



O'ZBEK TILINI IKKINCHI TIL SIFATIDA O'RGANUVCHILARNING YOZMA NUTQINI AVTOMATIK BAHOLASHNING LINGVISTIK ASOSLARI VA MEZONLARI

Quldashev Nizomiddin Axmadaliyevich

Farg'ona davlat universiteti filologiya fanlari
bo'yicha falsafa doktori (PhD), dotsent
kuldashev.nizomidin@mail.ru

Solijonova Mohinur Asrorjon qizi

Farg'ona davlat universiteti magistranti
solijonovam02@gmail.com

<https://doi.org/10.5281/zenodo.22870049>

Annotatsiya.

Maqolada o'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqini avtomatik baholash mezonlari tahlil qilinadi. Yozma nutqni baholashda grammatik xatolar bilan birga matn mazmuni, lug'aviy tarkibi, sintaktik murakkabligi, bog'liqligi va izchilligini hisobga olish zarurligi asoslanadi. O'zbek tilining agglutinatив xususiyatlari avtomatik tahlilda alohida yondashuvni talab qilishi ko'rsatiladi. Xorijiy tadqiqotlarda qo'llanilgan korpus tahlili, lingvistik ko'rsatkichlarni aniqlash, xatolarni tasniflash va yozma ish sifatini baholash usullari o'zbek tili materialini asosida ko'rib chiqiladi.

Abstract.

This article analyzes the criteria for the automatic assessment of the written production of learners of Uzbek as a second language. It argues that assessment should consider not only grammatical errors, but also text content, lexical composition, syntactic complexity, cohesion, and coherence. The agglutinative nature of Uzbek requires a specific approach to automatic analysis. The article also considers corpus analysis, linguistic feature measurement, error classification, and writing quality assessment methods used in international research and their application to Uzbek-language materials.

Аннотация.

В статье анализируются критерии автоматического оценивания письменной речи изучающих узбекский язык как второй язык. Обосновывается необходимость учитывать не только грамматические ошибки, но и содержание текста, лексический состав, синтаксическую сложность, связность и последовательность. Показано, что агглютинативный характер узбекского языка требует особого подхода к автоматическому анализу. Рассматриваются методы корпусного анализа, определения лингвистических показателей, классификации ошибок и оценивания качества письменной речи, применяемые в зарубежных



POLAND



POLAND

исследованиях, и возможности их использования на материале узбекского языка.

Kalit so'zlar: o'zbek tili, ikkinchi til, yozma nutq, avtomatik baholash, lingvistik mezonlar, o'rganuvchi korpusi, grammatik xatolar, sintaktik murakkablik.

Keywords: Uzbek language, second language, written production, automatic assessment, linguistic criteria, learner corpus, grammatical errors, syntactic complexity.

Ключевые слова: узбекский язык, второй язык, письменная речь, автоматическое оценивание, лингвистические критерии, корпус учащихся, грамматические ошибки, синтаксическая сложность.

Kirish

Bugungi kunda ikkinchi tilni o'qitishda o'rganuvchining tilni qay darajada egallaganini aniqlash masalasi tobora muhimlashib bormoqda. Ayniqsa, yozma nutqni baholashda o'qituvchining subyektiv qarashlarini kamaytirish, katta hajmdagi yozma ishlarni qisqa vaqt ichida tahlil qilish va o'rganuvchining rivojlanishini muntazam kuzatib borish zarurati mavjud. Shu sababli yozma matnlarni kompyuter yordamida tahlil qilish va baholash imkoniyatlarini o'rganish zamonaviy tilshunoslik hamda til o'qitish metodikasining muhim yo'nalishlaridan biriga aylanmoqda [1; 2].

Yozma nutq o'rganuvchining til imkoniyatlarini nisbatan to'liq ko'rsatib bera oladigan nutqiy faoliyat turlaridan biridir. Yozish jarayonida shaxs so'zlarni tanlaydi, ularni grammatik jihatdan shakllantiradi, gaplar tuzadi, fikrlar o'rtasidagi munosabatni ifodalaydi va umumiy mazmunni izchil bayon etadi. Demak, yozma matnni baholashda uning alohida qismlarini emas, balki bir-biri bilan bog'langan lingvistik xususiyatlarini hisobga olish lozim.

