Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 9, Number 2, pp: 1112-1127, December 2025
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v9i2.12521

MEDIATING ORAL COMPETENCE THROUGH GEN-AI IN GENRE-BASED SPEAKING INSTRUCTION FOR FRACTURED NARRATIVES

Herti Herawati, Emi Emilia, Finita Dewi
Universitas Pendidikan Indonesia, Bandung, Indonesia
E-mail: hertiherawati@upi.edu

Received: 2025-12-03 Accepted: 2025-12-18 Published: 2025-12-22

Abstract

Developing oral narrative skills in English as a Foreign Language (EFL) context is challenging. It is often made more difficult by limited opportunities for personalized feedback and the cognitive burden of real-time production. This study investigated the integration of GenAI tools, specifically Gemini and ChatGPT, within a Genre-Based Approach (GBA) to scaffold the speaking skills of 32 tenth-grade students in West Java, Indonesia. This study focuses on the creative genre of "fractured stories," which is a demanding task that requires progressive scaffolding. This study used a qualitative case study design to explore how GenAI tools functioned across the GBA cycle. Data were collected through classroom observations, student-AI chat logs, students' speaking recordings, and semi-structured interviews. The findings revealed that GenAI tools could improve students' organizational structure and pronunciation by effectively enacting Vygotskian scaffolding, functioning as a dynamic "More Knowledgeable Other," that progressed from a lexical support during the knowledge-construction stage to a personalized phonological coach in the phase of independent practice. Analysis of speaking recordings showed marked improvement in narrative structure across all proficiency levels, while lower-proficiency students demonstrated significant gains in pronunciation accuracy via multimodal rehearsal loops. The study concluded that when embedded in a rigorous genre-based framework, GenAI transforms from a mere text generator into a robust, affective-aware partner that democratizes access to high-quality oral feedback. Future research should examine long-term retention of skills beyond the intervention period and investigate GenAI integration across diverse oral genres.

Keywords: *fractured narratives; GenAI; Genre-Based Approach; speaking*

1. Introduction

The development of speaking proficiency represents a core challenge in English as a Foreign Language (EFL) context, often constrained by logistical limitations such as high student-to-teacher ratios and insufficient opportunities for substantive, personalized practice (Emilia et al., 2021). Traditional classroom instruction frequently struggles to provide the necessary immediate and tailored feedback crucial for refining oral skills, particularly complex components like pronunciation and fluency. Concurrently, the Genre-

Based Approach (GBA) has proven effective in structuring language instruction by focusing learning around the social purpose, schematic structure, and linguistic features of specific text types (Callaghan & Rothery, 1988; Emilia et al., 2021; Joyce & Feez, 2012; Martin, 2014).

The advent of highly advanced generative Artificial Intelligence (GenAI) tools, such as ChatGPT and Gemini, presents a significant technological integration opportunity for language education stakeholders (Warschauer, 2020). These tools, which utilize sophisticated Natural Language Processing (NLP) and Automated Speech Recognition (ASR) technology, are capable of offering personalized, versatile, and dynamic learning environments, effectively augmenting traditional instruction. This study investigates the pedagogical synergy created by integrating GenAI within the established cyclical framework of GBA, focusing specifically on a demanding oral production task: telling fractured stories.

The theoretical lens underpinning this investigation is rooted in Vygotsky's (1978) sociocultural theory, where GenAI acts as a dynamic mediator within the Zone of Proximal Development (ZPD). Unlike fixed forms of scaffolding, GenAI provides a responsive, adaptive support structure that guides students through the iterative stages of genre mastery. The effectiveness of GenAI stems from its capacity to offer instantaneous, data-driven feedback and adapt learning materials based on individual performance, fostering enhanced self-regulation and promoting learner autonomy. This technological scaffolding facilitates the move from collaborative learning to independent mastery of both genre requirements and necessary linguistic sub-skills (Christie & Derewianka, 2008).

Despite this potential, a critical gap remains in the empirical literature: How can AI-assisted GBA scaffold students' speaking development in telling fractured stories? While systematic reviews acknowledge the profound transformation brought about by GenAI in English Language Teaching, specific models integrating these advanced tools into the explicit four-stage GBA cycle (i.e., Building Knowledge of Field (BKOF), Modeling, Joint Construction, and Independent Construction) remain empirically underexplored, particularly in relation to fractured story genre. This genre is especially ideal for examining AI scaffolding, as it entails multidimensional support and demands a high level of cognitive integration involving content generation, structural knowledge, and delivery competence. Addressing this gap requires a nuanced examination of how AI functions dynamically across the pedagogical sequence to support the development of speaking skills.

This study makes three contributions:

a. It provides a concrete pedagogical model for integrating GenAI across the full GBA cycle for oral production.
b. It demonstrates how TTS and ASR technologies create targeted pronunciation practice loops; and
c. It offers empirical evidence of how AI scaffolding evolves from receptive support to productive coaching.

2. Literature Review

2.1 Genre-Based Approach and Oral Competence

The Genre-Based Approach (GBA), grounded in Systemic Functional Linguistics, views language as a resource for making meaning in social contexts and treats genres as recurrent, goal-oriented configurations of stages shaped by culture (Halliday, 1993; Martin & Rose, 2008). Pedagogically, GBA seeks to make the often "hidden" organization of texts visible through an explicit teaching–learning cycle, moving learners from supported participation to independent control (Callaghan & Rothery, 1988; Feez, 1998; Emilia, 2016). Key principles

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 9, Number 2, pp: 1112-1127, December 2025
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v9i2.12521

include explicit instruction in generic stages and linguistic features, systematic scaffolding of tasks, and a commitment to equity by providing all students access to powerful literacies (Christie & Derewianka, 2008; Gibbons, 2015; Singh, 2022).

