



<https://doi.org/10.70728/b.series.fil.v08.i06.021>

MASHINA TARJIMA TIZIMLARINING LEKSIKOGRAFIK MUAMMOLARI: SO‘ZMA-SO‘Z TARJIMADAN NEYRON TARJIMAGA O‘TISH JARAYONI HAQIDA

Abdujalilova Yoqutxon Alisher qizi

*O‘zbekiston davlat jahon tillari universiteti,
mustaqil izlanuvchi*

Uzswlu28@icloud.com

Ilmiy maslahatchi: Raxmanov Gulomjon Nematovich

*O‘zbekiston davlat jahon tillari universiteti,
katta o‘qituvchi*

e-mail: gulom1958@mail.ru

Annotatsiya Maqolada mashina tarjima tizimlaridagi muammolar hamda neyron mashina tarjimaga o‘tish jarayoni o‘rganiladi. Mazkur tarjima tizimlarining mustaqil ish metodlari va qiyoslash tizimi tahlil etiladi.

Kalit so‘zlar: mashina tarjimasi, neyron tarjima, leksikografiya, ko‘p ma’nolilik,, idiomalar, realiya, lakuna.

Abstract. The article deals with problems of machine translation systems and the process of transition from to neural translation. The independent working methods of translation and ways of comparing are studied.

Key words: machine translation, neural translation, lexicography, polysemy, homonymy, idioms, realia, lacuna.

KIRISH

Ma’lumki, tarjima jarayoni bosqichma- bosqich amalga oshiriladi. Avtomatik tarjimada bu bosqichlar esa aniq ajralib turadi. Bu analiz va sintez iborat bo‘lib ya’ni til materialini tanlash va tarjima qilishni o‘z ichiga oladi. Mazkur bosqichlarda avtomatik tarjima turli kodlar, formulalar, algoritmlar, tilning muayan birliklarini simvallashtiruvchi belgilar, statistik natijalar asosida amalga oshiriladi. Bu avtomatik tarjimaning sohasining o‘ziga xos ish uslubi, til materiallarini aniqlashtirish hamda qiyoslash tizimiga ega ekanligini bildiradi.

Matnni tillararo tarjima qilish va uning ma'nosini to'g'ri tushunish juda muhim. Biroq bu jarayon bir qator muammolarni keltirib chiqarishi tabiiy. Bularga, birinchi navbatda, tarjimada o'zga tildagi matn qayta yaratish masalalari kiradi; ikkinchidan, tarkibiy va leksik farqlar; va bundan tashqari, murakkab ko'p so'zli grammatik konstruksiyalar nazarda tutiladi. Bu masala mahalliy va xorijiy olimlar tomonidan o'rganilgan va yangi yonalishlarda tadqiq etilmoqda.

Ta'kidlash joizki, hozirgi kunda neyron mashina tarjimasi sohasida olib borilayotgan tadqiqotlarning asosiy qismi xorijlik olimlarning hissasiga to'g'ri keladi. Jumladan, A.B. Mironenko, S.M. Kalinin va N.G. Rtishcheva kabilar neyron mashina tarjimasining aniqligini oshirish va mashina tarjimasi tizimlarining sifatini baholashning joriy yondashuvlarini ko'rib chiqadi [7, 8,12].

Hozirgi zamonaviy neyron mashina tarjimasi esa kontekstni hisobga olgan holda ishlaydi va tarjima sifatini yaxshilashga xizmat qiladi.

ADABIYOTLAR TAHLILI VA METODOLOGIYA

Bugungi globallashuv jarayonida turli tillarda so'zlashuvchi insonlar o'rtasidagi muloqot tobora kuchayib bormoqda. Shu bois mashina tarjimasi va neyron tarjima tizimlari muammolari ko'plab tilshunos olimlar tomonidan tadqiq etilgan. Biroq ushbu dasturiy tarjima masalalariga doir muammolarning nazariy va amaliy jihatlarini tadqiq etish dolzarb mavzuga aylangan.

Xususan, Warren Weaver mashina tarjimasining dastlabki nazariy asoslarini yaratgan bo'lsa [11], Yehoshua Bar-Hillel avtomatik tarjimada semantik noaniqlik va kontekst muammolarini tahlil qilgan [2]. Zamonaviy neyron mashina tarjimasi masalalari esa Bahdanau, Cho va Bengio kabi olimlar tadqiqotlarida keng yoritilgan [1]. Daniel Jurafsky va James Martin esa nutq va matnga ishlov berish tizimlarining umumiy lingvistik modelini tizimlashtirgan [5]. Shu sababli mashina tarjima tizimlari tilshunoslik va axborot texnologiyalari sohasida muhim ahamiyat kasb etmoqda. Mashina tarjimasi — bu bir tildagi matn yoki nutqni avtomatik ravishda boshqa tilga o'girish jarayonidir.

