Content list available at http://ijltr.urmia.ac.ir

Iranian Journal of Language Teaching Research

ORIGINAL ARTICLE

Urmia University

AI-powered Image and Audio Generators for Very Young EFL Learners

Gulnara Sadykova [a], Albina Kayumova [a, *]

[a] *Kazan Federal University, Russia*

ABSTRACT

This paper focuses on AI-powered image and audio generators as tools for developing language skills of 4–6-year-old EFL learners. Drawing on the concept of digital screen mediation (Meskill, 2021, 2024) rooted in Vygotskian sociocultural paradigm, this mixed-method study explores young learners' experiences interacting with AI technologies mediated by teachers in preschool settings. Pre- and post-tests, questionnaires, reflection journals, interviews, and video-recorded classes involving 40 children speak in favor of AI-powered technologies as instructional tools that engage learners in highly motivating co-production activities. The findings suggest that teachers retain their leading role as mediators of knowledge who orchestrate conversations with, through and around AI tools, capitalize on young learners' feeling of pleasure when interacting with technologies, and counteract the unpredictability and instability that come with AI. The study shows real cases of successful AI-enhanced language classrooms, which might be of interest to early childhood educators and curriculum developers working in a bi- and multilingual contexts.

Keywords: artificial intelligence; EFL; early childhood education; language learning

ARTICLE HISTORY

Received: 13 Sept. 2024 **Revised version received:** 15 Jan. 2025

Accepted: 1 June 2025 **Available online:** 10 July 2025

* Corresponding author: Department of Romance and Germanic Philology, Institute of Philology and Intercultural Communication, Kazan Federal University, Kazan, Russia
Email address: alb1980@yandex.ru

10.30466/ijltr.2025.55564.2821

Introduction

Recent breakthroughs in the development of generative artificial intelligence (GenAI) and the growing accessibility of educational intelligent systems have made smart technologies a hot topic for educators. Foreign language teachers were among the first to embrace the potential of AI technologies thanks to their affordances for generating oral and written texts in various languages. While there is a clear interest in studying potentials of AI tools for teaching and learning, the academic community still has only limited experience in integrating generative AI technologies into language education. This is primarily due to the novelty of the technologies themselves that became widespread only in mid-2023. Moreover, existing research publications, as a rule, address the issue of using AI in higher education, but less so in secondary education. Research that examines intelligent technologies with preschool children is "at an embryonic stage" (Kewalramani et al., 2021, p. 356). While some researchers have raised the issue of using AI technologies in preschools and demonstrated their positive effect (Su & Yang, 2022), generally they have not been interested in AI-assisted development of language and speech in young children (focusing instead on the development of other competencies in young children – digital, social, emotional, etc). The purpose of this study is to partially close this research gap by examining the potential of generative AI for developing language skills of preschool learners. The study focuses on young learners' experiences with AI technologies and language teachers' mediating practices that facilitate these experiences. It examines how teachers capitalize on AI tools' affordances and counteract their pitfalls to build a safe and productive learning environment conducive to vocabulary growth among 4–6-year-old learners of English as a foreign language (EFL).

Review of Literature

Artificial Intelligence in Language Education

We currently observe a noteworthy increase in the number of studies that indicate the significant potential of AI technologies for developing language competencies of learners of different levels and ages. Most of the research is conducted at higher educational institutions (Chen et al., 2022; Keerthiwansha, 2018; Zou et al., 2020), less often in middle school (Holstein et al., 2018) and in elementary school (Ghoneim & Elghotmy, 2021; Underwood, 2017). Existing studies indicate that AI-tools could be successfully used to develop all four language skills – listening, speaking, reading, and writing.

AI tools have been shown to be effective for developing overall listening skills, as well as sound discrimination, vocabulary, comprehension skills, and grammar (Ghoneim & Elghotmy, 2021). In their study, Groneim and Elghotmy attribute the effectiveness of AI-supported classroom to highly interactive, pleasurable, and visually enhanced activities that engage students in meaningful and need-based tasks. AI tools enable students to easily manipulate video-based listening files, explore supporting hyperlinks, receive quick feedback, and use a 'reset' function to repeat the activities, thereby enriching students' linguistic output and providing a more individualized and autonomous learning experience (ibid.). Listening skills could also be developed by means of oral interaction with voice-based intelligent tutoring systems such as *Google Assistant* that provides multimodal, authentic, and playful activities as shown in Tai and Chen's (2022) experimental study of ninth-grade EFL students.

AI tools may equally benefit students' speaking performance, as was the case in the study by Junaidi et al. (2020) that involved 7th grade students who utilized *Lyra Virtual Assistant*. In another study (Kang, 2022), AI-driven software *Replica*, which featured an oral-visual interaction mode, was compared to human native speakers of English working via Zoom. The study demonstrated that

Business English college students who interacted with AI avatars in *Replica* outperformed their peers working with humans in four criteria of speaking measurement – fluency, accuracy, coherence, and interaction. This effect, the study shows, was particularly pronounced for low-level students, while high-level students benefitted more from interactions with human native speakers. In the study by Fathi et al. (2024), *Andy English Chatbot* was found to provide better speaking practice for language learners than peers in a face-to-face interaction.

Developing reading skills may also be the target of AI-supported language classroom. In Lee, Shin, and Noh's study (2023), young ESL learners engaged in reading activities that involved AI-based content generators. These activities proved to be effective for enhancing students' enjoyment and interest in reading. Behforouz and Ghaithi (2024) found that an interactive *What's Up* bot might be effective for developing EFL college students' reading skills. AI-powered tools were also found to be useful for facilitating social reading practice with 5th grade elementary school students (Liu et al., 2022). In Liu and his colleagues' study, the feeling of social connection with a chatbot supported learners' interest in reading.

Additionally, artificial intelligence has been shown to be applicable to developing writing skills. AI-powered writing assistants, such as *Wordtune*, may help language learners to formulate or rephrase their ideas, aid in selecting an appropriate register (casual or more formal), and shorten or expand original texts (Zhao, 2022). In an experimental study that involved adult EFL learners, the implementation of *Wordtune* activities resulted in enhanced vocabulary and more complex sentence structures (Mahmud, 2023). Improved writing skills were also reported when learners used *Writerly* and *Google Docs* (Wale & Kassahun, 2024). Similarly, an AI-based automated essay evaluation system was shown to provide feedback of higher quality compared to human evaluators (Wang, 2022). Tran (2024) reported teachers' and students' positive attitude towards AI and perceived benefits for academic writing tasks scoring higher in cohesion, coherence, lexical resources, grammatical range, and accuracy.

Thus AI-powered technologies have been used in language learning to provide language input and assess language output, to facilitate interaction, to enable individualized learning trajectories, to respond to students' affective needs, and to motivate students to keep on learning. Language teachers and learners utilize a variety of AI-powered technologies including intelligent tutoring systems, chatbots, machine translation, automated writing evaluation, and speech recognition systems. However, the field of intelligent computer-assistant language learning (iCALL) seems to be only in the early stages of examining the potentials of AI technology for (very) young language learners.

Technology for Young Language Learners

As technology users become younger and younger (Revealing Reality, n.d.), latest technological innovations find their way into their classrooms (Cizrelioğlu & Aydin, 2024; Derakhshan et al., 2024; Hua & Wang, 2023). While many early childhood educators (including parents) might not be among early innovators, some of them have nevertheless promptly responded to the demands of their technology-rich societies and introduced computer, mobile, VR (virtual reality), AR (augmented reality) and most recently AI technologies into the lives of young learners. Technologies have been shown to be effective for developing young learners' phonological awareness, letter naming skills, spelling, word comprehension, reading fluency, story sequencing and retelling, conversation quality, reading skills, as well as children's motivation, engagement, critical thinking, and interactional skills (Gençten & Aydemir, 2023). Research in EFL and ESL has demonstrated the potential of a range of digital applications and platforms applicable for young learners, including *Duolingo* (Suhendra et al., 2024) and *Reading Eggs* (Sadykova et al., 2016). Other

studies have explored AR and VR tools for simulating real-world language practice (Chen, 2020; Danaei et al., 2020; Yang et al., 2020).

