

O‘ZBEK TILI SHEVALARINI AVTOMATIK ANIQLASH VA TARJIMA TIZIMLARIGA INTEGRATSIYALASH

Xusaydinova Dilobar Mirxamidovna¹, Muxtarova Gulnora Xatamovna²

^{1,2}Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti «Raqamli texnologiyalar konvergentsiyasi» kafedrasida Toshkent, O‘zbekiston

¹e-mail: dilobarhusajdinova032@gmail.com, ²e-mail: gulnora6530@inbox.ru

Annotatsiya: Mazkur maqolada o‘zbek tili shevalarini, xususan Xorazm shevasini avtomatik aniqlash va ularni elektron tarjima tizimlariga integratsiyalash masalalari yoritiladi. O‘zbek tilining dialektal boyligi avtomatik tarjima tizimlari uchun muhim muammo bo‘lib, adabiy tilga moslashmagan sheva birliklari tarjima sifatining pasayishiga olib keladi. Tadqiqotda shevalarni avtomatik aniqlash usullari, lingvistik va statistik yondashuvlar, mashinali o‘rganish va neyron tarmoqlar asosida ishlovchi modellar tahlil qilinadi. Shuningdek, aniqlangan sheva xususiyatlarini elektron tarjima tizimlariga integratsiyalashning konseptual modeli taklif etiladi.

Kalit so‘zlar: o‘zbek tili dialektologiyasi, hududiy shevalar, Xorazm shevasi, dialektlarni avtomatik aniqlash, tabiiy tilni qayta ishlash (NLP), elektron tarjima tizimlari, neyron mashinali tarjima (NMT), lingvistik normallashtirish, mashinali o‘rganish, sun’iy intellekt.

KIRISH

So‘nggi yillarda sun’iy intellekt, mashinali o‘rganish va tabiiy tilni qayta ishlash (Natural Language Processing) texnologiyalarining jadal rivojlanishi elektron tarjima tizimlarining sifat va funksional imkoniyatlarini sezilarli darajada kengaytirdi. Biroq mavjud tarjima tizimlarining aksariyati asosan adabiy tilga yo‘naltirilgan bo‘lib, real nutqda keng qo‘llaniladigan shevalar va lahjalarni to‘liq qamrab olmaydi. Bu holat, ayniqsa, dialektal jihatdan boy bo‘lgan tillar, jumladan o‘zbek tili uchun dolzarb muammo hisoblanadi. O‘zbek tili hududiy shevalarga nihoyatda boy bo‘lib, ular fonetik, leksik va grammatik xususiyatlari bilan adabiy tildan sezilarli darajada farq qiladi. Xorazm shevasi esa o‘ziga xos talaffuz, so‘z yasalishi va lug‘aviy birliklari bilan ajralib turadi. Ushbu farqlar elektron tarjima tizimlarida noto‘g‘ri tarjima, semantik buzilishlar yoki matnning umuman tanib olinmasligiga olib kelishi mumkin. Natijada tarjima sifati pasayadi va tizimning amaliy samaradorligi cheklanadi.

Shu munosabat bilan o‘zbek tili shevalarini avtomatik aniqlash va ularni elektron tarjima tizimlariga integratsiyalash masalasi dolzarb ilmiy-amaliy vazifa sifatida namoyon bo‘lmoqda. Shevalarni aniqlash bilan bir qatorda, ularning lingvistik xususiyatlarini tarjima jarayoniga moslashtirish, normallashtirish va mos til modellari orqali qayta ishlash muhim ahamiyat kasb etadi. Ayniqsa, Xorazm shevasini misolida integratsiya mexanizmlarini ishlab chiqish o‘zbek tili uchun universal va kengaytiriladigan tarjima tizimlarini yaratishga xizmat qiladi.

