DO AI TOOLS MATTER? THE IMPACT OF THEIR USE ON UNIVERSITY STUDENTS' ENGLISH LANGUAGE LEARNING EFFICIENCY

[1]JELENA GUGIĆ, [2]IVA BLAŽEVIĆ

Faculty of Educational Sciences, Juraj Dobrila University of Pula
E-mail: [1]Jelena.gugic@unipu.hr, [2]iva.blazevic@unipu.hr

Abstract - This research studies the impact of artificial intelligence (AI) tools on the efficiency of English language learning among undergraduate students ofthe Juraj Dobrila University of Pula. As many as 42 female students were divided into an experimental group using AI tools and a control group employing traditional learning methods. Both groups covered identical course materials relevant to their future profession. Initial and final English proficiency tests assessed listening, reading, vocabulary, and grammar skills. Statistical analyses found no statistically significant differences in final test results between the two groups. As findings suggest, AI toolsdid not significantly enhance language learning efficiency compared to traditional methods. Potential reasons include limited intensity and duration of AI tool use and a narrow score range on final assessments, calling for future research with a larger number of participants and more comprehensive AI applications. In addition, the study underscores the importance of balanced teaching approaches that integrate both AI tools and traditional methods.

I. INTRODUCTION

The functioning of today's society is characterised by fast access to information enabled by artificial intelligence (AI). Such a rapid growth of AI tools introduced significant changes to the learning process as well. In a document published by the European Commission in 2026 AI is defined as a machine-based, autonomous system, adaptive to the students' needs which generates certain outputs after given certain inputs (EC, 2026).Teaching professionals, although afraid of the role AI tools could take within the learning process, have started to embrace it and use its full potential in adapting their lessons to suit modern generations of"tech-savvy learners" (Hashim, 2018, as cited in Tessensohn et al., 2025, p. 59). However, as Ghafar et al. (2023) state when discussing the role of AI in English language classes, class models including AI should at the same time include traditional teaching.

As widely known, language instructors have a difficult role of teaching a language by using languageper se, and since the complex nature of teaching a language with the aim to develop not only communicative skills, but also accuracy, demands high creativity in catering for different students' needs, it comes as no surprise that the AI tools industry has developed such a large number of tools directly addressing these needs(Shah et al., 2025; Khosravi et al., 2022). Studies have shown that the use of AI improves EFL teaching, especially developing the speaking skill (Jantakoon et al., 2025). Since AI tools offer a variety of activities with great motivational potential, and with the ability to reduce the shortcomings of traditional teaching, it has been suggested that the most appropriate teaching method is to blend both into classes in a balanced way (Qazi et al., 2023; Sahito et al., 2025; Jantakoon et al., 2025).

II. RESEARCH METHODOLOGY

The aim of this research was to determine whether there are differences in the efficiency of English language learning between a group of respondents who learned course materials using AI tools and a group of respondents who learned the same material using traditional learning methods, without using AI tools.

Sample of respondents

The research was conducted on a sample of 42 female students of the undergraduate study Early and Preschool Education at the Juraj Dobrila University of Pula (Croatia), who were divided into two groups during the winter semester, a control and an experimental group.

Research protocol

English language teaching took place over 15 weeks, or 12 teaching hours of 45 minutes each. Both groups used the same teaching materials, with the control group working on the material without using AI tools, while the experimental group used AI tools. Topics specific to the future profession of preschool teacher were covered – multiple intelligences, learning styles, famous educators, work protocols in kindergartens and planned grammar content. During the first lesson, the students were introduced to the aim and protocol of the research and gave their consent by signing an Informed Consent to participate in the research. The research was approved by the Ethics Committee at the Juraj Dobrila University of Pula.At the beginning of the semester, all respondents took a standardized English language proficiency test – a placement test

(Gugić, Malenica, Mustapić Malenica, n.d.) to determine the initial level of language competence and ensure comparability between groups.At the end of the teaching cycle, both groups took the same final knowledge exam which consisted of 40 tasks that tested the subjects' listening skills (10 questions), reading skills (10 questions), vocabulary (10 questions) and grammatical skills (10 questions).

Sample of variables

The sample of variables is constituted by the students' initial and final test results.