The recent and massive spread of AI technologies has also been picked up by early childhood educators. Several studies explored the effect of child-robot interaction on speech development; their findings suggest that learning activities involving social robots may enhance children's vocabulary in their native (Gordon & Breazeal, 2015; Kory-Westlund & Breazeal, 2019; Movellan et al., 2009) or foreign language (Gordon et al., 2016). An algorithm developed and implemented into a social robot by Gordon and Breazeal (2015) was able to accurately and efficiently assess children's word-reading skills. In their experiment, 4 to 8-year-old children played a novel story-creation tablet game with the robot whom they helped to learn how to read (as if the robot were their peer). Such a learning scenario was both empowering and pleasurable for children who developed their language skills in a playful environment. Improved (native) language production skills were also reported in Kory-Westlund and Breazeal (2019), where children aged 3 to 8 years played a story-telling game with a social robot. The study reported that child-to-robot interaction contributed to children's engagement and enjoyment, while also improving the child's capacity to retell a story. A social robot was also found to be an effective companion for learning a foreign (Spanish) language by 3 to 5-year-olds in Gordon et al. (2016). In their study, Gordon and his colleagues were able to demonstrate the positive impact of personalized motivational strategies exhibited by a robot both verbally and non-verbally.

Personalization and playfulness have therefore been shown to be of primary importance for building productive and enjoyable interaction between the child and AI-powered technologies. Moreover, research that focuses on very young learners emphasizes the paramount role of the teacher or parent as an intermediary between the technology and the child. When the child's interaction is closely guided by an adult, young learners develop most optimally (Barblett et al., 2023; Meskill et al., 2020; Meskill, 2024; Pifarré, 2019; Teichert, 2017). This idea has previously been discussed by educational researchers in application to more traditional technology. Considering the recent hype around generative AI, however, which arguably holds the potential for disrupting the entire educational system (García-Peñalvo, 2023), there is therefore an urgent need in empirical studies to develop and test new learning scenarios for effective and safe interaction between the child and AI-powered technologies. To date, there is a real dearth of research that examines very young learners' experiences with generative AI services and platforms. We need more studies outlining effective practices that teachers could use to enhance children's learning experiences through AI affordances while counteracting any potential risks associated with child-AI interaction. To our knowledge, no study has previously explored the integration of AI-powered visual and audio generators into foreign language programs for preschool children.

Theoretical Foundation

This research is based on the concept of *producerly pleasures* put forward by Meskill (2007) in application to televisual and electronic texts used for language learning. Meskill's concept was later employed to study bilingual preschool children's interaction with multimodal digital resources (Meskill et al., 2020). Meskill suggests that just as readers of printed texts may derive pleasures from meaning-making when reading paper books, so do consumers of electronic texts (the original idea of pleasures of reading through meaning-making belongs to philosopher and poststructuralist Roland Barthes).

According to Meskill (2007), producerly pleasures are derived through the act of *seeking* information (i.e., looking for particulars in a text and making choices) and the act of *identifying* with it (i.e., relating to the information one has sought out). The sources of producerly pleasures were later expanded to include the dimensions of *rewards* and *screen magic* that one may experience when interacting with

digital technology (Meskill et al., 2020). Thus, digital screens allow the child to seek information, to identify with (and relate to) it, to be enchanted by what is going on the screen, and to be rewarded for his efforts. .

For Meskill, the producerly pleasures result from our co-construction of meaning in tandem with what is happening on the screen. When applied to early childhood education contexts, producerly pleasures must be mediated by an adult who orchestrates the child's interaction with technology. Meskill's construct of *digital screen mediation* (Meskill, 2021, 2024), rooted in Vygotsky's (1978) sociocultural theoretical framework, is well suited to examine teachers' mediational practices in AI-supported language classroom. Vygotsky emphasized the significance of social interaction for the child's psychological development and the critical role of *more knowledgeable others* who serve as intermediaries between the learner and the content of learning. Meskill (2024) applies the concept of mediation to the technology-enhanced language classroom where teachers' mediational practices become critical for developing children's language skills "via active meaning-making with others *with*, *through* and *around* screens" (p. 173).

In Meskill's digital screen mediation (DSM) model, the teacher's and the student's agency are not emulated by technology; rather the teacher capitalizes on digital screen features and designs a conversation-rich classroom where the digital screen serves as an anchor for collaborative activities. While Meskill's DSM model can be applied to any technology-supported instruction, we will use it to understand learning "with, through and around" image- and audio-generating AI technologies in EFL classrooms. We will therefore consider AI tools as anchors that teachers use to mediate student-AI interaction resulting in producerly pleasure of seeking and identifying with digital content, and experiencing screen magic and rewards in the process (Figure 1).

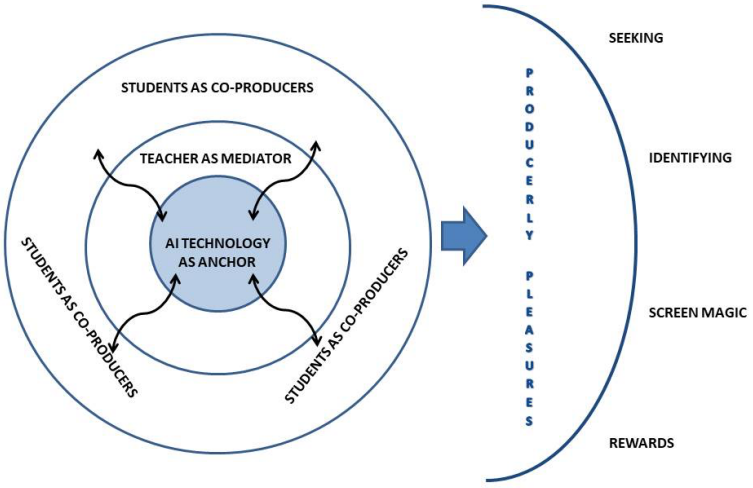

Figure 1. Visualization of the Conceptual Framework of the Study based on Meskill's Constructs of Digital Screen Mediation (DSM) and Producerly Pleasures

Method

Study Design

This mixed-method exploratory study involved the collection of quantitative data in a quasi-experimental pretest-posttest design format, as well as qualitative data from teachers' questionnaires, reflection journals, interviews, and video recordings of three 4-day AI-enhanced English language learning modules implemented in two preschools.

The study design was guided by the following research questions:

1. Do AI-powered image and audio generators have educational potential for 4–6-year-old EFL learners?

2. Do AI-powered image and audio generators enable 4-6-year-old EFL learners to experience producerly pleasures?

3. What mediational practices do EFL teachers employ with young learners in AI-enhanced classrooms?

Study Participants

Two key informants for the study were selected from participants of a survey that was previously conducted by the authors (Sadykova & Kayumova, 2024). Selecting key informants was based on the following criteria: the teacher needed to 1) work with children from 4 to 6 years old; 2) have regular EFL classes, 3) work according to state-regulated requirements in an early childhood education preschool. These criteria aimed at building cases that would be similar, thus transferable, to most educational contexts in the region where the study was conducted. Dinara, a 23-year-old female teacher in a private preschool (Site 1), and Guzel, a 41-year-old female educator in a state-funded preschool (Site 2), met all three of the criteria.