Mazkur maqolada o‘zbek tili shevalarini, xususan Xorazm shevasini avtomatik aniqlash va ularni elektron tarjima tizimlariga integratsiyalash usullari tahlil qilinadi hamda ushbu jarayon uchun konseptual yondashuv taklif etiladi. [1, 2, 3]

1. O‘zbek tili shevalarining tarjima jarayoniga ta’siri

O‘zbek tili shevalari tarjima jarayoniga bir nechta darajada ta’sir ko‘rsatadi. Birinchidan, shevaga xos leksik birliklar standart lug‘atlarda mavjud emas. Ikkinchidan, fonetik o‘zgarishlar nutqni avtomatik tanish bosqichida xatoliklarni keltirib chiqaradi. Uchinchidan, morfologik farqlar grammatik tahlil va sintaktik tuzilmani aniqlashni murakkablashtiradi. Xorazm shevasida uchraydigan ayrim fe‘l shakllari va so‘z birikmalari buni yaqqol ko‘rsatadi. Shu bois, shevalarni tarjima tizimiga integratsiyalash alohida modullar va moslashtirilgan algoritmlarni talab etadi.[6,7]

2. Shevalarni avtomatik aniqlashdan tarjimaga o'tish modeli

Shevalarni integratsiyalash jarayonining markaziy nuqtasi - aniqlash va tarjima o'rtasidagi bog'lovchi qatlamni yaratishdir. Ushbu qatlam quyidagi funksiyalarni bajaradi:

- kiruvchi matn yoki nutqning sheva turini aniqlash;
- shevaga xos birliklarni aniqlash va belgilash;
- ularni adabiy tilga mos normallashtirilgan shaklga o'tkazish;
- normallashtirilgan natijani tarjima yadrosiga uzatish.