Data processing methods

Descriptive statistics, T-test for independent samples and analysis of covariance (ANCOVA) were used to analyse the obtained data. Since the initial and final tests were not identical, but measured the same language competence with different tasks, analysis of covariance was used to analyse the effect of the experimental procedure. This served as a control for the influence of the students' initial level of knowledge on the final result, which enabled a more reliable assessment of the effect of the application of artificial intelligence tools.

III. RESULTS AND DISCUSSION

EXPERIMENTAL GROUP

	N	Mean	Minimum	Maximum	SD	Skewness	Kurtosis
PLACEMENT TEST	22	22.727	10.000	33.000	5.849	-0.220	-0.388
FINAL TEST	22	35.455	30.000	40.000	2.703	0.150	-0.494
CONTROL GROUP							
	N	Mean	Minimum	Maximum	SD	Skewness	Kurtosis
PLACEMENT TEST	18	23.611	15.000	34.000	5.112	0.263	-0.359
FINAL TEST	18	36.111	32.000	40.000	2.272	0.185	-0.535

Table 1. Descriptive indicators of the results of the initial and final tests in the experimental and control groups

N – number of participants; Mean – average result; Minimum – lowest result; Maximum – highest result; SD – standard deviation of the result from the arithmetic mean; Skewness – measure of the symmetry of a variable's distribution; Kurtosis – measure of 'peakedness' of a distribution

Descriptive data analysis showed that the experimental and control groups had very similar initial test results. The average score in the experimental group was 22.73 points out of a maximum of 40 points (SD=5.85), while in the control group it was 23.61 points (SD=5.11). The experimental group achieved an average score of 35.46 points (SD=2.70) on the final test, while the control group achieved a slightly higher average of 36.11 points (SD=2.27). The analysis of the normality of results distribution show that the values of skewness and kurtosis are within acceptable limits, which indicates an approximately normal distribution of the data and justifies the application of parametric statistical procedures.

	Mean CG	Mean EG	t-value	df	p	N CG	N EG	SD CG	SD EG	F-ratio Variances	p Variances
PLACEMENT TEST	23.611	22.727	0.503	38	0.618	18	22	5.112	5.849	1.309	0.578

Table 2. Results of the t-test between the experimental and control groups at the initial test

CG – control group; EG – experimental group; Mean – arithmetic mean; N – number of participants; SD – standard deviation; t – t-test value; df – degrees of freedom; F-ratio variances; p – significance level.

Table 2 shows the results of the t-test for independent samples on the results of the initial test with the aim of determining the initial comparability of the groups. The results showed no statistically significant difference between the groups (t(38)=0.50, p=0.618), which confirmed that the groups were uniform before the start of the experimental procedure.An analysis of covariance (ANCOVA) was performed to determine the effects of the application of artificial intelligence tools on the success of the test subjects in the final test, where the result of the final test was the dependent variable, group affiliation (experimental and control) was the independent variable, and the result of the initial test was the covariate.

	SS	Degr. of Freedom	MS	F	p
Intercept	2362.630	1	2362.630	374.038	0.000
PLACEMENT TEST	7.520	1	7.520	1.191	0.282
Group	3.371	1	3.371	0.534	0.470
Error	233.712	37	6.317		

Table 3. Results of covariance analysis (ANCOVA)

SS – sum of squares; df – degrees of freedom; MS – mean of square; F – F-testa value; p – statistical significance level; Intercept – model constant; PLACEMENT TEST – initial test result (covariate); Group – group affiliation (control or experimental); Error – unexplained variance (error)

The results of the analysis showed that the initial level of knowledge of the subjects was not a statistically significant predictor of the final test result (F=1.19, p=0.282). The results obtained did not show a statistically significant effect of group affiliation on the final test result (F=0.53, p=0.470), and the above indicates that there is no significant difference between the experimental and control groups after the research.

