

INTEGRATING GENERATIVE CHATBOTS IN SECONDARY ENGLISH TEACHING: A 2D3S4L MODEL FOR PEDAGOGICAL DESIGN

Wei Zhang, Dzintra Iliško, & Lipolelo Thamae
Educational Science, Daugavpils University (Latvia)

Abstract

This paper presents the 2D3S4L instructional model, designed to embed generative chatbots into secondary English classrooms with pedagogical coherence. The model aligns teaching and learning processes across two dimensions (teachers and students), three stages of instruction (pre-class, in-class, post-class), and four recursive loops (analyze, design, implement, evaluate). Grounded in sociocultural theory and the TPACK framework, the model was implemented in Chinese secondary schools using deep seek and Doubao AI to support writing, speaking, and reflective reading tasks. Findings show chatbots enhanced idea generation, linguistic accuracy, and learner autonomy, while teacher prompts and scaffolded feedback were critical for effective integration. Concrete examples illustrate how chatbots facilitate task execution and deepen reflection. The paper contributes a structured framework and practical guidance for future AI-enhanced language instruction.

Keywords: *Generative AI, secondary English, chatbot pedagogy, instructional design, task-based learning.*

1. Introduction

Generative artificial intelligence (AI) has opened new horizons for English language education. Tools like Deep Seek and Doubao AI offer learners real-time interaction, idea scaffolding, and language refinement. These capabilities reflect a broader trend in language education where AI increasingly supports diverse tasks from vocabulary practice to feedback generation (Zou & Wang, 2023; Wang&Dang, 2024). Despite these affordances, integration into secondary classrooms remains inconsistent, often driven by novelty rather than pedagogy. Without structured guidance, students may misuse chatbots as answer generators rather than dialogic partners. Teachers, in turn, lack frameworks to align chatbot capabilities with instructional goals. This paper responds by introducing the 2D3S4L model—a structured approach to pedagogically embed chatbots into language instruction. The model integrates two key actors (teachers and students), three temporal stages (pre-, in-, and post-class), and four recursive design loops (analyze, design, implement, evaluate). It is grounded in sociocultural theory and the TPACK framework, supporting intentional design and adaptive use of generative AI in classroom practice.

2. Objectives

This study explores how generative chatbots can support task-based English instruction through the 2D3S4L model. Objectives include: (1) to present a theoretically grounded and pedagogically actionable design framework; (2) to illustrate concrete teaching applications across writing, speaking, and reading tasks; and (3) to evaluate the roles chatbots play in supporting learner autonomy, language development, and metacognitive reflection.

3. Methods

To ensure transparency and rigor, the thematic analysis process drew upon Braun and Clarke's (2006) six-phase approach: familiarization, initial coding, theme generation, review, definition, and final reporting. Throughout this process, data from over 60 student artifacts, 18 student interviews, and 6 teacher observation logs were analyzed. All interviews and observation notes were transcribed verbatim. Student artifacts (e.g., written texts, chatbot interactions, annotated drafts) were anonymized and categorized by task type and grade level. We developed an initial coding framework based on Braun and Clarke's (2006) thematic analysis approach, beginning with open coding during familiarization. Two researchers

independently applied this framework to a subset of the data to calibrate consistency, followed by iterative refinement through discussion. Discrepancies were resolved through negotiated agreement. All remaining data were then manually coded using the agreed framework, with emergent themes identified inductively across the three task types. NVivo software was used to organize codes, support constant comparison, and ensure transparency in analytic decisions.

We employed a design-based research (DBR) methodology to develop and refine the 2D3S4L model in authentic classroom settings. Design-based research enables iterative, real-world intervention while maintaining theoretical grounding (Wang & Hannafin, 2005). The study involved three public secondary schools in China, comprising six English teachers and approximately 120 students across Grades 7 to 9. Teachers participated in a pre-study workshop on chatbot tools, prompt design, and pedagogical alignment. Tasks were co-constructed by researchers and teachers to embed chatbot interaction into curriculum-based learning goals. Three task types were implemented: opinion writing, tourist role-play speaking, and reflective reading. Chatbot use occurred during all instructional phases: idea activation, drafting, revision, feedback, and post-task reflection. Data included student texts, chatbot transcripts, classroom observations, and post-task interviews. A thematic analysis approach identified patterns in language outcomes, student attitudes, and pedagogical adaptations.

Two chatbot platforms were utilized with differentiated functions. DeepSeek was employed primarily for writing and reflection tasks, valued for its elaborative and coherent textual output. In contrast, Doubao AI supported spoken interactions, offering real-time pronunciation and fluency feedback suitable for dialogic practice.

