

Izvorni znanstveni rad
378.147:811.111
004.8:37
DOI
Primljeno: 9.6.2026.



Autori zadržavaju autorska prava i daju časopisu pravo prvog objavljivanja pod CC-BY 4.0 licencom koja dopušta drugima dijeljenje rada uz priznanje autorstva rada i početne objave u časopisu.

Comparison of AI-Supported and Traditional English for Specific Purposes Instruction among University Students

JELENA GUGIĆ ^{1*}, IVA BLAŽEVIĆ ¹

¹ Juraj Dobrila University of Pula, Faculty of Educational Sciences, Negrijeva 6, 52100
Pula, Croatia

e-mail: jelena.gugic@unipu.hr; iva.blazevic@unipu.hr

*Contact author: jelena.gugic@unipu.hr

Abstract This study compares the effectiveness of AI-supported learning activities and traditional teacher-led English language instruction among undergraduate students of the Juraj Dobrila University of Pula. A total of 42 female students were divided into an experimental group participating in AI-supported learning activities and a control group employing traditional learning methods. Both groups covered identical course materials relevant to their future profession. Initial and final English language 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, participation in AI-supported learning activities did not result in statistically significant differences in English language learning outcomes compared with traditional teacher-led instruction. Potential reasons include the limited intensity and duration of the intervention, the relatively small sample size, and the fact that the experimental condition combined AI tool use with increased learner autonomy and reduced teacher guidance. The study highlights the importance of carefully integrating AI-supported activities into language instruction and underscores the need for further research involving larger samples and more comprehensive implementation of AI-supported learning approaches.

Keywords: 1. artificial intelligence; 2. English for specific purposes; 3. AI-supported learning; 4. traditional learning methods; 5. university students

Usporedba nastave engleskog jezika za specifične namjene podržane umjetnom inteligencijom i tradicionalne nastave engleskog jezika za specifične namjene među sveučilišnim studentima

JELENA GUGIĆ ^{1*}, IVA BLAŽEVIĆ ¹

¹ Sveučilište Jurja Dobrile u Puli, Fakultet za odgojne i obrazovne znanosti, Negrijeva 6, 52100 Pula, Hrvatska

e-pošta: jelena.gugic@unipu.hr; iva.blazevic@unipu.hr

*Kontakt autor: jelena.gugic@unipu.hr

Sažetak Ovim istraživanjem uspoređuje se učinkovitost aktivnosti učenja uz pomoć umjetne inteligencije i tradicionalne nastave engleskog jezika koju vodi nastavnik među studentima prijediplomskog studija Sveučilišta Jurja Dobrile u Puli. Ukupno 42 studentice podijeljene su u eksperimentalnu skupinu koja je sudjelovala u aktivnostima učenja uz pomoć umjetne inteligencije i kontrolnu skupinu koja je koristila tradicionalne metode učenja. Obje skupine učile su koristeći iste nastavne materijale relevantne za njihovu buduću profesiju. Početni i završni testovi engleskog jezika procijenjivali su vještine slušanja, čitanja, vokabulara i gramatike. Statističkom analizom nisu pronađene statistički značajne razlike u rezultatima završnih testova između dviju skupina. Kao što nalazi sugeriraju, sudjelovanje u aktivnostima učenja uz pomoć umjetne inteligencije nije rezultiralo statistički značajnim razlikama u ishodima učenja engleskog jezika u usporedbi s tradicionalnom nastavom koju vodi nastavnik. Potencijalni razlozi tome uključuju ograničeni intenzitet i trajanje takvog oblika nastave, relativno malu veličinu uzorka i činjenicu da je eksperimentalni uvjet kombinirao korištenje alata umjetne inteligencije s povećanom autonomijom studenata i smanjenim vođenjem od strane nastavnika. Istraživanjem se ističe važnost pažljivog integriranja aktivnosti uz pomoć umjetne inteligencije u nastavu jezika i naglašava potrebu za daljnjim istraživanjima koja uključuju veće uzorke i sveobuhvatniju primjenu pristupa učenju uz pomoć umjetne inteligencije.

Ključne riječi: 1. engleski za posebne namjene; 2. sveučilišni studenti; 3. tradicionalne metode učenja; 4. učenje uz pomoć umjetne inteligencije; 5. umjetna inteligencija

1. Introduction

One of the defining features of today's society is fast access to information enabled by artificial intelligence (AI). Such a rapid development 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 students' needs, which generates certain outputs after being given certain inputs (European Commission, 2026). Teaching professionals, although afraid of the role AI tools could assume 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.