	Multiple R	Multiple R2	Adjusted R2	SS Model	df Model	MS Model	SS Residual	df Residual	MS Residual	F	p
FINAL TEST	0,219	0,048	-0,003	11,788	2	5,894	233,712	37	6,317	0,933	0,402

Table 4. Results of the significance test of the model for predicting the final test results

Multiple R – multiple correlation coefficient between the predictor and dependent variable; Multiple R^2 – coefficientof determination(share of the explained variance); Adjusted R^2 – adjusted coefficientof determination; SS Model – sum of model squared (explained variance); df Model – model degrees of freedom; MS Model – mean value of the squared model; SS Residual – sum of squares of the residual (unexplained variance); df Residual –degrees of freedom of the residual; MS Residual – mean value of the squared residual; F – F-testa value for the whole model; p – statistical significance level.

The analysis of the overall model did not show any statistical significance (F=0.93, p=0.402), with the model explaining 4.8% of the variance in the final test results (R^2=0.048). These results suggest that the included variables, i.e., initial knowledge and group affiliation, have a very limited contribution to explaining final test achievement.

CONCLUSION

The results of the study show that there are no statistically significant differences between the experimental and control groups in the overall result on the final test, although the control group achieved a slightly higher final test result. The results of this study suggest that the application of AI tools in this study did not have a statistically significant impact on the efficiency of listening, reading, vocabulary and grammatical skills development in the English language, nor did it influence the final test achievement. It is possible that the intensity and method of application of the tools were not sufficiently pronounced to lead to measurable changes, or that the time frame of the study was insufficient to develop significant effects. Moreover, the relatively small range of results on the final test may have limited the ability to detect differences between groups.Although the results obtained do not confirm the effectiveness of the application of artificial intelligence tools in this study, they provide valuable insight into the complexity of the impact of technology on learning English as a language of the profession and indicate the need for further research with larger samples and more elaborate models of application. However, this small-scale study can confirm the need for balanced lectures including both traditional teaching methods and AI tools use.

ACKNOWLEDGEMENTS

This research was supported by the Erasmus+ grant number 2024-1-RO01- KA220-HED-000248845, funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the AMPEU. Neither the European Union nor the AMPEU can be held responsible for them.

REFERENCES

[1] European Commission (2026). Guidelines on the ethical use of artificial intelligence and data in teaching and learning for educators. Publications office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/f692aa0b-17a7-11f1-8870-01aa75ed71a1/

[2] Ghafar, Z. N., Salh, H. F., Abdulrahim, M. A., Farxha, S. S., Arf, S. F., & Rahim, R. I. (2023). The role of artificial intelligence technology on English language learning: A

literature review. Canadian Journal of Language and Literature Studies, 3(2), 17-31.

[3] Hashim, H. U., Yunus, M. M., & Hashim, H. (2018). Language learning strategies used by adult learners of teaching English as a Second Language (TESL). TESOL Journal (Laguna), 13(4), 39–48.

[4] Jantakoon, T., Jantakun, T., Jantakun, K., Pongpanich, W., Pasmala, R., Wannapiroon, P., &Nilsook, P. (2025). The effectiveness of artificial intelligence in English instruction for speaking and listening skills: A meta-analysis. Contemporary Educational Technology, 17(4), Article ep596. https://doi.org/10.30935/cedtech/17310

[5] Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C, Tsai, Y-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S. &Gašević, D. (2022). Explainable Artificial Intelligence in education. Computers and Education: Artificial Intelligence3. https://doi.org/10.1016/j.caeai.2022.100074

[6] Qazi, W., Raza, S. A., & Waheed, M. (2023). Access, equity, and learning outcomes: The promise of AI tools in Pakistani classrooms. Journal of Educational Research and Practice, 13(1), 88–102.

[7] Sahito, J. K. M., Panwar, A. H. & Ramzan, I. (2025). Exploring the impact of artificial intelligence (AI) on the listening skills of English as a second language (ESL) learners. Journal of Applied Linguistics and TESOL (JALT), 8(1), 1059-1067

[8] Shah, S. H. R., Altaf, A. R. & Mughal, K. H. (2025). Exploring the impact of AI tools on reading proficiency among undergraduate students. The Critical Review of Social Sciences Studies, 3(3), 667-670.

[9] Tessensohn, T. C., Yunus, M. M. & Ismail, H. H. (2025). Using AI-powered tools in enhancing reading skills in the ESL classroom: A systematic review (2020-2024). International Journal of Academic Research in Progressive Education and Development, 14(2), 57-70

★★★