Chet tilini o'rganish bo'yicha tadqiqotlarda yozma nutq sifati ko'pincha aniqlik, murakkablik va ravonlik bilan bog'liq ko'rsatkichlar orqali tavsiflanadi. Keyingi yillarda bunga matnning lug'aviy boyligi, sintaktik tuzilishi va bog'liqligi kabi ko'rsatkichlar ham qo'shilmoqda. Xususan, X. Lu sintaktik murakkablikni korpus ma'lumotlari asosida avtomatik o'lchash imkoniyatlarini tadqiq etib, gap uzunligi, bo'laklar nisbati va sintaktik tuzilmalarning murakkabligi kabi ko'rsatkichlardan foydalanish mumkinligini ko'rsatadi [3]. Boshqa tadqiqotlarda esa lug'aviy murakkablikning yozma nutq sifatini aniqlashdagi ahamiyati sintaktik murakkablikka nisbatan yuqoriroq bo'lishi mumkinligi qayd etilgan [4].



O'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqini avtomatik baholash masalasida esa tayyor xorijiy mezonlarni bevosita qo'llash yetarli emas. Chunki o'zbek tilining grammatik qurilishi, so'z yasalishi va shakl yasash xususiyatlari boshqa ko'plab tillardan farq qiladi. O'zbek tili agglutinativ til bo'lgani sababli grammatik ma'nolar ko'pincha affikslarning ketma-ket qo'shilishi orqali ifodalanadi. Bu holat o'rganuvchilarning yozma nutqida o'ziga xos xatolar paydo bo'lishiga olib keladi [8; 9].

Shu jihatdan o'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqini avtomatik baholash modelini ishlab chiqishda, avvalo, mazkur nutqning lingvistik xususiyatlarini aniqlash, tipik xatolarni tasniflash va baholash mezonlarini belgilash talab etiladi. Mazkur maqolada ana shu masalalar tahlil qilinib, avtomatik baholash uchun muhim bo'lgan asosiy ko'rsatkichlar aniqlanadi.

Asosiy qism va tahlil

Yozma nutqni avtomatik baholash masalasiga murojaat qilganda, eng avvalo, «baholash» tushunchasining mazmunini aniqlab olish zarur. Avtomatik baholash matnda uchragan xatolarni sanash bilangina cheklanmaydi. Uning vazifasi o'rganuvchining yozma faoliyatini turli jihatlar bo'yicha tavsiflash va shu asosda uning tilni egallash darajasiga oid xulosa chiqarishdan iborat. Shu sababli baholash tizimida bir nechta ko'rsatkichning birgalikda ishlatilishi maqsadga muvofiqdir.

Xorijiy tadqiqotlarda o'rganuvchi yozma nutqini baholashda korpusga asoslangan yondashuv keng qo'llanadi. Bunda turli darajadagi o'rganuvchilarning yozma ishlari bir joyga jamlanadi, matnlar lingvistik jihatdan tahlil qilinadi va ularning baholari bilan turli ko'rsatkichlar o'rtasidagi bog'liqlik aniqlanadi [2; 5]. Masalan, 3440 ta yozma matn asosida olib borilgan tadqiqotda matn murakkabligi bilan yakuniy baho o'rtasidagi munosabat o'rganilib, ayrim lingvistik ko'rsatkichlardan yangi matnning taxminiy bahosini aniqlashda foydalanish mumkinligi ko'rsatilgan [5]. Bunday tajriba o'zbek tili uchun ham muhim metodik asos bo'la oladi.

Buning uchun dastlab o'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqidan iborat korpus yaratish kerak. Korpusga turli darajadagi o'rganuvchilarning matnlari kiritilishi, har bir matnning mavzusi, hajmi, o'rganuvchining til darajasi va topshiriq turi kabi ma'lumotlar qayd etilishi lozim. Shuningdek, matnlarda uchraydigan xatolar alohida belgilanishi va ularning turi ko'rsatilishi zarur. Bunday korpus keyinchalik avtomatik baholash modelini sinash va takomillashtirish uchun asos bo'lib xizmat qiladi.



