Young Learners and AI: Deciding When to Bot

Tatiana Kozlova

Sapienza University of Rome, Italy

Abstract

The AI-powered technologies, and chatbots in particular, are increasingly promoted as solutions to enduring challenges in English language teaching (ELT). While these tools are often praised for their capacity to deliver interactive and personalised learning experiences alongside traditional instruction, their rapid uptake in educational contexts has outpaced critical reflection on pedagogical readiness. The growing imperative to adopt AI raises fundamental questions about whether teachers and learners are sufficiently prepared - linguistically, cognitively, and emotionally - to engage productively with such technologies. Rather than assuming that earlier exposure is inherently beneficial, this study interrogates when *and* how *AI should be introduced to support meaningful language learning. Drawing on an on-site empirical investigation, the study examines the use of AI in English-medium Content and Language Integrated Learning (CLIL) classrooms, focusing on history and art in a bilingual school in Rome, Italy. It analyses interactions between primary (Grade 5) and lower secondary (Grades 6–7) students and AI-based tutoring tools on the SchoolAI platform. Through qualitative analysis of student–chatbot exchanges, the research critically explores how learners negotiate meaning with AI, revealing both the pedagogical affordances and the constraints of AI-mediated interaction in ELT settings. The presentation situates the study within current debates on AI in language education, challenging techno-optimistic narratives by foregrounding classroom realities. It outlines the research methodology and key findings, supported by illustrative excerpts from student-AI interactions, and concludes with a critical reflection on the implications for curriculum design and teacher decision-making*

Keywords: *AI in Education; CLIL; AI-powered tutors; Student-AI interaction; Personalised learning; Educational technology*

1. Introduction

The rapid integration of artificial intelligence into education has generated both enthusiasm and unease. AI-powered tools, especially chatbots, are increasingly framed as solutions to persistent pedagogical challenges, including learner autonomy and motivation, personalised feedback and scalable support [5], [6], [8], [20]. In English Language Teaching (ELT) and Content and Language Integrated Learning (CLIL) contexts, such technologies are often positioned as especially valuable, offering learners additional conversational practice and linguistic scaffolding while engaging with subject-specific content [1], [14].
However, the pace of technological adoption has arguably outstripped critical pedagogical reflection and there remains a lack of empirical research on the use of chatbots in ELT [16]. While AI tools are widely promoted as enhancing learning outcomes, less attention has been paid to the interaction process itself in order to understand whether these tools are genuinely pedagogically appropriate for young learners. Some concerns have already emerged regarding the standardisation of language, over-reliance on technology, data safety, ethics and responsibility [5]. This paper argues that the focus should shift from asking whether AI “works" for learning to investigating for whom, when and under what conditions it might be beneficial.
In Italy, the Guidelines for the Introduction of Artificial Intelligence in Educational Institutions recommend, while deploying AI in the support of students, to "take into account the different levels of development related to age" [15, p. 21]. This raises a central question: when are learners developmentally, linguistically and cognitively ready to benefit from AI-mediated interaction. Is it appropriate to introduce AI in primary school or is it preferable to wait until lower secondary school?
This study addresses this question through an empirical investigation conducted in a bilingual school in Rome, Italy. Focusing on primary (Grade 5) and lower secondary (Grades 6-7) CLIL classrooms, the research examines how learners interact with AI-powered tutors on the SchoolAI platform. Rather than evaluating AI solely in terms of efficiency or accuracy alone, the study adopts Vygotsky’s sociocultural perspective [17], [18] to explore the nature of artificial mediation and how meaning is negotiated in student-AI interaction. By analysing authentic classroom exchanges between learners and AI systems, this paper seeks to move beyond techno-optimistic narratives and instead foreground the complexities of AI integration in real educational contexts with young learners. It argues that the effectiveness of AI

in ELT is not determined solely by technical capability but by learners' developmental readiness and the pedagogical frameworks within which these tools are embedded.

2. Literature Review and Theoretical Framework

2.1. Growth of AI in Contemporary ELT Discourse

Recent large-scale research syntheses [7], [9], [13] confirm both the growth and diversification of AI implementation in second language education. A systematic review of AI-driven chatbots in this field [7] highlights an increase in empirical studies since 2020, reflecting the growing institutional and commercial interest in AI-supported learning environments. These tools are frequently associated with affordances such as increased learner autonomy and motivation, real-time feedback and extended opportunities for practice beyond the classroom [20].