In oral genres, GBA typically unfolds through four recursive stages (Building Knowledge of the Field (BKOF), Modeling, Joint Construction, and Independent Construction), each targeting different aspects of oral competence (Derewianka & Jones, 2016; Triastuti et al., 2022). For narratives, instruction focuses on deconstructing schematic structure (Orientation–Complication–Resolution) and language resources such as appraisal and causal connectors, then reconstructing them in learner production (Hyland, 2007; Rose & Martin, 2012). Empirical studies show that genre-based speaking instruction can improve discourse organization, accuracy, and confidence: Khatibi (2014) and Zhang (2018) found that GBA-based speaking lessons led learners to produce more coherent, stage-appropriate oral texts, while Nahid et al (2018) and Hakim et al (2025) reported gains in clarity and length of learners' spoken descriptions and expositions. However, these studies also highlight persistent challenges in providing individualized feedback and sufficient rehearsal time for all students, especially in large EFL classes (Burns, 2022; Wardani et al., 2021).

2.2 Generative AI as a Learning Tool for Speaking

To address limitations in feedback and practice opportunities, recent work has turned to Generative AI (GenAI) and related speech technologies as mediational tools in language classrooms (Hwang et al., 2020; Kessler, 2023). Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) systems can provide immediate, individualized feedback on learners' output and create accessible pronunciation models, thus expanding practice beyond classroom time (Eskenazi, 2009; Dizon, 2020). In this role, AI tools offer high-frequency, low-stakes opportunities for listening–speaking cycles that are difficult to achieve through teacher feedback alone, especially in resource-constrained contexts (Xu et al., 2021; Ambarita & Nurrahmatullah, 2024).

A growing body of empirical research documents AI's specific functional contributions to EFL speaking. Studies using ASR-based platforms show significant improvements in segmental accuracy and overall intelligibility, particularly for lower-proficiency learners who receive continuous feedback on mispronounced segments (Bashori et al., 2022; Sun, 2023; Abimanto & Sumarsono, 2024). Voice-based chatbots and AI speaking assistants have been found to enhance fluency, interactional confidence, and self-regulation by enabling repeated practice in a non-judgmental environment (Han, 2020; Kim et al., 2021; Wang & Young, 2023; Farooqi, 2025). In the Indonesian context, Ahyarudin and Jamilah (2024) demonstrate that AI-based instructional tasks embedded within a GBA framework improved students' speaking scores and engagement, suggesting that GenAI can effectively complement structured pedagogies. At the same time, concerns remain about technical bias and recognition gaps for non-native accents, which may lead to frustration or unequal support if AI outputs are not critically monitored (Koenecke et al., 2020; Qiao & Zhao, 2023).

2.3 Vygotskian Sociocultural Theory and Digital Scaffolding

Vygotskian sociocultural theory provides a useful lens for understanding how GenAI can mediate learning within a GBA cycle. Learning is seen as a socially mediated process in

which higher mental functions develop first on the interpsychological plane through interaction with others or tools, and then become internalized (Vygotsky, 1978). Central constructs include the Zone of Proximal Development (ZPD) (i.e., the space between what learners can do alone and what they can achieve with support) and the notion of the More Knowledgeable Other (MKO), who offers graduated assistance that is gradually withdrawn (Wood et al., 1976; Hammond & Gibbons, 2005).

In digital environments, AI systems can function as mediational tools that approximate an MKO by offering contingent, on-demand assistance aligned with learners' current needs (Holmes et al., 2019; Noroozi et al., 2024b). For speaking, features such as instant error detection, TTS models, and dialogic prompting can be interpreted as forms of scaffolding that reduce cognitive load and support movement through the ZPD: ASR feedback highlights the gap between learner output and target forms; TTS provides accessible models for imitation; and conversational agents offer guided questioning that helps learners plan and elaborate their texts (Dizon, 2020; Johnson & Lester, 2021; Liu & Ma, 2025). When embedded in a genre-based cycle, GenAI can thus act as a dynamic mediator (sometimes a resource bank, sometimes an analyst of structure, sometimes a collaborative co-author, and finally a pronunciation coach), mirroring the progression of scaffolding described in GBA and SFL-based instruction (Emilia & Hamied, 2015; Herazo, 2021; Sarmadi Kia et al., 2024).

At the same time, researchers caution that mediation and scaffolding must remain human-directed: AI can provide powerful assistance, but teachers need to orchestrate when and how learners rely on AI so that the technology amplifies, rather than replaces, dialogic teaching and peer interaction (Wardani et al., 2021; Hwang & Chen, 2023; Weng & Fu, 2025). This implies a design stance where AI is positioned explicitly as a scaffold within teacher-guided tasks, rather than as an autonomous instructor.

2.4 Research Gap

Existing GBA studies demonstrated that genre-based pedagogy improves learners' oral organization and speaking performance, but they rarely integrated advanced AI as a systematic mediator across the full teaching–learning cycle, particularly for creative genres like fractured narratives (Khatibi, 2014; Zhang, 2018; Puspitasari et al., 2024). Conversely, most GenAI-in-speaking research targets discrete skills (e.g., pronunciation or general chatbot practice) without embedding AI within a structured GBA framework that sequences BKOF, Modeling, Joint Construction, and Independent Construction (Bashori et al., 2022; Han, 2020; Wang & Young, 2023; Ahyarudin & Jamilah, 2024). Although sociocultural work increasingly conceptualizes AI as a mediational tool, there is still little empirical evidence showing how one GenAI system can act as a progressively evolving scaffold (from resource to analyst, collaborator, and coach) across a GBA cycle and how this shapes learners' speaking development in fractured narrative tasks; this study addresses that gap by examining GenAI as a dynamic scaffold within GBA for EFL students' oral production of fractured stories.