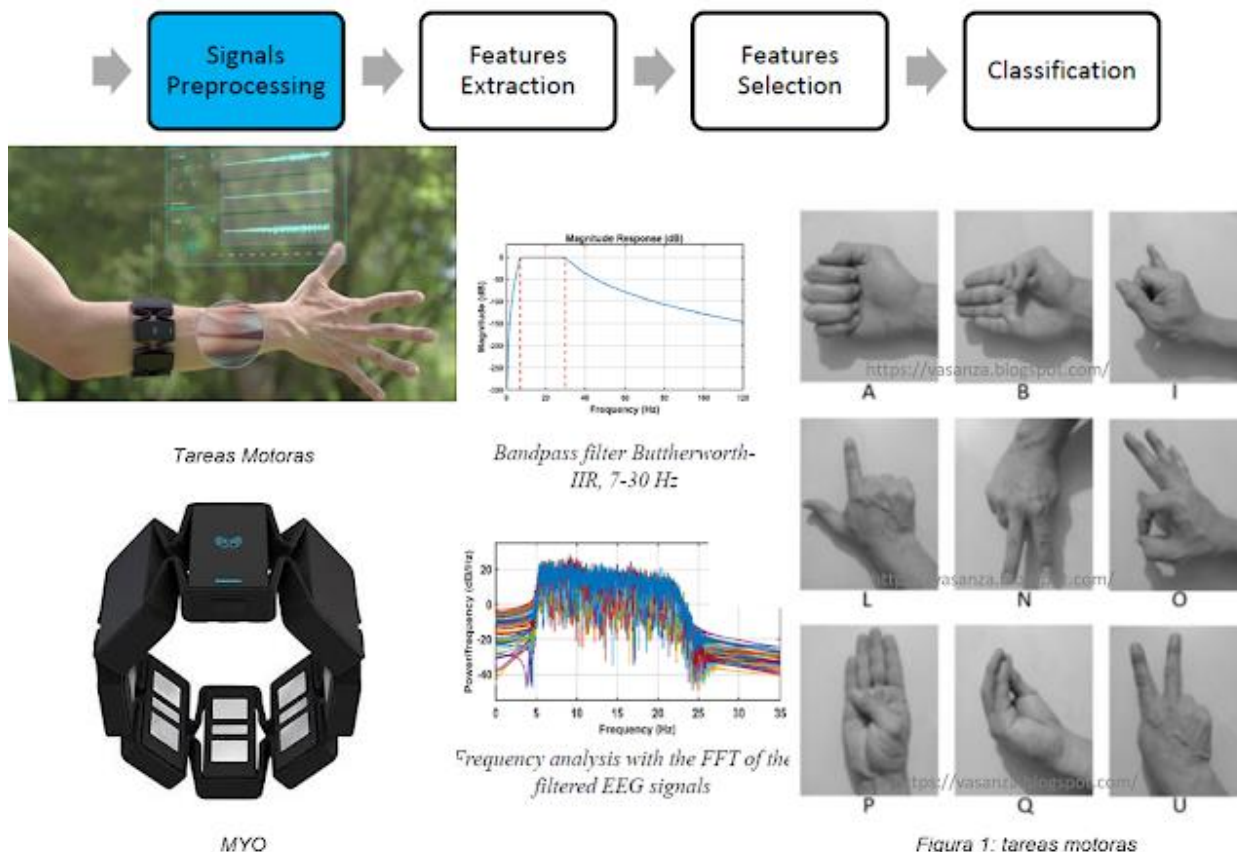
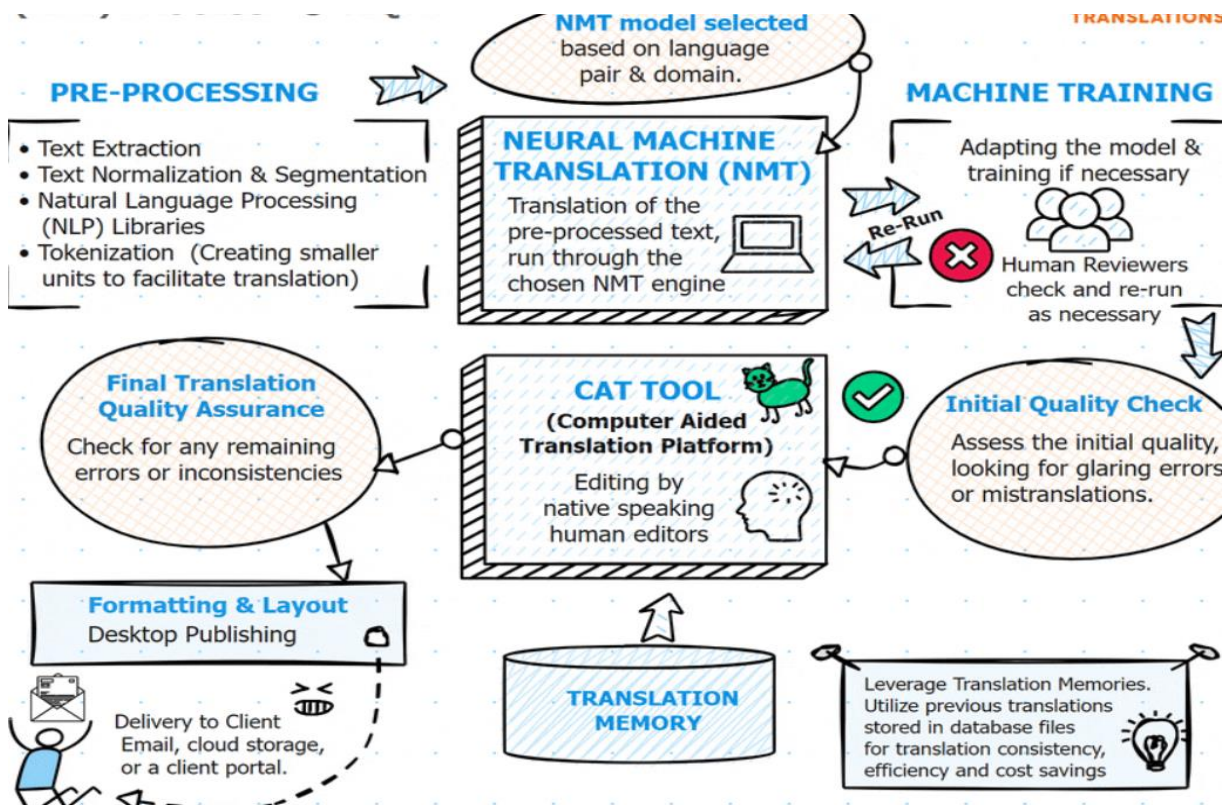


