Applied Language Learning

Action Research

Revolutionizing Language Learning: AI-Assisted Learning as a Catalyst for High Proficiency Attainment

Hanwei Tan, PhD
Defense Language Institute Foreign Language Center

This action research project investigates whether an AI-Assisted Learning Framework can help instructors and learners identify, track, and reduce recurring errors and non-native structures among intermediate and advanced Chinese learners at the Defense Language Institute Foreign Language Center. Recurring grammatical, lexical, and discourse patterns were coded as key forms when they met three criteria, distinguishing them from fatigue slips or one-off mistakes. Over a 19-week cycle, five learners engaged in an AI-enhanced intervention integrating five coordinated strategies: Error-Pattern Detection, Micro-Targeted Adaptive Exercises, Multimodal Feedback, AI-Mediated Conversational Tutoring, and Adaptive Review through Spaced Repetition. Data sources included pre- and post-test scores, logs, and learner journals. AI tools were used to assist instructors in identifying key forms across student output and generating targeted instructional materials, while all pedagogical decisions and validation remained instructor-mediated. Learners engaged in approximately 4.2 hours per week of AI-supported activities in addition to full-time classroom instruction. Findings indicate that three of the five learners demonstrated measurable ILR gains, while the frequency of key form errors declined by approximately 40%. Learner reflections suggest increased awareness of habitual error patterns and greater engagement with feedback processes. Although improvements cannot be attributed to AI use alone, the results suggest that structured, continuous diagnostic cycles may enhance the salience and treatment of persistent difficulties within high-intensity language programs. The study proposes an AI-assisted learning framework as a practical, instructor-guided approach for integrating diagnostic feedback, targeted practice, and longitudinal monitoring into intensive language instruction.

Keywords: *AI-Assisted Learning, Recurring Learner Patterns, Learner Autonomy, ILR Proficiency, Adaptive Feedback, Military Language Education*

BACKGROUND

Student Learning and the Role of Noticing

Language learners make errors for a variety of reasons. Some recurring learner errors may become stabilized over time despite repeated exposure and corrective feedback, meaning the learner has been taught an accurate form or pronunciation, yet continues to produce errors persistently despite continued exposure. Han (2004) conceptualized fossilization as developmental arrest, "a process whereby learning manifests a strong tendency toward cessation" that is both persistent and resistant to normal input cycles (p. 23). These patterns may become habitualized, persistent, and resistant to change. Such patterns frequently become automatized through repeated communicative success and may be reinforced by L1 transfer, early instructional routines, or strategic simplification under cognitive load (DeKeyser, 2007; Long, 1996). Other errors are developmental and a normal step in the learning process; these errors arise as learners test hypotheses about the target language system while restructuring their evolving grammatical representations (Corder, 1967; Selinker, 1972). Such errors are transitional and typically diminish as input, feedback, and restructuring mechanisms take effect.

Causes of Language Learning Errors and Non-Native Constructions

There are many reasons why learners produce errors. A main one is the influence of the native language. For example, English-speaking learners of Chinese often overgeneralize possessive constructions (e.g., 我的父亲 *my father* instead of 父亲 *father* in neutral contexts) or insert unnecessary determiners due to the influence of their L1. L1 influence can also manifest not as overt grammatical inaccuracy but as discourse-level or rhetorical deviation. Consider the learner production: 昨天我和我女朋友坐火车去旧金山**为了**吃中餐 (literal translation as: *Yesterday I and my girlfriend taking a train to San Francisco in order to eat Chinese food*). Although grammatically interpretable, the insertion of 为了 (*in order to*) mirrors English purposive logic rather than preferred Chinese verb-chaining sequencing: 昨天我和女朋友坐火车去旧金山吃中餐 (literal translation as: *Yesterday I and my girlfriend taking a train to San Francisco to eat Chinese food*). The former construction is grammatically interpretable but less target-like in natural Chinese discourse flow. These non-native forms may impede further language development and limit learners' ability to achieve higher levels of proficiency.