It is commonplace knowledge that language instructors are tasked with the demanding role of teaching a language by using language *per 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 industry has developed 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 when it comes to the development of the skill of speaking (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 instruction, 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).

2. AI in foreign language instruction and learning

When it comes to foreign language instruction and learning, technology has always played an important role. Even before its appearance, the methodology of foreign language teaching was rich in realia, prompts and various other nondigital tools. However, with the rise of the internet and technology, finding the appropriate means for facilitating language learning has become a much easier task for language instructors. As the main goal of foreign language learning is to develop communicative competence and intercultural understanding, foreign language classes focus on the development of the four

language skills along with grammar and specialised vocabulary knowledge. This fact was recognised by the creators of AI tools as a “gold mine” for their work. The first skill that needs to be mastered when learning a foreign language is listening. Being able to listen and understand what someone else is saying is crucial for the development of other language skills. In a study conducted by Sahito, Panwar and Ramzan (2025), the researchers examined the impact of AI-assisted language learning on the listening comprehension skills of ESL learners and analysed the effectiveness of AI-driven tools, such as speech recognition and virtual tutors, in enhancing ESL learners’ listening abilities. Similarly to the research presented in this paper, the researchers divided students into a control and experimental group and found out that the experimental group achieved better listening comprehension results than the control group. The AI tools supported activities, delivered independent and flexible language practice and led to better language processing and recognition of language patterns, as well as to a superior recognition of different accents. However, students’ overreliance on AI tools was also observed, leading to the conclusion that balance must be found between the introduction of technology and traditional teaching.

Ramadhani et al. (2025) drafted an overview of the research that studied the impact of AI tools on enhancing the speaking skill of users, and concluded that - contrary to traditional teaching methods, which provide limited practice - the majority of the research they examined (for instance, Suja et al., 2025; Agusniati et al., 2025; Rahmawati et al., 2025) referred to AI tools as offering immediate feedback, adapting the learning experience to increase confidence and fluency in students with different needs, while creating practice opportunities that made students feel at ease. When consistently implemented in the classroom, AI tools improve vocabulary mastery, pronunciation fluency, and overall communication skills. However, as already mentioned, the role of language instructors is still crucial due to the need for not only learning content, but also emotional support, which modern technology cannot provide. They emphasise the need for a balanced approach where AI technology is seen as a support to the work of language instructors.

As with other language skills, AI solutions can help improve students’ reading skills by customising reading texts, adapting reading comprehension tasks to the students’ needs and learning dynamics, providing interactive reading experiences, and giving immediate feedback (Alkhanbooli & Alshamsi, 2025). These authors also conducted a study comparing the results achieved by an experimental and by a control group, where a better reading score was achieved by the experimental group. The experimental group showed improved fluency and control and demonstrated better accuracy in reading tasks. Tessensohn et al. (2025) concluded that AI tools can be used for enhancing reading comprehension, accuracy, assessment of reading, critical thinking, and most

importantly, developing reading skills among special needs students. The use of AI tools offers personalised learning experiences by providing better clarification of texts, building confidence, and supporting the development of consistent reading habits (Shah et al., 2025).

Similar findings were reached for the writing skill as well. More precisely, AI tools improve students' writing skills by supporting their writing with personalised activities and tasks and offering more accurate feedback (Kızıldaş, 2025; Zhang et al., 2025). The use of automated writing evaluation tools improves students' writing performance greatly (Para & Calero, 2019). The use of AI writing tools can help students align their writing style to meet academic requirements (Amaria, 2025). What is seen as an issue to be dealt with is, however, the decline in critical thinking and personal writing style development which are the consequence of, again, an overreliance on AI tools (Budyono et al., 2025).

In line with the four areas tested in the experimental part of this research, many authors also studied the impact of AI tools on grammar acquisition. Burstein et al. (2004) state that AI offers immediate feedback on mistakes as well as explanations of grammatical issues (syntax, verb usage, articles, etc.). On the other hand, language instructors' feedback is delayed, thus slowing down the learning cycle. In a study conducted by Stevens (2025) in Sierra Leone, the researcher again divided students into a control and experimental group and concluded that thanks to the use of AI tools, the grammar skills of the experimental group significantly improved. However, some students reported difficulties in understanding AI-explained rules and expressed the need for additional instruction from a teacher. Students' enthusiasm for using AI tools was frequently expressed. Contrary to them, teachers expressed not only fear about inaccuracies in AI-delivered grammar rules and explanations, but also about the proper lack of training on how to implement such tools to achieve the best results.