The study also involved three groups of children (N=40) between the ages of 4 and 6: *Group 1* of 15 children aged 4-5 (10 girls and 5 boys); *Group 2* of 13 children aged 4-5 (8 girls and 5 boys) and *Group 3* of 12 children aged 5-6 (8 girls and 4 boys). All the children are bilingual and speak Russian and Tatar as their first or second languages. They also acquire English as a part of their multilingual preschool curriculum.

Study Context

Two study sites are located in the Republic of Tatarstan (Russia) that has two official languages (Russian and Tatar). *Site 1* is a private multilingual kindergarten and preschool with about 90 children from 3 to 7 years old and 35 educators. All of the children in this particular setting acquire Russian and Tatar as their first and second languages, English and Chinese as compulsory foreign languages, and Spanish as an optional additional language. The school curriculum is based on the state requirements and regulations; however, it is also designed to accommodate British requirements as a registered Cambridge International School.

Site 2 is a state-funded kindergarten and preschool with about 150 children from 3 to 7 years old and 23 educators. The main language of instruction is Tatar; however, all of the children study Russian as their second language and a state language. English is a compulsory foreign language. The school curriculum is based on the state requirements and regulations.

The research intervention at Site 1 was carried out twice. First, in accordance with the spring 2024 school curriculum , the teacher Dinara designed a 4-day English language module aimed at the development of vocabulary on the topic of 'Aquatic Animals'. After consultations with the research team, Dinara selected to employ *Fusion Brain*, a free AI-powered image generating platform using the *Kandinsky* neural network (https://fusionbrain.ai/en/) that enables users to generate still and motion pictures. For the second intervention period Dinara built a 4-day module on the topic of 'Tasty Week' with the help of *Suno AI* (https://suno.com/), a platform that generates songs based on texts and prompts given by users.

At Site 2 the intervention took place in the fall of 2024. The EFL teacher Guzel designed a 4-day English language module on the topic of 'Transport'. The educator utilized an AI-powered service *Songgenerator.io* (https://songgenerator.io/ru) that generates songs based on texts and prompts given by users.

Data Collection and Analysis

Data were collected before, during and after the interventions. The two EFL teachers Dinara and Guzel first completed a questionnaire that focused on participants' beliefs about, and experience with the use of AI technologies for language development of young learners. Then the two EFL teachers were asked to provide information about their EFL curriculum, as well as number, age and gender of their pupils. Based on this information, the team of researchers drafted a four-day AI-enhanced English language module for each of the three interventions. The drafts were revised several times and finalized by the teachers to ensure that they met EFL programs' requirements. Before the interventions, the teachers ensured to have received informed consent from parents of participating students, and conducted oral vocabulary pre-tests with the students using flashcards. All interventions were video-recorded and later sent to researchers for analysis. After the interventions, teachers conducted vocabulary post-tests and participated in two semi-structured audio-recorded interviews via a video-conference program. Two individual interviews with each of the two teachers (from 21 to 47 minutes each) were carried out by both researchers based on the video-recordings to clarify teachers' actions and intentions. Preschool websites were scrutinized to receive a better understanding of the study context.

The rich data collected were analyzed separately by the two researchers, then cross-checked and verified via three types of triangulation: methods triangulation (using both quantitative and qualitative data-collection methods), triangulation of sources (to ensure data consistency across questionnaires, interviews, video recordings, and documents), and analyst triangulation (using two trained analysts) (Patton, 2001). Pre- and post-test results underwent quantitative analysis (raw numbers and percentages) to assess vocabulary growth for each of the students and to obtain the average for each of the three intervention groups. The results were calculated and visualized in Microsoft Excel and further discussed with the participating teachers during interviews. All of the qualitative data were theme-coded by two researchers to identify recurrent themes related to 1) *producerly pleasures* (Meskill, 2007) potentially experienced by young learners and 2) *mediational practices* that the teachers used to initiate and carry out conversations with, through and around AI-generated audial and visual content.

Results

The rich data collected during three intervention periods underwent quantitative and qualitative analysis and yielded results grouped into three subsections below according to research questions (RQs).

RQ1: Do AI-powered image and audio generators have educational potential for 4–6-year-old EFL learners?

The study clearly supports the introduction of AI technologies into classrooms of very young EFL learners. The video-recorded classes, reflection journals, and teacher assessment given in post-intervention interviews, as well as pre- and post- vocabulary tests speak in favour of using AI tools such as *Fusion Brain, Suno AI* and *SongGenerator.io* for developing language skills, particularly for enriching vocabulary. When designing and implementing activities involving image and audio generating AI technologies, teachers found that these tools, despite their non-negligible learning curve (just as would have been the case with any other new tools), could be quickly integrated into existing teaching routines. The teachers reported that interactive and multimodal features of AI tools provoked young learners' involvement in conversations, increasing their language output, enhancing learners' engagement with the content, and enabling teachers to meet study goals. The educator at Site 1 reported that "all of the goals set were achieved despite some corrections [of the scenario] along the way." Impressed with the outcomes, the educator expressed her willingness to continue using AI tools in her work. The educator at Site 2 also noted that she was pleasantly surprised by working with AI-powered services and that almost all of her students learnt both words and patterns by heart in the course of interaction with AI.

Educators' assessment is supported by the data from pre- and post-tests: the mean percentage of correct responses to oral vocabulary tests given before and after the interventions show language gains in all three groups (Figure 2). In Group 1, at the pre-test stage an average participant was able to answer 61.6% of the questions correctly, while at the post-test stage the percentage of correct answers grew to 88.6% showing vocabulary gain of 27%. In Group 2, the average result went up from 50.2% at the pre-test stage to 86.1% at the post-test stage showing vocabulary gain of 35.9%. Group 3 made the most remarkable progress from 12.5% to 80.2% of correct responses demonstrating a 67.7% increase in the mean score.

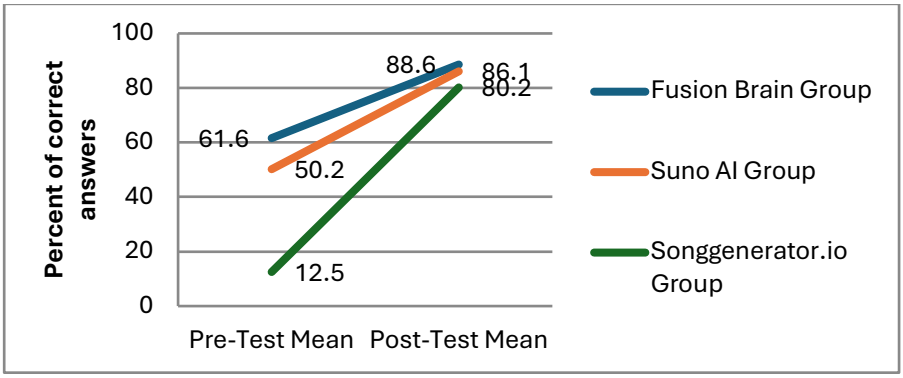

Figure 2. The Mean Percentage of Correctly Used Vocabulary in Pre- and Post-Tests for Three Interventions

Quantitative results of the three interventions speak in favour of learning activities designed around AI-powered platforms, and specifically in favour of image and audio generators tested in this study. While the small sample size and the absence of control groups limit interpretative power of the findings, the qualitative data discussed below provide further insight into the educational potential of AI.

RQ2: Do AI-powered image and audio generators enable 4-6-year-old EFL learners to experience producerly pleasures?

As discussed earlier, active children's engagement in technology-enhanced classroom is attributed to producerly pleasures children experience when interacting with digital tools (Meskill, 2007; Meskill et al., 2020). Our data demonstrates that AI-powered audial and visual generators possess unique qualities that enable young learners to derive producerely pleasures in all four dimensions: seeking information, identifying with it, experiencing screen magic, and deriving rewards from the experience. This subsection will illustrate these findings with examples from in-class activities, as well as comments provided by the teachers in their reflective journals and in the interviews.

