

BILINGV MATNLARDA KOD ALMASHINUVI (CODE-SWITCHING)NI ANIQLASHDA SUN'IY INTELLEKT YONDASHUVI

Hamroeva Shahlo Mirdjonovna

Filologiya fanlari doktori, professor

shaxlo.xamrayeva@navoiy-uni.uz

ToshDO'TAU

Ruziyeva Madina Ilhom qizi

tayanch doktoranti

madinaruziyeva67@gmail.com

ToshDO'TAU

Annotatsiya. Mazkur tadqiqot bilingv matnlarda kod almashinuvi (code-switching) hodisasini aniqlash va tahlil qilishda sun'iy intellekt yondashuvlarini qo'llashga bag'ishlangan. Tadqiqot doirasida Buxoro hududidan yig'ilgan 1000 ta kommentdan iborat bilingv nutq korpusi shakllantirildi. Korpus o'zbek–rus, o'zbek–ingliz va o'zbek–tojik til juftliklarini o'z ichiga oladi. Tahlil natijasida kod almashinuvi yuqori darajada, ayniqsa yoshlar nutqida faol qo'llanilishi aniqlandi. Sun'iy intellekt asosida kod almashinuvi holatlarini avtomatik aniqlash modeli taklif qilindi. Natijalar kod almashinuvi ijtimoiy moslashuv va identifikatsiyaning muhim ko'rsatkichi ekanligini tasdiqlaydi.

Kalit so'zlar: *bilingvizm, kod almashinuvi, sun'iy intellekt, NLP, korpus lingvistikasi*

Abstract. This study investigates the application of artificial intelligence methods for detecting and analyzing code-switching in bilingual texts. A bilingual corpus consisting of 1000 comments collected from the Bukhara region was compiled and manually annotated. The corpus includes Uzbek–Russian, Uzbek–English, and Uzbek–Tajik language pairs. The results show that code-switching is highly frequent, especially among young speakers. An AI-based model for automatic detection of code-switching is proposed. The findings confirm that code-switching serves as an important indicator of social adaptation and identity.

Keywords: *bilingualism, code-switching, artificial intelligence, NLP, corpus linguistics*

Kirish. Bilingv va multilingv muhitda kod almashinuvi fenomeni zamonaviy lingvistika va kompyuter lingvistikasining dolzarb masalalaridan biri hisoblanadi. Raqamli kommunikatsiya muhitida til aralashuvi keng tarqalgan bo'lib, uni avtomatik aniqlash zarurati ortib bormoqda. Mazkur tadqiqotning maqsadi – bilingv matnlarda kod almashinuvi holatlarini aniqlash va sun'iy intellekt asosida avtomatik model ishlab chiqish mexanizmini ko'rsatish.

Adabiyotlar sharhi. So'nggi yillarda bilingvizm va kod almashinuvi tilshunoslik hamda kompyuter lingvistikasida faol o'rganilmoqda. Tadqiqotlar ushbu hodisaning nafaqat lingvistik, balki ijtimoiy, madaniy va psixologik omillar bilan uzviy bog'liqligini ko'rsatadi [2:15–18]. Kod almashinuvi ko'pincha kommunikativ samaradorlikni oshirish va ijtimoiy moslashuvni ta'minlash vositasi sifatida qaraladi hamda nutq strategiyasi sifatida suhbatdoshlar o'rtasidagi yaqinlik va identifikatsiyani ifodalaydi [1:190–195]. Raqamli kommunikatsiyada, ayniqsa ijtimoiy tarmoqlarda, kod almashinuvi yanada faollashgani kuzatiladi [1:198–200]. Korpus lingvistikasi yondashuvi orqali bilingv nutqdagi kod almashinuvi chastotasi va funksional xususiyatlari tizimli ravishda tahlil qilinmoqda [4:425–430]. Shu bilan birga, sun'iy intellekt va NLP texnologiyalari yordamida kod almashuvni avtomatik aniqlash imkoniyati yuzaga kelib, mashinaviy va chuqur o'rganish modellari orqali yuqori aniqlikda tasniflash amalga oshirilmoqda [5:1398–1402]. O'zbek tilshunosligida bilingv nutq va til tanlovi asosan sotsiolingvistik nuqtai nazardan o'rganilib, uning ijtimoiy muhit, kommunikativ vaziyat va shaxsiy identifikatsiya bilan bog'liqligi ta'kidlanadi [3:47–50; 7:112–118]. Korpus lingvistikasi asosida til hodisalarini tizimli tahlil qilish imkoniyatlari ham ilmiy jihatdan asoslab berilgan [6:25–30]. Biroq o'zbek tilida kod almashinuvi va uning sun'iy intellekt asosida

aniqlanishi yetarli darajada o'rganilmagan bo'lib, mazkur tadqiqot ushbu ilmiy bo'shliqni to'ldirishga qaratilgan.