While this study did not include a formal control group, learning gains were inferred from pre- and post-task comparisons, supported by triangulated data from student self-reports, teacher observations, and rubric-based assessments.

While both DeepSeek and Doubao AI were employed, they supported different instructional goals. DeepSeek was used primarily for extended writing and reflective discussion due to its elaborative textual responses, whereas Doubao AI's real-time speech feedback was better suited for spoken language practice.

4. The 2D3S4L instructional model

The 2D3S4L model is grounded in two core theoretical frameworks. First, sociocultural theory (Vygotsky, 1978) emphasizes the role of social interaction and mediated tools in learning. In this model, the generative chatbot is viewed as a semiotic tool that scaffolds learners' cognitive and linguistic development through dialogic engagement. Second, the Technological Pedagogical Content Knowledge (TPACK) framework (Mishra & Koehler, 2006; Koehler & Mishra, 2009) provides a structure for integrating AI tools into curriculum design, ensuring that technology use supports both content mastery and pedagogical aims (Mishra, 2023). These two perspectives collectively inform the structure of the model and guide its classroom implementation.”

The model consists of three layers. First, the two dimensions: (1) teacher as designer and facilitator; (2) student as autonomous language generator. Second, the three stages: (a) pre-class chatbot priming (e.g., vocabulary activation, brainstorming); (b) in-class AI-supported performance (e.g., speaking drills, writing generation); (c) post-class metacognitive reflection (e.g., annotation, revision logs). Third, the recursive instructional loops—analyze learner needs, design chatbot prompts and activities, implement tasks, and evaluate both learning outcomes and chatbot behavior. Unlike tool-specific methods, 2D3S4L is platform-agnostic. Teachers adapt prompt complexity and chatbot roles to learner levels. The model views chatbots not as replacements for instruction but as dynamic tools that support negotiation of meaning, feedback literacy, and voice development.

5. Classroom implementation and case examples

5.1. Opinion writing task

In this writing task, students used the chatbot to generate outlines and retrieve relevant vocabulary during pre-class preparation. Teachers supported this by crafting guiding prompts and later encouraging students to annotate and reflect on AI-generated feedback. As a result, many students demonstrated improved fluency and coherence in their drafts. However, a common challenge emerged: several learners copied chatbot-generated phrases directly without contextual adaptation, prompting follow-up lessons on paraphrasing and voice development.

Teacher Strategy: Prompted prewriting with deepseek, and assigned reflection post-task.

Outcomes: Improved writing fluency and structure.

Challenges: Overreliance on copied phrases from chatbot.

Grade 7 students were assigned the task 'My Favorite Season.' Pre-class, the teacher crafted a prompt in ChatGPT that helped students generate outlines with reasons, examples, and vocabulary banks. During class, students expanded the outlines into full drafts with AI feedback on transitions, sentence fluency, and paragraph coherence. Post-class, they annotated chatbot edits, justifying revisions or rejections. While writing quality improved, several students copied phrases without contextual adaptation. Teachers responded by guiding discussions around language ownership and added a follow-up activity requiring students to rephrase chatbot suggestions in their own words. The task highlighted the model's reflection phase, encouraging metalinguistic awareness and critical engagement with AI suggestions (Eisenring et al., 2024).

5.2. Tourist role-play speaking task

For the speaking task, students interacted with Doubao AI to simulate tourist dialogues. The AI provided real-time feedback on pronunciation and sentence fluency. Teachers scaffolded these performances by reviewing recordings with students and highlighting moments of linguistic creativity. While the activity boosted learner confidence, it also revealed a tendency to mimic AI-generated structures, especially among less confident speakers.

Grade 8 students used Doubao AI to simulate tourist dialogues. Pre-task vocabulary review was followed by in-class chatbot interactions focused on giving directions and handling questions. Chatbot feedback included pronunciation correction and real-time fluency suggestions. Teachers observed increased confidence, but also noticed mimicry of AI structures. Post-task recordings were reviewed by students to evaluate conversational tone and complexity. In one case, a student modified AI-suggested sentences to include idiomatic expressions, demonstrating linguistic personalization. The teacher emphasized that AI outputs are scaffolds, not scripts, and required students to reflect on what made their version stronger or more authentic.

5.3. Reflective reading task

In the reading task, students engaged with chatbot-generated responses to questions about Mandela's biography. They were then encouraged to critique, modify, or extend the chatbot's answers through journaling and peer review. This led to thoughtful reinterpretations and deeper personal engagement. Still, some students initially accepted the chatbot's optimism uncritically, highlighting the need for explicit instruction in critical reading and stance formation.

