

<https://doi.org/10.53656/ped2025-1s.03>

USE OF ARTIFICIAL INTELLIGENCE IN FOREIGN LANGUAGE TEACHING

**Prof. Dr. Ekaterina Sofronieva,
Dr. Christina Beleva, Assist. Prof.,
Dr. Galina Georgieva, Assoc. Prof.**

Sofia University "St. Kliment Ohridski" (Bulgaria)

Abstract. The evolvement and implementation of generative artificial intelligence (AI) models in education calls for clearer outlining of AI benefits and limitations. This original research carried out in summer 2024 at Sofia University "St. Kliment Ohridski", aimed at testing the extent to which AI-generated texts can be used as a pedagogical tool in language instruction. A novel instrument was devised to measure how well students engage with and distinguish between AI-generated and human texts across two literary genres. Altogether, 134 university students from several majors in the field of education took part in the study. Findings revealed that participants overestimated their knowledge of generative AI models as less than half of them managed to identify AI-generated texts correctly. As hypothesized, certain subject variables were meaningfully related to the students' discernment of AI texts. The study also investigated the effectiveness of AI output for students' language skills development.

Keywords: generative artificial intelligence; AI-generated texts; literary fiction; language teaching

Studies on the use of generative AI applications in educational practices have increased abundantly over the last years, reflecting concerns and endeavours to make learning experience more effective, engaging and universally accessible.

The present research aims at gaining insights of the extent to which undergraduate and postgraduate students in the field of language education have acquired the skills to differentiate between AI-generated and human texts, as well as how they perceive the usability of these two text categories for their own progress in language learning. This study offers a novel angle on how university students engage with AI-generated texts across genres.

1. Key aspects of research topic

1.1. AI application in education: benefits, constraints and ethical considerations

The Institute for Ethical AI in Education in the UK provides guidance to the safe and ethical AI use in educational settings and issued in 2021 ‘The Ethical Framework for AI in Education’¹. The document highlights AI potential to mitigate issues related to assessment systems and social immobility. The capacity of AI systems to equip learners with higher control over their own learning is acknowledged. The Framework calls attention to the need for greater awareness of AI societal and ethical implications.

Current AI tools in education fall into two broad categories: rule-based applications, such as intelligent tutoring systems which offer more personalised learning, and machine-based learning applications for automated scoring of student output (Murphy 2019). Intelligent tutoring systems are considered as supportive of students and teachers alike (Akgun & Greenhow 2022) and easily adapted to suit learners’ abilities (Chang & Kidman 2023). Automated scoring systems lessen teachers’ workload (Perin & Lauterbach 2018) and extend their productive capacity (Akgun & Greenhow 2022). In general, teachers have positive attitudes to the integration of AI-based tools in their practice (Bakracheva & Mizova 2024).

A further advantage that discerning use of AI technologies highlights is fostering cross-cultural interactions (Chang & Kidman 2023) and enhancing learning through increased interactivity in human-machine conversations (Chiu 2023). Chatbots are found not only to respond in an increasingly human-like way (Baidoo-Anu & Owusu Ansah 2023) but also to employ a conversational tone (Akgun & Greenhow 2022).

A particularly fruitful consideration for educational practices is the use of generative AI to encourage interdisciplinary teaching (Chiu 2023). This line of research is consistent with the understanding that a properly designed learning activity which incorporates AI technology can encourage reflection (Chang & Kidman 2023) and divergent thinking (McKnight 2021).

The limitations of AI applications in educational context, identified so far, are as varied. Most prominent among them is AI bias which may further widen the educational divide (Pack & Maloney 2024) and digital exclusion¹. Further to that, the output of automated assessment systems may increase bias rates due to algorithms’ interpretative limitations (Perin & Lauterbach 2018). Some studies infer that machine-learning algorithms generate output which is biased in terms of race and gender (Murphy 2019) or societal stereotypes (Akgun & Greenhow 2022).

Another constraint of generative AI models is their inability to reach the level of human-to-human interaction (Chang & Kidman 2023), thus affecting adversely those learners who only thrive in personal connection with teachers (Baidoo-Anu & Owusu Ansah 2023). Other studies indicate that intelligent tutoring systems are less effective than human-to-human educational interaction (Luckin et al. 2022).

Learning-wise, it should be noted that AI applications are more suitable for devising predictable and unvaried tasks (Murphy 2019). More complex tasks,

aimed at encouraging learners' critical thinking skills (Chang & Kidman 2023), require human-designed pedagogical interventions.

In educational practices, tailored to support learner's wellbeing and learning progress, the role of AI technologies remains supplementary. Teaching professionals are advised to use AI applications responsibly (Luckin et al. 2022) and stay aware of their ethical implications (Pack & Maloney 2024). The scope of AI applications in educational contexts may be also impacted by considerations related to student data privacy concerns¹. Both educators and students alike should be provided with training on how to use AI technologies (Chiu 2023) and offered guidance in responsible interactions with AI applications (Chang & Kidman 2023).