Metodologiya. Mazkur tadqiqotda bilingv nutqda kod almashinuvi hodisasini o'rganish uchun 1000 ta kommentdan iborat korpus shakllantirildi. Ma'lumotlar Buxoro shahri va tumanlaridagi ijtimoiy tarmoqlardan yig'ilib, qo'lda annotatsiya qilindi. Korpus strukturasining namunaviy ko'rinishi 1-jadvalda keltirilgan.

1-jadval.

Salom, “magazinni adresi” qayerda?	O'zbek-rus		
“Bravo” gap yo'q.		O'zbek-ingliz	
Buncha “xushro'y” arzon tet.			O'zbek-tojik
13 “layk” Buxorolik shogirdlarizdan akama “gap nes”.		O'zbek-ingliz	O'zbek-tojik
Buxorom dunyoba yagona. Buxoro “nomer odin”.	O'zbek-rus		O'zbek-tojik
“Ok”. Tushundim. “Rahmati “kalon”.		O'zbek-ingliz	O'zbek-tojik

Kod almashinuvi hodisasini tahlil qilish uchun korpus lingvistikasi, sotsiolingvistik tahlil va statistik metodlar qo'llanildi. Shuningdek, Likert shkalasi asosida til ishlatish darajasi baholandi. Bundan tashqari, kod almashinuvi holatlarini avtomatik aniqlash uchun sun'iy intellektga asoslangan model taklif qilindi. Model kirish sifatida matnni qabul qilib, chiqishda kod almashinuvi mavjudligini va uning turini aniqlaydi. Tahlil qilingan nutq birliklarida kod-aralashuvning uch asosiy shakli – intrasentensial (gap ichidagi), intersentensial (gaplararo), va tag switching (so'zning affikslari orqali aralashuv) holatlari kuzatildi. Har bir aralashuv holati alohida ijtimoiy kontekstda ko'rib chiqildi va shaxslarning ijtimoiy identifikatsiyasi, muhitga moslashuvi, ijtimoiy rolga kirishi kabi omillar bilan bog'lanib tahlil qilindi. Quyidagi jadvalda bilingval nutq korpusida uchragan kod-aralashuv holatlari va

ularning ijtimoiy kontekstdagi izohi keltirilgan. Misollar Buxoro shahri va tumanlaridagi real muloqotlardan olingan. Kod almashinuvi turlari 2-jadvalda keltirilgan.

2-jadval

№	Nutq birligi	Kod-aralashuv turi	Izoh
1	Mana pulni perevod qilib yubordim	Intrasentensial	Ruscha “perevod” so‘zi O‘zbek gap strukturasi ichida ishlatilgan – shaxs iqtisodiy kontekstdagi moslashuvni aks ettiradi.
2	Qachon kelasan? – Навэрhone, ertaga	Intersentensial	Gaplararo almashuv – javob beruvchi shaxs o‘z pozitsiyasini ikki til yordamida aniq ifodalamoqda.
3	Lineykani olgan	Tag switching	Ruscha “lineyka” so‘ziga o‘zbekcha “-ga” qo‘shimchasi qo‘shilgan – fonologik va grammatik moslashuv ko‘zga tashlanadi.
4	Men uni WhatsApp’da blok qildim	Intrasentensial	Raqamli kontekstda ishlatiladigan terminlar til aralashuvini kuchaytiradi – texnologik muhitga ijtimoiy xoslanish.

