновение области машинного перевода в результате интеграции областей компьютерных технологий и переводоведения, направления этой области и проблемы корпусного машинного перевода, то есть несоответствия в переводах фразеологизмов в произведениях искусства на другой язык и использования произведений искусства при решении этих задач подчеркивается лингвистическое значение создания двуязычных или многоязычных параллельных корпусов и создания на их основе словаря фразеологизмов.

Ключевые слова: фразеологический словарь, машинный перевод на основе правил (MT), корпусный машинный перевод (Corpus-based MT), гибридный машинный перевод (Hybrid Machine Translation (HMT)).

Bugungi kunda kompyuter texnologiyasi, avtomatik tarjima, mashina tarjimasi, google tarjimon, amazon tarjimon, microsoft tarjimon, DeepL, parallel korpus kabi atamalarning kundalik iste’molda tez-tez takrorlanishi, nafaqat tarjimashunoslik, balki mashina tarjimasi (XX asr 50-yillaridan kompyuter va tilshunoslik sohalarining integratsiyasidir) sohalarida ko‘plab ishlar amalga oshirilganligidan dalolat beradi. Ko‘pchilikka ma’lumki, bu boradagi birinchi tajriba

1954-yil AQSHda Jorjtaun universitetida o'tkazilgan bo'lib, 1960-yilda esa mustaqil fan sifatida shakllangan [Rahimov A. 2011: 160].

Jahonda mashina tarjimasi sohasini rivojlantirish uchun ko'plab olimlar ilmiy izlanishlar olib borgan. N.D. Andreev, I.A. Melchuk, I.I. Revzin, V.Yu. Rozensveyg, Yu.N. Marchuk, R.G. Piotrovskiy, Yu.A. Motorin, K.B. Bektaev, A.N. Belyaev, I.K. Belskiy, A.V. Zubov, G.E. Miram, L.L. Nelyubin, V.I. Perebiynos, V.A. Chijakovskiy, Ye.A. Shingarev, G.G. Belonogov, R.G. Kotov, Babushkina N.V, Z.Shalyapina, O.Yu. Mansurova, A.S. Panina, A.A. Xoroshilov (Rossiya); A.But, R. Richans, J. Xatchins, J.Allen, P. Braun (AQSh); M. Nagao (Yaponiya); A.Vaxer (Estoniya); Frantsiyada J.Astrouni; R.Sinha, A.Jain (Hindiston); B.Bleyzer, U.Shvol, A.Storrer (Germaniya) singari tadqiqotchilarning ishlari shular jumlasidandir [N.Abduraxmonova. 2012]. Shu bilan birgalikda, D.O.Dobrovolskiy, Yu. Tao, A.A.Kokoeva va V.Zaxarov kabi olimlar parallel korpusning tuzilishi, tarkibi va imkoniyatlari bo'yicha tadqiqotlar olib borib, ko'plab asarlar yaratishdi. Ularning fikricha, mashina tarjimasi sohasida parallel korpuslar yuaratish muhim ahamiyatga ega, chunki unga asl matn va tarjima matn biteks ko'rinishida kiritiladi. Bu esa matn tajrimasida konteksga qarab matn mazmunini aniqlash va boshqa tilga tarjima qilishga semantik muvofiqini topish hamda parallel matnlarning semantik munosabatiga oid ko'plab kerakli ma'lumotlarni o'z ichiga olish imkonini beradi.

Shuningdek, respublikamizda ham ushbu sohani rivojlantirishga qaratilgan jiddiy tadqiqotlar amalga oshirilgan hamda chet tildagi matnlarni o'zbek tiliga va yoki buning aksini tarjima qiluvchi dasturning lingvistik ta'minotini yaratishga ham alohida e'tibor qaratilgan. Ushbu sohaga dasturchi va matematiklar ham e'tibor qaratib muayyan tadqiqotlar olib borganligi eslatish ham bu sohaning dolzarbligini anglatadi [Хакимов М. 2008: 297].

Tarjimashunoslikda frazeologik birliklarga oid masalalar G'.Salomov, Q.Musaev, Yu.Po'latov, S.Abidxodjaeva, G.K.Odilova, I.G'afurov, O.Mo'minov, N.M.Qambarov, M.Xolbekov, Sh.Safarov kabi olimlarning tadqiqotlarida atroflicha o'rganilgan. Zamonaviy tilshunoslikda frazeologik birliklarni boshqa tillar bilan qiyosiy o'rganib, ularning milliy-madaniy xususiyatlarini tadqiq qilish keng tus oldi. Bu jihatdan L.Kulieva, K.Musaev, I.Stepanova, E.Arsenteva, I.N.Isabekov singari tilshunos olimlarning ishlari bir-biriga qardosh bo'lmagan ikki til doirasida olib borilgan tadqiqotlar yuzasidan ahamiyatlidir.