Helping Students Learn and Minimize Errors

DeKeyser (2007) argues that practice strengthens existing mental representations; therefore, if learners repeatedly practice inaccurate forms, those forms may become further proceduralized rather than corrected (p. 2). Without mechanisms that destabilize entrenched routines, repeated communicative success may consolidate non-target-like forms, increasing their resistance to change. These errors require more than corrective feedback; they necessitate diagnostic identification and structured destabilization to reopen the interlanguage system to further development.

Autonomous Language Learning

A key strategy for helping students minimize persistent errors is the development of learner autonomy and self-directed learning. When learners become more aware of their own recurring weaknesses, actively monitor their language use, and take responsibility for reviewing and correcting their own output, they are more likely to notice non-target-like patterns and engage in sustained improvement. Benson (2011) associates learner autonomy with learners' capacity to take control of their learning through planning, monitoring, and evaluating their progress (pp. 58–63). Similarly, Little (1991) defines learner autonomy as a capacity for "detachment, critical reflection, decision-making, and independent action" (p. 4). In second language acquisition, Schmidt's (1990) Noticing Hypothesis proposes that conscious attention to linguistic forms is a necessary condition for learning, suggesting that learners must notice problematic forms before they can modify their language use (pp. 140–141). Therefore, fostering autonomy and self-directed learning is not only a motivational strategy, but also an essential condition for reducing recurring stabilized learner errors and supporting sustained proficiency growth.

AI Tools in Language Teaching and Learning

Much has been written recently on the ways in which AI tools can benefit the language teaching and learning process. In her article focusing on DLIFLC students, Amini Harsin (2026) showcases case studies on how teachers can use these tools to create interactive, relevant materials tailored to individual learning needs, noting the importance of “crafting precise and pedagogically sound prompts” (p. 32). She concludes that these “tools offer a valuable partnership for language educators at DLIFLC, where innovative instructional methods are essential due to the program’s fast-paced nature” (p. 49). Wallace and Lima (2025) also note the importance of well-written prompts in using these tools. They propose the *CTPSC Prompting Framework* (Context, Task, Purpose, Sample, Clarification) and offer sample prompts that would be applicable to teaching any language. Students and teachers alike must learn how to use AI tools, and not trust them blindly (Slamet, 2024); training and practice in addition to the critical review of output are crucial.

DLIFLC Teaching and Learning Challenges and Opportunities

DLIFLC instructors in Intermediate and Advanced courses must provide fast-paced instruction in a short period of time, with students being required to take a high-stakes test (the DLPT) at the end of 19 weeks of study. To support students’ learning, instructors in these team-teaching courses use data from a number of pre-course assessments (termed Diagnostic Assessment, DA) in order to plan individualized instruction for each learner. These include online and in-person proficiency tests to pinpoint areas of weakness and strength and also various inventories (i.e., the Motivated Strategies for Learning Questionnaire, Barsch Learning Style Inventory, and Myers-Briggs Type Indicator). These tools are not intended to categorize learners rigidly; rather, they provide instructors with contextual information about processing tendencies, learning strategies (such as reliance on memorized patterns, avoidance of complex structures, or overdependence on translation), and affective factors such as confidence, motivation, and anxiety that may influence performance. Instructors use this information to create instructional plans and materials, but implementing these plans and providing students with sufficient feedback can be challenging due to time limitations and workload expectations.

As an instructor of Chinese at DLIFLC, I was interested in how AI might be able to help students identify and correct their errors and non-native forms. I developed a framework integrating continuous error-pattern detection, micro-targeted adaptive practice, multimodal reinforcement, guided AI-mediated production, and instructor-managed spaced review. Rather than replacing instructor expertise, this framework is designed to enhance diagnostic precision and feedback consistency within existing instructional structures, thereby creating conditions more conducive to noticing and learning, in addition to allowing students to develop as lifelong autonomous learners.

Research Questions

Consistent with the exploratory and practice-based orientation of action research, this action research project addresses the following questions:

1. In what ways can AI tools help teachers and students as they move through the steps of error-pattern detection, micro-targeted drills, multimodal feedback, conversational tutoring, and adaptive spaced review as related to key forms?
2. How do learners describe their experience of AI-supported diagnostic and feedback processes in terms of clarity, usefulness, engagement, and their awareness of key forms?