3. Research Method

3.1 Research Design

This study adopted a qualitative case study design (Cohen et al., 2018; Hancock, 2002; McMillan & Schumacher, 2011; Yin, 2011). A qualitative case study was chosen to gain

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 9, Number 2, pp: 1112-1127, December 2025
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v9i2.12521

a deep, holistic understanding of the process of AI scaffolding and the mechanism of mediation within a real-world, bound context (the GBA class) (Berg, 2007; Creswell & Poth, 2018; Lincoln & Guba, 1985; Merriam & Tisdell, 2016).

3.2 Research Participants

The study was conducted at a public senior high school in Subang Regency, West Java, Indonesia, a site that represents typical Indonesian EFL classroom conditions while providing sufficient digital infrastructure for AI-mediated activities. The broader participant cohort consisted of 32 tenth-grade students (12 male, 20 female, aged 15–17) from predominantly middle to low socioeconomic backgrounds. They owned personal mobile devices and were familiar with basic digital/AI applications, ensuring access to the technological components of the intervention. For the intensive qualitative analysis, purposive stratified sampling was used to select nine focal students (three high, three mid, and three low achievers) based on previous semester English scores and teacher evaluations of speaking performance and classroom engagement, so that they represented a maximum-variation profile of proficiency and likely interaction with the AI tools and thus provided information-rich cases for examining how AI-assisted GBA scaffolding operates across different ability levels within a case study design (Stake, 1995); all participants and guardians provided informed consent, and pseudonyms (S1–S9) were used to protect anonymity.

3.3 Data collection

Data were gained from three main sources: (1) classroom observations, (2) student artifacts (AI chat logs and pre-/post-speaking recordings), and (3) semi-structured interviews. The observations spanned eight ninety-minute sessions focused on teaching the fractured story genre and concluded with a final speaking evaluation. The teaching was structured around the four cyclical stages of GBA: BKOF, Modeling, Joint Construction, and Independent Construction, as suggested by Callaghan and Rothery (1988), Joyce & Feez (2012), Martin (2014), and Emilia et al (2021).

3.4 Data analysis

The data were analyzed using a multi-layered approach. Classroom observation data were analyzed descriptively by mapping instructional activities onto the Genre-Based Approach (GBA) cycle (Callaghan & Rothery, 1988; Joyce & Feez, 2012; Martin, 2014; Emilia et al, 2021). For each stage, the analysis identified the specific instructional activities and the distinct roles performed by the GenAI tools. Student artifacts and interviews were analyzed thematically using Braun and Clarke's (2006) six-step procedure. AI chat logs were open-coded to identify recurring functions (e.g., lexical clarification, genre explanation, guided idea generation, form-focused feedback, pronunciation rehearsal), then clustered into higher-order themes such as Lexical Support, Genre Mediation, and Performance Coaching, which were interpreted through Vygotskian notions of mediation, MKO, and ZPD to show how AI support shifted with learner needs (Vygotsky, 1978; Hammond & Gibbons, 2005). Interview data were coded similarly, generating themes like Reduced Anxiety, Increased Agency, and Critical Engagement with AI linked to sociocultural perspectives on affect and learner agency. Speaking performance was rated with Emilia's (2016) GBA-aligned

rubric across three dimensions (narrative structure, pronunciation, and fluency) to track development over the intervention.

3.5 Gen AI Tools and Prompting

This study used two GenAI tools: ChatGPT (free mobile version, GPT-4o) and Google Gemini (free mobile app). Regarding the prompting, the teacher modelled using the prompt to ask AI as a learning tutor in the joint construction stage by asking the AI to guide them using questions to create a fractured story. In Independent Construction, students used ChatGPT/ Gemini speech features for practice (e.g., "Listen to my story and give feedback on my pronunciation, pausing, and intonation."), combining TTS/ASR feedback with their own oral rehearsal.

3.6 Trustworthiness

Trustworthiness was supported through triangulation and member checking. Classroom observations, student artifacts (AI chat logs, written texts, ASR/TTS records), and interviews were systematically cross-checked so that, for example, reported gains in pronunciation or confidence were examined against log evidence and rated speaking performance, helping separate robust patterns from isolated claims. Member checking was conducted by sharing interview transcripts and brief thematic summaries with focal participants for confirmation or correction, enhancing the credibility and authenticity of the interpretations.

4. Results and Discussion

4.1 Improvement in oral performance

Analysis of the pre- and post-observation speaking scores showed that all nine focal students improved in at least one of the three assessed dimensions (narrative structure, pronunciation, fluency), with the most consistent gains observed in narrative organization and pronunciation. Narrative structure scores indicated that all students reached the "Advanced" level by the end of the intervention, suggesting that the combination of explicit genre instruction and AI-supported drafting successfully supported their control of orientation–complication–resolution sequencing.

This pattern reinforces evidence that explicit genre teaching can move learners from fragmented speech to more coherent, purposeful narratives by making schematic structure visible and repeatedly practiced in scaffolded cycles (Derewianka & Jones, 2016; Tran, 2021; Zhang, 2018). The following table shows students' improvement in organizational structure.