Grade 9 students engaged with Mandela's biography and used deepseek to explore questions like 'What would Mandela advise modern leaders?' Students responded by editing, challenging, or extending the chatbot's answer. While most AI-generated responses emphasized optimism, several students critiqued this stance and included more nuanced socio-political contexts. One teacher commented that 'the AI helped launch discussion, but the real depth came from disagreement.' This task illustrated the chatbot's role as a dialogic partner, not an authority. It activated reflective reading and interpretation, with students refining their stances through post-task journaling and peer review.

To provide a clearer picture of how the 2D3S4L model functioned across the implemented tasks, Table 1 summarizes key contrasts in student-AI engagement, instructional strategies, observed learning outcomes, and implementation challenges.

Table 1. Comparative Summary of AI-integrated Tasks.

Task Type	Student Use	Teacher Strategy	Outcome	Challenge
Opinion Writing	Generate outline, vocabulary	Prompt design + revision reflection	Improved fluency	Copying AI phrases
Role-Play Speaking	Dialogue with AI, pronunciation feedback	Scaffold performance + playback	More confidence	Structural mimicry
Reflective Reading	Debate AI interpretation	Journaling + peer review	Stronger critical thinking	Adopting AI views uncritically

5.4. Thematic findings from classroom data

Thematic analysis revealed four recurring themes across the three tasks and grade levels. These themes captured students' cognitive, emotional, and behavioral engagement with chatbot-supported learning. Table 2 summarizes each theme, a brief description, representative evidence, and observed frequency.

Table 2.

Theme	Description	Sample Evidence	Frequency
AI as Confidence Booster	Chatbots lowered students' anxiety and encouraged language risk-taking.	"It doesn't laugh at my mistakes." (G8 student)	Reported by 13 of 18 interviewed students
Surface Dependence on AI Output	Students initially copied AI-generated text without personal adaptation.	Copied 'Spring is the season of rebirth...' from chatbot.	Found in 9 of 20 early writing samples
Importance of Teacher Scaffolding	Structured prompts and modeling led to better AI engagement.	Revised 'Mandela was a great leader' into more nuanced version.	Observed in all 3 classrooms
Reflection Promotes Personalization	Post-task reflection led to critique and deeper understanding.	"The AI missed the political side." (Student journal)	Found in 10+ annotated journals

6. Discussion

Findings suggest the 2D3S4L model supports meaningful chatbot integration in secondary English classes. Across tasks, students were observed to use chatbots for idea generation, draft expansion, and reflective evaluation. This aligns with recent research showing that generative AI enhances writing fluency and accuracy while also posing challenges for voice and authenticity (Wang & Dang, 2024). Teachers served as AI mediators—designing prompts, interpreting outputs, and helping students reframe responses. The cyclical structure promoted feedback literacy and learner agency. Though lacking a control group, progress was gauged via before/after trends and triangulated data.

In addition to qualitative reflections, selected student work was rated using a standardized rubric (adapted from CEFR writing descriptors). Among 24 writing samples, 18 showed improvements of at least one band level in coherence and lexical range, with average writing scores increasing from 3.1 to 4.2 (out of 5). In oral tasks, peer-assessed fluency ratings improved for 72% of participants. Post-task surveys indicated that 86% of students agreed AI feedback helped them notice weaknesses in their argumentation or organization.

Notably, a minority of participants (3 out of 24) demonstrated limited gains despite full access to AI tools. Follow-up interviews revealed contributing factors including digital unfamiliarity, lack of motivation, and passive engagement. This reinforces the need for adaptive onboarding, ongoing teacher monitoring, and personalized scaffolding strategies.

However, success depended on metacognitive scaffolding. Without explicit post-task analysis, students risked copying AI-generated structures. Metacognitive scaffolding is essential to help learners internalize AI suggestions rather than reproduce them mechanically (Mason et al., 2023; Barkley, & Major (2024). Classroom norms and feedback tasks played a vital role in transforming chatbot use from passive assistance to co-construction. The model proved adaptable across language skills and grade levels, affirming its utility in scaffolded language learning environments.

As AI becomes a classroom co-participant, its integration demands thoughtful pedagogical design. Observed classrooms showed students not only used AI output but learned to question and revise it. This triadic engagement—between self, peer, and machine—indicates a shift toward collaborative meaning-making. Institutionally, this evolution requires teacher training, cross-departmental planning, and policy support for sustainable practice.