At the same time, meta-analytic evidence [13], [19] provides a more detailed picture of their impact. In a comprehensive meta-analysis of chatbot-assisted language learning, Wang *et al.* [19] report a moderate positive effect on language learning performance (g=0,484), based on 70 effect sizes across 28 studies. However, the analysis also identifies substantial variability across contexts, suggesting that effectiveness is contingent on multiple interacting factors, including learner characteristics and instructional design. Following this line of reasoning Lyu *et al*. [13, p. 834], suggest that "a more refined perspective is needed when discussing the benefits of the use of chatbots in language learning process".

2.2. Evidence of Effectiveness

While a growing body of research report generally positive outcomes associated with chatbot use, these findings must be interpreted with caution. The meta-analysis by Wang *et al.* [19] demonstrates that chatbot-assisted learning can enhance language performance overall; however, the strength of this effect is moderated by several key variables, such as educational level, language level, interface design and interaction capability. This suggests that AI tools do not operate as universally effective pedagogical solutions but rather as context-dependent resources. Similarly, systematic reviews [7], [9], [13] emphasise that improvements are often observed in specific domains, particularly speaking, writing and listening, where opportunities for practice and immediate feedback are crucial. At the same time, these studies note that many empirical investigations are conducted in controlled or experimental settings, which may not fully reflect the complexities of real classroom environments.

A key limitation emerging from this body of research is its predominant focus on measurable outcomes, such as test scores or task completion, rather than on the processes through which learning occurs. In particular, relatively little attention has been paid to how developmental factors shape these interactions. As a result, claims regarding effectiveness often remain insufficiently grounded in an understanding of how learners actually use and interpret AI-generated input.

2.3. Chatbots as Interactional Tools

Within ELT, chatbots are frequently conceptualised as conversational partners capable of simulating authentic interaction. In their study Belda-Medina and Calvo-Ferrer [2] describe chatbots as AI-driven interlocutors that can engage learners in dialogue, thereby creating opportunities for communicative practice that might otherwise be limited in traditional classroom setting. This framing aligns with communicative and interactionist approaches to language learning, which emphasise the importance of meaningful interaction is developing linguistic competence.

However, the notion of chatbots as "partners" requires closer scrutiny. Unlike human interlocutors, AI systems do not possess genuine communicative intent. While they can generate contextually appropriate responses, these are ultimately based on algorithmic patterns rather that a shared understanding. As a result, interaction with chatbots may differ qualitatively from human interaction, particularly in terms of feedback and co-construction of meaning. This distinction becomes especially relevant when considering younger or less proficient learners whose limited metapragmatic and critical awareness may render them especially susceptible to forms of semiotic regulation embedded in AI-mediated discourse.

2.4. The CLIL Context and AI Integration

The application of AI in CLIL contexts remains relatively underexplored, despite its potential relevance. CLIL environments which integrate content and language learning, place additional cognitive and linguistic demands on learners, making them a particularly interesting site for examining the role of AI-mediated support.

Existing studies on AI chatbots in CLIL settings indicate that such tools can support both content comprehension and language development by providing explanations, prompts and feedback tailored to subject-specific tasks [1], [14]. However, these studies tend to focus primarily on the design and affordances of AI systems rather than on the process of their actual use in classroom interaction.The extent to which AI can effectively scaffold this dual demand remains unclear, particularly for young learners whose cognitive and linguistic resources are still developing.

2.5. Critical Gap

Despite the growing body of research on AI in ELT. Several critical gaps remain. First, there is a notable lack of attention to developmental factors. While studies such as Wang et al. [19] identify educational level as a moderating variable, few investigations explicitly examine how learners' cognitive, linguistic and metacognitive readiness shapes their ability to engage productively with AI systems. This is particularly important in research involving young learners. Where developmental differences may significantly influence interactional outcomes.

