



(REVIEW ARTICLE)



AI-ASSISTED CO-CREATIVE MULTIMODAL STORYTELLING IN ENGLISH: ENGAGEMENT, MOTIVATION AND VOCABULARY USE IN A SIXTH-GRADE PRIMARY SCHOOL INTERVENTION

Evangelia Tzagkourni ¹ and Christos Simos ^{2,*}

¹ Department of Preschool Education, University of Crete, Greece.

² Faculty of Social Theology and Religious Studies, National and Kapodistrian University of Athens; Department of Greek Philology, Democritus University of Thrace, Greece.

* Corresponding Author

Global Journal of Engineering and Technology Advances, 2026, 29(01), 121–125

Publication history: Received on 29 August 2026; revised on 02 October 2026; accepted on 05 October 2026

Article DOI: <https://doi.org/10.30574/gjeta.2026.29.1.0281>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

This paper presents an educational intervention of artificial intelligence (AI)-assisted co-creative multimodal storytelling, implemented in the English course with 48 sixth-grade students at a public primary school. Students composed short descriptive paragraphs about imaginary creatures, received narrative expansions of their texts from an AI tool, and created multimodal artworks that incorporated words and phrases from these expansions. The evaluation of the intervention was based on structured observation, organized around four dimensions of motivation. The records indicate increased engagement and creative initiative, more intensive peer collaboration and the spontaneous use of new descriptive vocabulary in students' multimodal products and oral presentations. As no assessment of written production was conducted before and after the intervention, the findings are limited to observed engagement and language use.

Keywords: Artificial intelligence, Multimodal learning, Motivation, Engagement, English vocabulary

1 INTRODUCTION

The development of writing skills in the teaching of English as a Foreign Language (EFL) remains a significant pedagogical challenge in primary education (Shin & Crandall, 2014). In Greek schools, although English instruction begins in kindergarten, students in the upper grades often engage only to a limited extent with writing activities and draw on a limited vocabulary in their written production (Dendrinos, 2013).

The present paper describes and evaluates an educational intervention that combines two fields rarely brought together within a single framework in the Greek public primary school, namely co-creative storytelling with large language models (LLMs) and multimodal meaning-making. The aim of the intervention was to use AI not as a tool for correction or automation, but as a supportive creative partner. Its evaluation focuses on two questions: how students' observed engagement and motivation change across the successive phases, and how students use the vocabulary of the AI expansions in their multimodal products and in their oral language.

1.1 Research Framework

AI techniques have a long history in education, notably in student modeling (Drigas et al., 2009; Pappas & Drigas, 2015), and AI-driven agents such as social robots have been used to foster metacognitive skills and inclusion (Mitsea et al., 2020). More recently, the literature has highlighted that LLMs can function as interactive partners in language learning, broadening lexical input and supporting motivation to write (Bonner et al., 2023; Darmawansah et al., 2024; Liu et al., 2024; Wang, 2025). More specifically, co-creative storytelling systems, in which children and AI construct narratives collaboratively, such as StoryBuddy (Zhang et al., 2022), Mathemyths (Zhang et al., 2024) and Tinker Tales (Choi et al., 2026), show that AI can provide open-ended questions and text expansions without undermining children's agency and creative "ownership", especially when the interaction design is developmentally appropriate (Kurian, 2025).

Multimodal learning, grounded in social semiotic theory (Kress, 2010), recognizes that meaning-making takes place through the orchestration of multiple semiotic resources (Jewitt et al., 2016). The systematic review by Lim et al. (2022) highlights the contribution of multimodal pedagogies to the language classroom, while in primary education these approaches appear to broaden meaning-making opportunities through the variety of stimuli they offer (Kulju et al., 2018). More specifically, linking visual art creation with language instruction (Thulasivanthana, 2020) allows students to connect words with images (Chung, 2023): when they translate textual descriptions into visual representations and incorporate written language into their artworks, working with the concrete supports access to the abstract, enhancing vocabulary acquisition (Mahmoudi & Mahmoudi, 2012). This process also draws on mental imagery, the support of which through digital applications has been associated with improved learning outcomes, including for learners with learning difficulties (Drigas et al., 2020; Drigas & Kostas, 2014;).