Action Plan

Participants and Context

This action research was conducted over a 19-week instructional cycle in two Chinese courses at DLIFLC—one Intermediate (entry approximately ILR 2 range) and one Advanced (entry approximately ILR 2+ range). Five learners (three Intermediate and two Advanced) were selected as participants. On average, participants engaged in AI-supported tasks for approximately 4.2 hours per week (both in- and out-of-class). Engagement time was recorded weekly to maintain relative consistency across participants and to contextualize performance trends.

Data Sources

Data collection included each learner's:

- Most recent DLPT score

- Online Proficiency Test administered during Week 1
- In-Person Proficiency Test conducted in Week 2
- Online Proficiency Test administered in Week 12
- Final DLPT score at course completion
- Reflective Journals

The pre- and post-DLPT and OPI scores were used to indicate overall proficiency development and whether each participant's listening and reading scores increased, decreased, or remained stable during the 19-week instructional cycle. The Week 1 Online Proficiency Test and Week 2 In-Person Proficiency Test served as baseline diagnostic measures. These assessments were reviewed by the instructor-researcher to identify recurring grammatical, lexical, and discourse-level patterns across learner output. Reflective journals offered insight into learners' awareness of their own recurring patterns and their perceptions of AI-mediated feedback; these reflections were reviewed during scheduled one-on-one conferences and were used to refine individualized intervention plans throughout the course.

A linguistic structure was classified as a key form only if it met three criteria: (1) it recurred across multiple tasks and modalities, (2) it persisted despite prior instruction or corrective feedback, and (3) it was present in both the Week 1 Online Proficiency Test and the Week 2 In-Person Proficiency Test. Isolated mistakes, fatigue errors, and newly introduced developmental errors were excluded. To determine the number of key form occurrences, the number of confirmed key-form occurrences was divided by the total number of words produced and multiplied by 100 to calculate a rate per 100 words. This allowed comparison of recurring error density before and after the intervention across syntactic, lexical, and discourse domains.

The five strategies described below were designed to operationalize the research questions within the constraints of action research, integrating detection, targeted intervention, learner reflection, and performance monitoring into a unified instructional cycle.

The 5-Step AI-Assisted Learning Framework

To meet students' needs and streamline instruction, I created the 5-step AI-Assisted Learning Framework: Error-Pattern Detection**,** Micro-Targeted Adaptive Exercises**,** Multimodal Feedback**,** AI-Mediated Conversational Tutoring**,** and Adaptive Review through Spaced Repetition. I used this framework throughout the course. These components operated as an integrated diagnostic–instructional loop rather than isolated interventions. Each phase generated data that informed the next. A wide range of learner production—including daily quizzes, translation assignments, reflective journals, weekly research-based presentations, debate performances, and transcribed speaking samples—served as the continuous data stream informing this cycle. All AI tool use was conducted in accordance with DLIFLC and Department of War policy. Microsoft Copilot served as the primary institutionally approved tool used on secure systems; no proprietary or sensitive material was input into AI tools.

Step 1. Detection of Key Forms and Patterns (Diagnostic Phase)

Identifying Key Forms

The researcher used Copilot to streamline the process for identifying key forms. I input learner-created texts (essays, translations, and speaking transcripts) into Copilot. The analysis targeted aspect-marker usage (such as omission or misplacement of 了, 过, 着), word-order deviations (including adverbial positioning and 把/被 constructions), discourse cohesion patterns (for example, overuse or omission of connectors such as 所以 *so*, 因此 *therefore*, 但是 *but*), and lexical precision in formal contexts. I used standardized prompt templates to ensure methodological consistency.

- Sample prompt for specific key forms, learner texts were submitted with instructions such as: "Analyze the following Chinese text and identify recurring grammatical or structural errors. Focus on aspect markers (了, 过, 着), word order, connective devices, and register. Count the frequency of each error type and indicate whether the pattern recurs across clauses."
- Sample prompt for spoken transcripts: "Review this transcript and flag repeated omissions or misuses of aspect markers or connectors. Highlight patterns that occur more than once and may indicate systematic rather than incidental errors."