Table 1

Students' development in organizational structure

Student	Structure Level in Pre-observation	Structure Level in Post-observation
S1	Basic	Advanced
S2	Advanced	Advanced
S3	Intermediate	Advanced
S4	Intermediate	Advanced
S5	Intermediate	Advanced
S6	Intermediate	Advanced
S7	Basic	Advanced

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 9, Number 2, pp: 1112-1127, December 2025
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v9i2.12521

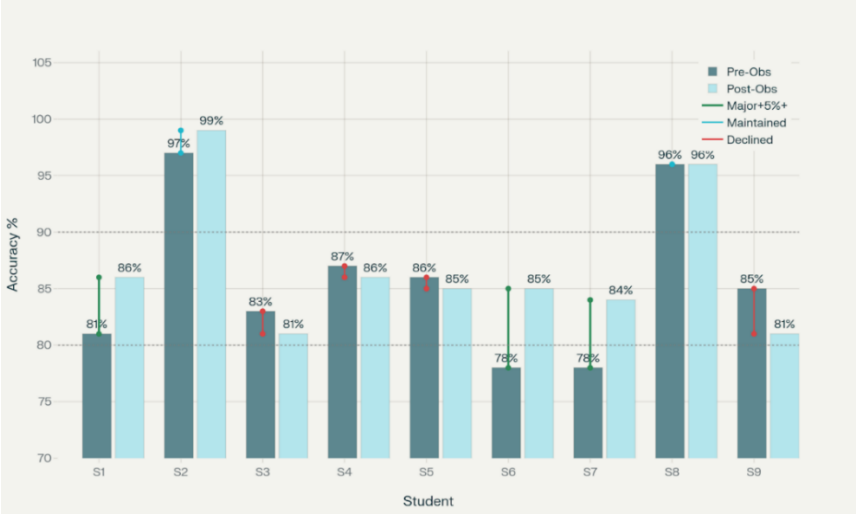
S8	Advanced	Advanced
S9	Intermediate	Advanced

All students' narratives contained the required story elements (orientation, complication, climax, resolution). Differentiation among levels was determined by the quality and sophistication of narrative coherence. Basic proficiency indicated complete but simple narratives lacking substantial elaboration or sophisticated linguistic development. Intermediate proficiency demonstrated straightforward, competent narratives with standard coherence across all stages. Advanced proficiency reflected narratives that seamlessly integrated all elements with sophisticated language choices and thorough elaboration, demonstrating superior genre mastery.

Lower-proficiency students showed marked pronunciation gains, with several moving from poor intelligibility to scores above 84%. Student S3 reflected: "I listened to the AI voice many times and practiced words it flagged as wrong. It helped because I was shy to ask the teacher to repeat." This illustrates how pronunciation improvement depends on repeated, private engagement. This is something the non-judgmental AI context facilitates. These gains align with research showing ASR-mediated pronunciation practice accelerates segmental accuracy for lower-proficiency learners (Abimanto & Sumarsono, 2024; Bashori et al., 2022; Gottardi et al., 2022).

Figure 1

Students' pronunciation improvement (pre- to post-observation)



Fluency scores showed smaller but meaningful improvements, especially in reduced pausing and fillers as students became more familiar with their fractured stories through repeated AI-mediated rehearsal. Such incremental fluency gains are consistent with research indicating that repeated, low-stakes rehearsal with AI tools helps learners reduce dysfluencies over time by stabilizing lexical retrieval and phrasing, even when formal fluency training is not the primary focus (Barrot, 2021; Farooqi, 2025).

Taken together, these results suggest that when AI is embedded in a genre-based cycle, it can function as a mediating tool that supports movement through the Zone of Proximal Development by combining explicit structural guidance with frequent, individualized feedback on form and delivery (Christie & Derewianka, 2008; Vygotsky, 1978).

4.2 Theme 1: AI as a Lexical and Structural MKO

Across the Building Knowledge of the Field, Modeling, and Joint Construction stages, AI functioned as a "More Knowledgeable Other" that mediated learners' access to lexico-grammatical resources and genre structure. In the BKOF stage, students used ChatGPT or Gemini as a digital glossary and information-retrieval tool to clarify unfamiliar vocabulary and cultural references encountered in the fractured stories, thereby reducing lexical cognitive load and enabling deeper engagement with the narrative content. This immediate access to lexical support ensured that students entered the subsequent stages with sufficient content knowledge and vocabulary to manipulate the genre meaningfully, echoing research that identifies AI's preparatory role in early learning phases. This preparatory function resonates with sociocultural views of mediation, where tools and artifacts help manage cognitive load so that learners can allocate more attention to meaning-making rather than basic decoding (Halliday & Matthiessen, 2014; Lantolf & Thorne, 2006).

During the Modeling stage, AI's role shifted from simple look-up support to that of an analytical assistant and exemplar provider. Students requested models of fractured stories based on familiar folktales and then prompted AI to analyze their schematic structure and language features, gaining metalinguistic awareness of how orientation, complication, and resolution are realized linguistically. Chat logs from this phase show learners asking AI to identify specific stages in sample narratives and to highlight key linguistic resources, while interview data (e.g., S5) confirmed that this process clarified expectations about how a story "should start with the orientation and move to the complication and finally the resolution." In this sense, GenAI partially assumes the traditional teacher role in GBA by making "hidden" genre rules explicit and negotiable, thereby extending earlier work on modeling and deconstruction in oral narrative pedagogy (Emilia et al., 2021; Joyce & Feez, 2012; Martin, 2014).

AI's capacity to deconstruct text models bridges educational divides by providing individualized linguistic analysis (Weng & Fu, 2025). Student S5 stated: "AI helped me understand the story structure (orientation, complication, resolution)." By enabling on-demand structural analysis, AI functions as a dynamic MKO accessible anytime, not just during teacher-led episodes (Lee, 2025; Wang, 2025). Chat logs from the Modeling stage (S8-Chat Log) further demonstrate this: when asked to identify schematic structure, S8 requested, "Can you show me where the orientation ends and the complication starts?" This reveals how AI's availability prompted metalinguistic inquiry beyond traditional whole-class interaction.