2 THE EDUCATIONAL INTERVENTION

2.1 Context and Participants

The intervention was implemented with 48 sixth-grade students (aged 11–12) at a primary school in Athens, within Unit 3 ("Imaginary Creatures") of the official school coursebook, which focuses on the use of adjectives for descriptive writing. The class teacher designed and implemented the intervention and also carried out the observation. Informed parental consent and student assent were obtained, with the explicit option to withdraw without consequences; no personal student data were entered into the AI tool, and all data were recorded anonymously.

2.2 Description of the Intervention

The intervention consisted of three successive phases implemented over four teaching periods.

- **Phase 1: Initial writing activity.** Students were asked to write, individually, a short paragraph of up to 150 words describing an imaginary creature of their own invention. In line with the unit's thematic focus on adjectives, they were encouraged to provide as many descriptive details about the creature's appearance as possible.
- **Phase 2: AI-generated expansion.** The students' texts were entered by the teacher, without any personal data, into an LLM tool (OpenAI ChatGPT, GPT-3.5). Students had no direct access to the tool. Standardized prompts requested the expansion of each narrative with regard to the creature's origin, its extended appearance, its habitat, its personality traits and the language it speaks. The responses were checked by the teacher for language level and child-appropriateness before being distributed.
- **Phase 3: Multimodal production.** Students studied the expanded descriptions and created artworks depicting the creature and its environment, incorporating words and phrases from the AI text as collage elements. This phase required students to understand the textual information and "translate" it into visual representations, while simultaneously handling elements of written language. The presentation of the artworks in class concluded the intervention.

2.3 Evaluation of the Intervention

The evaluation was based on structured classroom observation across all phases, with field notes recorded after each session. The observation focused on engagement behaviors, collaborative interactions, the spontaneous use of vocabulary from the AI texts and the extent to which textual and visual elements were integrated into the final products.

To organize the observation, an observation framework of motivational indicators was developed, drawing on Dörnyei's (2001) framework, across four dimensions (Table 1). These indicators were not used to score individual students, and no scores were recorded per student or per session; instead, they served as focal axes for class-level observation during each phase.

Table 1: Observation framework of motivational indicators

Dimension	Observable indicators
Task Engagement	Autonomous initiation of work, duration of focus, need for reminders
Creative Initiative	Voluntary extension of work beyond the minimum requirements, original ideas
Collaborative Orientation	Spontaneous sharing of ideas, peer feedback
Linguistic Exploration	Use of new vocabulary, experimentation with descriptive language, incorporation of the AI text

The field notes were analyzed through thematic analysis procedures (Braun & Clarke, 2006) by the teacher-researcher. Initial codes were developed inductively and subsequently examined in relation to the four dimensions of the observation framework. As no numerical records exist, no distributions or mean values are reported. The findings that follow are descriptive in nature and derive from a single observer, a limitation discussed below.

3 EVALUATION AND FINDINGS

Three themes concerning engagement and language use during the intervention emerged from the analysis of the observations.

First, regarding vocabulary use, students selected specific adjectives and descriptive phrases from the AI text to include in their collages, demonstrating deliberate selection rather than passive copying (field notes: “S7 underlines the word *glowing* in the AI text, cuts it out, places it next to the creature's eyes in his collage and asks: ‘Miss, is it OK if I use two more words from here?’”). It was also noted that, during the oral presentation of their artworks, several students used lexical items that did not appear in their initial paragraphs but were present in the AI expansions (field notes: “S22, whose initial paragraph was three lines long, describes orally where her creature lives and why it chooses to stay there based on its appearance — elements that were not in her original text”). The oral descriptions in Phase 3 were richer in detail than the initial written paragraphs. However, since these are different “final products” and different modes of expression, this observation does not substantiate an improvement in written production.

In addition, regarding engagement, the Phase 1 notes are dominated by records of intermittent focus and repeated reminders to keep attention on the writing activity, whereas in Phases 2 and 3 the corresponding records mainly concern autonomous initiation and sustained work (field notes: “Two students discuss for two minutes where to place the phrase lives in the frozen caves and end up cutting it in two, on either side of the cave”). Regarding creative initiative, while in Phase 1 students generally adhered to the minimum requirements, in Phase 3 repeated instances were recorded of voluntary elaboration of the artworks, requests for additional information and experimentation with different arrangements of the textual elements.

Finally, regarding collaborative orientation and linguistic exploration, Phases 2 and 3 were accompanied by spontaneous sharing. Students showed the AI descriptions to their classmates, compared creature profiles and suggested ideas for each other's artworks. At the same time, whereas in Phase 1 they relied mainly on familiar, high-frequency vocabulary, after their exposure to the AI texts instances were recorded of experimentation with less familiar words and more complex descriptive structures, through tactile and spatial interaction with written language in the collage.