The final important area for mastering a language correctly is the knowledge of vocabulary, especially specialised, profession-related vocabulary. AI tools have proven to be effective in this field, as well. Being developed on large language models and having the ability to collect all the "materials" accessible online, AI tools are seen as the best solution for vocabulary acquisition. Comparing two groups of students (control and experimental), Chingakham and Tamuk (2024) found out that using AI is more effective than using traditional methods of English vocabulary learning among grade 12 high school learners in India. The research indicates a positive impact on students' academic achievement in English vocabulary development. Changing the usual methodology and using a narrative inquiry approach, Hingpis Aruta and Madeja (2026) explored the potential impact of AI chatbots in enhancing the vocabulary of English language students at the Eastern Visayas State University. Their findings are in line with

recent studies. More precisely, AI tools provide a platform for the expansion of their vocabulary. By offering features such as instant feedback and exposure to new vocabulary and ideas, they have a direct effect on the development of vocabulary.

3. Research methodology

3.1. Research aim

The aim of this research was to determine whether there are differences in English language learning efficiency between undergraduate students who learned course materials using artificial intelligence tools and students who learned the same materials through traditional learning methods.

3.2. 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 (a control and an experimental group) during the winter semester of the 2025/2026 academic year.

3.3. Research protocol

English language instruction 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. 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 of the Juraj Dobrila University of Pula.

At the beginning of the semester, all respondents took a standardised 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. This short proficiency test¹ was created following similar examples of online tests such as the Cambridge University Press and the

¹ The proficiency test was designed as a part of a larger project *Procesiranje višerječnih izraza u engleskom kao stranom jeziku - PROVIZENG* (Eng. *Processing multiword expressions in English as a foreign language*) funded by the University of Zadar (IP-UNIZD-2023).

Assessment placement test. The test covers four categories – vocabulary, spelling, reading comprehension and grammar. There are 10 questions per each category and in the research conducted as part of the PROVIZENG project mentioned in the footnote, there was a time limit per section (3 minutes per section, 12 minutes in total) due to the fact that prior to the proficiency test students had to solve a lexical decision task. Each section of the test needed to be registered separately. In the present research the time to solve the short proficiency test was not limited. The 40 items are in the form of multiple-choice questions where the three incorrect options are logical, but impossible options. The original test was implemented using a Microsoft Forms platform because it enabled the time limitation option for tests, but in this research study it was redesigned as a Google Forms test since the researchers wanted to have a unique test which needed a single registration rather than the four registrations needed in the original short proficiency test.

At the end of the teaching cycle, both the control and experimental group students took the same final knowledge exam which consisted of 40 tasks that tested their listening skills (10 questions), reading skills (10 questions), vocabulary (10 questions) and grammatical skills (10 questions). The original form of the short proficiency test was retained for comparability reasons, but the spelling section was replaced with a listening section because listening comprehension is the language skill necessary to achieve the main goal of the subject course English language I: Systematic refreshment of previously acquired knowledge of the English language with emphasis on intensive training of all four language skills (listening, speaking, reading and writing).

When it comes to the teaching material itself, topics specific to the future profession of preschool teachers were covered – multiple intelligences, learning styles, famous educators, work protocols in kindergartens and planned grammar content. The lessons were designed to cover the skills and knowledge which would be tested at the end of the semester: listening and reading comprehension, and knowledge of vocabulary and grammar. For instance, lessons dedicated to listening comprehension were covered by both groups, but in a different manner. The control group followed a traditional scheme: the lesson was divided into pre-listening, while-listening and post-listening activities coordinated and led by the language instructor. The experimental group was exposed to the same vocabulary in the pre-listening phase. However, they had to use an AI tool to discover the meaning of the given vocabulary. They were next divided into smaller groups where each had to create a text (using AI) containing the given vocabulary and then transform this text into speech (again using an AI tool designed to transform text to speech). The final phase of the pre-listening part was dedicated to creating listening comprehension questions. In the while-listening phase, each group exchanged the recordings they had

created in the previous phase of the lesson, whereas in the post-listening phase students were required to solve comprehension questions. A similar protocol was used to practice reading, grammar and vocabulary and the main difference in the lessons was that, for the control group, the language instructor prepared the teaching material in advance and led the activities, explaining and instructing students what to do, while students of the experimental group worked more independently: they were given initial instructions or topics and they were told what was expected from them, but how they were going to reach the targeted result was up to them.