O'zbek tili materiali uchun birinchi navbatda grammatik to'g'rilik mezonini belgilash muhim. Grammatik to'g'rilik so'z shakllarining, affikslarning, fe'l shakllarining, kelishik va egalik qo'shimchalarining hamda gap qurilishining adabiy til me'yorlariga mosligini qamrab oladi [8; 9]. O'rganuvchi matnida grammatik xatolarning mavjudligi uning tilni egallash darajasi haqida muhim ma'lumot beradi. Biroq xatolar sonining o'zi yetarli emas. Xatoning qaysi birlikda va qanday shaklda yuz bergani ham aniqlanishi kerak.

Masalan, o'zbek tilini o'rganuvchi kelishik qo'shimchasini noto'g'ri tanlashi, egalik shaklini tushirib qoldirishi yoki fe'l zamonini kontekstga mos kelmaydigan shaklda qo'llashi mumkin. Ayrim hollarda bitta so'z tarkibida bir necha grammatik xato ham uchraydi. O'zbek tilining affiksial tuzilishi sababli bunday xatolarni avtomatik aniqlash uchun morfologik tahlil muhim o'rin tutadi. Shuning uchun baholash tizimida grammatik xatolarni morfologik va sintaktik xatolarga ajratish maqsadga muvofiq.

Xatolarni tasniflashning yana bir muhim tomoni ularning o'rganuvchining ona tili bilan bog'liqligini aniqlashdir. Ikkinchi tilni o'zlashtirish jarayonida o'rganuvchi ba'zan o'z ona tilidagi grammatik yoki sintaktik modelni yangi tilga ko'chiradi. Bunday holat interlingval ta'sir bilan bog'liq bo'lib, yozma nutqda so'z tartibi, so'z tanlash, grammatik shakllar va gap tuzilishida ko'rinishi mumkin [6]. Shu sababli o'zbek tilining o'rganuvchi korpusida xatolarning kelib chiqishiga oid ma'lumotlarni ham qayd etish keyingi tahlillar uchun foydali bo'ladi.

Yozma nutq sifatini belgilaydigan yana bir muhim ko'rsatkich lug'aviy boylikdir. O'rganuvchi matnida ishlatilgan so'zlarning xilma-xilligi uning til imkoniyatlarini baholashda muhim ahamiyatga ega. Lug'aviy boylikni aniqlash uchun takrorlangan va takrorlanmagan so'zlarning nisbati, sinonimik birliklardan foydalanish, mavzuga mos so'zlarni tanlash kabi belgilarni hisobga olish mumkin. Leksik xilma-xillikni o'lchashda type-token ratio kabi ko'rsatkichlardan foydalanish mumkin, biroq matn hajmining natijaga ta'sirini ham e'tiborga olish zarur [4].

O'zbek tilida bu masala yana bir jihatdan murakkablashadi. Bir leksema turli grammatik qo'shimchalarni qabul qilgan holda bir necha shaklda uchrashi mumkin. Shuning uchun avtomatik tizim so'zning har bir shaklini alohida birlik sifatida hisoblamasdan, zarur hollarda uning lug'aviy asosini ham aniqlashi kerak. Buning uchun morfologik tahlil va lug'aviy tahlil o'zaro bog'langan holda olib borilishi lozim.

Yozma nutqning rivojlanganlik darajasini ko'rsatadigan mezonlardan biri sintaktik murakkablikdir. Oddiy gaplardan foydalanish boshlang'ich bosqich



POLAND



POLAND

uchun tabiiy bo'lsa, tilni yaxshi egallagan o'rganuvchi fikrini qo'shma gaplar, ergash konstruksiyalar va turli sintaktik bog'lanishlar orqali ifodalashi mumkin. Shu sababli gapning o'rtacha uzunligi, gaplar tarkibidagi predikativ birliklar soni, qo'shma gaplarning ulushi va ergash konstruksiyalardan foydalanish darajasi yozma nutqni tavsiflashda qo'l keladi [3].