Dastlabki mashina tarjima tizimlari asosan so'zma-so'z tarjima qilish usuliga (Rule-Based va Statistical) asoslangan edi [4]. Biroq bu yondashuv kontekst va ma'noni hisobga olmagan uchun ko'pincha noto'g'ri va tabiiy bo'lmagan tarjimalarga olib kelgan. Sun'iy intellekt rivojlanishi bilan neyron mashina tarjimasi (Neural Machine Translation – NMT) keng qo'llonila boshlandi.

Ushbu zamonaviy yondashuv butun gapni neyron tarmoqlar (Deep Learning) va kontekst asosida tarjima qilish imkonini beradi [3]. Shunga qaramay, har bir til o'ziga xos semantik, pragmatik va madaniy xususiyatlarga ega bo'lgani sababli, leksikografik muammolar hanuzgacha avtomatik tarjimaning eng murakkab jihatlaridan biri bo'lib

qolmoqda. Ayniqsa, ko‘p ma’noli so‘zlar, idiomatik birliklar, metaforik ifodalar hamda madaniy realiyalar mashina tarjima tizimlari uchun jiddiy qiyinchilik tug‘diradi.

NATIJALAR

Ta’kidlash joizki, polysemy (ko‘p ma’nolilik) hodisasi mashina tarjima tizimlaridagi eng muhim leksikografik muammolardan biridir. Bunday holatlarda tarjima tizimi so‘zning aynan qaysi ma’noda qo‘llanilganini aniqlash uchun matndagi kontekstni to‘g‘ri tahlil qilishi kerak [6]. Masalan, ingliz tilidagi "bank" so‘zi ikki xil ma’noda ishlatilishi mumkin: I deposited money in the bank. Men pulni bankka qo‘ydim – Bu yerda moliyaviy muassasa joylashgan hudud, zhoy ma’nosida kelgan. They sat on the bank of the river – Ular daryo qirg‘og‘ida o‘tirishdi – Bu gapda esa daryo qirg‘og‘idagi zhoy, hudud ma’nosida ishlatilgan. Rus tilida ham polysemiya lingvistik hodisa sifatida keng tarqalgan. Masalan, "ключ" ko‘p ma’onali so‘z hisoblanib umumiste’molda qujidagi ma’nolalarda faol qo‘llaniladi: kalit, ocqich; o‘t oldirish kaliti; kalit, uch, yo‘l (biror masalani tushunishga, yechishga yordam beradigan omil, yo‘l; kalavanning uchi; test kaliti va h.k.z. Я потерял ключ от дома – Men uy kalitini yo‘qotdim – Bu yerda kalit ma’nosida. Мы нашли ключи к решению задачи – Biz kalavani uchini topdik.

An’anaviy so‘zma-so‘z tarjima tizimlari bunday farqlarni ko‘pincha ajrata olmaydi. Neyron mashina tarjimasi esa kontekstual vektorlar (embeddings) yordamida polysemic birliklarni ancha aniqroq tarjima qila oladi.

Idiomatik iboralar va frazeologizmlar ham mashina tarjima tizimlari uchun murakkab leksikografik muammo hisoblanadi. Idiomalar so‘zma-so‘z tarjima qilinmaydi, chunki ularning ma’nosi alohida so‘zlarning lug‘aviy ma’nosidan emas, balki umumiy ifodadan kelib chiqadi [9]. Masalan, ingliz tilidagi "It’s raining cats and dogs" iborasi so‘zma-so‘z tarjima qilinganda "mushuk va itlar yog‘moqda" degan ma’noni beradi. Aslida esa bu ibora juda kuchli yomg‘ir yog‘ayotganini anglatadi. Yana bir mashhur misol — "Break a leg" iborasi bo‘lib, literal tarjimada "oyog‘ingni sindir" ma’nosini bildiradi. Biroq real kommunikativ vaziyatda ushbu ibora "kingadir omad tilash" uchun qo‘llaniladi.