1.2. Incorporating AI in foreign language teaching (FLT): specific educational aspects

Post-modern perspective on FLT recognises "contextualised discourse as the locus of language" (Dodigovic 2005, p. 4). This viewpoint becomes increasingly relevant when we consider the impact of the changing technological context on how learners and educators relate to evolving digital tools and perceive themselves as agents (Kennedy & Levy 2009). The learner in the age of AI is regarded as "a being composed of [...] networked human and non-human presence" (McKnight 2021, p. 447).

In line with viewing the learning process as outstepping the boundaries of the purely cognitive or individual (Miller 2012), current studies (Perkins 2023, Godwin-Jones 2022) show that human and AI co-creation in writing tasks will play a greater role in future language classrooms. Thus, AI use in FLT may be viewed from the angle of "collective, participatory, situated and dialogic" (McKnight 2021, p. 452) construction of knowledge. The improved linguistic coherence of output by AI tools, based on large language models, renders them useful for conversational practices (Godwin-Jones 2024) as well. Shifts towards asynchronous learning modalities in post-pandemic times prove AI technologies as contributive to self-directed learning (Baidoo-Anu & Owusu Ansah 2023). Furthermore, AI-assisted error correction in FLT is viewed as a pedagogical technique offered "in a socially non-threatening way" (Dodigovic 2005, p. 2).

Cambridge University Press & Assessment issued in 2023 a paper, titled 'English language education in the era of generative AI: our perspective'² that provides insights on how generative AI applies to English language teaching, learning and assessment. Among the educational benefits of AI technologies the paper points out the personalised and interactive learning experiences that enhance learners' progress, as well as the production of custom-made tests to suit various test-takers' needs. The paper specifically relates responsible use of generative AI to in-depth understanding of AI strengths and limitations.

The role of discerning approaches to AI use in education¹ and FLT in particular² supports the claim for cultivating a better grasp of AI (Luckin et al. 2022). Our

previous study on algorithm literacy and interactions with generative AI models among students of education (Sofronieva et al. 2024) established positive correlation between language skills and awareness of AI models.

Practical constraints of AI use for lower performing learners (McKnight 2021) and less privileged students (Godwin-Jones 2024), social implications of plagiarism (Perkins 2023) and academic integrity (Waltzer, Cox & Heyman 2023) need also be duly considered.

Literary fiction in language classrooms: a brief rationale

If the application of AI technologies is a fairly recent development in FLT, using literary texts to support language learning is a universally incorporated approach. Literary fiction texts provide an engaging contextual framework for language practice. Poetic and narrative genres constitute authentic linguistic resources, enhancing the receptive vocabulary range and transforming it into productive language knowledge (Collie & Slater 1987). Fictional texts introduce learners to diverse language use and promote cultural awareness (Bland 2013).

Discursive and communicative dimensions of using literary texts in FLT (Widowson 2013) are corroborated by the understanding of fiction as an instructional tool which fosters emotive (Nikolajeva 2019), imaginative (Nussbaum 2010) and social (Collier 2021) skills.

1.3. Creative AI vs human creativity in text generation: what makes us human

Generative AI technologies are regarded as fluent in their output (Perkins 2023) and able to produce grammatically and semantically coherent texts of different genres (Köbis & Mossink 2020). The understanding of creativity, traditionally viewed as an inherently human ability, undergoes reassessment given AI's capacity to generate original content (Koivisto & Grassini 2023) and the rapid pace at which its capabilities for creative output evolve (Messingschlager & Appel 2022).

A study comparing human and AI ability for creative thinking (Koivisto & Grassini 2023) suggests that AI outputs are at least at the same level, if not higher, as those of average human beings. Such insights are congruent to findings on whether humans can distinguish between AI and human literary output. Using GPT-2 – a natural language generation algorithm which was a precursor to ChatGPT, N. Köbis and L. Mossink (2020) studied if human poetic texts can be differentiated from algorithm-generated ones, as well as how appealing both text types are. Their findings convincingly show that humans cannot distinguish reliably human from artificial poetry. In terms of preference, human poems were preferred to AI-written ones.

In their turn Gunser et al. (2021), who also used GPT-2 for the generation of both poems and short narratives, demonstrated that AI-written texts cannot always be correctly distinguished by human texts. These findings are particularly intriguing given the professional background in literature of the study participants.

Further to that, research on how information about alleged human or AI authorship of fiction stories impacts reader engagement (Messingschlager & Appel 2022) provides other valuable insights. Respondents reported similar engagement rates with human or AI-generated science fiction stories. Engagement with contemporary fiction stories was stronger with human-written ones.

To conclude this section, findings from a study on the capacity of high-school teachers to differentiate between ChatGPT-generated essays, and student-written ones (Waltzer, Cox & Heyman 2023) show that teachers found it difficult to decide between human or AI authorship in the case of well-written student essays.

2. Research method

2.1. Research questions

The present study tried to address the following research questions:

Q1: Can we use AI-generated texts as an effective pedagogical tool in language teaching and learning?