It is important to emphasise that the experimental condition differed from the control condition not only in the use of AI tools, but also in the degree of learner autonomy and teacher guidance. Therefore, the findings should be interpreted as reflecting the effects of an AI-supported learning approach rather than the isolated effect of AI tools use.

3.4. Sample of variables

The sample of variables is constituted by the students' initial and final test results. The initial (short proficiency) test was conducted at the beginning of the 2025/2026 academic year, winter semester, while the final test was conducted at the end of the same semester.

3.5. Data processing methods

Descriptive statistics, a 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.

4. Results and discussion

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

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

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 shows 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.

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

	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

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 show no statistically significant difference between the groups ($t(38)=0.50$, $p=0.618$), which confirms that the groups were uniform before the start of the experimental procedure. Given that the initial and final tests were not identical, but assessed the same linguistic competence with different tasks, analysis of covariance (ANCOVA) was performed to determine the effects of the experimental procedure. The result of the final test was used as the dependent variable, group affiliation (control or experimental) was the independent variable, while the initial test result was introduced as the covariate to statistically monitor the effect of initial language competence on the final result.

Table 3: Results of covariance analysis (ANCOVA)

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

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 show that the initial level of knowledge of the subjects is not a statistically significant predictor of the final test result ($F=1.19$, $p=0.282$). The results obtained do 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.

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

	<i>Multiple R</i>	<i>Multiple R²</i>	<i>Adjusted R²</i>	<i>SS Model</i>	<i>df Model</i>	<i>MS Model</i>	<i>SS Residual</i>	<i>df Residual</i>	<i>MS Residual</i>	
<i>FINAL TEST</i>	0,219	0,048	-0,003	11,788	2	5,894	233,712	37	6,317	0,3

Multiple R – multiple correlation coefficient between the predictor and dependent variable; Multiple R^2 – coefficient of determination (share of the explained variance); Adjusted R^2 – adjusted coefficient of 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 does also 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.

The results suggest that participation in AI-supported autonomous learning activities did not result in measurable differences in achievement compared with traditional teacher-led instruction. This does not necessarily mean that artificial intelligence tools do not have educational value, but may indicate the complexity of the foreign language learning process and the fact that the effectiveness of technology depends on the way it is integrated into the teaching process. It is possible that the intensity of the application of AI-supported learning activities during the research was not sufficiently pronounced to produce an effect that could be statistically confirmed.

An important methodological consideration is that the experimental and control group instruction differed in several aspects beyond the use of AI tools. Students in the experimental group were more autonomous during lessons, generated learning materials independently and relied less on direct teacher guidance. Consequently, these findings should be understood as the result of an AI-supported learning approach, rather than the independent impact of the AI technology alone.

One possible explanation for the results obtained relates to the way in which artificial intelligence is implemented in teaching. Research in the field of educational technologies shows that the educational effect of digital tools does not stem solely from their presence in the teaching process, but from the quality of their pedagogical integration into teaching activities. The effectiveness of technology is related to the way it supports learning, individualisation of instruction and active involvement of students in teaching activities (Consoli et al., 2025). Systematic reviews of research on AI in education highlight that its positive effects depend on thoughtful implementation, teacher competence and alignment with learning outcomes, while the presence of technology alone does not guarantee better educational outcomes (Garzón, Patiño & Marulanda, 2025; Tan, Cheng & Ling, 2025). It is therefore possible that the used AI tools and AI technology-supported activities were not sufficiently focused on developing the competencies assessed by the final test or that their application was not intensive enough to achieve measurable effects.

The results also show that the initial level of knowledge was not a statistically significant predictor of success on the final test, which may indicate a relatively homogeneous group of respondents who had similar opportunities to acquire the teaching content during the teaching process. The uniformity of the groups was also confirmed by the results of the initial testing, which ensured the methodological justification of the comparison of the experimental and control groups.

When interpreting the results, it is necessary to take into account certain limitations of the research. The research was conducted on a relatively small sample of respondents, which reduces the statistical power of the applied tests and makes it difficult to detect minor effects of the experimental procedure. The final results of both groups were relatively high and with low variability, which could have further reduced the possibility of determining differences between the groups.