Second, much of the existing literature is not grounded in sociocultural approaches to learning. Drawing on Lev Vygotsky's work, learning is understood as a mediated process that occurs through interaction with more knowledgeable others within the learner's Zone of Proximal Development. From this perspective, the effectiveness of any instructional tool, including AI, depends on its capacity to provide appropriate scaffolding. However, current research rarely interrogates whether chatbots can genuinely function as mediational agents comparable to human interlocutors. This concern stems from sociocultural approaches to language learning [10], [11] [12], which argue that language development emerges through socially situated interactions rather than through exposure alone, placing a particular emphasis on the quality of interaction

Finally, there is a lack of fine-grained classroom discourse analysis. Much of the literature relies on quantitative measures or self-reported data, offering limited insight into the micro-level dynamics of student AI-interaction. As a result, important questions remain unanswered: How do learners formulate queries to AI? How do they interpret and respond to AI generated feedback? To what extent do these interactions support or hinder the negotiation of meaning?

Addressing these gaps requires a shift in focus from outcomes to processes and from technological affordances to pedagogical realities. The present study responds to this need to examining student-AI interaction in CLIL classrooms through a sociocultural lens, with particular attention to developmental readiness and the conditions under which AI can meaningfully support learning.

3. Methodology

The study adopts a qualitative classroom-based design, which focuses on interactional process. A qualitative approach was selected because the study attempted to understand not simply whether AI "works" but how AI-mediated interaction unfolds in authentic educational setting. The participants interacted with AI-tutors once a week for 30 minutes during four weeks in November, the school year 2025-2026.

Data were analysed using reflexive thematic analysis [3], [4], in which recurring interactional patterns and learner behaviours were identified across the dataset. MAXQDA software was used to facilitate coding and codes were subsequently grouped into broader analytical themes reflecting the affordances and limitations of AI-mediated interaction within CLIL environments.

The analysis was informed by a sociocultural theory [17], [18] and its application in second language development [10], [11] [12], Consequently, coding focused on themes related to mediation, scaffolding, meaning negotiation and interaction within the ZPD.

The analysis proceeded through six recursive phases:

1. Familiarisation with the data;
2. Initial coding of student -AI interactions:
3. Identification of recurring interactional patterns, mediation episodes, misunderstanding, scaffolding, prompt formulation strategies;
4. Development and refinement of themes;
5. Thematic interpretation in relation to sociocultural theory;

6. Final synthesis in relation to sociocultural theory.

Rather than examining interaction solely at the linguistic level, the analysis aimed to understand how learners engaged with AI tutors during CLIL tasks and how developmental readiness influenced the quality of interaction.

3.1. Research Context and Participants

The study was conducted in a bilingual school in Rome where English is used as a medium of instruction for selected CLIL subjects. The participants included 36 students from one primary (Grade 5) and two lower secondary classrooms (Grades 6 and 7), representing learners between 9 and 13 years of age. Their linguistic level was from A1 to B1.

3.2. AI Platform and Learning Tasks

Students interacted with AI history tutor and AI art tutor on School AI platform during classroom activities designed by the teacher. The chatbot platform was used to support content exploration, answer questions, provide explanation and assist learners in completing subject related tasks in English.
Activities included asking chatbot for explanation of historical or artistic concepts, summarising information, generating ideas and clarifying vocabulary. The AI tool functioned as a supplementary instructional resource rather than a replacement for teacher-led interaction.

3.3. Data Collection

Data were collected over four classroom sessions through chat logs, classroom observation notes and focus groups. The primary dataset consisted of chat logs. These transcripts, which documented exchanges between learner and the chatbots, allowed for closer examination of how students formulated prompts, interpreted AI responses and interactional dynamics that emerged from that. Classroom observations were used to contextualise the interactional data by recording learner behavior, moments of confusion, patterns of engagement and teacher intervention during activities. Focus groups were used to clarify the learners' perception of these activities.

4. Findings and Analysis

The analysis of the chat logs reveals a significant mismatch between AI conversational design and young learners' developmental readiness within CLIL environments. Several recurring themes emerged from the data.