Q2: What are the students' AI and language self-assessed skills?

Q3: To what extent can students successfully identify AI-generated poetry and narrative texts from human-written ones?

Q4: What are the demographic characteristics which are meaningfully related to the students' abilities to correctly identify AI-generated texts?

2.2. Participants

One hundred and thirty-four students in pedagogy and language teaching at the Faculty of Educational Studies and the Arts of Sofia University "St. Kliment Ohridski" in Bulgaria, participated in the present research. The mean age of the group was 27.23 (S.D. = 9,69; range 19 – 56). Of all students, 64.2 % were under 25 years of age. Tabular presentation of data is offered below.

Table 1. Demographic characteristics of the participants:
gender, major and year of studies

	Response categories	N	%
Gender	Male	8	6.0
	Female	126	94.0
Students' university major	Preschool education and FLT	70	52.2
	Media education and art communication (and FLT)	18	13.4
	Postgraduate qualification for teachers of English	46	34.3
Year of studies	Year 1	27	20.1
	Year 2	28	20.9
	Year 3 postgraduate level	33	24.6
		46	34.3
Total		134	100

2.3. Procedures

The study was conducted in the summer months of 2024 when university students took part on a voluntary basis.

A three-part survey instrument was developed to test respondents' skills in differentiating between human and AI-generated texts, as well as in completing language tasks.

Two contemporary works of literary fiction were selected - the poem 'Hour' by Carol Ann Duffy (Duffy 2005, p. 7) and the short story 'On the train' by Lydia Davis (Davis 2014, p. 144) – as representative of authentic, artistic language use in present day British and US literature. These two fictional texts were selected on the basis of brevity, relevance to participants' proficiency level in English, and topic relevance – both in terms of comprehensibility and pertinence to students' personal socio-emotional experience.

The two AI-generated texts – a poem and a short story – resulted from an uninterrupted interaction with ChatGPT-3.5 that took place on 11-th March 2024 from 10:35 hrs to 10:48 hrs local time. Initial prompts had been considered beforehand, yet during the interaction some additional prompts were entered so that the solicited responses of the AI model would match better the purposes of the research.

2.4. Instruments

A novel instrument was designed to gather information on three sections.

The first part of the survey collected general information about participants' gender, age, university major and year of studies. Moreover, it comprised two questions about students' self-reported AI usage skills and English language fluency. Both skills were assessed on a 5-point Likert scale (1 – basic; 2 – average; 3 – good; 4 – very good; 5 – excellent).

The second part of the instrument focused on the use of poetry in FLT. It comprised two poetry pieces – one written by a human author and the other one generated by AI (see appendix A). There were 5 questions which aimed at finding information about students' preferences, development of language skills, abilities to identify AI-generated poems and confidence in their choices.

The third part of the instrument focused on the use of narratives in FLT. It comprised two “fill in the gaps” language tasks based on two short stories, which differed in origin. One was AI-generated and the other - written by a human author (see appendix B). At the end, there were 4 questions in relation to the narrative texts and the tasks based on them.

All supplementary questions on the instrument about text linguistics and discourse analysis have been excepted from this study since they are part of a separate review and discussion.

3. Research results

The Statistical Package for the Social Sciences (SPSS) 23.0 was used in the analysis of data. There was a rich constellation of findings. Herein below are presented some of the main results, related to the posed research questions in this article.

3.1. AI skills

Students' skills to use AI are illustrated in Figure 1.

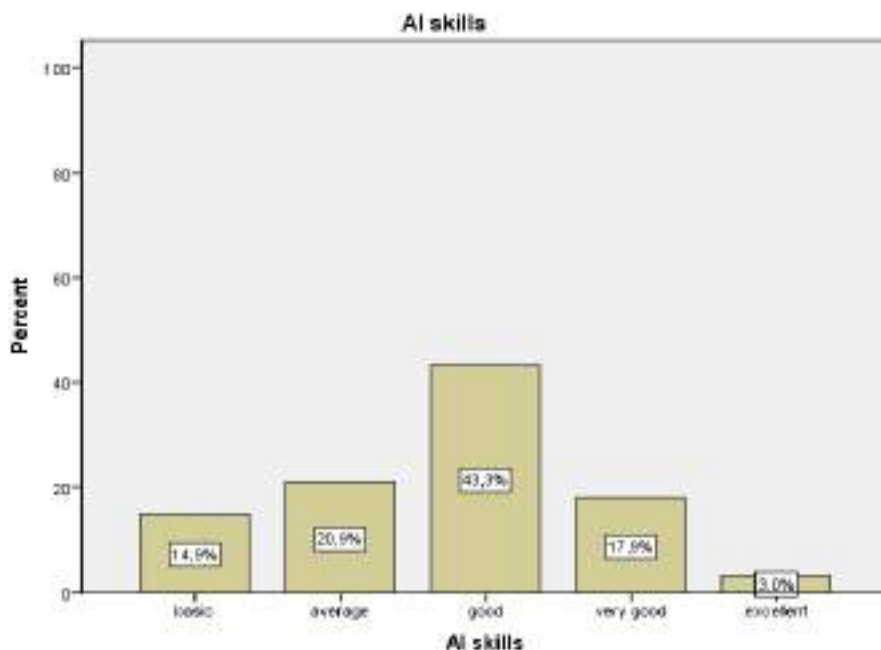


Figure 1. Participants' self-reported skills to use AI

As is seen, university students exhibited a relatively high confidence in their AI skills. The majority assessed them as “good”, and on total 61.2% assessed them as “good” or “very good”. It is also worth noting that only a very small percentage of all students (3%) believed that they had “excellent” AI skills.