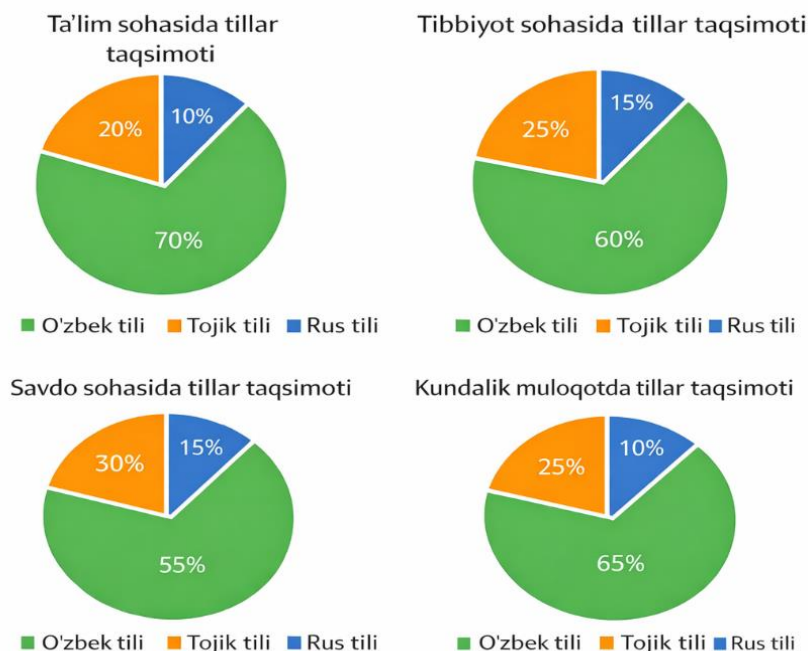
Yuqoridagi misollar shuni ko‘rsatadiki, kod-aralashuv holatlari shunchaki til aralashuvi emas, balki shaxslarning ijtimoiy muhitga, kommunikativ vaziyatga va identitetga bog‘liq holda til tanlovi orqali o‘zini ifodalash vositasidir.

Natijalar. Kod-almashinuvi eng ko‘p yoshlar orasida kuzatildi (78%), ayniqsa ijtimoiy tarmoqlarda. Kod almashinuvi bilingv nutqda faqat lingvistik hodisa emas, balki ijtimoiy va kommunikativ strategiya sifatida namoyon bo‘ladi. Sotsiolingvistik yondashuvlarga ko‘ra, til tanlovi shaxsning ijtimoiy identifikatsiyasi, kommunikativ maqsadi va muhitga moslashuvi bilan bevosita bog‘liqdir. Kod almashinuvi orqali nutq egasi bir vaqtning o‘zida bir nechta ijtimoiy

rollarni namoyon qilishi, suhbatdoshga moslashishi va muloqot samaradorligini oshirishi mumkin. Ayniqsa, raqamli kommunikatsiya sharoitida bu jarayon yanada faollashib, til birliklarining erkin kombinatsiyasi orqali yangi diskursiv shakllar yuzaga keladi. Shu nuqtai nazardan, kod almashinuvi bilingv nutqda ijtimoiy moslashuv, identifikatsiya va kommunikativ optimallashtirishning muhim mexanizmi sifatida qaraladi. Korpus natijalari shuni ko'rsatdiki cod-switching ,asosan, ta'lim, savdo, kundalik muloqot va tibbiyotda ko'p uchraydi. (Qarang: 3-jadval)

Soha	Aralashuv turi	Namuna
Ta'lim	Ruscha terminlar	Natija “vernii” chiqdi
Savdo	Tojikcha so'zlar	Arzon kunet
Tibbiyot	Pragmatik birliklar	“Izvinite”, ahvolingiz yaxshimi
Kundalik nutq	Aralash	Xudoba shukur. Bugun hammasi “normalniy”

Buxoro shahri hududi misolida kod-almashuv taqsimoti quyiagi sohalarda eng ko'p uchrashi aniqlandi. Mazkur diagrammalar tahlili shuni ko'rsatadiki, Buxoro shahri hududida kod-almashuv turli ijtimoiy sohalarda turlicha darajada namoyon bo'lib, ayniqsa kundalik muloqot va savdo sohalarida yuqori faollik kuzatiladi.



1-rasm. Ta'lim, tibbiyot, savdo va kundalik muloqot sohalarida tillarning foiz taqsimoti

Mazkur tadqiqot natijalari asosida bilingv matnlarda kod almashinuvi hodisasini aniqlash va undan samarali foydalanish bo'yicha quyidagi amaliy tavsiyalar ishlab chiqildi:

1. Ta'lim tizimi uchun tavsiyalar:

Diagramma natijalari kod almashinuvi ayniqsa ta'lim va kundalik muloqot sohalarida faol ekanligini ko'rsatadi. Shu bois:

- bilingv muhitda o'qitishda kod almashuvni salbiy emas, balki kommunikativ strategiya sifatida qabul qilish;
- o'quv dasturlariga *translanguaging* va *code-switching* asosidagi mashqlarni kiritish;
- talabalar nutqini rivojlantirishda ikki tillilikdan foydalangan holda interaktiv metodlarni qo'llash tavsiya etiladi.

2. Sun'iy intellekt va NLP tizimlari uchun:

- avtomatik tarjima va chatbot tizimlarida kod almashuvni aniqlovchi modul joriy etish;

- tilni aniqlash (language detection) algoritmlarini token darajasida ishlashga moslashtirish;
- ko‘p tilli korpuslar asosida mashinaviy o‘rganish modellari (ML/DL)ni rivojlantirish zarur.