7. Conclusions

The 2D3S4L model offers a robust framework for integrating chatbots into language instruction with pedagogical coherence. It aligns technology use with student needs and reflective learning cycles. Classroom implementation across three task types demonstrates the model's adaptability and potential to foster fluency, autonomy, and critical thinking. Future work should explore longitudinal impact, student-AI interaction styles, and cross-linguistic applicability. Building teacher capacity through targeted training in prompt design and AI critique remains essential.

8. Limitations and ethical considerations

This study was limited by sample size and geographical scope. Findings may not generalize to schools with different digital access or cultural contexts. In addition, AI novelty may have temporarily boosted student motivation. Ethical considerations also emerged: students occasionally treated chatbot output as truth without question, underscoring the need for AI literacy. Privacy concerns remain, particularly regarding data handling in cloud-based platforms. Ethical frameworks such as those proposed

by UNESCO (2021) and the OECD (2024) emphasize the importance of transparency, informed consent, and digital equity when deploying AI in educational contexts. Educators must critically assess tools before adoption and foster dialogue around AI responsibility and reliability in learning contexts. Additionally, although triangulation of data was applied, the coding process was conducted by researchers immersed in the Chinese education context, which may introduce interpretive bias. Student English proficiency levels varied across classes, potentially influencing the depth of chatbot engagement. These factors limit generalizability and highlight the need for replication in diverse educational settings.

Beyond individual performance, the model encouraged collaborative dialogue about AI. Students often discussed chatbot limitations, such as vague topic development or unnatural phrasing, and collectively improved on them. This positioned AI as a third-party reviewer—neither teacher nor peer—whose presence provoked both critique and creativity. Such interactions reflect emerging AI literacy and emphasize the value of critical digital citizenship in education.

Another significant insight relates to the role of teacher agency. While chatbots provided consistent feedback, it was the teacher's intervention—through prompt design, scaffolding, and feedback interpretation—that transformed AI interaction into genuine learning opportunities. Teachers noted that modeling how to question and negotiate with chatbot output helped students move beyond surface-level dependence. Ultimately, the model empowered teachers to reposition AI not as a replacement, but as a co-facilitator within a carefully orchestrated learning environment.

Acknowledgments

This article was completed during a time of deep emotional difficulty. At a point when finding inner peace seemed impossible, the café Ezīša Kofišops became a place of calm and comfort. I would also like to express heartfelt gratitude to my friend, Tarun. As fellow INFPs, he resonated with everything I expressed. Lastly, thank my supervisor, Dzintra for her encouragement and thoughtful revisions of the thesis, and my dear colleague, Lipolelo Thamae. Cooperating with you is greatly joyful.

References

- Chandel, P., & Lim, F. V. (2024). Generative AI and literacy development in the language classroom: A systematic review of literature. *Ubiquitous Learning: An International Journal*, 18(2), 31-49.
- Eisenring, M. A., et al. (2024). The use of chatbots in English language teaching to promote modern language learning: A literature review. *International Journal of Indonesian Education and Learning*, 8(1), 127-139.
- Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60-70.
- Wang, H., & Dang, A. (2024). Enhancing L2 writing with generative AI: A systematic review of pedagogical integration and outcomes. *Preprint posted on Research Gate, October 2024*. Retrieved May 30, 2025, from: <https://www.researchgate.net/publication/384732061>
- Mason, J., et al. (2023). Metacognitive strategies to grow students' independent thinking. *Turnitin Blog*. Retrieved April 30, 2025, from <https://www.turnitin.com/blog/metacognitive-strategies-to-grow-students-independent-thinking>
- Mishra, P. (2023). *TPACK in the age of ChatGPT and generative AI*. Retrieved April 20, 2025, from <https://punyamishra.com/wp-content/uploads/2023/08/TPACK-in-the-age-of-ChatGPT-and-Generative-AI.pdf>
- OECD. (2024). *The potential impact of artificial intelligence on equity and inclusion in education*. Retrieved April 30, 2025, from <https://www.oecd.org/education/the-potential-impact-of-artificial-intelligence-on-equity-and-inclusion-in-education.pdf>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. Retrieved April 30, 2025, from <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53(4), 5-23.
- Barkley, E. F., & Major, C. H. (2024). Scaffolding AI as a learning collaborator: *Integrating artificial intelligence in college classes*. K. Patricia Cross Academy. Retrieved April 30, 2025, from <https://kpcrossacademy.ua.edu/scaffolding-ai-as-a-learning-collaborator-integrating-artificial-intelligence-in-college-classes/>
- Zou, B., & Wang, D. (2023). Artificial intelligence in language education: A review of recent developments. *Computer Assisted Language Learning*, 36(2), 123-140.