Materiallar to'plami · Conference Proceedings
"ZAMONAVIY TADQIQOTLAR VA BARQAROR RIVOJLANISH" respublika ilmiy-amaliy konferensiyasi
04 August 2026
UDK 811.111
FILOLOGIYA

AN AI-AGENT-BASED DIAGNOSTIC MODEL FOR DETECTING INTERFERENCE ERRORS IN THE ENGLISH LANGUAGE OF UZBEK-SPEAKING STUDENTS: A LANGUAGE-SPECIFIC APPROACH

Munzira Yusupova*
Yangi O'zbekiston Universiteti magistri
✉ munzirayusupova4@gmail.com

ABSTRACT

This thesis is aimed at identifying the errors that arise in the acquisition of English by Uzbek-speaking students. In addition, this work proposes an AI-agent-based diagnostic model. The study relies on Contrastive Analysis and Interlanguage theories and attempts to predict the types of errors arising from the typological differences between the two languages. The purpose of the research is to compare the error-detection accuracy of the AI agent with human-expert evaluation, and to show in which error categories it is strong and in which it is weak. The results of the proposed model can serve as a practical basis for adapting English-language teaching methodology and AI tools to local conditions.

Keywords

interlanguage, contrastive analysis, error analysis, AI agent, language-specific diagnostics, English language teaching.

DOI: 10.5281/zenodo.21743169

INTRODUCTION

In recent years, artificial intelligence tools have been actively introduced into the foreign-language teaching process. Such tools give students the opportunity to practice independently and receive immediate feedback. However, their error-detection accuracy has mainly been studied in the context of English, Chinese, or Arabic, while it has been almost unexplored for Uzbek students. At the same time, Uzbek and English differ fundamentally in typological terms. Uzbek has an agglutinative structure and an SOV word order, whereas English has an analytic structure and an SVO order. Furthermore, the category of articles does not exist in Uzbek at all. These systemic differences give rise to recurring errors among students[2]. If AI tools do not recognize these language-specific errors with sufficient accuracy, the automatic feedback they provide may be incorrect or misleading. For this reason, systematically pre-testing the diagnostic accuracy of AI agents is a necessary condition for their responsible introduction into the educational process.

The purpose of this thesis is to propose a clear and verifiable research design for studying how accurately existing AI agents recognize the error types characteristic of Uzbek students.

This research relies on two classical theories. One is the Contrastive Analysis Hypothesis proposed by the American linguist Robert Lado (1915–1995). According to his theory, systematic differences between two languages make it possible to predict errors in language learners[2]. The other is the theory of error analysis founded by the American linguist Larry Selinker. According to this theory, a learner's linguistic system develops as an independent, systematic stage between the native language and the target language, and the errors at this stage are not accidental but a lawful phenomenon[1]. In particular, based on contrastive analysis, the main types

of errors expected in Uzbek students' acquisition of English and their linguistic causes are presented in Table 1. This typology serves as the basis for separately assessing the AI agent's accuracy in each category.

Table 1

Error type	Linguistic cause (contrastive)	Error example	Expected difficulty for AI
Article errors	The category of articles does not exist in Uzbek	”I saw car” (instead of a car)	Medium — context-dependent
Preposition errors	Preposition systems do not correspond; expressed through suffixes	”depend of” (instead of on)	Medium
Verb tense-aspect errors	Perfect/Continuous forms are not grammatically expressed in Uzbek	”I have seen him yesterday”	Medium-low
Word order (SOV trace)	The SOV order of the native language interferes with the SVO structure	”I to school go”	High - structure is clearly violated
Subject-verb agreement	A rule-based, universal error type	”She go” (instead of goes)	High accuracy expected
Lexical-semantic / collocation	Calque from the native language, high context-dependency	”make a homework” (instead of do)	Low - the most difficult category for AI

Table 1. Typology of interference errors expected among Uzbek students

As can be seen from the table, some errors are rule-based and clearly manifest at the structural level. For this reason, they are expected to be relatively easy for AI to detect. Other errors, however, are deeply context-dependent. It is precisely in these categories that the AI agent's accuracy may decline. This is one of the central hypotheses of the conducted research.

The proposed diagnostic model is based on a four-stage cycle consisting of corpus collection, manual annotation, AI analysis, and statistical comparison. The content of each stage, the method/tool used, and the expected outcome are summarized in Table 2.

Table 2

Stage	Content of the activity	Method / tool	Result
1. Corpus collection	Written works (argumentative essay genre) of 60–80 students are collected	Written test assignment, CEFR B1–B2 level	Raw error corpus
2. Manual annotation	Two independent expert teachers categorize the errors	Manual coding, Cohen's Kappa control	”Gold standard”
3. AI analysis	The AI agent independently analyzes the same texts	AI agent (language-model-based diagnostic tool)	AI error annotation
4. Comparison	AI results are compared with the gold standard	Calculation of Precision, Recall, F1-score	Accuracy indicators by category

Table 2. Stages of the diagnostic model

In the first stage of the research, the error rate was examined based on the written works of 60–80 students at CEFR B1–B2 level. In the second stage, two independent expert teachers manually review the errors, and this is accepted as the ”gold standard.” In addition, the level of agreement between the experts is checked using Cohen's Kappa coefficient; that is, a low level of agreement would require revising the annotation criteria. In the third stage, the AI agent independently analyzes the same texts. In the fourth stage, the AI results are compared with the gold standard, and separate indicators are calculated for each error category.

Based on the results of the conducted research, the following conclusions were reached:

1. The AI agent shows high precision and recall for rule-based error types that are clearly defined at the structural level,
2. Accuracy is lower for article and preposition errors, because these categories have no systematic equivalent in Uzbek and are highly context-dependent,

3. The AI makes the most errors in detecting lexical-semantic errors, because this type of error is identified not through purely grammatical rules but through broad contextual knowledge.

Based on the conclusions above, it would be incorrect to say that AI agents are equally reliable across all error types when introducing them into the educational process. At the same time, teacher oversight should be strengthened in the higher-risk categories. The results can also be practically useful in several respects. Accordingly, the list of the most common error typology for Uzbek students can be used in preparing textbooks and exercise materials; the AI agent's level of reliability by category can show teachers when to fully trust the AI's assessment and when to strengthen their own oversight; and this can also create a scientific basis for developing AI tools adapted to local conditions.

CONCLUSION

The proposed research is aimed at systematically measuring, for the first time, the diagnostic accuracy of AI agents for the Uzbek-English language pair. The results not only reveal the practical limitations of AI tools but can also serve as a scientific basis for creating language-specific AI solutions adapted for Uzbek students. In future research, the model can also be extended to other genres and other language levels.

REFERENCES

1. Selinker L. Interlanguage // IRAL – International Review of Applied Linguistics in Language Teaching. – 1972. – Vol. 10, No. 1–4. – P. 209–231.

2. Lado R. Linguistics Across Cultures: Applied Linguistics for Language Teachers. – Ann Arbor: University of Michigan Press, 1957. – 141 p.

3. Almusharraf N., Alotaibi H. An Error-Analysis Study from an EFL Writing Context: Human and Automated Essay Scoring Approaches // Technology, Knowledge and Learning. – 2023. – Vol. 28. – P. 1015–1031.

4. Mamatova D.B. Talabalarining ingliz tilidagi og'zaki nutqni rivojlantirishda AI texnologiyalaridan foydalanish metodikasi // The Lingua Spectrum. – 2025. – Vol. 5. – P. 357–368.