In the Joint Construction stage, AI operated as a collaborative writing guide and form-focused feedback provider that linked structural analysis to productive language use. Teacher-led demonstrations of prompts showed students how to co-construct fractured stories with AI, after which pairs or small groups used similar prompts on their own devices to draft narratives. As students composed, AI-generated guiding questions, suggested lexical choices, and provided immediate feedback on grammar and clarity, helping them refine lexico-grammatical accuracy before moving to independent oral production. This iterative, form-focused feedback reflects findings that learners often perceive automated corrective feedback as objective and consistent, which can make them more willing to make and revise ambitious language choices (Barrot, 2021, 2023; Algaraady & Mahyoob, 2023; Lee, 2025).

Overall, evidence from chat logs, classroom observations, and interviews suggests that in these early and mid-cycle stages, AI effectively assumed a lexical and structural MKO

Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 9, Number 2, pp: 1112-1127, December 2025
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v9i2.12521

role, taking on much of the analytical and corrective work that would otherwise fall solely on the teacher. This progressive mediation illustrates a classic scaffolding trajectory, where support is initially heavy and explicit but gradually shifts toward learner control as performance becomes more independent, aligning with Vygotskian accounts of MKO-guided development in the ZPD and with GBA's emphasis on moving from joint to independent construction (Gibbons, 2015; Lantolf & Thorne, 2006; Martin, 2014).

4.3 Theme 2: AI as a Phonological and Delivery Coach

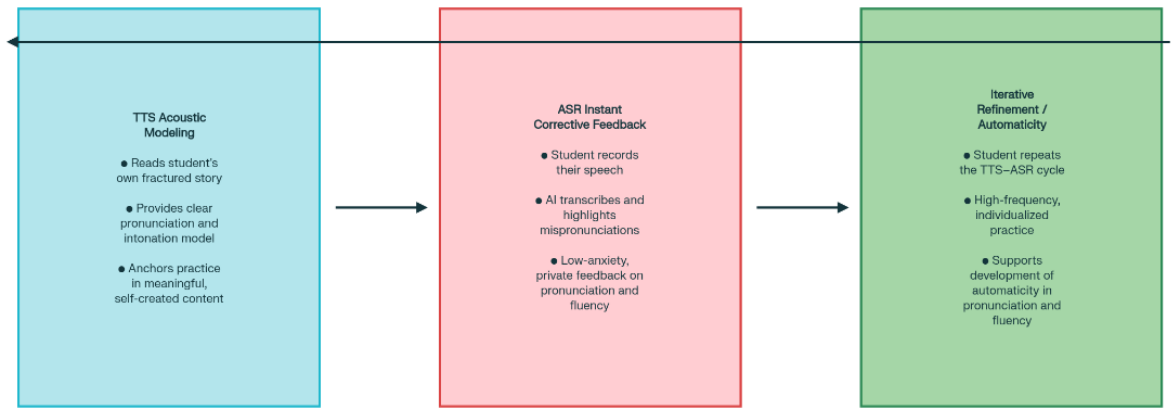
In the Independent Construction stage, AI's role evolved into that of a phonological and delivery coach, supporting students' transition from written drafts to oral performance. After using AI to help generate and refine their fractured stories, students converted their texts into audio using text-to-speech tools, which provided consistent acoustic models for pronunciation, rhythm, and intonation. They then engaged in a multimodal practice loop in which they listened repeatedly to the AI-generated audio, imitated the pronunciation, and used automated speech-recognition (ASR) features in ChatGPT or Gemini to receive feedback on their oral production. The TTS–ASR loop implemented here mirrors designs in which synthetic models and automatic evaluation jointly create a closed feedback cycle, a configuration that has been shown to promote phonological awareness and more precise self-monitoring in L2 learners (Belda-Medina & Calvo-Ferrer, 2022; Dizon, 2020; Aryadoust, Zakaria, & Jia, 2024).

The TTS–ASR loop proved particularly effective for lower-proficiency learners, enabling significant pronunciation improvement through private, consequence-free rehearsal. Students engaged in iterative listening and targeted practice on words identified by the system as requiring attention. Interview data (S7) reveal that AI use evolved from vocabulary and grammar checking to pronunciation practice, fostering confidence through autonomous, non-evaluative repetition. This exemplifies the progression from lexical to structural to phonological scaffolding.

Similar to findings by Al-Ghezi et al. (2023) and Abimanto and Sumarsono (2024), the present data suggest that frequent, privately accessible practice can be particularly empowering for students who might otherwise avoid oral participation in face-to-face settings. The following figure illustrates the pronunciation loop when students practiced with GenAI.

Figure 2

AI pronunciation practice loop



Fluency gains, while more modest, appeared linked to this repeated rehearsal, as students became more comfortable with the pacing and wording of their stories and thus reduced hesitations during the final assessment.

At the same time, the data revealed clear limitations in AI-based phonological support, particularly the reduced accuracy of automatic speech recognition (ASR) for lower-proficiency accents. Some students (e.g., S5) reported that AI failed to respond appropriately during speaking practice, initially interpreting this as a personal pronunciation deficit, which poses a risk to learner self-efficacy. However, S5's subsequent decision to discontinue AI use in these situations reflected emerging critical AI literacy, as the learner attributed the breakdown to technological constraints rather than individual inadequacy.