X. Lu tadqiqotlarida sintaktik murakkablikni avtomatik o'lchash uchun turli lingvistik ko'rsatkichlardan foydalanish imkoniyati ko'rsatilgan. Bunday yondashuvda gap va bo'laklarning uzunligi, predikativ birliklarning nisbati hamda sintaktik tuzilmalarning murakkablik darajasi alohida ko'rib chiqiladi [3]. Keyingi tadqiqotlarda ham yuqori baholangan yozma ishlarda ayrim murakkab sintaktik konstruksiyalar ko'proq uchrashi aniqlangan [7].

Biroq sintaktik murakkablikni faqat gap uzunligi bilan o'lchash to'g'ri natija bermasligi mumkin. Uzun gap grammatik jihatdan noto'g'ri yoki mazmunan chalkash bo'lishi mumkin. Aksincha, qisqa gaplardan tashkil topgan matn mazmunan aniq va izchil bo'lishi mumkin. Demak, sintaktik murakkablik grammatik to'g'rilik va mazmuniy yaxlitlik bilan birgalikda baholanishi kerak.

Yozma nutqning yana bir muhim jihati fikrning mazmunan rivojlantirilishidir. O'rganuvchi berilgan mavzu doirasida asosiy fikrni bayon qilishi, uni izohlashi, dalillar keltirishi va zarur hollarda qo'shimcha ma'lumot bilan boyitishi kerak. Fikrni ortiqcha takrorlarsiz ixcham ifodalash ham yozma nutq mahoratining muhim belgisi hisoblanadi. Shuning uchun matn hajmini oddiy hisoblash bilan cheklanib qolmasdan, unda yangi mazmuniy birliklar qanchalik ko'p va izchil berilganini ham hisobga olish zarur.

Mazmunning rivojlanganligini aniqlash avtomatik baholashning murakkab tomonlaridan biridir. Bunda gaplar soni, matnning umumiy hajmi, takroriy birliklarning ulushi, asosiy fikrni rivojlantiruvchi gaplar soni, sabab-oqibat va izohlash munosabatlarining mavjudligi kabi ko'rsatkichlardan foydalanish mumkin. Bunday yondashuv yozma ishning faqat grammatik jihatdan to'g'ri yoki noto'g'riligini emas, balki berilgan topshiriqqa mazmunan javob berish darajasini ham aniqlashga yordam beradi.

Matnning bog'liqligi va izchilligi ham baholash mezonlaridan biri sifatida qaralishi kerak. Yozma matndagi gaplar bir-biridan uzilib qolmasdan, yagona fikriy yo'nalishda rivojlanishi lozim. Bunda bog'lovchilar, olmoshlar, takroriy va sinonimik birliklar, sabab-oqibat munosabatini bildiruvchi vositalar hamda vaqt va makon ko'rsatkichlari muhim rol o'ynaydi.

Cohesion va coherence tushunchalariga asoslangan bunday yondashuv xorijiy yozma nutq tadqiqotlarida keng qo'llanadi. Zamonaviy tadqiqotlarda



leksik boylik, sintaktik murakkablik va matnning bog'liqligi bir-biridan butunlay ajralgan ko'rsatkichlar emas, balki yozma nutq sifatini birgalikda shakllantiradigan omillar sifatida qaralmoqda [4; 10]. Shu sababli o'zbek tilining ikkinchi til sifatidagi yozma nutqini baholashda ham mazkur ko'rsatkichlarni birgalikda ko'rish maqsadga muvofiq.