4 DISCUSSION AND PEDAGOGICAL IMPLICATIONS

The records indicate that AI-assisted co-creative multimodal storytelling can support engagement, motivation and the use of new vocabulary in the English course, without, however, substantiating an improvement in writing performance, which was not assessed. The observed use of descriptive vocabulary is consistent with research showing that LLMs provide rich and contextually appropriate linguistic input (Bonner et al., 2023), while the physical handling of textual elements through the collage added a tactile dimension to their processing, in line with the principles of multimodal learning (Kress, 2010; Jewitt et al., 2016). Likewise, the shift from intermittent focus in Phase 1 to self-initiated, sustained work in the later phases can be read in light of models of attention that treat sustained and selective attention as prerequisites for learning (Drigas & Karyotaki, 2019, Drigas & Karyotaki, 2019b; Karyotaki & Drigas, 2015).

The positive patterns across all four motivational dimensions are consistent with Dörnyei's (2001) framework, which recognizes the novelty of the activity, autonomy and opportunities for creative expression as key motivational factors. Playful, hands-on activities have likewise been associated with increased engagement even among very young learners

and learners with special educational needs (Kokkalia et al., 2016). At a pedagogical level, the intervention shows that AI can be integrated into the public school as a collaborative support tool within a scenario that remains student-centered and hands-on, with the teacher mediating access to the tool and checking the appropriateness of the generated content.

The findings require careful interpretation. The increased engagement cannot be attributed specifically to AI, as the medium, the nature of the activity and the opportunities for social interaction changed simultaneously between the phases. The phases are therefore not directly comparable, and the findings concern the intervention as an integrated combination of elements.

4.1 Limitations

The evaluation concerns a single grade in one school, without a control group, and was based on a single observer, who was also the person implementing the intervention. Moreover, there was no second observer and no independent check of the coding. The motivational indicators were not recorded as numerical values per student; therefore, it is not possible to report distributions or mean values per dimension and phase. Finally, no writing activity was conducted after the intervention, so no conclusions are drawn regarding writing performance. Future implementations would benefit from pre- and post-intervention assessment of writing using a common rubric, and from systematic scoring of the motivational indicators per student using an analytic scale.

5 CONCLUSIONS

The intervention shows that AI-assisted co-creative storytelling, combined with hand-made multimodal production, can serve as a feasible teaching scenario for the English course in primary school. The records indicate increased engagement and creative initiative, more intensive collaboration and the spontaneous use of new vocabulary, highlighting AI as a potentially useful creative partner rather than a correction tool. Given the limitations of the evaluation, the approach is proposed as a direction for further testing and more systematic evaluation.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The Authors would like to thank the SPECIALIZATION IN ICTs AND SPECIAL EDUCATION: PSYCHOPEDAGOGY OF INCLUSION Postgraduate studies Team, for their support.

Disclosure of conflict of interest

The Authors proclaim no conflict of interest.

REFERENCES

- [1] Bonner, E., Lege, R., & Frazier, E. (2023). Large language model based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology*, 23(1), 23–41.
- [2] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- [3] Choi, N., Hong, J., Cyebukayire, P., Choi, I., & Choi, J. D. (2026). Tinker Tales: A tangible dialogue system for child–AI co-creative storytelling. In *Proceedings of the 27th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (pp. 25–50). Association for Computational Linguistics. <https://aclanthology.org/2026.sigdial-1.3/>
- [4] Chung, D. T. K. (2023). The efficacy of visual aids in enhancing vocabulary acquisition in EFL classes. *International Journal of TESOL & Education*, 3(4), 1–15.
- [5] Darmawansah, D., Hwang, G. J., Chen, M. R. A., & Liang, J. C. (2024). Generative AI and EFL learning: A systematic review of literature. *Smart Learning Environments*, 11, 35.
- [6] Dendrinis, B. (2013). Testing and teaching mediation. *Glossologia*, 21, 75–96.
- [7] Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge University Press.
- [8] Drigas, A. S., Argyri, K., & Vrettaros, J. (2009). Decade review (1999–2009): Artificial intelligence techniques in student modeling. In *World Summit on Knowledge Society* (pp. 552–564). Springer.
- [9] Drigas, A., Dede, D. E., & Dedes, S. (2020). Mobile and other applications for mental imagery to improve learning disabilities and mental health. *International Journal of Computer Science Issues (IJCSI)*, 17(4), 18–23.