AI-generated outputs provided frequency summaries and example sentences. All outputs were manually verified, ensuring that AI functioned as a co-analyst rather than an autonomous evaluator.

Identifying Patterns of Key Forms

I categorized recurring patterns into syntactic, lexical, or discourse domains. For instance, repeated omission of 了 was categorized as a syntactic aspect-marking pattern; overreliance on 想 (*want*) in formal argumentative contexts was categorized as lexical overgeneralization; and fragmented clause sequencing without cohesive markers was classified as discourse-level instability. Only patterns that appeared in at least two task types and aligned with Week 2 findings were retained as key forms.

Utilize Contrastive Modeling

To support instructional planning and targeted revision, instructors used AI tools to generate contrastive comparisons between learner-produced sentences (from assignments or assessments) and more target-like forms. These comparisons helped both instructors and learners see exactly how meaning, grammar, and discourse structure differed, with all AI-generated suggestions reviewed and validated by the instructor before being used in class.

Instructors and students used prompts such as: “Compare this learner-created text with a more target-like Chinese version. Identify any grammatical or discourse-level differences, explain why the original form is less natural or has errors, and provide a short teaching explanation.” As an example, one learner had produced the following sentence in an assignment, which was flagged by AI: 我吃饭以后去图书馆 (*I eat and then go to the library.*). Copilot created the revised form: 我吃**了**饭以后去图书馆 (*I ate and then went to the library.*) and generated an explanation that the aspect marker 了 indicates a completed action in past-time narration, and that without 了, the sentence sounds less complete and less natural in this context. The contrast helped the student notice that the issue was not vocabulary, but aspect marking and temporal clarity; the instructor was available to provide follow-up information as needed.

For discourse-level revision, instructors and students used prompts such as: “Revise this learner paragraph to improve logical cohesion and formal register. Focus on connectors, sentence flow, and argument structure. Explain the specific changes made.” As an example, one student produced fragmented logical connections in this sentence, flagged by Copilot: 经济发展很快, **所以**政策变化很大 (*The economy developed quickly, so policy changes were large.*) Copilot suggested the following revision: **由于**经济发展很快，**因此**政策变化也很大 (*Because the economy developed quickly, policy changes also became significant.*). The revised version strengthened logical cohesion by using a more formal cause-and-effect structure (由于 *because*… 因此 *so*…) that is preferred in advanced academic and professional Chinese.

Through this structured sequence—identification, categorization, and contrastive modeling—the process remained systematic, transparent, and pedagogically controlled. With this information, I created a Learner Profile for each student. These individualized profiles directly informed subsequent planning. Using Copilot for analysis and explanations streamlined the pedagogical process, saving me time to focus on other tasks.

Advanced Group Examples: Students K and S

- **Student K**: This student exhibited discourse cohesion errors characterized by repeated clause fragmentation, omission of logical connectors such as 因此, and incomplete syntactic closure. Across a three-week period, 27 instances of such fragmentation were documented across writing and speaking modalities. Because this pattern appeared longitudinally and aligned with Week 2 In-Person Proficiency Test findings, it was classified as a key form.

- **Student S**: This student demonstrated aspect-marker errors, with a 43% recurrence rate of 了 omission in completed events. This syntactic instability co-occurred with lexical overgeneralization in formal contexts, including frequent reliance on 想 (*want*) and 在 (*at*). These patterns were classified as key forms.

Step 2. Micro-Targeted Exercises (Adaptive Practice Phase)

Grounded by Schmidt's (1990) Noticing Hypothesis, I designed Micro-Targeted Adaptive Exercises focusing on each learner's key forms and Learner Profile. This phase included two steps:

1) Task Creation by Instructor: I used AI tools to produce narrowly focused drills targeting a single domain at a time (e.g., aspect marking, discourse connectors, modifier placement). Prompts were standardized to enhance replicability (e.g., "Generate five sentence-ordering exercises requiring appropriate use of 因此 (*therefore*), 然而 (*however*), 不过 (*but*), with model answers" or "Create six fill-in-the-blank items requiring correct use of 了 in past-time narration"). All AI-generated materials were reviewed and refined before implementation to ensure linguistic accuracy and alignment with instructional goals.
2) Student Participation: Learners actively completed drills and revised responses; they also orally rehearsed corrected forms with their peers or the instructor. The instructor curated AI-generated explanations and provided targeted clarification. Key forms that learners produced accurately and consistently in subsequent communicative tasks were entered into instructor-managed review logs for later reinforcement. Key forms that remained unstable re-entered the micro-targeting cycle.