It is important to emphasise that the results of this research do not lead to the conclusion that the application of artificial intelligence tools is less effective than the traditional learning approach. The results obtained show that under the conditions of the conducted research, no statistically significant advantage of the use of artificial intelligence tools was determined, so their potential value can be highlighted in other aspects of the teaching process, such as increasing the availability of teaching content, providing immediate feedback, and individualizing student learning and work (Azhar, 2024).

The results obtained indicate the need for further research that would include larger samples of respondents, a longer period of application of AI-supported learning activities, and more detailed monitoring of various aspects of the foreign language learning process. Future research should, in addition to achievements in knowledge tests, also include variables such as motivation for learning, satisfaction with the teaching process, frequency of use of artificial intelligence tools in various learning activities, and perception of their usefulness in learning English (Garzón, Patiño & Marulanda, 2025; Li et al., 2025; Tan, Cheng & Ling, 2025).

Another limitation concerns the experimental design. Since the experimental condition combined AI tool use with increased learner autonomy caused by diminished teacher guidance, it is difficult to determine the isolated effect of AI tools. Future studies should attempt to control these variables more carefully in order to identify the impact AI technologies have on language learning outcomes.

5. Conclusion

The aim of this study was to determine whether there were differences in the efficiency of English language learning between female students who learned the course material using artificial intelligence tools and female students who learned the same material using traditional learning methods. In order to assess the effect of the application of artificial intelligence, the results of the experimental and control groups on the final test were compared, with the analysis of covariance allowing control of the influence of the initial level of language competence of the subjects.

The results of the study did not show statistically significant differences between the experimental and control groups on the final test, and group affiliation did not prove to be a significant predictor of success on the final test after controlling the initial level of knowledge. Based on the results obtained, it can be concluded that participation in AI-supported learning activities did not result in statistically significant differences in English language achievement compared with traditional teacher-led instruction.

The results obtained should be viewed in the context of the characteristics of the research conducted. The failure to establish statistically significant differences does not necessarily mean that artificial intelligence tools do not have the potential to improve the process of learning foreign languages, but rather indicates the fact that their effectiveness depends on numerous factors, such as the method of integration into the teaching process, the intensity and frequency of use, the quality of teaching activities, and the characteristics of the respondents themselves. It is possible that the duration of the research and the scope of application of AI-supported learning activities were not sufficient to develop effects that could be clearly manifested in the results of the final testing. The relatively small sample of respondents also represents one of the limitations of the research and may affect the ability to detect minor differences between groups. Since all respondents were students of the same study program and similar educational characteristics, the ability to generalise the results to a wider population of students is limited.

Despite the above, the research provides a valuable contribution to the understanding of the application of artificial intelligence tools in English language learning in higher education. The results indicate the need for further research that would include larger and more heterogeneous samples of respondents, a longer period of application of artificial intelligence tools, and more detailed monitoring of various aspects of the learning process. Future research should, in addition to achievement or knowledge tests, also include variables such as motivation for learning, satisfaction with the teaching process, the level of student engagement, and the frequency and manner of using

artificial intelligence tools. Such an approach would enable a more comprehensive understanding of the potential and limitations of artificial intelligence in the field of learning and teaching foreign languages.

The results of this research show that the AI-supported learning approach implemented in this study did not lead to statistically significantly different English language learning outcomes compared with traditional teacher-led instruction. Given the rapid development of artificial intelligence technologies and their increasing presence in education, it is necessary to continue to study their pedagogical capabilities in order to determine the conditions under which they can make the greatest contribution to the quality and efficiency of the learning process.

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.

Conflict of interest statement

The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Agusniati, A., Wahid, A., Nur, R., & Asdar, A. (2025). A Need Analysis MOOC for Based Instructional Material in Flipped English Speaking Class at Indonesian Higher Education. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 10(1), 125. <https://doi.org/10.21093/ijeltal.v10i1.1968>
- Alkhanbooli, B. & Alshamsi, M. K. (2025). The impact of employing AI programs in developing reading skill among first-cycle students. *Iris J of Edu & Res.* 5(4). DOI: 10.33552/IJER.2025.05.000615
- Amaria, F. (2025). The Impact of AI Writing Tools on Academic Writing Skills in English Learning Classroom. *International Journal of Language and Education Research*, 7(3), 159-177.
- Azhar, K. A. (2024). *Artificial intelligence and self-regulated language learning: A systematic review*. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.5658651>