- [10] Drigas, A., & Karyotaki, M. (2019). Attention and its role: Theories and models. *International Journal of Emerging Technologies in Learning (ijET)*, 14(12), 169–182.
- [11] Drigas, A. S., & Karyotaki, M. (2019b). A layered model of human consciousness. *International Journal of Recent Contributions from Engineering, Science & IT (ijES)*, 7(3), 41–50. <https://doi.org/10.3991/ijes.v7i3.11117>
- [12] Drigas, A., & Kostas, I. (2014). On line and other ICTs applications for teaching math in special education. *International Journal of Recent Contributions from Engineering, Science & IT (ijES)*, 2(4), 46. <http://dx.doi.org/10.3991/ijes.v2i4.4204>
- [13] Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.
- [14] Karyotaki, M., & Drigas, A. (2015). Online and other ICT applications for cognitive training and assessment. *International Journal of Online Engineering (ijOE)*, 11(2), 36–42.
- [15] Kokkalia, G., Drigas, A., & Economou, A. (2016). The role of games in special preschool education. *International Journal of Emerging Technologies in Learning (ijET)*, 11(12), 30–35.
- [16] Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- [17] Kulju, P., Kauppinen, M., Kankkunen, T., & Laukka, S. (2018). A systematic review of multimodal pedagogies in primary classrooms. *Education Inquiry*, 9(2), 195–216. <https://doi.org/10.1080/20004508.2017.1394148>
- [18] Kurian, N. (2025). Once upon an AI: Six scaffolds for child–AI interaction design, inspired by Disney. *International Journal of Child-Computer Interaction*, 46, Article 100788. <https://doi.org/10.1016/j.ijcci.2025.100788>
- [19] Lim, F. V., Toh, W., & Nguyen, T. T. H. (2022). Multimodality in the English language classroom: A systematic review of literature. *Linguistics and Education*, 69, Article 101048. <https://doi.org/10.1016/j.linged.2022.101048>
- [20] Liu, Z. M., Hwang, G. J., Chen, C. Q., Chen, X. D., & Ye, X. D. (2024). Integrating large language models into EFL writing instruction: Effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2024.2389923>
- [21] Mahmoudi, S., & Mahmoudi, A. (2012). An evaluation of the effects of art on vocabulary learning of EFL learners. *Procedia: Social and Behavioral Sciences*, 47, 1817–1821. <https://doi.org/10.1016/j.sbspro.2012.06.906>
- [22] Mitsea, E., Lytra, N., Akrivopoulou, A., & Drigas, A. (2020). Metacognition, mindfulness and robots for autism inclusion. *International Journal of Recent Contributions from Engineering, Science & IT (ijES)*, 8(2), 4–20.
- [23] Pappas, M. A., & Drigas, A. S. (2015). ICT based screening tools and etiology of dyscalculia. *International Journal of Engineering Pedagogy (ijEP)*, 3, 61–66.
- [24] Shin, J. K., & Crandall, J. A. (2014). *Teaching young learners English: From theory to practice*. National Geographic Learning.
- [25] Thulasivanthana, U. (2020). Incorporating visual arts into English language teaching. *Shanlax International Journal of English*, 8(4), 52–56. <https://doi.org/10.34293/english.v8i4.3320>
- [26] Wang, Q. (2025). EFL learners' motivation and acceptance of using large language models in English academic writing: An extension of the UTAUT model. *Frontiers in Psychology*, 15, Article 1514545. <https://doi.org/10.3389/fpsyg.2024.1514545>
- [27] Zhang, C., Jeon, S., Liu, X., Yu, C. L., Ziska, K., & Xu, Y. (2024). Mathemyths: Leveraging large language models to teach mathematical language through child–AI co-creative storytelling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Article 128). ACM. <https://doi.org/10.1145/3613904.3642647>
- [28] Zhang, Z., Xu, Y., Wang, Y., Yao, B., Ritchie, D., Wu, T., Yu, M., Wang, D., & Li, T. J. J. (2022). StoryBuddy: A human–AI collaborative chatbot for parent–child interactive storytelling with flexible parental involvement. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (Article 218). ACM. <https://doi.org/10.1145/3491102.3517479>