Advanced Group Examples: Students K and S

Advanced-level cases illustrate the process.

- **Student K**: Challenges: discourse fragmentation involved missing connectors and illogical sequencing. AI Exercises: drills emphasizing paragraph reconstruction and connector selection. When errors persisted, prompt parameters were adjusted to reduce lexical load and isolate cohesion variables.

- **Student S**: Challenges aspect omission (particularly missing 了) and lexical overgeneralization recurred across tasks. AI Exercises: structured practice progressed from recognition to controlled production and then to transfer tasks requiring time-sequenced narratives.

Learner reflections indicated that isolating one feature per cycle reduced cognitive overload and increased metacognitive noticing of previously automatized forms. By narrowing focus to entrenched patterns rather than addressing all weaknesses simultaneously, the exercises supported targeted interlanguage destabilization while preserving communicative fluency.

Step 3. Multimodal Feedback (Reinforcement Phase)

In the third stage of the AI-Assisted Learning Framework, students received multimodal feedback to reinforce corrected language forms through synchronized text, audio, and visual input. The purpose of this stage was to help students move beyond simply recognizing errors toward accurately producing corrected forms in listening, speaking, and writing. Students processed the same target form through multiple channels—seeing it, hearing it, and producing it themselves. This approach aligns with multimedia learning theory (Mayer & Moreno, 2003), which suggests that coordinated visual and auditory input improves retention, deeper processing, and long-term procedural control. Learners engaged in structured comparison tasks. For example, when using LTEA Transcript Trainer 5.0 (a DLIFLC-supported learning platform), students first listened to the audio and transcribed the sentence exactly as spoken. They then compared their original utterances with instructor-approved reformulations, listened to native-speaker recordings of corrected forms at adjustable speeds, and reviewed color-coded overlays indicating omissions, misplacements, or structural deviations. Afterwards, they were asked to explain why the corrected version was more natural, particularly in terms of aspect markers, discourse connectors, collocations, or word order differences. Forms that students produced accurately and consistently across multiple subsequent tasks were marked as stabilized. These stabilized forms were then entered into the instructor-managed spaced review system for periodic reactivation in future lessons. Forms that remained unstable were returned to the earlier micro-targeted practice stage for additional focused work.

Step 4. AI-Mediated Tutoring and Role Play (Guided Production Phase)

AI-Mediated Tutoring extended the restructuring process into real-time text-based production. The goal was to help students transfer stabilized language patterns from controlled practice into functional communication. This stage was conducted twice a week through instructor-guided AI interaction sessions. Instructors first designed prompts based on each learner's diagnostic profile. AI tools functioned as simulated interlocutors, generating immediate responses and follow-up questions that sustained interaction and forced learners to negotiate meaning in real time.

Instructors gave students prompts that required the use of targeted structures. For example, for a student who frequently omitted aspect markers, the prompt for Copilot might be: "I am an advanced student studying Chinese Mandarin. I want your help with practicing grammar. I will describe a major decision I made in my military career and explain how that experience has changed my current professional goals. I will use completed actions and resulting states. I want you to identify any errors and explain them." After this practice activity with AI, students were ready to give oral presentations in class.

If students attempted to avoid a targeted structure during interaction, instructors redirected the conversation through follow-up prompts that required its use. This selective intervention prevented cognitive overload while ensuring that practice remained focused, measurable, and transferable to future spontaneous speaking performance.

Step 5. Adaptive Review (Instructor-Guided Spaced-Repetition Phase)

The final component, Adaptive Review through Spaced Repetition, is intended to help learners with long-term retention of key forms. While the earlier phases emphasized identifying errors, correcting them, and practicing them in guided production, this final stage ensured that improvements became stable habits rather than temporary performance gains.