Seeking

Producerly pleasures through seeking information were evident when children were guided to 1) listen to songs and identify words that they asked *Suno AI* and *SongGenerator.io* to use when creating a song and 2) identify and name objects in images generated by *Fusion Brain* based on children's prompts. In all three groups, teachers purposefully directed children's attention to audio and visual content produced by AI-powered services, creating a learning situation in which children had to pick out new words and then to produce them orally.

The interaction from Intervention 1 below demonstrates how attentive the students were to visual content:

> **Teacher** [in Russian]: Let's have a look [points at a newly-generated motion picture of a swimming seahorse].
>
> **Students** [look at the screen, some stand up for better viewing]
>
> **Student B** [exclaims in Russian]: It is swimming!
>
> **Teacher** [in Russian] Well, we asked for [switches to English] a big yellow seahorse…
>
> **Student C** [interrupts the teacher in Russian emotionally]: It's incorrect! It's not jumping!
>
> **Teacher** [repeats in Russian]: Indeed, it's not jumping. [switches to English] It's not jumping.
>
> **Students** [echo in English]: It's not jumping.

Later, in Group 1 the educator engaged the students in the process of listening by giving a well-thought instruction:

> **Teacher [**in English] Take your seats. [switches to Russian to restore discipline] We should sit down while Suno is creating a song. She needs time to compose a song. And we should sit down. We should get our ears ready. [the songs are ready and the teacher asks in English] Let's start?
>
> **Students** [in English]: Yes.
>
> **Teacher** [in Russian]: Listen for your fruit and vegetables.

Suno [starts the song]

Students [fall silent]

Both video recordings and interview responses demonstrate that the level of students' interest in AI-generated content was higher than in traditional audiovisuals. "The students listened to a co-generated song more attentively, with greater interest. It was obvious that they wanted to catch familiar words," shared Guzel in the interview about Intervention 3.

Identification

The data showed that students treated AI tools as animated smart magic friends with whom they can interact and identify. Children learned that these friends can take commands and generate images or songs based on their prompts. In the first intervention, when asked to create a prompt, one of the students came up to the screen and addressed the screen, not the teacher, saying "Draw a shark, please." In the second intervention the educator encouraged the students to speak up when giving a prompt to *Suno AI* and one of the children asked the teacher, "Did she hear me?"

Keeping in mind the objectives of the EFL classes, on occasion the teachers emphasized that these smart friends understand only English. When the learners tried to address an AI tool in their native language, teachers could ask them to switch to English. In the following example, teacher Dinara accentuated the idea that their magic friend followed only the children's commands in order to encourage the children to build sentences independently:

Teacher [says in Russian]: Now I want *you* to give a command to the computer. It follows only children's commands. Listen to me. Then repeat. [switches to English] Big pink shark swims in the sea [stretches arms forward as a sign of encouragement to speak] ...

Students [start saying in English] Big...

Teacher [nods in approval; shows a flashcard of the colour 'pink']

Students [in English]: pink...

Teacher [nods in approval; shows a flashcard of 'a shark']

Students [in English]: shark...

Teacher [pretends to be swimming]

Students [pick up] swim...

Student C [in English without support]: in the sea

On Day 4 of the second intervention, the teacher also exploited the students' belief that an AI tool has human-like feelings and needs their commitment.

Teacher [says in Russian]: Shall we listen to the song?

Students [enthusiastically]: Yes!

> **Teacher** [in Russian] Let's ask Suno to play the song. All together, [switches to English] 'Suno, play the song!'
>
> **Most of the students** [in English]: Suno, play the song!
>
> **Teacher** [in Russian in a sad voice]: No, she is not playing it. That's because some of you didn't ask her to play. Let's ask her again, all together... [adds a head nod to help students start in sync]
>
> **Students** [in a loud voice in chorus in English]: Suno, play the song!

In these examples, the students were not only able to relate to technologies as their 'smart friends' but also to identify themselves as English language speakers, which seemed to add to the pleasure of their interaction with the AI tools.

Screen Magic

The enjoyment that children received from production activities designed around AI-powered platforms was found to be one of the most prominent themes in this study. The data demonstrated that the magic effect of the screen undoubtedly increased the students' situational interest and, consequently, made a significant contribution to classroom dynamics. Classroom video recordings, reflective journals and post-intervention interviews revealed that there was genuine pleasure in co-experiencing and co-creating the AI-generated content.

On Day 3 of the first intervention the teacher exploited the screen magic effect to capture students' attention:

> **Teacher** [says in Russian mysteriously]: I'll tell you a secret. I have a magic friend. A smart friend. You can ask it to create a picture. It will do it.
>
> **Student B** [in a curious voice, speaks Russian]: Like *Alisa*?
>
> **Teacher** [explains in Russian]: Yes, like *Alisa*. But *Alisa* speaks. And this one makes pictures.

Here the teacher introduced the new tool *Fusion Brain* for the first time and created an image of a powerful and mysterious friend while relying on children's previous knowledge about a similar tool – a voice assistant *Alisa*. The magic effect doubled while the image was being created and children were observing twinkling stars on the darkened screen. Video recordings captured students' anticipation of the magic and excitement when it happened, i.e. when the screen showed still images of aquatic animals that the children wished to create. The '*wow*'-effect intensified on Day 4 when children saw an animated video they created with the help of *Fusion Brain*. Some children jumped up from the carpet and ran to the screen showing genuine excitement, pleasure, and engagement (Figure 3).

Figure 3. AI-Inspired 'Wow'-Effect

Rewards

The screen magic dimension of *producerly pleasures* is clearly intertwined with rewards derived from success with the AI tool. When the students correctly produced the prompt in English, they were rewarded by the image appearing on the *Fusion Brain* screen or by the songs played by *Suno AI* and *SongGenerator.io*. The pleasures of these linguistic and digital successes were very obvious. In the first intervention the children jumped up with excitement, ran to the computer screen, and literally squealed with delight. At this very moment, one of the students that remained seated started complaining, "I can't see! I can't see!" demonstrating the FOMO (fear of missing out) effect in relation to the AI tool and illustrating the children's genuine interest in seeing how their co-created product takes on physical form.

In the second and third interventions, students were similarly interested in, and pleased with the result of writing songs together with *Suno AI* and *SongGenerator.io.* Some were especially excited when catching the lines of their authorship in the song, shouting out "This is my [line]!"

Thus, the data provides ample evidence that shows that preschool children are highly attracted by visual and audio content generated by AI technologies. Evidently, they feel exited, rewarded and empowered by co-production activities they are able to carry out in AI-supported language learning environment. The mediational role of a teacher comes forward in such classrooms. In the sections below, we will analyse mediational practices that teachers employed to engage young learners into conversations with, though and around AI technology.

RQ3. What mediational practices do EFL teachers employ with young learners in AI-enhanced classrooms?

Our research data re-emphasizes the significance of the mediating role of the teacher. AI image and audio generating tools are verbally silent and do not interact with a user unless the latter types in a prompt. Considering students' pre-literate stage, the teachers needed to serve as the nexus between the young humans and AI. Teachers Dinara and Guzel introduced the AI tools to the students, provided access to them,, helped students navigate around the platforms, assisted students with writing and re-writing the prompts, and managed time, thus overall mediating children's interaction with the technology. Most importantly, these were teachers – the more knowledgeable others –

who encouraged learners to engage into conversations with, through and around audio and visual generating tools.

The mediating role of the teacher was significant throughout all three stages of GenAI-supported language activities:

- *prompt-writing stage,* when the teacher and the students create a prompt for an AI tool;
- *product-generation stage,* when the AI tool takes time to produce an image or a song based on a given prompt and the teacher fills in the time gap by other activities;
- *post-production stage,* when the teacher facilitates conversations based on generated images or initiates song-listening and singing activities.