Shuningdek, ingliz tilidagi "Spill the beans" iborasi ham so‘zma-so‘z tarjimada "loviyalarni to‘kib yubormoq" deb tushunilishi mumkin, ammo uning haqiqiy ma’nosi "sirni oshkor qilish" ma’nosini anglatadi. O‘zbek tilida ham shunga o‘xshash idiomatik birliklar mavjud. Masalan, "ko‘ngli tog‘dek ko‘tarildi" iborasi so‘zma-so‘z tarjima qilinganda tabiiy eshitilmaydi, asl ma’nosi esa "juda xursand bo‘ldi" deganidir. Yana "burnidan ip o‘tkazib bo‘lmaydi" iborasi literal tarjimada tushunarsiz bo‘lishi mumkin, biroq u juda mag‘rur yoki qaysar insonga xos bo‘lgan fe’l-atvor belgilarni tasvirlaydi. Bu kabi birliklarni qo‘llashda nafaqat lingvistik bilimlar, balki milliy-madaniy yondashuvlar ham talab etiladi.

Realiya va lakuna muammosi ham leksikografiyaning eng nozik jihatlaridan biridir. Realiya – bu ma’lum bir xalqning madaniyati, turmush tarzi yoki ijtimoiy hayotiga xos

bo‘lgan, boshqa tillarda aynan bir so‘z bilan ifodalanmaydigan tushunchalardir. [10] Masalan, o‘zbek tilidagi mahalla, sumalak, sovchi kabi realiyaga oid so‘zlar boshqa tillarga tarjima qilinganda odatda ularning to‘liq ekvivalentlarini berish imkoniyati bo‘l maydi.

Jumladan, o‘zbek tilidagi “mahalla” tushunchasi ingliz tiliga neighbourhood, “quarter”, “district” so‘zilari, yoki “local community”, “local authority” kabi so‘z birikmalari yordamida tarjima qilinishi mumkin, biroq ushbu tarjima varianti faqatgina “mahalliy hudud” ma’onisini anglatishi bilan kifoyalanadi. Bunda "mahalla" so‘zining ijtimoiy boshqaruv va an’anaviy tizimiga xos bo‘lgan milliy-madaniyatimizga xos tushunchalar esa mazkur tarjimada yo‘qolib ketadi. Aynan sumalak so‘zini "traditional food" deb berilishi ham ushbu so‘z dadagi Navro‘z bayrami bilan bog‘liq milliy- madaniyatga xos ma’noni to‘liq aks ettirmaydi.

MUHOKAMA

Shuningdek, har bir tilga xos bo‘lgan lakunalar esa bir tildada mavjud bo‘lgan tushunchaning o‘zga tilde adekvat ifoda vositasini yo‘qligini anglatadi. Masalan, rus tilidagi “rocka” so‘zini ingliz tiliga “melancholy” yoki “longing” deb tarjima qilish mumkin, lekin bu kabi tarjima variantlari uning chuqur ruhiy va hissiy ma’nosini to‘liq bera olmaydi. Xuddi shunday, o‘zbek tilidagi “hashar” tushunchasini "collective volunteer work" deb izohlash mumkin, lekin undagi ijtimoiy birdamlik va o‘zaro yordam ruhi tarjimada to‘liq aks e‘tmaydi.

Ma’lumki, omonimiya hodisasi ham mashina tarjimasi uchun jiddiy to‘siqdir. Omonimiya — shakli bir xil, ammo ma’nolari o‘zaro bog‘liq bo‘lmagan so‘zlardir. Masalan, ingliz tilidagi "light" so‘zini quyidagi gaplardagi faol qo‘llaniladigan ma’nolarini olaylik: light – chiroq, yengil, oson kabi. Please, turn on the light – Chiroqni yoqib yuboring. Bu yerda "light" chiroq ma’nosida qo‘llangan. This bag is very light – Bu sumka juda yengil. Bu gapda esa "light" yengil ma’nosini bildiradi. It is very light task – Bu juda oson topshiriq.

Neyron mashina tarjimasi sun’iy neyron tarmoqlari va kontekstual diqqat (Attention Mechanism) mexanizmi orqali ushbu omonim va polisemiya muammolarini sezilarli darajada kamaytirishga muvaffaq bo‘ldi [1]. Biroq murakkab badiiy matnlarda yoki falsafiy kontekstlarda ushbu zamonaviy tizimlar tarjima jarayonida hamon xatolarga yo‘l qo‘ymoqda. Masalan, Attention Mechanism vositasida hajmi jihatidan katta bo‘lgan matnlarni tarjima qilishda kontekstda diqqatni susayishiga olib keladi.