In this phase, instructors—not students—used AI tools to track learner progress across time. After students completed speaking tasks, writing assignments, translation exercises, and classroom discussions, instructors used AI-assisted analysis to review recurring error patterns and compare current performance with earlier samples. This process allowed instructors to identify whether previously corrected forms had become stable or whether they were still likely to reappear. In this sense, “longitudinal aggregation” refers to collecting language samples from different weeks and across multiple skill areas (listening, reading, speaking, and writing) in order to monitor long-term progress rather than judging performance from a single assignment.

The principle behind this stage was spaced repetition: important language forms should be reviewed repeatedly over time rather than intensively only once. Ebbinghaus (1913) demonstrated that forgetting occurs rapidly without distributed review, and later research confirmed that spaced retrieval strengthens long-term retention more effectively than massed practice (Cepeda et al., 2006). In practical teaching, this meant that instructors intentionally brought back previously corrected forms in later lessons, discussions, and assessments to test whether learners could still use them accurately under new conditions.

Instructors monitored four practical indicators of stabilization. First, they checked recurrence frequency across modalities—whether the same error appeared repeatedly in speaking, writing, listening, or translation tasks. Second, they considered learner self-reported confidence during regular student–teacher conferences, asking students whether they felt the form had become natural or still required conscious attention. Third, they observed response latency in spontaneous production—whether students could use the form quickly and automatically or whether they paused noticeably before producing it. Fourth, they examined cross-modal consistency to see whether students could use the corrected form accurately not only in writing but also in spontaneous speaking and listening interpretation. Only forms that remained accurate and automatic across time and context were removed from intensive review cycles. This prevented students from repeatedly practicing structures they had already mastered while allowing instructors to focus attention on forms most likely to re-stabilize as recurring errors.

This final phase completed the AI-Assisted Learning Framework. Diagnostic identification led to targeted intervention; intervention outcomes were monitored across time, and reactivation occurred whenever instability reappeared. In this way, short-term correction was systematically transformed into sustained interlanguage restructuring rather than temporary performance adjustment.

FINDINGS AND DISCUSSION

Three of the five learners (students M, B, K) improved in their DLPT/OPI scores from the beginning to the end of the 19-week course (see Table 1). One student (J) dropped from 2 to 1+ in their reading DLPT score and another student (S) stayed at the same scores. Student S remained at ILR 2+/2+ in Listening and Reading despite substantial improvement in structural accuracy, discourse organization, and control of recurring patterns throughout the instructional cycle. Similarly, Student J achieved ILR 2+ in Listening but did not reach ILR 2+ in Reading. Classroom observations and diagnostic assessments suggested that student J's listening comprehension developed more rapidly than reading fluency and processing speed, particularly when interpreting complex written texts under time constraints. Diagnostic records nevertheless documented reductions in recurring error frequency for both learners across writing samples, guided speaking tasks, and reflective journals. Several students also demonstrated improved self-monitoring and greater awareness of their own recurring patterns, even when these changes were not fully reflected in final DLPT scores. These findings suggest that standardized proficiency outcomes may not fully capture ongoing developmental progress occurring during intensive language instruction.

Table 1 also shows the scores students received on the other proficiency tests they took at the beginning and end of the course. These scores reveal some of the challenges and successes students experienced as they moved through the coursework; dips in scores were used by the instructors to inform instruction. While the AI-assisted learning framework may have contributed to any gains made in DLPT/OPI scores, it should be noted that many other classroom-based and learner-based factors undoubtedly affected learners' scores as well.