Below we describe the mediational practices at each of these stages.

Mediational Practices at Different Stages of Child-AI Interaction

Mediation at Prompt-Writing Stage. When the researchers and the participating teachers teamed up to create AI-enhanced learning modules, it became evident that unlike games (mobile or computer), AI image and audio generators provide users with open-ended opportunities, adapting to any content input and empowering users to become co-authors of generated products. Therefore, the intervention scenarios were designed to capitalize on these critically important affordances of AI technologies.

When working with *Fusion Brain*, it was up to the teacher and the students to decide what to draw, whether the image was going to be still or animated, in pencil or in oil, etc. Similarly, *Suno AI* and *SongGenerator.io* enabled the young learners and their educators to select the style and the tempo of the songs as well as the singers' gender. The capabilities of new AI-powered technologies allow users to be creative, apply a range of diverse artistic tools and personalize the final product. In all three intervention periods, the young participants were constantly given power to choose the characteristics of the final products they were designing:

> **Teacher** [asks in English about the tempo of the song]: Fast or slow? [switches to Russian] Do you remember the game that we play? [in English] Fast and Slow.
>
> Or:
>
> **Teacher** [asks in Russian]: What else will we ask to draw?
>
> **Student B** [shouts out in English]: Octopus!
>
> **Teacher** [asks in English]: What colour will it be?
>
> **Student H** [shouts out]: Green!

Here the students are given decision-making power. Even the students who did not feel eager to give a prompt were still active participants in the learning process. They repeated words in chorus after the teacher or echoed their peers:

Teacher [types in a prompt and dictates it orally]: A seahorse lives in the sea.

Student I [chants in English clapping her hands]: Seahorse… seahorse… seahorse

Student D [picks up the chant and claps his hands]: Seahorse… turtle… octopus

The episodes above illustrate the significant mediating role of the educator who applies AI tools in a way that empowers learners and encourages their active participation in co-authoring a final product. Through modelling, questioning, and eliciting students' answers the teachers in the study created ample opportunities for students to hear and comprehend new words and then to speak in a target language when completing a prompt-writing task. Thus, the young learners developed their language skills while being engaged in meaningful conversations where they used new words in a fun and authentic language environment.

Mediation at Product-Generation Stage. Depending on the speed of the Internet connection, equipment used, and AI technologies involved, it may take several minutes before children and teachers are able to observe or listen to a generated product. For novice users, as was the case with the participants in this study, the delay may cause confusion and frustration. In the first intervention, low Internet speed slowed down the generation of a short motion video with *Fusion Brain* and Dinara gave children a chance to direct their energy into moving around the classroom and interacting with each other. Such a break between language activities was seen as appropriate for young learners. This waiting time, however, could also be used to carry out non-AI language learning activities, as was the case in Intervention 3. While the AI tool was generating a song, the students were coloring pictures of different forms of transport. The teacher circulated around the classroom and elicited key vocabulary items from the coloring pages (e.g. "What is the boy riding?", "What color is his bike?" etc.). Thus at this stage, teachers could continue with instructional conversations in which they direct students' attention to new words and structures, allow children to practice pronuncingnew vocabulary, and support the process of learning via off-AI activities appropriate for preschool children (such as drawing or moving around).

Mediation at Post-Production Stage. In the post-production stage, the teachers maximized the students' language output by using newly generated images and songs as anchors for conversations with, through and around audio- and image-generating tools. In the example below, Dinara skillfully questions children about what is happening on the screen:

Teacher [asks in English pointing to the newly created video]: Is it a shark?

Students [in English]: Shark!

Teacher [asks in English] Is it big [<u>says loudly with arms to the sides</u>] or small [<u>says quietly with hands close together</u>]?

Student F [in English]: Big!

Teacher [echoes in English]: Big. Very good! [asks in English] What color is it?

Student G [in English]: Pink!

This polylogue happened right after *Fusion Brain* generated the image and the teacher used children's excitement by the magic produced on the screen to encourage language comprehension and production. Dinara elicited students' answers through questioning and used non-verbal cues (voice

and gestures) to assist their comprehension. The '*wow*'-effect of the generated image helped the teacher to keep learners on task for several minutes.

In Intervention 3, at the post-production stage the teacher also provided non-verbal support for students to better understand the lyrics of a newly-generated song. While the song about riding, driving, flying and rowing different means of transport was being played, she asked the students to get up and repeat movements after her. After everyone listened to the song, Guzel asked, 'What is the song about?' The students named some recently learned means of transportation, such as 'car', 'plane' or 'train'. However, as the objective of the lesson was to introduce *verb + noun* collocations (e.g. 'fly a plane'), the teacher made a prompt decision to initiate an additional flashcard activity in order to achieve the desired effect:

> **Teacher** [shows a flashcard of a child riding a bike and says in English]: Ride a bike. [adds movements] Ride a bike.
>
> **Students** [repeat words and movements]: Ride a bike, ride a bike.
>
> **Teacher** [in Russian]: Now let's do it this way. I start and you finish. Okay? [switches to English] Ride a …[pauses]
>
> **Students** [in English]: Bike!
>
> **Teacher** [in English]: Yes, ride a bike.
>
> **Students** [echo in English]: Ride a bike.

Thus, in all three stages of AI-supported co-production activities, the teachers orchestrated conversations with, through and around AI-tools. These activities were further enhanced by off-AI tasks that involved the use of printed flashcards, physical objects (e.g., toys) or kinesthetic and hands-on activities - moving, drawing, or singing along. Learners' comprehension and language output was supported by verbal cues, such as modeling, questioning and calling attention to forms, as well as by non-verbal means of communication such as gestures, visuals, and voice.

Mediation to Counteract AI Pitfalls

When it comes to early childhood education, teachers' role in counteracting potentially harmful impact of technology becomes paramount. The data of this study reemphasize the critical significance of teacher mediation in cases where encounters with AI might hold risks for children's emotional and psychological well-being. This study revealed that image generating AI tools such as *Fusion Brain* may misunderstand a prompt, distort reality or produce images that hyperbolize negatively-charged actions or feelings. For example, on Day 3 of the first intervention, it took three attempts to generate a still picture of two objects ('green octopus swimming in the sea near pink jellyfish'). The image was generated incorrectly twice, first showing a bicolored creature that resembled both an octopus and a jellyfish (Figure 4). The teacher used this AI failure as an excellent way for showing children that even technology can have deficiencies and that people might have to deal with them. Moreover, this case illustrated that 4-6-year-old children are quick to understand how 'silly' some technologies could be, as evidenced by the following extract:

> **Teacher** [speaks Russian and switches to English when using key vocabulary]: Look! It didn't draw it the right way. We asked to draw a '*green octopus* and a '*pink jellyfish*'. And it drew the head of a jellyfish and the legs of an octopus. It might happen like that. The

computer can misunderstand us. Let's write it in a different way [starts writing down a prompt].

Student D [addresses fellow students in Russian]: This computer isn't a real helper. It is silly, right?

Teacher [repeats the instruction in Russian and switches to English]: Let's try to do it in a different way. *'Green octopus swimming in the sea NEAR pink jellyfish'.*

Figure 4. AI-Generated Image of 'a Green Octopus and a Pink Jellyfish Swimming in the Sea'

While green-and-pink jelly-octopus might be fun to observe, other 'silly' pitfalls of some AI-powered services should be anticipated and prevented by carefully mediated teacher's moves. This study revealed that motion images created by *Fusion Brain* might be scary when showing negatively-charged actions or feelings. When testing this platform before the class, we found that applying the prompt "the girl got scared when the cat jumped at her", *Fusion Brain* generated a short video that would be appropriate for an adult horror cartoon rather that for usage with preschool children. Knowing about that, Dinara had to suppress students' splash of fantasy to avoid AI's exaggerated visualization of potentially horrifying actions:

Teacher [in Russian]: What will the shark do?