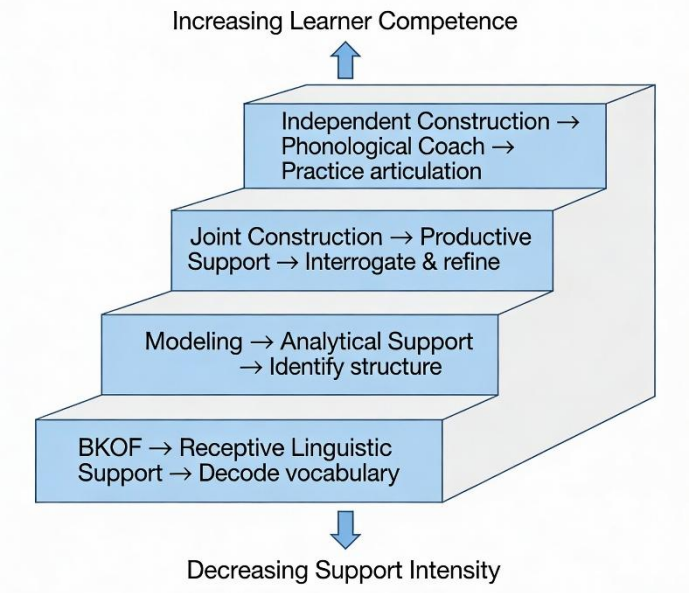
Evidence confirms AI functioned as a lexical and structural MKO (Emilia et al., 2021). Unlike class-bound teachers, GenAI enabled "on-demand deconstruction" of genre features, which is instantaneous rather than scheduled. This is a capability not previously documented in GBA. This progression through look-up, analysis, and refinement mirrors Vygotsky's (1978) ZPD scaffolding, where support decreases with competence, except here the mediator is digital and self-paced.

4.4 The Progressive Evolution of AI Scaffolding Across the GBA Cycle

The data reveal a systematic evolution in AI's mediational role that exemplifies progressive scaffolding within the Zone of Proximal Development. In BKOF, the AI operates at the level of receptive linguistic support: students consult it to decode unfamiliar vocabulary and cultural content, a function that requires minimal cognitive engagement beyond recognition. In Modeling and Joint Construction, AI's role escalates to analytical and productive support: students now interrogate the tool to understand generic structure, suggest alternatives, and refine grammar. These tasks require metalinguistic awareness and deliberate choice-making. Finally, in Independent Construction, AI becomes a phonological coach, where learners must practice articulation, receive automatic feedback, and iteratively adjust their motor execution. The following figure illustrates the systematic evolution in AI-assisted GBA.

Figure 3

AI scaffolding in the GBA cycle



Language Literacy: Journal of Linguistics, Literature, and Language Teaching
Volume 9, Number 2, pp: 1112-1127, December 2025
e-ISSN: 2580-9962 | p-ISSN: 2580-8672
DOI: https://doi.org/10.30743/ll.v9i2.12521

This four-stage progression (from receptive linguistic support to phonological coaching) represents a classical scaffolding trajectory where each stage presupposes competence from the previous one (Lantolf & Thorne, 2006). Rather than a single mediator withdrawing support, AI's role transforms in response to evolving learner needs, maintaining appropriate cognitive load while progressively shifting responsibility. This finding extends Vygotsky's (1978) original conception by showing that technological mediators can be dynamically repurposed within a single pedagogical cycle, something a static human interlocutor cannot easily do.

4.5 Limitations and Future Research

This study contributes to understanding AI-assisted GBA for oral development, yet several constraints limit generalizability. The sample (nine students from a single school in West Java) restricts applicability across contexts. The intervention focused exclusively on fractured narratives; applicability to other oral genres (e.g., speeches, debates) remains untested. The eight-session duration precludes assessment of long-term retention or deepening of critical AI literacy. Additionally, ASR and TTS limitations with non-native accents suggest the need to compare different platforms (e.g., ELSA, Speechling, Duolingo). Finally, a more detailed discourse analysis of AI–student interactions could illuminate mediation mechanisms.

Future research should: (1) test the model across multiple contexts and proficiency levels; (2) examine other oral genres and task types; (3) employ longitudinal designs to measure retention and transfer; (4) compare AI platforms to optimize configurations for different learner profiles; and (5) investigate how teachers mediate between AI and human feedback while fostering critical evaluation of algorithmic error detection.

5. Conclusion

This study demonstrates that Generative AI functions as a dynamic, progressive mediator for oral competence, bridging instructional theory and technological application. The findings validate AI as an adaptable, More Knowledgeable Other capable of fulfilling Vygotsky's Zone of Proximal Development by transforming its role across the pedagogical cycle (from lexical support in the BKOF to phonological coaching in Independent Construction). This extends classical Vygotskian scaffolding into the digital age.

All nine focal students improved in narrative structure, with lower-proficiency learners showing the most substantial gains. The TTS–ASR pronunciation practice loop proved particularly effective by providing low-anxiety, high-frequency rehearsal, while students' increased confidence highlights AI's democratizing potential and affective benefits.

For practitioners, the implications are actionable: educators must design high-utility prompts that strategically shift AI's role across the GBA cycle (from resource provider to performance coach). Rather than an add-on, AI should be embedded in teacher-guided frameworks where technology amplifies, not replaces, pedagogical decision-making.

While robust, the findings are constrained by a single-context design and focus on one genre. Longitudinal research, comparative studies with control groups, investigation across oral tasks, and examination of optimal prompt-engineering for different proficiency levels would strengthen claims about generalizability and long-term efficacy.

This study illustrates a human-AI partnership grounded in pedagogy rather than technological enthusiasm. When designed within established theoretical frameworks, AI expands access to personalized feedback, particularly for under-resourced learners. The challenge lies not in whether to integrate AI, but in doing so while honoring learner autonomy and sustaining human relationships in language education.