3.2. Language skills

As regards students' language skills, results were quite expected. Bearing in mind that all students on the three programmes were being trained for language teachers at preschool and primary school level, they were believed to have a fairly high level of proficiency in English. Findings confirmed that the majority of them assessed their skills as “good” or “very good” (79.9%); 14.2% rated them as “excellent” and only 6% rated them as “average” or “basic”.

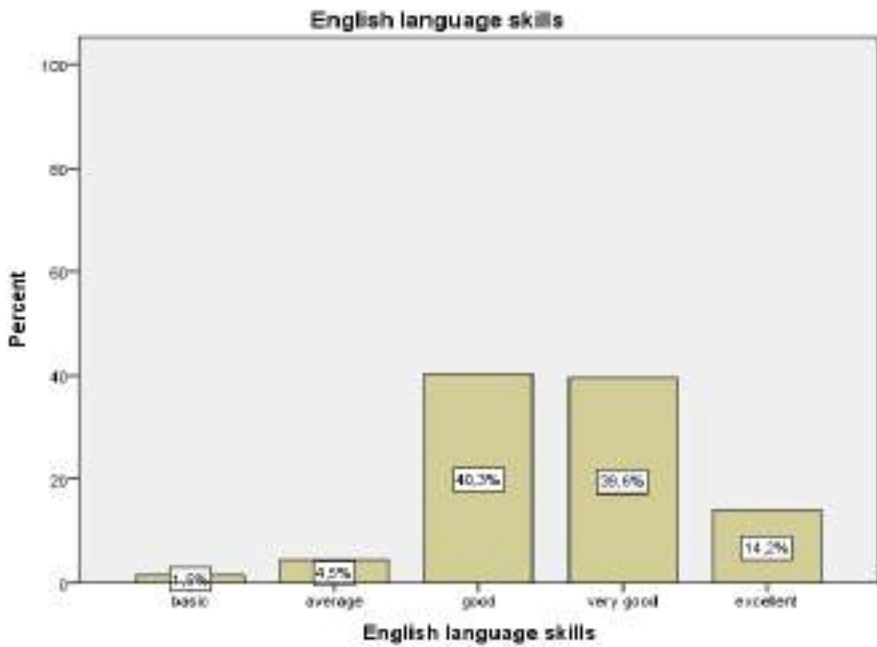


Figure 2. Participants' self-reported language skills

3.3. Use of poetry in FLT

The questions on the second part of the survey instrument, entitled “Artificial or human intelligence: analysis of poems” are listed below.

Which of the two poems:

- affects you more? (Q1);
- is closer to the poetic genre? (Q2);
- contributes to a greater extent to the development of your language skills? (Q3);
- has been generated by artificial intelligence (AI)? (Q4);

To what extent are you confident that you have identified correctly the text, generated by AI? (Q5).

Data on students' answers to Qs 1 – 4 are presented below.

Table 2. Frequency and percent distribution of participants' responses

Q	Poetry piece	Frequency	Percent
Q1 – general effect	Poem A	25	18.7
	Poem B	93	69.4
	Both poems	16	11.9
Q2 – poetic genre	Poem A	24	17.9
	Poem B	106	79.1
	Both poems	4	3.0
Q3 – language skills	Poem A	27	20.1
	Poem B	62	46.3
	Both poems	45	33.6
Q4 – AI-generated poem identification	Poem A	89	66.4
	Poem B (AI poem)	43	32.1
	Neither	2	1.5
Total		134	100

The AI-generated poem was way preferred to the human poem when it came to its total effect on the readers and stylistic closeness to the poetic genre.

Nearly half of the students (46.3%) believed that the poem generated by AI contributed greater to the development of their language skills, whereas only 20.1% stated that it was the human poem they found more useful language-wise. The rest (33.6%) found both poems equally useful.

As is visible, two thirds of all students (67.9%) did not succeed in identifying the AI-generated poem. This was a surprising finding, bearing in mind that over 60% of the students assessed their abilities to use AI from “good” to “excellent” on the self-reported scale.

Finally, we analyzed data gathered from the responses on the last question in this section (Q5). They were rated on a 5-point Likert scale (from 0 to 4).