Student I [in Russian]: It… it… it will eat everyone up.

Teacher [says in Russian giving a disgusted look]: Maybe we won't draw a shark eating everyone up? [switches to English] It will smile [gives a smile]. Okay? Smile [gives a smile and points at it].

Student I [repeats in English]: Smile [gives a smile and points at it].

Teacher [voices the prompt in English]: Blue shark smiles.

Some students [repeat in English]: Blue shark smiles.

AI technologies that voice text prompts, including song generators *Suno AI* and *SongGenerator.io*, can also distort the pronunciation of words and sentences. For example, teacher Dinara reflected

that the first version of the song they created with the students was not good enough because it had "an unnatural sentence stress in the lines." The second version, in its turn, sounded more "authentic" and "appropriate".

Thus, mediational practices employed by teachers either capitalized on AI affordances related to co-production capabilities or served to counteract AI pitfalls associated with its unpredictable open-ended nature that could distort reality. Our data analysis revealed that at no point did the teachers feel like they lost control over the classroom and allowed unmediated direct children-technology interaction. In their interviews, both teachers confirmed that they saw their role as being paramount when it comes to the use of AI technologies by very young learners. "Children of this age need emotional contact, social interaction, and the development of communication skills that AI cannot provide", Guzel reported. Dinara reiterated the same idea saying that "AI can be a great tool to support pedagogical work, but it is the teacher who creates an atmosphere of trust and comfort, which is especially important for effective learning in the early stages of language development."

Discussion

The data collected during three intervention periods revealed that AI-powered image and audio generators enabled two participating EFL teachers to involve children in highly engaging and pleasurable co-production activities that provoked target language conversations with, through and around AI-powered tools. Supporting Meskill's concept of producerly pleasures (2007), the study illustrated how willingly children became engaged in language activities while seeking for and relating to information, receiving rewards and experiencing screen magic, all while being involved in playful and enjoyable tasks appropriate for very young learners. Evidently, co-production activities done with audio and image generators resonate perfectly with young learners as they appeal to children's multisensory world perception and enable teachers to create playful learning environment. Playfulness is at the core of early childhood education (Singer, 2013), while multisensory (multimodal) learning matches well young children's cognitive and emotional needs (Broadbent et al., 2018; Buckmayer et al., 2024). Not surprisingly, game-based language learning with technology has long been practiced by teachers (Godwin-Jones, 2014), including early childhood educators (AlNatour & Hijazi, 2018; Redondo et al., 2020); and multimodality has been described as one of the most appealing affordances of digital technologies for young learners (Kayumova & Sadykova, 2019; Yelland, 2018). Therefore, playful tasks designed around image and audio AI-powered generators should enrich teachers' repertoires working with very young learners.

Both quantitative and qualitative data speak for the language gains derived from GenAI-enhanced learning environment. Students learned the language through listening and speaking activities embedded into co-production tasks. They found themselves in a conversation-rich environment where they had to provide verbal input in the target language, first, to create a prompt for generating an image or a song, and then, to assess and possibly correct the AI-generated product. These findings support the DSM (digital screen mediation) model by Meskill (2021, 2024) and promote active co-production of songs and still and motion pictures, which, with the advent of generative AI, has become more than doable for non-experts working in common educational settings.

The data also highlighted the critical role of the teacher as a mediator of learner-AI interaction in both orchestrating meaningful conversations and in harnessing 'untamed' powers of generative AI technologies. The participating teachers used a range of instructional conversation techniques to maximize language gains including modelling, questioning, eliciting students' answers, and calling attention to forms. Being "the essential venue for teaching and learning" (Meskill & Anthony, 2007, p. 7), conversations that are framed around a technology tool are of primary importance and may require some training and experience from educators (ibid.). Without teacher mediation, no

language learning might have happened in these given settings. This supports previous findings focused on technology-enhanced language classroom for young learners (Christ et al., 2019; Kervin et al., 2015; Korat et al., 2013; Meskill et al., 2020).

The feature that was unique to our study, however, was the fact that the teachers were faced with previously inconceivable power of technologies, i.e. their human-like capabilities to generate the content. This generative power both expanded teachers' instructional repertoires and presented additional responsibilities for them, making them accountable for children's interaction with a non-human with anthropomorphic qualities. Generally speaking, this study provided evidence in support of those who have some reservations about the integration of AI in early childhood education. At the current stage of development, AI technology similar to *Fusion Brain* may generate the content inappropriate for young learners as the images could distort reality and exaggerate negatively-charged actions and feelings like fear. Questions related to AI's potential risks to students raised previously (Alm & Watanabe, 2023; Holmes, 2024) should certainly be further explored, while educators and policy makers should consider teaching young children about flaws and biases imbedded into AI (Su & Zhong, 2022). This study illustrates that some of these flaws could be and should be counteracted by the educator when testing, probing, and playing with AI technology *before* the class and by taking a moment to point at these issues while the teacher mediates child-AI interactions *in* class. Our findings point to several important implications of integrating generative AI tools into language classroom for very young learners. First of all, educators should capitalize on the generative power of AI by developing co-production activities. It is the co-authoring of the prompts that was found to be most productive for engaging students in conversations with, through and around AI-powered technologies. Second, GenAI tools should be treated as intelligent assistants ('smart friends') that 'speak' in the target language and require children's language output to co-produce a picture, a video, a song or some other content. This enhances students' identification as L2 speakers and makes interaction with technology more pleasurable and meaningful. Third, teachers should embed interactions with GenAI within off-AI activities, such as moving, drawing, or singing along, to account for children's developmental needs and to ensure their active engagement while AI takes time to generate the content. Finally, it is the responsibility of the adult (the teacher or the parent) to predict the unpredictable and counteract any GenAI generated content that puts at risk children's well-being. Learning and probing technologies beforehand and teaching children about AI's limitations have become as important as never before.

Conclusion

Since AI is supposed to mimic the human brain, one can argue that AI-tools will never be fully understood and learned. However, the uncertainty that comes with the technology should reemphasize the critically important role of the teacher as the mediator of the child-AI interaction. While previous research in CALL has numerously demonstrated that the effectiveness of instructional technology depends not so much on *what* is used but on *how* it is used, today educators need to reiterate and extend this idea in application to AI. We must also accentuate the significance of AI literacy among early childhood educators who have started gradually introducing AI as teaching and learning devices. Examining more cases of how educators adopt AI tools for teaching languages would provide the academic community with clearer understanding of effective practices for harnessing AI potential for the benefit of new generations of learners.

This research has several limitations that affect transferability of the study findings. Research was carried out in two sites and involved only two EFL teachers. The pre- and post-test results were not statistically tested due to the low sample size. The interventions involved the use of three AI tools – *Fusion Brain, Suno AI* and *Songgenerator.io* – and young learners' interactions with these tools were implemented via a single computer during three 4-day modules. These limitations were

mitigated by the application of both qualitative and quantitative data-collection methods, by the richness of data obtained through multiple sources, and by employing two trained analysts (researchers). All of this ensured the three types of triangulation mentioned above (methods triangulation, triangulation of sources and analyst triangulation), thus contributing to the study's validity.