- Budyono, H., Marzuki, Pudjaningsih, W., Prastio, B. & Maulidina, A. (2025). Exploring long-term impact of AI writing tools on independent writing skills: A case study of Indonesian language education students. *International Journal of Information and Education Technology*, 5(15), 1003-1013. doi: 10.18178/ijiet.2025.15.5.2306
- Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated Essay Evaluation: The Criterion Online Writing Service. *AI Magazine*, 25(3), 27. <https://doi.org/10.1609/aimag.v25i3.1774>
- Chingakham, D. S. & Tamuk, K. (2024). Using artificial intelligence (AI) in English vocabulary development for higher secondary school students. *Research Square*, DOI: <https://doi.org/10.21203/rs.3.rs-4500329/v1>
- Consoli, T., Schmitz, ML., Antonietti, C. et al. Quality of technology integration matters: Positive associations with students' behavioral engagement and digital competencies for learning. *Educ Inf Technol* 30, 7719–7752 (2025). <https://doi.org/10.1007/s10639-024-13118-8>
- 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/>
- Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic Review of Artificial Intelligence in Education: Trends, Benefits, and Challenges. *Multimodal Technologies and Interaction*, 9(8), 84. <https://doi.org/10.3390/mti9080084>
- 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.
- Gugić, J., Malenica F., Mustapić Malenica, E. (N.D.). Assessing the Reliability of Short Language Proficiency Tests. *Discover Education*. (In review).
- 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.
- Hingpis Aruta, M. G. & Madeja, J. B. (2026). Exploring the impact of AI chatbots on vocabulary development of teacher education students: a narrative inquiry. *International Research Journal of Multidisciplinary Scope (IRJMS)*, 7(2), 606-615. DOI: 10.47857/irjms.2026.v07i02.09629
- 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>

- 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 Intelligence* 3. <https://doi.org/10.1016/j.caeai.2022.100074>
- Kızıltaş, Y. (2025). Integration of artificial intelligence (AI) into primary school students' writing skills: the impact of ChatGPT on creative writing and writing self-efficacy. *Journal of Educational Computing Research*, 63(7–8), 1717–1747. <https://doi.org/10.1177/07356331251365187>
- Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence*, 9, Article 100445. <https://doi.org/10.1016/j.caeai.2025.100445>
- Parra G., L., & Calero S., X. (2019). Automated Writing Evaluation Tools in the Improvement of the Writing Skill. *International Journal of Instruction*, 12(2), 209-226. <https://doi.org/10.29333/iji.2019.12214a>
- 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.
- Rahmawati, Y., Fitriati, S. W., Widhiyanto, & Wahyuni, S. (2025). A Four-Stranded Approach to AI-Assisted Pre-Class Learning Modules: EFL Teachers' Practices in Designing Daily Conversation Content for Flipped Classrooms. *Language Teaching Research Quarterly*, 49, 90–112. <https://doi.org/10.32038/ltrq.2025.49.05>
- Ramadhani, I. S., Khotimah, K., Putri, R. D. & Arisya, S. (2025). AI tools in improving speaking skills: a literature review. *JITERA – journal in teaching and education area*, 2(3), 349-362. doi: <https://doi.org/10.69673/n7ebwk13>
- 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
- 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.
- Stevens, J. P. (2025). Artificial intelligence in the teaching of English grammar. *International Journal of Research and Innovation in Applied Science (IJRIAS)*, 10(6), 878-896, DOI: <https://doi.org/10.51584/IJRIAS.2025.10060066>
- Suja, A., Muassomah, M., Sutaman, S., Hamid, M. A., & Rosyidi, A. W. (2025). Guided vs. Unguided social media-assisted language learning: An in-depth analysis of speaking performance, willingness to communicate, and social emotional learning. *Australian Journal of Applied ...*, 8(3), 110–121.

- Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*, 8, Article 100355. <https://doi.org/10.1016/j.caeai.2024.100355>
- 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
- Zhang, Z., Wang, C., & Shang, Y. (2025). Are AI tools effective in reducing cognitive load or boosting writing self-efficacy for L2 postgraduates? *The Asia-Pacific Education Researcher*. <https://doi.org/10.1007/s40299-025-01031-8>