Table 1

Pre- and Post-Intervention Scores for Intermediate and Advanced Participants (N=5)

		Pre-Intervention Scores			**Post-Intervention Scores**	
Group	**Student**	**Entry DLPT/OPI (L/R/S)**	**Week 1 Test (L/R/S)**	**Week 2 Test (L/R/S)**	**Week 12 Test (L/R)**	**Final DLPT/OPI (L/R/S)**
INT	**M**	2+/2/1+	2+/3/1+	2/2/1+	1+/3	2+/2+/2
	J	2/2/1+	2/1/1+	2/1+/1+	2+/2+	2+/1+/2
	B	2+/2/2	1+/2/1+	2/2+/2	2/2+	3/2+/2
ADV	**S**	2+/2+/2	1+/2+	2 Low/2 Low	–	2+/2+/2
	K	2+/2+/2	1+/2+	2/2/2	–	3/3/2+

Research Question 1

The first research question asked how AI tools can assist teachers and students with error detection and correction. Table 2 shows the number of key form errors at the beginning and the ending of the 19-week program.

Table 2
Number of Key Form Errors, Before and After Use of AI-Assisted Learning Framework

Error Domain	Pre-Intervention Key Error Rate (/100 words)	Post-Intervention Key Error Rate (/100 words)	% Reduction
Syntactic	0.41	0.24	−41 %
Lexical	0.33	0.19	−42 %
Discourse	0.28	0.15	−46 %
Mean Reduction	-	-	≈ 43%

The approximately 43% mean reduction reflects longitudinal tracking of previously confirmed recurring key-form occurrences. Because learners concurrently received full-time instruction, the reduction cannot be attributed to AI tools in isolation. Instead, the pattern is consistent with the hypothesis that structured diagnostic review cycles—more frequent identification, targeted activation, and scheduled re-checking of stabilized forms—can increase the likelihood that previously persistent routines are destabilized and replaced by more target-like representations. The largest reductions were observed in aspect marking, discourse connector usage, and formal-register precision. Students who participated more consistently in targeted review and feedback activities generally demonstrated stronger gains in structural accuracy and control of recurring patterns.

Table 3 synthesizes representative individual cases and aggregate trends emerging from this cycle.

Table 3

Representative AI-Assisted Learning Framework Interventions on Recurring Learner Errors

Student / Focus Area	Diagnostic Findings (from DA / AI logs)	AI-Generated Micro-Tasks	Post-Intervention Improvement
Student S — Aspect-Marker & Lexical Precision	• 43 % recurrence of aspect-marker omission (了) • Overuse of high-frequency verbs (想 *want,* 在 *at*) in formal writing • Connector overuse (但是 *but*)	**Micro-drill sequence:** 1. Recognition → choose correct aspect marker 2. Controlled → fill-in or translation drills 3. Contextual → summarize short news text inserting appropriate markers 4. Transfer → compose 150-character paragraph with time sequencing	• Aspect-marker accuracy +44 %. • Reduction in 想 *(want)* / 在 (*at*) misuse across 3 weeks. • Student journal (Week 6): "I could see where '在' (*at*) doesn't belong — it finally clicked."
Student K — Discourse Cohesion & Word-Order Accuracy	• 27 instances of fragmented discourse; missing connectors (所以 *so*, 因此 *therefore*) • Modifier/adverbial misplacement; 把 /被 errors • Limited cohesion in paragraph summaries	**Adaptive sequence:** • Sentence ordering to enforce logical flow • Connector selection tasks (然而 *however*, 因此 *therefore*, 不过 *but*) • Sentence combining using 把/被 patterns • Inference-cue simplification when AI detected repeated error	• Discourse cohesion errors −38 %. • Word-order errors −41 %. • Listening summary scores +0.5 ILR.

Students S and K (Table 3) exemplify how iterative targeting of aspect marking and discourse cohesion, respectively, corresponded with measurable declines in recurrence and parallel improvement in performance indicators. These associations do not establish causation but illustrate how focused destabilization of entrenched patterns may align with upward proficiency movement. Taken together, the findings suggest that when key forms are systematically identified, repeatedly activated under communicative demands, and revisited through spaced monitoring, learners may demonstrate both reduced recurrence and improved performance stability. Within high-intensity military language programs, AI-assisted diagnostic support may enhance targeted learning and feedback processes.

Research Question 2

The second research question explored students' perceptions of the use of these AI tools. Weekly reflective journals provided insight into learners' perceptions of the intervention and their evolving awareness of recurrent error patterns. Journal prompts invited learners to identify persistent difficulties and explain how feedback informed their revisions and subsequent performance.