While limited in transferability, the study shows real cases of successful AI-enhanced language classrooms and demonstrates mediating practices that might be adopted by other early childhood educators working in a bi- and multilingual contexts. However, further research is needed to develop and test other scenarios of AI integration into curricular of children of different age groups and diverse educational settings. In addition, research should focus on AI potential for home schooling and informal learning of a second or a foreign language, which will highlight parents' role as knowledge mediators. Moreover, as AI technologies are continuously evolving, researchers should keep an eye on new developments, explore their affordances and provide practitioners with research-based recommendations for their productive and safe integration into young learners' classrooms.

Acknowledgments

We thank the schools' administration and the key informants Dinara and Guzel whose participation made this study possible.

Funding

The research was funded by Russian Science Foundation grant #24-28-01129, https://rscf.ru/project/24-28-01129/.

References

Alm, A., & Watanabe, Y. (2023). Integrating ChatGPT in Language Education: A Freirean Perspective. *Iranian Journal of Language Teaching Research*, *11*(3 (Special Issue)), 19-30. https://doi.org/10.30466/ijltr.2023.121404

AlNatour, A. S., & Hijazi, D. (2018). The impact of using electronic games on teaching English vocabulary for kindergarten students. *US-China Foreign Language*, *16*(4), 193-205. http://doi.org/10.17265/1539-8080/2018.04.001

Barblett, L., Akifeva, R., Lavina, L. & Boylan, F. (2023). *Digital literacy: A review of literature*. Edith Cowan University. https://doi.org/10.25958/wb7c-g264

Behforouz, B., & Ghaithi, A. A. (2024). Investigating the effect of an interactive educational chatbot on reading comprehension skills. *International Journal of Engineering Pedadogy, 14*(4), 139-154. http://doi.org/10.3991/ijep.v14i4.48461

Broadbent, H. J., White, H., Mareschal, D., & Kirkham, N. Z. (2018). Incidental learning in a multisensory environment across childhood. *Developmental science, 21*(2), e12554. https://doi.org/10.1111/desc.12554

Buckmayer, K., Nicolau, H., & Pires, A. C. (2024). Storytelling with ALMA: A Multisensory Approach to Enhance Kindergarteners' Emotion Regulation. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference (IDC '24)*, 887-891. http://doi.org/10.1145/3628516.3659412

Chen, C. (2020). AR videos as scaffolding to foster students' learning achievements and motivation in EFL learning. *British Journal of Educational Technology*, *51*(3), 657-672. http:/doi.org/10.1111/bjet.12902

Chen, Y., Hsu, C., Lin, C., & Hsu, H. (2022). Robot-Assisted Language Learning: Integrating Artificial Intelligence and Virtual Reality into English Tour Guide Practice. *Education Sciences, 12*(7), 437. https://doi.org/10.3390/educsci12070437

Christ, T., Wang, X. C., Chiu, M. M., & Strekalova-Hughes, E. (2019). How App Books' Affordances Are Related to Young Children's Reading Behaviors and Outcomes. *AERA Open, 5*(2), 233285841985984. https://doi.org/10.1177/2332858419859843

Cizrelioğlu, İ., & Aydin, S. (2024). Using artificial intelligence in EFL learning among multilingual young learners. *In M. Sanmugam et al.* (eds.), *Fostering inclusive education with AI and emerging technologies* (pp. 303-322). IGI Global. https://doi.org/10.4018/979-8-3693-7255-5.ch013

Danaei, D., Jamali, H. R., Mansourian, Y., & Rastegarpour, H. (2020). Comparing reading comprehension between children reading augmented reality and print storybooks. *Computers & Education*, *153*, 103900. http://doi.org/10.1016/j.compedu.2020.103900

Derakhshan, A., Teo, T., Robat, E. S., Enayat, M. J., & Jahanbakhsh, A. A. (2024, May 1). Robot-Assisted Language Learning: a Meta-Analysis. *Sage Journals*. Retrieved from https://journals.sagepub.com/doi/10.3102/00346543241247227

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System, 121*, 103254. https://doi.org/10.1016/j.system.2024.103254

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico. *Education in the Knowledge Society, 24*, e31279. https://doi.org/10.14201/eks.31279

Gençten, V. Y., & Aydemir, F. (2023). Technology-Assisted Interactive Reading Activities In Early Childhood Education. *International Journal of Education Technology and Scientific Researches, 24,* 2390-2416. https://doi.org/10.35826/ijetsar.676

Ghoneim, N. M. M., & Elghotmy, H. E. A. (2021). Using an artificial intelligence based program to enhance primary stage pupils' EFL listening skills. *Sohag University International Journal of Educational Research, 83*(1), 1-32. https://doi.org/10.21608/edusohag.2021.140694

Godwin-Jones, R. (2014). Games in language learning: Opportunities and challenges. *Language Learning & Technology 18*(2), 9-19. Retrieved from http://llt.msu.edu/issues/june2014/emerging.pdf

Gordon, G., & Breazeal, C. (2015). Bayesian Active Learning-Based Robot Tutor for children's Word-Reading skills. *Proceedings of the . . . AAAI Conference on Artificial Intelligence, 29*(1). https://doi.org/10.1609/aaai.v29i1.9376

Gordon, G., Spaulding, S., Westlund, J. K., Lee, J. J., Plummer, L., Martinez, M., Das, M., & Breazeal, C. (2016). Affective personalization of a social robot tutor for children's second language skills. *Proceedings of the . . . AAAI Conference on Artificial Intelligence, 30*(1). https://doi.org/10.1609/aaai.v30i1.9914

Holmes, W. (2024, March 22). *The Unintended Consequences of Artificial Intelligence and Education*. Education International. Retrieved from https://eiie.io/2023AInEdu

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student learning benefits of a Mixed-Reality Teacher Awareness tool in AI-Enhanced classrooms. In C. Penstein Rosé et al. (eds.), *Lecture notes in computer science* (pp. 154-168). https://doi.org/10.1007/978-3-319-93843-1_12

Hua, C., & Wang, J. (2023). Virtual reality-assisted language learning: A follow-up review (2018–2022). *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1153642

Junaidi, J., Hamuddin, B., Julita, K., Rahman, F., & T. Derin. (2020). Artificial Intelligence in EFL Context: Rising Students' Speaking Performance with Lyra Virtual Assistance. *International Journal of Advanced Science and Technology, 29*(05), 6735-6741. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/17726

Kang, H. (2022). Effects of Artificial Intelligence (AI) and Native Speaker Interlocutors on ESL Learners' Speaking Ability and Affective Aspects. *Multimedia-Assisted Language Learning, 25*(2), 9-43. https://doi.org/10. 15702/mall.2022.25.2.09

Kayumova, A., & Sadykova, G. (2019). Implementing a multimodal online program for very young learners of Russian: Educators' perspective. *Indonesian Journal of Applied Linguistics, 9*(1). 1-13. https://doi.org/10.17509/ijal.v9i1.13323

Keerthiwansha, N. W. B. S. (2018). Artificial Intelligence Education (AIEd) in English as a Second Language (ESL) Classroom in Sri Lanka. *International Journal of Conceptions on Computing and Information Technology, 18*(6), 31-36.