4.1 AI Sustains Interaction but Not Necessarily Learning

One of the most prominent findings concerns the chatbot's strong ability to sustain interaction. The AI consistently maintained conversations, redirected students and prolonged engagement; however, these interactions often became disconnected from curricular objectives.
For example:
Student: "do you like pizza"
AI tutor: "How do you think simple pleasures… can be likened to art?"
Student: "do you like spaghetti"

Although the interaction continued linguistically, it drifted pedagogically away from CLIL goals. Conversational continuity was therefore frequently mistaken for meaningful educational engagement.
This illusion of "meaningful dialogue" appeared repeatedly throughout the dataset. The chatbot generated sophisticated, coherent and reflective responses, yet learners often did not engage with the conceptual content, ignored reflective prompts, shifted topics rapidly and produced minimal responses.
For example**:**
Student: "do you like tik tok "
AI tutor: "I appreciate the curiosity and creativity it inspires. How do you think artists can use it to share their creations? "
Student: "tik tok on the clock nanana youuu youuu"

These patterns suggest surface-level interaction rather than negotiated meaning. Interpreted through the lens of Sociocultural Theory, the findings indicate that interaction alone does not necessarily guarantee mediation or learning.

4.2 Developmental Readiness and Prompting Ability

The data also suggest that productive AI interaction depended significantly on learners' metacognitive and communicative maturity. Older or more reflective learners were more likely to sustain topics, elaborate ideas, engage conceptually and connect disciplines. By contrast, younger or less focused learners tended to test boundaries, ask random or highly social questions, produce fragmented prompts or disengage from the educational framing of the activity. These findings indicate that developmental readiness strongly influenced the quality and educational value of AI-mediated interaction.

4.3 Redirection

Another recurring theme concerns the AI's capacity for unlimited patience and continuous redirection. The chatbot never rejected interaction, when deviations from the topic occurred it reframed conversation educationally and persistently redirected students toward reflective discussion.
Examples of inappropriate or off-task behaviour by pupils that required redirection included profanity, nonsense text, trolling behaviour and repetitive off-task questioning. Nevertheless, the AI continually attempted to redirect students toward topics related to art, history, creativity, or reflection. This characteristic created a low-anxiety environment that encouraged participation and emotional safety. However, while this persistence represents an important affordance, the AI was unable to meaningfully distinguish between productive engagement and merely performative participation.

4.4 Teacher Mediation Remains Crucial

The need for teacher mediation emerged consistently across all interactions. The AI could not accurately assess understanding, frequently overestimated engagement, rewarded superficial participation and repeatedly generated generic reflective prompts. Furthermore, the automated summaries generated for teachers often described students as demonstrating "high engagement", "interdisciplinary thinking " or "apprentice artist" qualities, even in cases where students trolled, avoided the task, disengaged, demonstrated misunderstanding. AI-generated evaluations therefore frequently interpreted conversational participation as evidence of conceptual learning. These findings reinforce the continued necessity of teacher interpretation, scaffolding and more fine-grained assessment within AI-supported learning environment.

4.5. Productive Moments of AI-Mediated Learning

Despite these limitations, the data also contained several productive moments of AI-mediated learning. In such cases, the AI successfully sustained exploratory dialogue and encouraged interdisciplinary thinking. However, these successful interactions generally occurred when learners already possessed conversational stamina, reflective capacity and were genuinely interested in the topic.
Another significant affordance concerned multilingual learners. In several interactions, students communicated using imperfect English, occasional code-switching, and simplified grammatical structures, yet they were still able to sustain lengthy intellectual conversations because the AI adapted flexibly and nonjudgmentally without harsh correction.
These findings suggest that conversational AI may reduce anxiety for language learners by providing interaction that is patient, responsive and supportive.
Additionally, students frequently used the AI not only for factual discussion but also for emotional reflection. This tendency became particularly evident in some students, where conversations during art-related discussions evolved into topics such as anger management, emotional expression through painting, colour psychology and self-image.
For example, the student stated:
"I tried to paint to let my anger out and it WORKED!"
This is significant because it demonstrates the AI functioning not merely as an informational tool, but also as an emotional mirror and reflective conversational partner. Such interactions suggest the potential for AI-mediated dialogue to extend beyond factual learning into deeper forms of personal reflection and self-expression.