Yozma nutqni avtomatik baholashda yuqoridagi mezonlarning har biri alohida natija berishi mumkin. Biroq yakuniy baho ularning umumiy ko'rsatkichlari asosida shakllantirilsa, o'rganuvchining yozma nutqini aniqroq tavsiflash imkoniyati paydo bo'ladi. Shu nuqtayi nazardan o'zbek tili uchun quyidagi mezonlar tizimini taklif qilish mumkin:

Baholash mezonlari	Asosiy ko'rsatkichlar
Grammatik to'g'rilik	morfologik va sintaktik xatolar, affikslarning qo'llanishi, fe'l shakllari
Lug'aviy boylik	so'zlarning xilma-xilligi, takrorlanish darajasi, leksik tanlov
Sintaktik murakkablik	gap uzunligi, predikativ birliklar, qo'shma va ergash gaplar
Mazmunning rivojlanganligi	asosiy fikrni izohlash, dalillash, qo'shimcha ma'lumot berish
Bog'liqlik	bog'lovchi va boshqa aloqa vositalarining qo'llanishi
Izchillik	fikrlarning mantiqiy ketma-ketligi, mavzuning rivojlanishi
Kommunikativ moslik	topshiriq mazmunini ochish, mavzuga moslik, fikrning aniqligi

Mazkur mezonlar o'zaro bog'langan holda qo'llanganda yozma nutqning ancha to'liq manzarasini beradi. Masalan, grammatik jihatdan to'g'ri, lekin juda qisqa va mazmunan rivojlanmagan matn yuqori baholanmasligi kerak. Xuddi shuningdek, murakkab gaplardan foydalangan, ammo ko'plab grammatik xatolarga ega bo'lgan matnni ham faqat sintaktik murakkablik asosida yuqori darajaga kiritish mumkin emas.

Shu jihatdan avtomatik baholash modelida har bir mezonga alohida vazn berish mumkin. Vaznlar esa real o'rganuvchi korpusidagi ma'lumotlar asosida belgilanadi. Masalan, turli darajadagi yozma ishlar tahlil qilinib, qaysi lingvistik ko'rsatkichlar baholangan matnlarda ko'proq uchrashi aniqlanadi. Keyinchalik statistik usullar yordamida ushbu ko'rsatkichlarning yakuniy bahoga ta'siri



hisoblab chiqilishi mumkin. Xorijiy tadqiqotlarda ham aynan korpus ma'lumotlari asosida lingvistik belgilarni tanlash va ularning yozma ish sifatini bashorat qilishdagi imkoniyatlarini tekshirish usulidan foydalanilgan [4; 5].

Bu jarayonda inson baholovchining o'rni ham muhim. Avtomatik tizim uchun dastlabki korpusdagi matnlar mutaxassislar tomonidan baholanishi va belgilanishi kerak. Bir nechta baholovchi tomonidan qo'yilgan natijalarni taqqoslash orqali umumiy mezonlar aniqlanadi. Shundan keyin avtomatik tizim ushbu baholar bilan lingvistik ko'rsatkichlar o'rtasidagi munosabatni o'rganadi. Shu tariqa inson bahosi va avtomatik tahlil o'rtasida bog'liqlik yuzaga keladi.

O'zbek tili uchun bunday modelni ishlab chiqishda eng katta vazifalardan biri sifatli o'rganuvchi korpusini yaratishdir. Korpusda nafaqat to'g'ri yozilgan matnlar, balki turli darajadagi xatolarga ega yozma ishlar ham bo'lishi lozim. Har bir xatoning turi, joylashuvi va imkon qadar uning sababi qayd etilishi avtomatik tahlilning aniqligini oshiradi. Bundan tashqari, matnlarning turli mavzularda bo'lishi, turli topshiriqlar asosida yozilishi va turli til darajalarini qamrab olishi ham muhimdir.

Bunday korpus asosida o'zbek tiliga xos lingvistik ko'rsatkichlarni aniqlash mumkin bo'ladi. Natijada xorijiy tillar uchun ishlab chiqilgan mezonlarni aynan takrorlash o'rniga o'zbek tilining grammatik va leksik xususiyatlariga mos keladigan baholash tizimini shakllantirish imkoniyati paydo bo'ladi. Ayniqsa, affiksial morfologiya, so'z shakllari, gap bo'laklarining joylashuvi va bog'lovchi vositalarning qo'llanishi kabi xususiyatlar modelda alohida hisobga olinishi kerak.