Figura 1: tareas motoras

Mazkur yondashuv elektron tarjima tizimining asosiy algoritmini o'zgartirmagan holda, unga moslashuvchanlik qo'shish imkonini beradi. [4,5]

3. Shevalarni elektron tarjima tizimlariga integratsiyalash usullari

Shevalarni integratsiyalashning bir nechta samarali usullari mavjud:

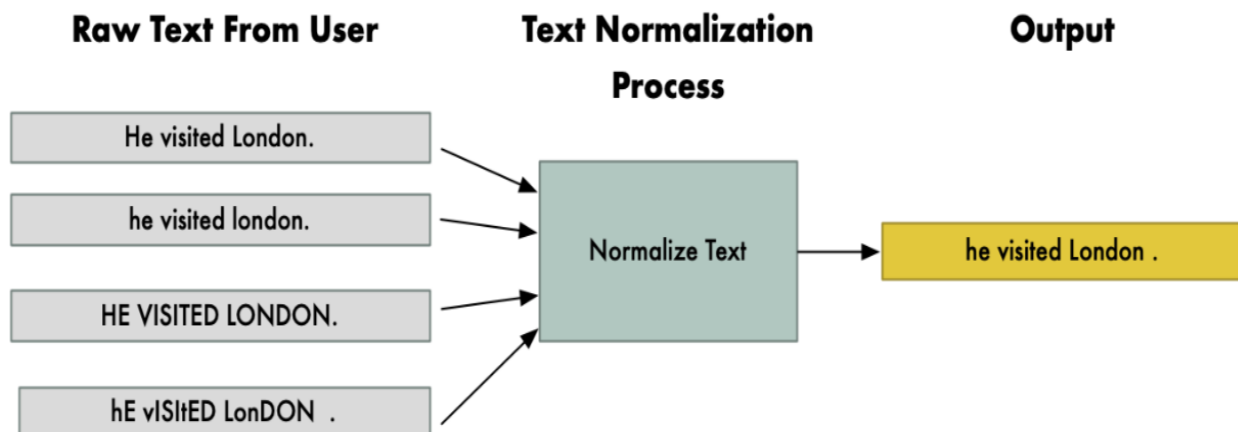


3.1. Oldindan normallashtirish asosidagi integratsiya-bu usulda shevada kiritilgan matn avval maxsus normallashtirish modulida qayta ishlanadi. Xorazm shevasiga xos soʻz va shakllar adabiy til ekvivalentlariga almashtiriladi. Shundan soʻng matn standart tarjima tizimiga uzatiladi. Ushbu usulning afzalligi - tarjima sifatining sezilarli oshishi va tizimni qayta oʻqitish zaruratining kamayishidir.

3.2. Koʻp variantli tarjima modeli-ushbu yondashuvda tarjima tizimi bir nechta til variantlari bilan ishlaydi: adabiy til va alohida shevalar. Sheva aniqlangach, mos til modeli faollashtiriladi. Bu usul katta hisoblash resurslarini talab qilsa-da, yuqori aniqlikni taʼminlaydi.

3.3. Neyron tarjima tizimlarida integratsiya-neyron mashinali tarjima (NMT) tizimlarida shevalarni integratsiyalash uchun qoʻshimcha belgilash (tagging) usuli qoʻllaniladi. Masalan, Xorazm shevasiga oid matn maxsus belgi bilan markalanadi va model ushbu belgiga mos tarjima strategiyasini tanlaydi. Bu yondashuv shevalararo farqlarni avtomatik oʻrganishga imkon beradi. [10,3]