Results showed that 47% of all students were “very confident” or “confident” that they had correctly identified the poem generated by AI. Only a small percentage (14.2%) were “not confident” or “not confident at all” that they had correctly identified the AI poem. The rest (38.8%) could not assess the degree of their confidence and opted for the “I am not sure” choice of response.

3.3.1. Demographics and group differences

Next, we analyzed differences between the group of students who correctly identified the AI poem and the group who did not in relation to their demographic characteristics, as well as to AI and language skills. Chi-square (χ^2) analyses were conducted to assess these differences. Cramer's V post test was used to identify the strength of the relationship between the set variables (at level of statistical significance $p \leq 0.05$).

No statistically significant differences were found in relation to the ability to properly identify AI generated poems and students' age or university major.

There were no differences between the two groups of students and their self-reported abilities to use AI. In fact, the students (N=4) who evaluated their skills to use AI as "excellent", failed to identify the AI-generated poem. Similarly, 24 students assessed their skills as "very good" but only 8 of them managed to identify the AI poem.

Differences between the two groups were found to be statistically significant in relation to students' gender ($\chi^2 = 4.02$, $df = 1$, Cramer's $V = 0.17$, $p = 0.05$) and their level of language proficiency ($\chi^2 = 11.05$, $df = 4$, Cramer's $V = 0.29$, $p = 0.03$).

Gender

A surprising finding when gender differences were tested, was the fact that all 8 male students gave wrong answers and could not identify the AI-generated poem.

Language skills

Data confirmed our assumption that students with higher command of English would be more likely to differentiate between an AI and a human poem. Thus, the great majority of the students (72.1%) who managed to identify the AI poem, reported that they had "excellent" or "very good" language skills.

3.4. Use of narratives in FLT

This section presents the results found in relation to the third part of the survey instrument, entitled "Artificial or human intelligence: analysis of narratives and tasks based on them".

First, participants were asked to complete two "fill in the gap" language tasks based on two kinds of narratives: a short story by a human author and an AI-generated short story. Then, they were invited to answer the following questions:

Q1: Which of the two tasks did you find more difficult?

Q2: Which of the two tasks did you enjoy doing more?

Q3: The text of which task has been generated by AI?

Q4: To what extent are you confident that you have identified correctly the text, generated by AI?

3.4.1. Task completion

The analysis of participants' correct answers in the language tasks and their responses to Q1-Q3 are given in the two tables below.

As is seen, students generated more correct answers in task 2 which was based on a human short story, and fewer in task 1, based on an AI short narrative.

Congruently, they found task 1 slightly more difficult than task 2 and reported that they took more joy in completing task 2.

Around 40.2 % percent of all participants managed to successfully identify the task based on an AI-generated text.

Table 3. Participants’ correct answers: mean and standard deviation

Task	Mean	Std. Dev.	Range	Minimum	Maximum
Task 1 (AI text)	3.65	2.15	7	0	7
Task 2	4.05	1.84	7	0	7

Table 4. Participants’ responses to Q1 – Q3: frequency and percent distribution

Q	Task based on narratives	Frequency	Valid Percent
Q1 – task difficulty	Task 1	50	37.9
	Task 2	41	31.1
	Both tasks	41	31.1
	Missing data	2	-
Q2 – joy	Task 1	41	30.8
	Task 2	62	46.6
	Both tasks	30	22.6
	Missing data	1	-
Q3 – AI text identification	Task 1 (AI text)	53	40.2
	Task 2	71	53.8
	Both tasks	4	3.0
	Neither tasks	4	3.0
	Missing data	2	-

Finally, Q4 generated answers on a 5-point Likert scale. Only one third of the students (33.3%) were “very confident” or “confident” that they had correctly identified the AI-generated text; 24.3% were “not confident” or “not confident at all” and almost half of the group (42.4%) could not assess the degree of their confidence and opted for the “I am not sure” choice of response. On the whole, participants showed less confidence when identifying AI – generated fiction narratives than poetry pieces. However, in reality students were better at identifying the narratives’ origin.

3.4.2. Demographics

Our next task was to test the differences between the two groups (who identified correctly or wrongly the origin of the narrative texts) in relation to their demographic factors.

The χ^2 analysis revealed no statistically significant group differences in relation to students’ gender or language proficiency level.

Group differences were found significant for students’ age, university major and self-reported level of AI skills. For age (under or above 25 years) $\chi^2 = 3.79$, $df = 1$, Cramer’s $V = 0.17$, $p = 0.05$; for major $\chi^2 = 10.57$, $df = 2$, Cramer’s $V = 0.28$, $p = 0.01$; and for AI skills $\chi^2 = 15.87$, $df = 4$, Cramer’s $V = 0.35$, $p = 0.00$.

Age

As anticipated, younger students had greater interest and experience in AI models compared to older professionals. The mean age of the group (N=53) who identified the AI text was 24.70 years (S. D. = 6.3) and the mean age of the group who failed in this (N=79) was 29.06 (S. D. = 11.1). Altogether, 73.6% in the group who identified the AI text were students under 25 years.