3. Ijtimoiy tarmoqlar tahlili uchun:

- ijtimoiy media monitoring tizimlarida kod almashuvni indikator sifatida ishlatish;
- yoshlar nutqidagi lingvistik tendensiyalarni aniqlash uchun AI asosidagi analizatorlardan foydalanish;
- sentiment va identifikatsiya tahlilida kod almashuvni muhim parametr sifatida kiritish.

4. Lingvistik tadqiqotlar uchun:

- kod almashuvni ijtimoiy moslashuv indikatorlaridan biri sifatida o‘rganish;
- korpus lingvistikasi asosida real nutq materiallarini kengaytirish;
- bilingv nutqni kvantitativ va diskursiv metodlar orqali kompleks tahlil qilish.

Kod almashuvni aniqlash uchun sun’iy intellekt modeli arxitekturası

Bosqich	Komponent	Tavsif	Qo‘llanadigan usullar
1	Ma’lumotlar yig‘ish	Bilingv kommentlar korpusi	1000 ta (uz–ru, uz–en, uz–tj)
2	Preprocessing	Matnni tozalash	Tokenizatsiya, normalizatsiya, stop-so‘zlar
3	Tilni aniqlash	Token darajasida til aniqlash	N-gram, FastText
4	Detection Layer	Kod almashuvni aniqlash	LSTM, BiLSTM, BERT
5	Classification	Kod almashuv turini aniqlash	Inter, intra, tag-switching
6	Output	Natijani chiqarish	CS mavjudligi, foizlar

Mazkur jadvalda taqdim etilgan model arxitekturasida kod almashuvni aniqlash jarayonini bosqichma-bosqich tashkil etib, matnni dastlabki qayta ishlashdan tortib, chuqur o'rganish asosida tasniflashgacha bo'lgan barcha jarayonlarni qamrab oladi. Ushbu yondashuv kod almashuvni nafaqat aniqlash, balki uning turlarini avtomatik ravishda klassifikatsiya qilish imkonini beradi.

Xulosa. Mazkur tadqiqot bilingv matnlarda kod almashuv hodisasi keng tarqalganligini va ayniqsa kundalik muloqot hamda yoshlar nutqida faol qo'llanilishini ko'rsatdi. Model token darajasida tilni aniqlash va chuqur o'rganish algoritmlari yordamida kod almashuv holatlarini aniqlash imkonini beradi, bu esa lingvistik tahlillarni avtomatlashtirishda muhim ahamiyat kasb etadi. Shuningdek, natijalar kod almashuvning ijtimoiy moslashuv, identifikatsiya va kommunikativ kompetensiya bilan chambarchas bog'liqligini tasdiqlaydi. Ayniqsa, bilingv nutqda ijtimoiy identifikatsiya elementlarining ustunligi ushbu hodisaning kollektiv xarakterga ega ekanligini ko'rsatadi. Kelgusida ushbu yo'nalishda katta hajmdagi ko'p tilli korpuslar yaratish, sun'iy intellekt asosidagi modellarni takomillashtirish hamda real vaqt rejimida ishlovchi tizimlarni ishlab chiqish muhim ilmiy va amaliy vazifalardan biri hisoblanadi.

Foydalanilgan adabiyotlar ro'yxati

1. Androutsopoulos, J. (2015). *Networked multilingualism and digital discourse*. International Journal of Bilingualism, 19(2), 185–205. <https://doi.org/10.1177/1367006915583314>
2. García, O., & Li, W. (2018). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan. <https://doi.org/10.1057/978-1-137-57078-5>
3. Hamroyeva, Sh. (2022). *Bilingv nutqda ijtimoiy omillar va til tanlovi*. Filologiya masalalari, 3(2), 45–52.



4. Kim, S., & Kim, J. (2019). *Corpus-based approaches to code-switching analysis*. Linguistic Research, 36(3), 421–443.
<https://doi.org/10.17250/khisli.36.3.201909.001>
5. Li, Y., & Fung, P. (2020). *Code-switch language modeling with neural networks*. Proceedings of the ACL, 1395–1405.
<https://doi.org/10.18653/v1/2020.acl-main.123>
6. Mahmudova, N. (2020). *Sotsiolingvistika asoslari*. Toshkent: Fan va texnologiya.
7. Rahmatullayev, Sh. (2017). *O'zbek tilining izohli lug'ati va semantikasi*. Toshkent: O'zbekiston nashriyoti.