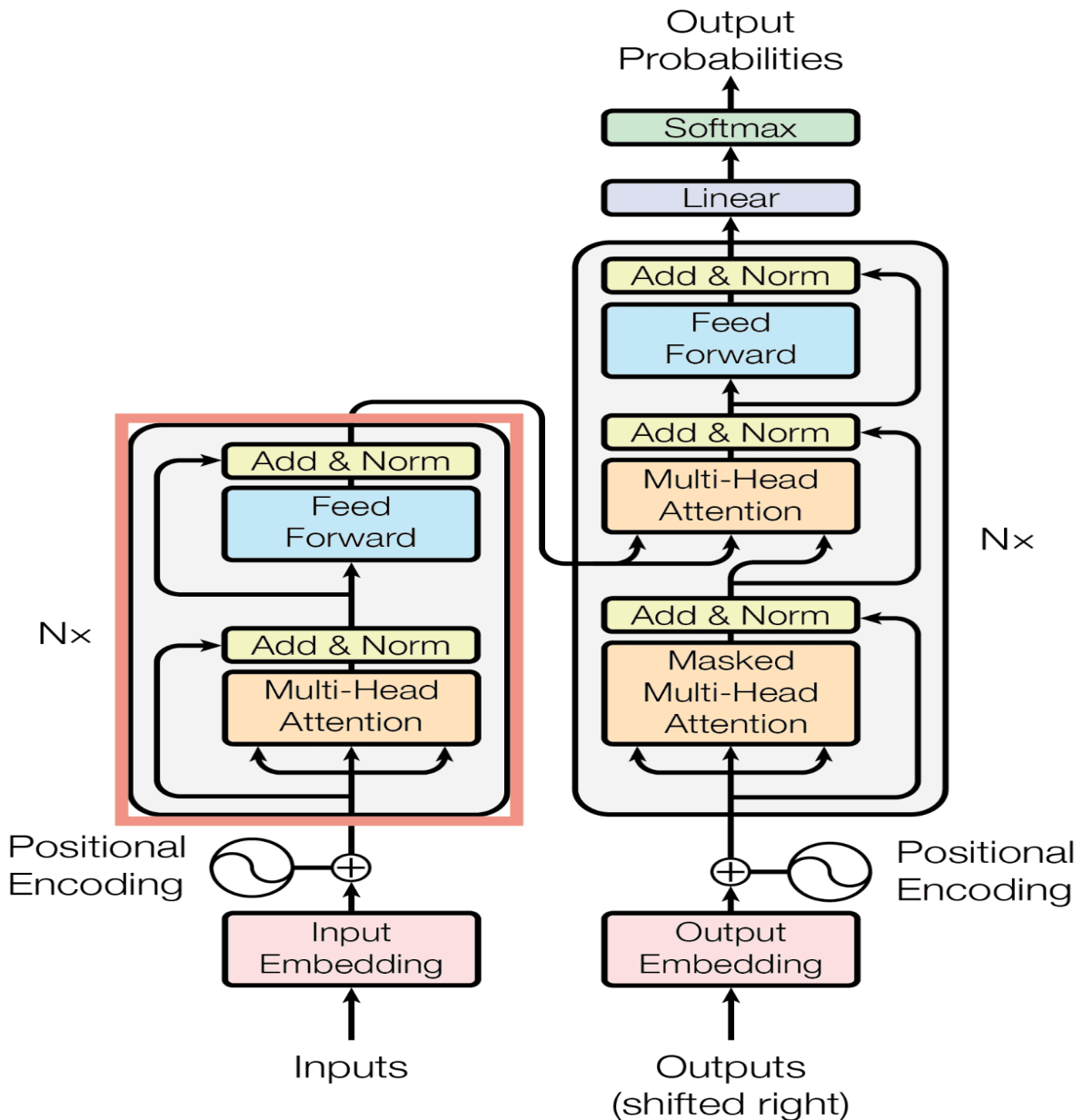
4. Xorazm shevasi misolida integratsiya jarayoni



Xorazm shevasi asosida olib borilgan tahlillar shuni koʻrsatadiki, normallashtirish lugʻatlari va morfologik moslashtirish qoidalari tarjima sifatiga bevosita taʼsir koʻrsatadi. Tajriba natijalariga

ko'ra, sheva integratsiyasi qo'llanilgan tizimda semantik moslik va tarjima aniqligi an'anaviy tizimlarga nisbatan yuqoriroq bo'ladi.[1]

5. Taklif etilayotgan konseptual integratsiya modeli



Tadqiqotda taklif etilgan konseptual model quyidagi modullardan iborat:

- shevalarni avtomatik aniqlash moduli;
- shevaga xos birliklarni normallashtirish bloki;
- moslashtirilgan elektron tarjima yadrosi;
- natijani baholash va teskari aloqa mexanizmi.