University major

Students of preschool education (67.9%) predominated in the group who gave correct answers. In the group who failed, the older students on the post graduate qualification programme predominated. They constituted 40.5% of that group. In this light, findings in relation to students' age and university major may be interrelated and interpreted as coexisting.

AI skills

About two thirds of the students (69.8%) who identified the AI-generated short story had "good" or "very good" AI skills. Students (72.2%) who had "average" or "good" skills predominated in the other group. It is interesting, that all students who self-assessed their AI usage skills as being "excellent" (3% of all students) failed to identify the origin of the texts.

4. Conclusions and implications for future studies

The study on language learners' engagement with and discernment between AI-generated and human texts of different genres contributed to educational and research endeavours, aimed at pinpointing efficient ways of incorporating AI technologies in education.

As regards correct identification of AI texts, findings of this research are fully consistent with results from previous studies (Gunser et al. 2021, Köbis & Mossink 2020), indicating that humans find it difficult to discern between AI and human literary output. As expected, all studied demographic factors were related to participants' abilities to identify the origin of literary works.

Furthermore, the study established that students overrate their knowledge and awareness of AI models as only 32.1 % of all identified correctly the AI poem and 40.2% - the AI short story. This result is in line with findings on overconfidence in algorithm understanding (Köbis & Mossink 2020), as well as with the notion that previous experience with AI is not related to the ability to distinguish AI texts correctly (Waltzer, Cox & Heyman 2023).

In terms of engagement with AI and human texts, participants' preferences were stronger for the AI poem and for the human short story. This finding is consistent with the observation that humans engage differently with AI-generated texts across genres (Messingschlager & Appel 2022).

Altogether, the study confirmed that AI-generated texts can be successfully used as a pedagogical tool in FLT. Students found both the poetic and narrative texts and

the task, based on the narrative text, useful for the development of their language skills. We fully support suggestions that AI applications should be used responsibly and ethically in teaching practices (Pack & Maloney 2024, Luckin et al. 2022) as supplementary educational tools (Chang & Kidman 2023).

We also endorse the proposed merit of further studies on textual level differences across genres (Messingschlager & Appel 2022) between human and AI texts.

Appendix A

Artificial or human intelligence: analysis of poems

POEM A	POEM B
<p>Love's time's beggar, but even a single hour, bright as a dropped coin, makes love rich. We find an hour together, spend it not on flowers or wine, but the whole of the summer sky and a grass ditch. For thousands of seconds we kiss; your hair like treasure on the ground; the Midas light turning your limbs to gold. Time slows, for here we are millionaires, backhanding the night so nothing dark will end our shining hour, no jewel hold a candle to the cuckoo spit hung from the blade of grass at your ear, no chandelier or spotlight see you better lit than here. Now. Time hates love, wants love poor, but love spins gold, gold, gold from straw.</p>	<p>In the garden of youth, they find their bloom Two hearts entwined, in love's tender room His laughter, a melody, her smile, a dance Their souls intertwined in a passionate trance In whispers of twilight, they share their dreams Underneath the moon's soft gleam Hand in hand, they stroll the night Wrapped in each other's gentle light. Their love, a tapestry woven with care In the quiet moments, they find solace there Through stormy seas and calmest tides They navigate life's winding rides For youth may fade and time may flee But their love, a beacon, eternally In the garden of youth, they find their grace Two souls ablaze, in love's sweet embrace.</p>

Notes (this information was not disclosed to study participants):

* Poem A – 'Hour' by Carol Ann Duffy (Duffy 2005, p. 7)

** Poem B – Poem, generated by ChatGPT3.5

Appendix B

Artificial or human intelligence: analysis of narratives and tasks

<p>Read carefully the two short text tasks. An example is given at the beginning of each task. Fill in the blanks from (1) to (7) with ONE suitable word.</p>	
<p>TASK I</p> <p>On the train, I sit (0)...by.... the window, watching the landscape blur past. (1)..... passing moment is a flicker of movement, a tableau (2)..... transient lives. A man in (3)..... faded cap is lost in a paperback. A child with tousled hair presses her nose against the glass, eyes wide (4)..... wonder. The rhythmic clatter of (5)..... on tracks forms a backdrop to my thoughts as I ponder the fleeting nature of human connection. Across from me, an elderly (6)..... holds hands, their weathered fingers intertwined, speaking volumes in silence. In this fleeting encounter, I glimpse the beauty of transience, knowing that every journey, (7)..... life itself, is but a passing moment in time.</p>	<p>TASK II</p> <p>We (0)...are... united, he and I, though strangers, (1)..... the two women in front of us talking so steadily and audibly (2)..... the aisle to each other. Bad manners. Later in (3)..... journey I look over at him (across the aisle) and he is picking (4) nose. (5)..... for me, I am dripping tomato from my sandwich on to my newspaper. Bad habits. I (6)..... not report this if I were the one picking my nose. I look again and he is still at it. As for the women, they are now sitting together side by side and quietly (7)....., clean and tidy, one a magazine, one a book. Blameless.</p>