Bugungi kunda mashina tarjimasining quyidagi yo'nalishlari mavjuddir. Bu yo'nalishlarni aniqlashda mashina tarjimasini lingvistik aspektda o'rgangan Jon Xatchinsning xizmatlari kattadir. Olim mashina tarjimasi tizimlarida lingvistik tahlil bosqichlarini quyidagicha tasniflaydi [N.Abduraxmonova. 2012]:

Qoidaga asoslangan mashina tarjimasi (MT):

- transfer tizim (transfer-based machine translation);
- interlingval tizim (interlingual machine translation);
- lug'at asosidagi tizim (dictionary-based machine translation).

Korpus asosidagi mashina tarjimasi (Corpus-based MT):

- statistikaga asoslangan MT (statistics-based);
- namuna asosidagi MT (example-based).

Gibrid mashina tarjimasi (Hybrid machine translation (HMT)).

Hozirda bu yo‘nlishlardan eng keng tarqalagani sifatida korpusga asoslangan mashina tarjimasini keltirish mumkin. Bunda parallel korpusga kiritilgan bitekslarning (inglizcha – o‘zbekcha matn) har biri avtomatik tahlil qilinadi va tarjima uchun muayyan xulosalarga kelinadi[Abdurakhmonova, N. 2021,2022]. Mohir tarjimonlar yordamida tajima qilingan parallel matnlardagi so‘zlar, frazalar konteksga qarab ma’nolari aniqlanadi.

Tilshunosligimizda o‘zbek tili frazeologiyasini qarindosh bo‘lmagan ingliz tili frazeologiyasi bilan solishtirish ikkala tilning o‘ziga xos xususiyatlarini yaxshiroq anglashga, tillararo frazeologik birliklarning mohiyatini aniqlashga imkon beradi. O‘zbek va ingliz tillarining frazeologik birliklari o‘rtasidagi o‘xshashlik va farqlar nimada, ular tilda, asosan, qanday namoyon bo‘ladi; ular qanday intralingvistik va ekstralingvistik omillar bilan belgilanadi, frazeologik birliklarning tillararo ekvivalentligi darajasi qanday; qardosh bo‘lmagan tillarning frazeologiyasini qanday moslash mumkin degan savollarga javob topish nafaqat tilshunosligimizning, balki mashina tarjimasini sohasining istiqbollari ham belgilaydi[Abdurakhmonova, N. 2021,2022].

A.V.Kunin frazeologizmlarni quyidagicha ta’riflagan: «fazeologik birlik – qisman yoki to‘laligicha ko‘chma ma’noli so‘zlarning barqaror birikmasi» A.Kunin frazeologik birliklardagi turg‘unlik belgisini ularning leksikasida ishlatilishi, semantikasi, morfologik, sintaktik tuzilishida mavjudligi orqali sharhlaydi [Nurmuhammedov Y. 2022: 90]. Bundan ko‘rinib turibdiki, har bir tilning frazeologiyasi o‘sha tilning madaniyati, yashash tarzi, geografik joylashuvi, odatlariga bog‘liq bo‘lganligi uchun ularni qardosh bo‘lmagan boshqa tilga tarjima qilganda o‘zining anglashilgan ma’nosiga putur yetdi. Agar bunday birliklar berilgan asarlarni boshqa tilga tarjima qilsak, asarning mazmun-mohiyati, g‘oyaviy ahamiyati va tarjima sifatiga zarar yetadi. Chunonchi, inglizlarda “A little bird told me” frazema mavjud bo‘lib, agar ma’lumotni oldindan eshitgan yoki bilgan, lekin kimdan eshitgani aytishni xohlamasangiz, shu iboradan foydalanishingiz mumkin. Bu ibora o‘zbek tilidagi “ teshik quloq eshitadi” iborasiga semantik jihatdan mos keladi. Ushbu ibora asarda Did you hear that Ann is going to marry? Where did you find this information? A little bird told me tarzida qo‘llanilgan bo‘lib, mashina tarjimasini orqali tarjima qilinganda, Enn turmushga chiqmoqchi ekanligini eshitdingizmi? Bu ma’lumotni qayerdan topdingiz? Kichkina qush aytdi kabi tarjima qilinib, ma’no jihatdan uslubiy g‘alizlikni keltirib chiqaradi. Shuningdek, “Aching heart” idiomasi o‘zbek tilida “ko‘ngli yarim yoki yuragi yarim” iborasiga to‘g‘ri keladi. Lekin tarjima dasturlar orqali tarjima qilinsa, “yaralangan yurak yoki og‘riyotgan yurak” deb tarjima qilinadi. Bu tarjima esa hamma konteks mazmuniga ham muvoqil kelavermaydi. Yoki “Able to take (can take)” iborasi so‘z o‘zbek tilida semantik jihatdan “Tosh kelsa kemirmoq, suv kelsa simirmoq” iborasiga to‘g‘ri keladi. Eng achinarlisi, bu ibora tarjima dasturlar orqali tarjima qilinsa, “olish mumkin” deb tarjima qilinadi. Matn tarjimasida bunday xatoliklarni ko‘p bora takrorlanishi asl matn mazmunini keskin o‘zgartiradi. Shunday ekan, badiiy asarlar asosidagi frazemalar lug‘atini yaratib, parallel korpusga biteks ko‘rinishida kirtilsa va doimiy boyitib borilsa, tarjima jarayonida har qanday badiiy matnning asliyatiidagi ma’nosini saqlab qolish imkoniyati yaratiladi