Shunday qilib, o'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqini avtomatik baholash uchun birgina mezon asosida ishlaydigan tizimdan ko'ra, bir nechta lingvistik ko'rsatkichlarni birlashtirgan yondashuv maqsadga muvofiqdir. Bunda xatolar tahlili, lug'aviy boylik, sintaktik murakkablik, mazmunning rivojlanganligi, bog'liqlik, izchillik va kommunikativ moslik bir-birini to'ldiradi. Mazkur ko'rsatkichlarni real o'rganuvchi korpusi asosida tekshirish esa kelgusida avtomatik baholash modelining ishonchliligini oshirishga xizmat qiladi.

Xulosa

O'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqini avtomatik baholash uchun, avvalo, yozma nutqning asosiy lingvistik xususiyatlarini aniq belgilash zarur. Tahlil shuni ko'rsatadiki, yozma ish sifatini faqat grammatik xatolar soni bilan baholash o'rganuvchining real nutqiy imkoniyatlarini to'liq aks ettirmaydi. Grammatik to'g'rilik bilan bir qatorda



lug'aviy boylik, sintaktik murakkablik, mazmunning rivojlanganligi, matnning bog'liqligi va izchilligi ham hisobga olinishi kerak.

O'zbek tilining agglutinativ tuzilishi avtomatik baholash tizimiga alohida talablar qo'yadi. Grammatik ma'nolarning affiksalar yordamida ifodalanishi so'z shakllarini avtomatik tahlil qilishda morfologik yondashuvning ahamiyatini oshiradi. Shuning uchun o'zbek tili uchun ishlab chiqiladigan baholash modeli umumiy til mezonlari bilan bir qatorda o'zbek tilining o'ziga xos grammatik xususiyatlarini ham qamrab olishi lozim.

Xorijiy tadqiqotlarda qo'llanilgan korpus asosidagi tahlil, sintaktik murakkablikni o'lchash, lug'aviy boylikni aniqlash va lingvistik ko'rsatkichlar yordamida yozma ish sifatini bashorat qilish usullari o'zbek tili uchun ham muayyan nazariy va metodik asos bo'la oladi [3; 4; 5]. Biroq ularni o'zbek tiliga moslashtirish real o'rganuvchi materiallari asosida amalga oshirilishi kerak.

Shu nuqtayi nazardan, o'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqi korpusini yaratish, undagi tipik xatolarni tasniflash, lingvistik belgilarni aniqlash va ularning yozma nutq sifati bilan bog'liqligini tekshirish keyingi tadqiqotlarning muhim yo'nalishlaridan hisoblanadi. Mazkur bosqichlar asosida shakllantirilgan mezonlar keyinchalik o'zbek tilini ikkinchi til sifatida o'rganuvchilarning yozma nutqini avtomatik baholash modelining lingvistik poydevorini yaratishi mumkin.

Foydalanilgan adabiyotlar ro'yxati

1. Bachman L. F., Palmer A. S. *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press, 1996.
2. Granger S. Learner corpus research: Past, present and future // *Language Teaching*. 2002. Vol. 35. No. 4. P. 247–256.
3. Lu X. Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment // *Language Testing*. 2017. Vol. 34. No. 4. P. 528–548.
4. Youn H., Crossley S. A. What linguistic features distinguish and predict L2 writing quality? A study of examination scripts written by adolescent Chinese learners of English in Hong Kong // *System*. 2021. Vol. 97. Article 102461.
5. Automated assessment of learner text complexity // *Assessing Writing*. 2021. Vol. 49. Article 100529.
6. Ellis R. *The Study of Second Language Acquisition*. 2nd ed. Oxford: Oxford University Press, 2008.
7. syntactic complexity and writing quality in assessed first-year L2 writing // *Journal of Second Language Writing*. 2019. Vol. 44. P. 51–62.



POLAND

CURRENT APPROACHES AND NEW RESEARCH IN MODERN SCIENCES

International scientific-online conference



POLAND

8.Hojiev A. *O'zbek tili morfologiyasi, morfemikasi va so'z yasalishi*. Toshkent: Fan, 2002.

9.Rahmatullaev Sh. *Hozirgi adabiy o'zbek tili*. Toshkent: Universitet, 2006.

Council of Europe. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*. Strasbourg: Council of Europe Publishing, 2020.