Kervin, L. K., Verenikina, I., & Rivera, M. C. (2015). Collaborative onscreen and offscreen play: examining meaning-making complexities. *Digital Culture & Education, 7*(2), 228. Retrieved from https://ro.uow.edu.au/sspapers/1946/

Kewalramani, S., Palaiologou, I., Dardanou, M., Allen, K., & Phillipson, S. (2021). Using robotic toys in early childhood education to support children's social and emotional competencies. *Australasian Journal of Early Childhood, 46*(4), 355-369. https://doi.org/10.1177/18369391211056668

Korat, O., Shamir, A., & Heibal, S. (2013). Expanding the boundaries of shared book reading: Ebooks and printed books in parent–child reading as support for children's language. *First Language, 33*(5), 504-523. https://doi.org/10.1177/0142723713503148

Kory-Westlund, J. M., & Breazeal, C. (2019). Exploring the effects of a social robot's speech entrainment and backstory on young children's emotion, rapport, relationship, and learning. *Frontiers in Robotics and AI, 6*. https://doi.org/10.3389/frobt.2019.00054

Lee, J. H., Shin, D., & Noh, W. (2023). Artificial Intelligence-Based Content Generator technology for young English-as-a-Foreign-Language learners' reading enjoyment. *RELC Journal, 54*(2), 508-516. https://doi.org/10.1177/00336882231165060

Liu, C., Liao, M., Chang, C., & Lin, H. (2022). An analysis of children' interaction with an AI chatbot and its impact on their interest in reading. *Computers and Education/Computers & Education, 189*, 104576. https://doi.org/10.1016/j.compedu.2022.104576

Mahmud, F. A. (2023). Investigating EFL Students' writing Skills Through Artificial Intelligence: Wordtune application as a tool. *Journal of Language Teaching and Research, 14*(5), 1395-1404. https://doi.org/10.17507/jltr.1405.28

Meskill, C. (2007). Producerly texts: Implications for language in education. *Journal of Language and Education, 21*(2), 95-106.

Meskill, C. (2021). *Digital Screen Mediation in Education: Authentic and Agentive Technology Practices for Teaching and Learning* (1st ed.). Routledge. https://doi.org/10.4324/9781003013594

Meskill, C. (2024) Digital screen mediate on in language learning and teaching. In *The Bloomsbury Handbook of Language Learning and Technology* (pp 173-184). Bloomsbury.

Meskill, C., & Anthony, N. (2007). Learning to orchestrate online instructional conversations: A case of faculty development for foreign language educators. *Computer Assisted Language Learning, 20*(1), 5-19. https://doi.org/10.1080/09588220601118487

Meskill, C., Sadykova, G., & Kayumova, A. (2020). Mediating digital screens with very young emerging bilinguals. *Bilingual Research Journal, 43*(2), 137-156. https://doi.org/10.1080/15235882.2020.1743383

Movellan, J. R., Eckhardt, M., Virnes, M., & Rodriguez, A. R. (2009). Sociable robot improves toddler vocabulary skills. *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction (HRI '09)*. https://doi.org/10.1145/1514095.1514189

Patton, M. Q. (2001). *Qualitative research & evaluation methods : integrating theory and practice*. CiNii Books. Retrieved from http://ci.nii.ac.jp/ncid/BB18275167

Pifarré, M. (2019). Using interactive technologies to promote a dialogic space for creating collaboratively: A study in secondary education. *Thinking Skills and Creativity*, *32*, 1-16. https://doi.org/10.1016/j.tsc.2019.01.004

Redondo, B., Cózar-Gutiérrez, R., González-Calero, J.A., & Sánchez Ruiz, R. (2020). Integration of Augmented Reality in the Teaching of English as a Foreign Language in Early Childhood Education. *Early Childhood Education Journal, 48*, 147-155. https://doi.org/10.1007/s10643-019-00999-5

Revealing Reality. (n.d.). Children's Media Lives Wave 10. In *Children's Media Lives* (pp. 1–71) [Report]. https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/children/children-media-use-and-attitudes-2024/childrens-media-lives-2024-summary-report.pdf?v=367549

Sadykova, G., & Kayumova, A. (2024). Educators' perception of artificial intelligence as instructional tool. *TEM Journal*, 3194-3204. https://doi.org/10.18421/tem134-54

Sadykova, G., Gimaletdinova, G., Khalitova, L., & Kayumova, A. (2016). Integrating mobile technologies into very young second language learners' curriculum. In *CALL communities and culture – short papers from EUROCALL 2016* (pp. 408–412). https://doi.org/10.14705/rpnet.2016.eurocall2016.597

Singer, E. (2013). Play and playfulness, basic features of early childhood education. *European Early Childhood Education Research Journal*, *21*(2), 172-184. https://doi.org/10.1080/1350293X.2013.789198

Su, J., & Yang, W. (2022). Artificial intelligence in early childhood education: A scoping review. Computers and Education. *Artificial Intelligence, 3*, 100049. https://doi.org/10.1016/j.caeai.2022.100049

Su, J., & Zhong, Y. (2022). Artificial Intelligence (AI) in early childhood education: Curriculum design and future directions. *Computers and Education. Artificial Intelligence, 3*, 100072. https://doi.org/10.1016/j.caeai.2022.100072

Suhendra, N.N., Antik Sari, N.T., & D. Wahyudin. (2024). The Effect of Using the Duolingo Application on Elementary School Students' Mastery of English Vocabulary. *Educational Studies and Research Journal*, *1*(2), 93-98. https://doi.org/10.37147/eltr.v4i2.71

Tai, T., & Chen, H. H. (2022). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, 1-28. https://doi.org/10.1080/09588221.2022.2040536

Teichert, L. (2017). To Digital or not to Digital: How Mothers are Navigating the Digital World with their Young Children. *Language and Literacy*, *19*(1), 63. https://doi.org/10.20360/g22p5w

Tran, T. T. H. (2024). AI tools in teaching and Learning English Academic writing skills. *Proceedings of the AsiaCALL International Conference, 4*, 170-187. https://doi.org/10.54855/paic.23413

Underwood, J. (2017). Exploring AI language assistants with primary EFL students. In *CALL in a climate of change: adapting to turbulent global conditions – short papers from EUROCALL 2017* (pp. 317-321). https://doi.org/10.14705/rpnet.2017.eurocall2017.733

Vygotsky, L. (1978). *Mind in Society: the development of higher psychological processes*. Retrieved from https://ci.nii.ac.jp/ncid/BA03570814

Wale, B. D., & Kassahun, Y. F. (2024). The Transformative Power of AI Writing Technologies: Enhancing EFL Writing Instruction through the Integrative Use of Writerly and Google Docs. *Human Behavior and Emerging Technologies, 2024*(1), 9221377. https://doi.org/10.1155/2024/9221377

Wang, Z. (2022). Computer-assisted EFL writing and evaluations based on artificial intelligence: a case from a college reading and writing course. *Library Hi Tech*, *40*(1), 80-97.

Yang, G., Chen, Y., Zheng, X., & Hwang, G. (2020). From experiencing to expressing: A virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement. *British Journal of Educational Technology*, *52*(2), 807-823. https://doi.org/10.1111/bjet.13056

Yelland, N. J. (2018). A pedagogy of multiliteracies: Young children and multimodal learning with tablets. *British Journal of Educational Technology, 49*(5), 847-858. https://doi.org/10.1111/bjet.12635

Zhao, X. (2022). Leveraging Artificial intelligence (AI) technology for English Writing: Introducing Wordtune as a Digital writing assistant for EFL writers. *RELC Journal*, *54*(3), 890-894. https://doi.org/10.1177/00336882221094089

Zou, B., Liviero, S., Hao, M., & Wei, C. (2020). Artificial Intelligence Technology for EAP Speaking Skills: Student perceptions of opportunities and challenges. In *Springer eBooks* (pp. 433–463). Springer International Publishing. https://doi.org/10.1007/978-3-030-34212-8_17

Gulnara Sadykova: PhD, Associate Professor, Department of Romance and Germanic Philology, Kazan Federal University (Russia). Research interests include technology-assisted language teaching/learning, artificial intelligence in language education, early childhood education, and language acquisition. She is the Principal Investigator of a grant project on the use of AI in preschool language education.

Albina Kayumova: Associate Professor, Department of Romance and Germanic Philology, Kazan Federal University (Russia). Research interests include technology-assisted language teaching/learning, artificial intelligence in language education, and early multilingualism. She is the head of the master's program *Multilingual Technologies of Early Childhood Development* at Kazan Federal University.