Notes (this information was not disclosed to study participants):

* Task A – short story, generated by ChatGPT3.5

** Task B – ‘On the train’ by Lydia Davis (Davis 2014, p. 144)

Correct answers:

Task 1 – Each, of, a, with, wheels, couple, like

Task 2 – against, across, the, his, As, would, reading

NOTES

1. THE INSTITUTE FOR ETHICAL AI IN EDUCATION, 2021. Final Report: The Ethical Framework for AI in Education. Available from: <https://www.buckingham.ac.uk/wp-content/uploads/2020/02/The-Institute-for-Ethical-AI-in-Education-Interim-Report-Towards-a-Shared-Vision-of-Ethical-AI-in-%20Education.pdf> [viewed 2024-06-03].

2. CAMBRIDGE UNIVERSITY PRESS & ASSESSMENT, ENGLISH RESEARCH GROUP, 2023. English language education in the era of generative AI: our perspective. Available from: <https://www.cambridgeenglish.org/english-research-group/generative-ai-for-english/> [viewed 2024-05-27].

Acknowledgements and Funding

This study is financed by the European Union – NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project SUMMIT № BG-RRP-2.004-0008-3.3.

REFERENCES

- AKGUN, S.; GREENHOW, C., 2022. Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI Ethics*, vol. 2, pp. 431–440 [viewed 25 January 2024]. Available from: <https://doi.org/10.1007/s43681-021-00096-7>.
- BAIDOO-ANU, D.; OWUSU ANSAH, L., 2023. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI*, vol. 7, no. 1, pp. 52 – 62 [viewed 27 October 2023]. Available from: <https://dergipark.org.tr/en/pub/jai/issue/77844/1337500>.
- BAKRACHEVA, M.; MIZOVA, B., 2024. Personality predictors of teachers' attitudes towards digital tools and artificial intelligence, *EDULEARN24 Proceedings*, pp. 1973 – 1979. ISSN 2340 –1117.
- BLAND, J., 2013. *Children's Literature and Learner Empowerment: Children and Teenagers in English Language Education*. London: Bloomsbury Academic. ISBN: 978-1-441-14441-6
- CHANG, C. H.; KIDMAN, G., 2023. The rise of generative artificial intelligence (AI) language models – challenges and opportunities for geographical and environmental education. *International Research in Geographical and Environmental Education*, vol. 32, no. 2, pp. 85 – 89 [viewed 25 May 2024]. Available from: <https://doi.org/10.1080/10382046.2023.2194036>.
- CHIU, T. K. F., 2023. The impact of Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. *Interactive Learning Environments*, pp. 1 – 17 [viewed 10 May 2024]. Available from: <https://doi.org/10.1080/10494820.2023.2253861>.

- COLLIE, J.; SLATER, S., 1987. *Literature in the Language Classroom: A Resource Book of Ideas and Activities*. Cambridge: Cambridge University Press. ISBN: 978-0-521-31224-0.
- COLLIER, P., 2021. *Teaching Literature in the Real World: A Practical Guide*. London: Bloomsbury Publishing Plc. ISBN: 9781350195066.
- DAVIS, L., 2014. *Can't and Won't*. New York: Farrar, Straus and Giroux. ISBN 978-0-374-71143-6 (ebook).
- DODIGOVIC, M., 2005. *Artificial Intelligence in Second Language Learning: Raising Error Awareness*. Clevedon: Multilingual Matters Ltd. ISBN 978-1853598302.
- DUFFY, C.A., 2005. *Rapture*. London: Picador. ISBN 978-0-330-43391-4.
- GODWIN-JONES, R., 2024. Distributed agency in second language learning and teaching through generative AI. *Language Learning & Technology*, vol. 28, no. 2, pp. 5 – 31 [viewed 30 July 2024]. Available from: <https://hdl.handle.net/10125/73570>.
- GUNSER, V. E., et al., 2021. Can Users Distinguish Narrative Texts Written by an Artificial Intelligence Writing Tool from Purely Human Text? *International Conference on Human-Computer Interaction*, pp. 520 – 527 [viewed 17 March 2024]. Available from: https://doi.org/10.1007/978-3-030-78635-9_67.
- KENNEDY, C.; LEVY, M., 2009. Sustainability and computer-assisted language learning: factors for success in a context of change. *Computer Assisted Language Learning*, vol. 22, no. 5, pp. 445 – 463 [viewed 27 June 2024]. Available from: <https://doi.org/10.1080/09588220903345218>.
- KÖBIS, N.; MOSSINK, L. D., 2020. Artificial Intelligence versus Maya Angelou: Experimental Evidence that People Cannot Differentiate AI-generated from Human-written Poetry. *Computers in Human Behavior*, vol. 114, article ID 106553 [viewed 2 August 2024]. Available from: <https://doi.org/10.1016/j.chb.2020.106553>.
- KOIVISTO, M.; GRASSINI, S., 2023. Best humans still outperform artificial intelligence in a creative divergent thinking task. *Scientific Reports*, vol. 13, no. 13601 [viewed 30 June 2024]. Available from: <https://doi.org/10.1038/s41598-023-40858-3>.
- LUCKIN, R., et al., 2022. Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, vol. 3, no. 100076 [viewed 27 May 2024]. Available from: <https://doi.org/10.1016/j.caeai.2022.100076>.
- MCKNIGHT, L., 2021. Electric Sheep? Humans, Robots, Artificial Intelligence, and the Future of Writing. *Changing English*, vol. 28, no. 4, pp. 442 – 455 [viewed 4 June 2024]. Available from: <https://doi.org/10.1080/1358684X.2021.1941768>.