XULOSA

Alqissa, yuqorida ta’kidlab o‘tilganidek avtomatik tarjima tizimlarining ish jarayoni analiz va sintez, tanlash va tarjima qilish bosqichlaridan iborat bo‘ladi. Shunday qilib, avtomatik tarjima jarayonida til strukturasining qiyosiy tahlili natijalaridan ham foydalanishni taqozo etadi. Chunki, tilni strukturasini sistemali tahlil qilmasdan avtomatik tarjimani amalga oshirib bo‘lmaydi.

Ma'lumki, avtomatik tarjima til tizimini modellashtish, ya'ni algoritmallashtirish orqali amalga oshiriladi. Biroq polisemiya, omonimiya, idiomatik birliklar, realiya va lakuna kabilar bilan bog'liq muammolar hanuzgacha avtomatik tarjima tizimlarida leksikografik tadqiqotlar uchun dolzarb masala biri bo'lib qolmoqda. Ushbu muammolarni bartaraf etish qilishda tarjima tizimlarini nafaqat sof lingvistik qoidalar, balki madaniy, ensiklopedik va pragmatik bilimlar bazasi bilan ham boyitish zarur.

Shunday bo'lasada, neyron tarjimada ham idiomalar, metaforalar va madaniyatga xos realiyalarni tarjima qilishda muammolar mavjudligini quyidagicha izohlash mumkin: 1) milliy madaniyatga xos kontekstni anglay olish; 2) milliy tushunchalarning o'ziga xos xususiyatlarni aks ettiruvchi leksik birliklarning tillararo muqobillarini topish yoki yaratish; 3) matn uslubini imkon darajada saqlab qolish; 4) neyron mashina tarjimasining sifatini oshirishda pragmatic tadqiqotlarni olib borishni tizimlashtirishdan iborat.

Demak, mashina tarjima tizimlari dastlab so'zma-so'z tarjima tamoyiliga asoslangan bo'lsada, zamonaviy neyron mashina tarjimasi esa kontekstni hisobga olish orqali tarjima sifatini sezilarli darajada yaxshiladi.

Maqolada aynan neyron mashina tarjimasi tizimlarida leksikografik birliklarning turli kontekstlarda qo'llanishiga oid muammolar va ularning mashina tarjima sifatiga ta'siri muallif tomonidan qiyosiy tahlil qilindi.

Kelajakda neyron mashina tarjimasini yanada rivojlantirish uchun parallel korpuslar, raqamli izohli lug'atlar va madaniy ontologiyalardan keng foydalanish muhim ahamiyat kasb etadi.

ADABIYOTLAR RO'YXATI

1. Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural Machine Translation by Jointly Learning to Align and Translate. In International Conference on Learning Representations (ICLR), 45 p.
2. Bar-Hillel, Y. (1960). The Present Status of Automatic Translation of Languages. *Advances in Computers*, 1, pp. 191-163.
3. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press, 800 p.
4. Koehn, P. (2009). *Statistical Machine Translation*. Cambridge University Press, 450 p.
5. Jurafsky, D., & Martin, J. H. (2020). *Speech and Language Processing* (3rd ed. draft). Pearson, 1025 p.
6. Manning, C. D., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press, 680 p.

7. Калинин С.М. Актуальные подходы к улучшению точности нейронного машинного перевода / Kalinin С.М. – URL: <https://cyberleninka.ru/article/n/problemny-mashinnogo-perevoda>(дата обращения: 28.02.2019).

8.Мироненко А.Б. Проблемы машинного перевода / Мироненко А. Б, Омельченко П. А. – URL: <https://cyberleninka.ru/article/n/problemny-mashinnogo-perevoda> (дата обращения: 28.02.2019).

9. Nida, E. A. (1964). Toward a Science of Translating. E.J. Brill, 331 p.

10. Влахов, С., & Флорин, С. (1980). Непереводимое в переводе. Международное отношения, 342 с.

11. Weaver, W. (1949). Translation. Machine Translation of Languages, pp.

12.Ртищева Н.Г. Оценка качества работы систем машинного перевода / Ртищева Н.Г. – URL: <https://cyberleninka.ru/article/n/otsenka-kachestva-raboty-sistem-mashinnogo-perevoda> (дата обращения: 28.02.2019). –