References

Abimanto, D., & Sumarsono, W. (2024). Improving English pronunciation with AI speech-recognition technology. *Acitya: Journal of Teaching & Education*, *6*(1), 146–156. https://doi.org/10.30650/ajte.v6i1.3810

Algaraady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journal*, *9*, 3–17. https://doi.org/10.24093/awej/call9.1

Al-Ghezi, R., Voskoboinik, K., Getman, Y., Von Zansen, A., Kallio, H., Kurimo, M., & Hildén, R. (2023). Automatic speaking assessment of spontaneous L2 Finnish and Swedish. *Language Assessment Quarterly*, *20*(4-5), 421–444. https://doi.org/10.1080/15434303.2023.2292265

Aryadoust, V., Zakaria, A., & Jia, Y. (2024). Investigating the affordances of OpenAI's large language model in developing listening assessments. *Computers and Education: Artificial Intelligence*, *6*, 100204. https://doi.org/10.1016/j.caeai.2024.100204

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY**:** Freeman.

Barrot, J. (2021). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, *36*(4), 584–607. https://doi.org/10.1080/09588221.2021.1936071

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2022). 'Look, I can speak correctly': learning vocabulary and pronunciation through websites equipped with automatic speech recognition technology. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2022.2080230

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using Chatbots as AI conversational partners in language learning. *Applied Sciences*, *12*(17), 8427. https://doi.org/10.3390/app12178427

Berg, B.L. (2007) Qualitative Research Methods for the Social Sciences. Boston, MA: Pearson.

Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology*, *23*(1), 23–41. https://doi.org/10.56297/bkam1691/wieo1749

Bui, N. M., & Barrot, J. S. (2024). ChatGPT as an automated essay scoring tool in the writing classrooms: How it compares with human scoring. *Education and Information Technologies*, *30*, 2041–2058. https://doi.org/10.1007/s10639-024-12891-w

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Callaghan, M., & Rothery, J. (1988). *Teaching. Factual Writing: A genre based approach*. Marrickville, N.S.W.: DSP. Literacy Project, Metropolitan East.

Christie, F., & Derewianka, B. (2008). *School discourse: Learning to write across the years of schooling.* London, UK**:** Continuum.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). London, UK**:** Routledge.

Coskun, T., & Alper, A. (2024). Evaluating the evaluators: A comparative study of AI and teacher assessments in higher education. *Digital Education Review*. https://doi.org/10.1344/der.2024.45.124-140

Dai, K., & Liu, Q. (2024). Leveraging artificial intelligence (AI) in English as a foreign language (EFL) classes: Challenges and opportunities in the spotlight. *Computers in Human Behavior*, *159*, 108354. https://doi.org/10.1016/j.chb.2024.108354

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Derewianka, B., & Jones, P. (2016). *Teaching language in context* (2nd ed.). Oxford University Press.

Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies*, *29*, 14151–14203. https://doi.org/10.1007/s10639-023-12402-3

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology, 24*(1), 16–26. https://doi.org/10.10125/44705

Emilia, E., Gunawan, W., & Kurniawan, E. (2021). *Pembelajaran berbasis genre: Konsep, praktik dan kajian kritis Pengajaran Bahasa*. Bandung, Indonesia**:** UPI Press.

Farooqi, S. U. H. (2025). Efficacy of AI-generated feedback by SmallTalk2Me for improving speaking skill of Saudi EFL learners. *Forum for Linguistic Studies*, *7*(3), 714–728. https://doi.org/10.30564/fls.v7i3.8294

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, *6*(1), 57–63. https://doi.org/10.37074/jalt.2023.6.1.22

Geethanjali, K. (2025). AI prompting and the growth of metalinguistic skills in educational contexts. *International Journal of Creative Research Thoughts (IJCRT)*, *13*(9), f189–f192. Retrieved from http://www.ijcrt.org

Gibbons, P. (2015). *Scaffolding language, scaffolding learning (2nd ed.)*. Portsmouth, NH: Heinemann.

Gottardi, W., Almeida, J. F., & Tumolo, C. H. S. (2022). Automatic speech recognition and text-to-speech technologies for L2 pronunciation improvement: Reflections on their affordances. *Texto Livre: Linguageme Tecnologia*, *15*, e36736. https://doi.org/10.35699/1983-3652.2022.36736

Grab, M. O. (2025). Strategies for integrating artificial intelligence (AI) to improve and assess English oral communication skills. *International Journal of Research in Education and Science (IJRES)*, *11*(3), 533–552. https://doi.org/10.46328/ijres.1312

Guo, K., Pan, M., Li, Y., & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education*, *63*, 100962. https://doi.org/10.1016/j.iheduc.2024.100962

Halliday, M. A. K., & Matthiessen, C. (2014). *Halliday's introduction to functional grammar (4th ed.)*. Oxon: Routledge.

Hancock, B. (2002). *An introduction to qualitative research*. Nottingham, UK: Trent Focus Group.

Huang, X., Xu, W., Li, F., & Yu, Z. (2024). A meta-analysis of effects of automated writing evaluation on anxiety, motivation, and second language writing skills. *The Asia-Pacific Education Researcher*, 1–20.

Hyland, K. (2007). Genre pedagogy: Language, literacy, and L2 writing instruction. *Journal of Second Language Writing, 16*(3), 148–164. https://doi.org/10.1016/j.jslw.2007.07.005

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: young learners' experiences and perspectives. *Computer Assisted Language Learning*, *37*(1–2), 1–26. https://doi.org/10.1080/09588221.2021.2021241

Jia, F., Sun, D., Ma, Q., & Looi, C. (2022). Developing an AI-based learning system for L2 learners' authentic and ubiquitous learning in English language. *Sustainability*, *14*(23). https://doi.org/10.3390/su142315527

Jin, Y., & Fan, J. (2023). Test-taker engagement in AI technology-mediated language assessment. *Language Assessment Quarterly*, *20*, 488–500. https://doi.org/10.1080/15434303.2023.2291731

Joyce, H. de S., & Feez, S. (2012). *Text-based language and literacy education: Programming and methodology*. Sydney, Australia: Phoenix Education.