Ushbu model bosqichma-bosqich kengaytirilishi va boshqa o'zbek shevalarini ham qamrab olishi mumkin.[13,18]

XULOSA

O'zbek tili shevalarini, jumladan Xorazm shevasini elektron tarjima tizimlariga integratsiyalash tarjima sifatini oshirishning muhim sharti hisoblanadi. Tadqiqot natijalari shuni ko'rsatadiki,

shevalarni faqat aniqlash bilangina cheklanmasdan, ularni normallashtirish va tarjima jarayoniga moslashtirish zarur.

Maqolada taklif etilgan shevalarni elektron tarjima tizimlariga integratsiyalash konsepsiyasi normallashtirish, moslashuvchan til modellaridan foydalanish va neyron tarjima tizimlarida shevaga xos belgilarni hisobga olishga asoslanadi. Xorazm shevasi misolida olib borilgan tahlillar shuni ko'rsatadiki, shevaga oid leksik va morfologik birliklarni adabiy til ekvivalentlariga moslashtirish tarjima aniqligi va semantik yaxlitlikni sezilarli darajada oshiradi. Bu esa tarjima tizimlarining real muloqot sharoitlarida samarali ishlashiga xizmat qiladi.

Shuningdek, ishlab chiqilgan konseptual integratsiya modeli o'zbek tili uchun moslashuvchan va kengaytiriladigan yechim sifatida baholanadi. Ushbu modelni bosqichma-bosqich rivojlantirish orqali boshqa hududiy shevalarni ham qamrab olish, ko'p shevali va ko'p tilli elektron tarjima tizimlarini yaratish imkoniyati mavjud. Tadqiqot natijalari nafaqat elektron tarjima tizimlari uchun, balki nutqni avtomatik tanish, dialektologiya va amaliy tilshunoslik sohalari uchun ham muhim ilmiy-amaliy ahamiyatga ega.

Xulosa qilib aytganda, o'zbek tili shevalarini, jumladan Xorazm shevasini elektron tarjima tizimlariga integratsiyalash masalasi zamonaviy til texnologiyalarini rivojlantirishda ustuvor yo'nalishlardan biri bo'lib, mazkur tadqiqot ushbu yo'nalishda keyingi ilmiy izlanishlar uchun mustahkam nazariy va amaliy asos yaratadi.

FOYDALANILGAN ADABIYOTLAR

1. Jurafsky D., Martin J. *Speech and Language Processing*. - Pearson, 2023.
2. Koehn P. *Neural Machine Translation*. - Cambridge University Press, 2020.
3. Abdurahmonov G'. *O'zbek tili dialektologiyasi*. - Toshkent: Fan, 2018.
4. Goldsmith J. *The Handbook of Phonological Theory*. - Wiley-Blackwell, 2019.
5. Suyunova D. *O'zbek tili shevalarining lingvistik tahlili*. - Toshkent, 2021.
6. Jurafsky D., Martin J. *Speech and Language Processing*. - Pearson, 2023.
7. Vaswani A. et al. *Attention Is All You Need // Advances in Neural Information Processing Systems (NeurIPS)*. - 2017.
8. Goldsmith J. *The Handbook of Phonological Theory*. - Wiley-Blackwell, 2019.
9. Zisserman A., Bishop C. *Pattern Recognition and Machine Learning*. - Springer, 2016.
10. Alotaibi Y., Hussain A. *Comparative analysis of dialect identification techniques for Arabic speech // Applied Acoustics*. - 2018.
11. Sennrich R., Haddow B., Birch A. *Improving Neural Machine Translation Models with Monolingual Data // ACL*. - 2016.
12. Huang X., Acero A., Hon H.-W. *Spoken Language Processing: A Guide to Theory, Algorithm, and System Development*. - Prentice Hall, 2014.
13. O'zbekiston Respublikasi Prezidentining qarori "O'zbek tilining davlat tili sifatidagi nufuzi va mavqegini oshirish chora-tadbirlari to'g'risida". - Toshkent, 2019.