Xulosa

Ta’kidlash joizki, bugungi kunda mashina tarjimasini sohasida juda ko‘plab tadqiqotlar amalga oshirildiki, buning natijasi o‘laroq, so‘zma-so‘z tarjima usulidan voz kechildi, tarjima matn ancha mahalliyashtirildi, ilmiy va rasmiy matnlarning bir tildan boshqa tildagi tarjimalari mukammallashtirildi. Lekin har bir tilning o‘ziga xos lingvistik xususiyatlari, leksik va frazeologikbirliklari hamda so‘zlarining ma’no nozikliklari boshqa tilga to‘laonli mos kelmasligi

sababli badiiy matnlarni boshqa tilga tarjima qilinganda, asl matnning nafaqat g'oyasi, balki mazmuni ham keskin o'zgarishga uchramoqda. Shu sababli, badiiy asarlarning ikki tilli biteks matnlarini hamda ular asosidagi frazeologik birliklar lug'atini yaratish va ushbu sohaga tatbiq qilish ushbu sohani yanada rivojlantirish va tarjima sifatini yaxshilashga xizmat qiladi. Bu esa, o'z navbatida, yuqorida ko'zda tutilgan muammolarning qisman hal qililishiga olib keladi.

Adabiyotlar ro'yhati:

1. Abduraxmonova N. Inglizcha matnlarni o'zbek tiliga tarjima qilish dasturining lingvistik ta'minoti (sodda gaplar misolida).2018.
2. Abdurahmonova N. Mashina tarjimasi lingvistik asoslari. Qo'llanma. T.: Akademnashr, 2012.
3. Абдурахмонова, Н., & Абдувахобов, Г. (2021). О 'quv lug 'atini tuzishning nazariy metodologik asoslari. *Международный журнал искусство слова*, 4(6).
4. Abdurakhmonova, N. (2021). Formal-Functional Models of The Uzbek Electron Corpus. *ANGLISTICUM. Journal of the Association-Institute for English Language and American Studies*, 10(8), 59-66.
5. Abdurakhmonova, N., Alisher, I., & Toirova, G. (2022, September). Applying Web Crawler Technologies for Compiling Parallel Corpora as one Stage of Natural Language Processing. In *2022 7th International Conference on Computer Science and Engineering (UBMK)* (pp. 73-75). IEEE.
6. Abdurakhmonova, N., Shakirovich, I. A., & O'G'Li, K. N. S. (2022). Morphological analyzer (morfoAnalyse) Python package for Turkic language. *Science and Education*, 3(9), 146-156.
7. Nurmammedov Y. Frazeologiyani dolzarb masalalari. O'quv-uslubiy qo'llanma. – S.: SamDCHTI, 2022. – 90 b.
8. Rahimov A. Kompyuter lingistikasi asoslari. – T.: Akademnashr, 2011. – 160b.
9. Хакимов М.Х. Расширяемый входной язык математического моделирования естественного языка для многоязычной ситуации машинного перевода // ЎЗМУ хабарлари, 2009. – № 1. – С.75-80;