Kaliisa, R., Misiejuk, K., López-Pernas, S., & Saqr, M. (2025). How does artificial intelligence compare to human feedback? A meta-analysis of performance, feedback perception, and learning dispositions. *Educational Psychology*. https://doi.org/10.1080/01443410.2025.2553639

Kartal, G., & Yeşilyurt, Y. E. (2024). A bibliometric analysis of artificial intelligence in L2 teaching and applied linguistics between 1995 and 2022. *ReCALL*, *36*(3), 359–375. https://doi.org/10.1017/S0958344024000077

Kenshinbay, T., & Ghorbandordinejad, F. (2024). Exploring AI-driven adaptive feedback in the second language writing skills prompt. *EIKI Journal of Effective Teaching Methods*, *2*(3). https://doi.org/10.59652/jetm.v2i3.264

Kopuz, E., & Kartal, G. (2025). AI-assisted L2 assessment: A biblio-systematic analysis. *PASAA*, *70*, 340–370. Retrieved from https://files.eric.ed.gov/

Korsmeyer, H., & Pink, S. (2025). Special issue introduction: Everyday digital energy futures. *Futures*, *170*, Article 103598. https://doi.org/10.1016/j.futures.2025.103598

Kuzman, T., Mozetič, I., & Ljubešić, N. (2023). Automatic genre identification for robust enrichment of massive text collections: Investigation of classification methods in the era of large language models. *Machine Learning and Knowledge Extraction*, *5*(3), 1149–1175. https://doi.org/10.3390/make5030059

Lee, J. (2025). EFL adolescents' engagement in artificial intelligence and peer interaction. *Language Learning & Technology*, *29*(1), 1–21. https://doi.org/10.64152/10125/73643

Li, S. (2025). Generative AI and second language writing. *Digital Studies in Language and Literature*, *2*(1), 122–152. https://doi.org/10.1515/dsll-2025-0007

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Liu, J., & Ma, Q. (2025). Supporting low-proficiency L2 learners' vocabulary learning with custom GPT-scaffolded corpus-based language pedagogy: a case study. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2025.2539152

Martin, J. R. (2014). Evolving systemic functional linguistics: beyond the clause. *Functional Linguistics*, *1*(3). https://doi.org/10.1186/2196-419X-1-3

McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry* (6th ed.). Boston, MA: Allyn and Bacon.

Meniado, J. C. (2023). The impact of ChatGPT on English language teaching, learning, and assessment: A rapid review of literature. *Arab World English Journal*, *14*(4). https://dx.doi.org/10.24093/awej/vol14no4.1

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

Mirowski, P., Mathewson, K. W., Pittman, J., & Evans, R. (2022). Co-writing screenplays and theatre scripts with language models: An evaluation by industry professionals. *arXiv preprint arXiv:2209.14958*

Monib, W. K., Qazi, A., Apong, R. A., Azizan, M. T., De Silva, L., & Yassin, H. (2024). Generative AI and future education: A review, theoretical validation, and authors' perspective on challenges and solutions. *PeerJ Computer Science*, *10*, e2105. https://doi.org/10.7717/peerj-cs.2105

Nazari, N., Shabbir, M., & Setiawan, R. (2021). Application of artificial intelligence-powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, *7*. https://doi.org/10.1016/j.heliyon.2021.e07014

Nejadghanbar, H., Salman, A. R., Ebadi, S., & Khosravi, H. (2025). Exploring the impact of generative AI on peer review: Insights from journal reviewers. *Journal of Academic Ethics*, *23*, 1383–1397. https://doi.org/10.1007/s10805-025-09604-4

Singh, R. (2022). Systemic functional linguistics and inclusivity in the classroom. *ERL Journal, 7*(1), 75–85. https://doi.org/10.36534/erlj.2022.01.07

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: a mixed methods investigation. *Frontiers in Psychology*, *14*:1210187. https://doi.org/10.3389/fpsyg.2023.1210187

Tran, M. T. T. (2021). Use of self-regulated learning strategies in paragraph writing at Van Lang University. International Journal of TESOL & Education, 1(3), 1-13. https://i-jte.org/index.php/journal/article/view/80

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wang, C., & Wang, Z. (2025). Investigating L2 writers' critical AI literacy in AI-assisted writing: An APSE model. *Journal of Second Language Writing*, *67*, 101187. https://doi.org/10.1016/j.jslw.2025.101187

Wang, M. (2025). Artificial intelligence assisted dual-teacher model constructing practices. *European Journal of AI, Computing & Informatics*, *1*(2), 110–117. Retrieved from https://pinnaclepubs.com/index.php/EJACI

Warschauer, M. (2020). The paradoxical future of digital learning. *Learning, Media and Technology, 45*(1), 1–13. https://doi.org/10.1080/17439884.2020.1707229

Weng, Z., & Fu, Y. (2025). Generative AI in language education: Bridging divide and fostering inclusivity. *International Journal of Technology in Education*, *8*(2), 395–420. https://doi.org/10.46328/ijte.1056

Yin, R. K. (2011*). Qualitative research from start to finish*. New York, NY: Guilford.

Zhang, L. (2018). Genre-based approach and EFL learners' speaking skills. *System, 78*, 158-169. https://doi.org/10.1016/j.system.2018.08.009

Zhang, Z., Aubrey, S., Huang, X., & Chiu, T. K. F. (2025). The role of generative AI and hybrid feedback in improving L2 writing skills: A comparative study. *Innovation in Language Learning and Teaching*. Published online: 13 May 2025. https://doi.org/10.1080/17501229.2025.2503890