Discussion

From a socio-cultural perspective, the findings suggest that the AI frequently failed to operate effectively within learners' Zone of Proximal Development (ZPD), generating lengthy philosophical inquires that did not resonate with young learners. Although the chatbot maintained continuous interaction and generated sophisticated reflective prompts, these prompts often exceeded learners' developmental readiness and communicative capacity. As a result, students frequently ignored questions, shifted topics abruptly or disengaged from conceptual exploration.
Unlike human teachers, the AI was unable to dynamically recognise misunderstanding, confusion, or cognitive overload in ways that could meaningfully adjust the interactional trajectory. Instead, conversational persistence frequently replaced genuine scaffolding. The chatbot continued producing elaborated reflective responses regardless of whether learners demonstrated comprehension, conceptual engagement or readiness to sustain the discussion. This finding is significant because socio-cultural theory emphasises that learning occurs not through interaction alone, but through appropriately mediated interaction within the learner's developmental capacity. In many cases within the dataset, the AI maintained linguistic participation without successfully supporting conceptual progression. Consequently, the interaction often simulated educational dialogue without fully achieving scaffolded learning.

Conclusion

The findings therefore suggest that conversational AI currently lacks the adaptive pedagogical sensitivity required to consistently function as an effective scaffold within learners' Zone of Proximal Development, particularly for younger learners or students with limited metacognitive and communicative maturity. The older the learners are the more likely they are to benefit from the interaction with AI. While AI may support low-anxiety interaction, exploratory dialogue and sustained conversational participation, human teacher mediation remains essential for recognising learner readiness and transforming interaction into meaningful educational development.

REFERENCES

[1] Andujar A., Spratt M., "Using AI to support CLIL teacher language", Journal of Research in Applied Linguistics, pp. 7–19, 2023.
[2] Belda-Medina, J., Calvo-Ferrer, J. R. Using Chatbots as AI Conversational Partners in Language Learning. Applied Sciences, 12(17), 8427, 2022.
[3] Braun V., Clarke V., "Using thematic analysis in psychology", 2006.
[4] Braun V., Clarke V., "Thematic analysis: A practical guide", 2021.
[5] Creely, Edwin. Exploring the role of generative AI in enhancing language learning: opportunities and challenges. International Journal of Changes in Education, 2024.
[6] Crompton, H., Edmett, A., Ichaporia, N., Burke, D. AI and English language teaching: affordances and challenges. British Journal of Educational Technology, 5(6), 2024.
[7] Jeon J., Lee S., Choe H., "Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning", Computers & Education, 2023.
[8] Kim N.-Y., "Chatbots and language learning: Effects of the use of AI chatbots for EFL learning", Eliva Press, 2020.
[9] Koç F.Ş., Savaş P., "The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024", ReCALL, pp. 4–21, 2025.
[10] Lantolf J.P. (ed.), "Sociocultural theory and second language learning", Oxford, Oxford University Press, 2000.
[11] Lantolf J.P., Poehner M.E., "Sociocultural theory and the pedagogical imperative in L2 education", London, Routledge, 2014.
[12] Lantolf J.P., Thorne S.L., "Sociocultural theory and the genesis of second language development", Oxford, Oxford University Press, 2006.
[13] Lyu B., Lai C., Guo J., "Effectiveness of chatbots in improving language learning: A meta-analysis of comparative studies", International Journal of Applied Linguistics (35), pp. 834-851, 2024.
[14] Mageira K., Pittou D., Papasalouros A., Kotis K., Zangogianni P., Daradoumis T., "Educational AI chatbots for content and language integrated learning", Applied Sciences, pp. 32–39. 2022.

[15] Ministero dell'Istruzione e del Merito, Linee guida per l'introduzione dell'Intelligenza Artificiale nelle Istituzioni scolastiche, Versione 1.0, Roma, 2025.
[16] Polakova P., Klimova B., "Implementation of AI-driven technology into education – a pilot study on the use of chatbots in foreign language learning", Cogent Education, 2024.
[17] Vygotsky L., "Mind in society: The development of higher psychological processes", Cambridge, Harvard University Press, 1978.
[18] Vygotsky L., "Thought and language", Cambridge, MIT Press, 1986.
[19] Wang F., Cheung A.C.K., Neitzel A.J., Chai C.-S., "Does chatting with chatbots improve language learning performance? A meta-analysis of chatbot-assisted language learning", Review of Educational Research, 2024, pp. 623–660.
[20] Wei, L. Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. Frontiers in Psychology, 14, 2023.