- MESSINGSCHLAGER, T. V.; APPEL, M., 2022. Creative Artificial Intelligence and Narrative Transportation. *Psychology of Aesthetics, Creativity and the Arts*, vol. 18, no. 5, pp. 848 – 857 [viewed 18 February 2024]. Available from: <http://dx.doi.org/10.1037/aca0000495>.
- MILLER, E., 2012. Agency, Language Learning and Multilingual Spaces. *Multilingua – Journal of Cross-Cultural and Interlanguage Communication*, vol. 31, no. 4, pp. 441 – 468 [viewed 11 July 2024]. Available from: <https://doi.org/10.1515/multi-2012-0020>.
- MURPHY, R., 2019. *Artificial Intelligence Applications to Support K-12 Teachers and Teaching: A Review of Promising Applications, Challenges, and Risks*. Santa Monica, CA: RAND Corporation [viewed 12 June 2024]. Available from: <https://doi.org/10.7249/PE315> <https://www.rand.org/pubs/perspectives/PE315.html>.
- NIKOLAJEVA, M., 2019. What is it Like to be a Child? Childness in the Age of Neuroscience. *Children's Literature in Education*, vol. 50, pp. 23 – 37 [viewed 5 April 2024]. Available from: <https://doi.org/10.1007/s10583-018-9373-7>.
- NUSSBAUM, M., 2010. *Not for Profit: Why Democracy Needs the Humanities*. Princeton: Princeton University Press. ISBN 978-0-691-14064-3.
- PACK, A.; MALONEY, J., 2024. Using Artificial Intelligence in TESOL: Some Ethical and Pedagogical Considerations. *TESOL Quarterly*, vol. 58, no. 2, pp. 1007 – 1018 [viewed 3 August 2024]. Available from: <https://doi.org/10.1002/tesq.3320>.
- PERIN, D.; LAUTERBACH, M., 2018. Assessing text-based writing of low-skilled college students. *International Journal of Artificial Intelligence in Education*, vol. 28, no. 1, pp. 56–78 [viewed 2 August 2024]. Available from: <https://doi.org/10.1007/s40593-016-0122-z>.
- PEYTCHEVA-FORSYTH, R.; ALEKSIEVA, L.; YOVKOVA, B., 2018. The impact of prior experience of e-learning and e-assessment on students' and teachers' approaches to the use of a student authentication and authorship checking system. In: L. CHOVA, A. MARTINEZ & I. TORRES (Eds.). *EDULEARN18 Proceedings*, pp. 2311 – 2321. Valencia: IATED Academy. ISBN 978-84-09-02709-5.
- PERKINS, M., 2023. Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, vol. 20, no. 2 [viewed 30 May 2024]. Available from: <https://doi.org/10.53761/1.20.02.07>.

- SOFRONIEVA, E., et al., 2024. Artificial Intelligence, Algorithm Literacy, Locus of Control, and English Language Skills: A Study among Bulgarian Students in Education. *Pedagogika–Pedagogy*, vol. 96, no. 5, pp. 579–599 [viewed 3 June 2024]. Available from: <https://doi.org/10.53656/ped2024-5.01>.
- WALTZER, T.; COX, R.; HEYMAN, G., 2023. Testing the Ability of Teachers and Students to Differentiate between Essays Generated by ChatGPT and High School Students. *Human Behavior and Emerging Technologies*, vol. 2023, no. 923981 [viewed 1 June 2024]. Available from: <https://doi.org/10.1155/2023/1923981>.
- WIDDOWSON, H.G., 2013. *Stylistics and the Teaching of Literature*. Abingdon: Routledge. ISBN: 978-0-582-55076-6.

✉ **Prof. Dr. Ekaterina Sofronieva**

ORCID iD: 0000-0002-1774-6197

✉ **Dr. Christina Beleva, Assist. Prof.**

ORCID iD: 0000-0002-8818-1787

✉ **Dr. Galina Georgieva, Assoc. Prof.**

ORCID iD: 0000-0001-9493-7336

Sofia University “St. Kliment Ohridski”

Sofia, Bulgaria

E-mail: e.sofronieva@fppse.uni-sofia.bg

E-mail: hnbeleva@uni-sofia.bg

E-mail: g.georgieva@fppse.uni-sofia.bg