Overall, students reported that they valued the AI-assisted activities and intended to continue using them independently. Many described incorporating AI into an iterative revision cycle. For example, one student wrote, “This time, I want to focus on connectors, collocations, transitional expressions, and contrastive structures.” Such comments indicate increasing metacognitive awareness and intentional control over learning targets and that they use the tools for different goals. Also, learners consistently described AI-mediated feedback as pattern-oriented and cumulative across tasks. Unlike live instructor feedback—which is necessarily selective due to time constraints—AI-assisted analysis repeatedly highlighted the same forms across multiple drafts and transcripts. This recurrence appeared to increase salience. One learner observed, “I didn’t realize how often I overuse ‘在’ (*zai*) until it kept appearing in the corrections. I finally saw the pattern” (Student S, Week 6).

Importantly, learners distinguished between unfamiliar grammar and entrenched usage habits. Several reported that the intervention was most effective for forms they “already knew but kept using incorrectly,” suggesting that AI-assisted feedback was particularly useful for addressing recurring patterns rather than introducing new structures. One student commented, “It’s great at pointing out when you have a habit of using grammar wrong” (Student, Week 11, P55LJY).

Multimodal reinforcement also emerged as a salient theme. Learners reported that synchronized text–audio comparison, color-coded overlays, and adjustable playback speeds facilitated more precise self-monitoring. In particular, students described how comparing their own output with corrected versions across visual and auditory channels helped them identify subtle differences in word order, collocation, and discourse flow. This aligns with the design of the intervention, in which learners repeatedly processed corrected forms through multiple modalities to strengthen retention and proceduralization.

Finally, several learners indicated that they planned to continue using AI tools beyond the course, demonstrating a shift toward greater learner autonomy, with AI functioning not only as a feedback mechanism but also as a self-directed learning partner. More importantly, their reflections suggest that AI was not used passively, but rather as a tool for critical evaluation and informed, self-directed decision-making. One student explicitly described accepting and rejecting AI feedback based on communicative goals: “The words highlighted in yellow are corrections made by the AI; they are more accurate than my original expressions. The words highlighted in green are suggestions I chose to reject. I rejected them for two reasons: first, my original expression was also acceptable; second, although the AI’s version was more formal, my goal was not formality but correctness.” The student further provided examples of rejected suggestions, such as replacing 但是 (*but*) with 然而 (*however*), and 不管 (*no matter*) with 无论 (*regardless of*), indicating an awareness of stylistic nuance and register choice.

This type of reflection demonstrates that learners were not only internalizing corrected forms but also developing the ability to evaluate language choices based on context, purpose, and audience. In this sense, AI functioned not merely as a corrective tool but as a catalyst for higher-level metalinguistic awareness and autonomous control over language use. Taken together, the journal data indicate that the intervention fostered not only pattern recognition and accuracy, but also independent, strategic engagement with language learning beyond the classroom.

Learner reflections suggested increased awareness of recurring language patterns and greater ability to explain corrections metalinguistically. Several learners distinguished between unfamiliar grammar and long-standing habitual errors, indicating increased self-monitoring and attention to form. These observations are consistent with Schmidt's (1990) Noticing Hypothesis, which emphasizes the importance of conscious attention in language development.

Learners also responded positively to multimodal feedback activities, including synchronized text-audio comparison, color-coded visual feedback, and slowed audio playback. These activities appeared to support more precise self-monitoring and reinforced corrected forms through multiple channels of input. This finding aligns with multimedia learning research (Mayer & Moreno, 2003), which suggests that coordinated visual and auditory input may improve retention and procedural control. The integration of AI-supported comparison and feedback tools helped instructors provide these multimodal learning experiences more consistently within the constraints of an intensive instructional schedule.

Pedagogical Implications

Although this study was conducted within a DLIFLC cohort, the framework may also be relevant for other language programs characterized by intensive schedules, rotating instructors, and performance-based assessment systems. In such environments, instructors often face challenges tracking recurring learner patterns across multiple tasks, modalities, and instructional cycles. Without structured support, feedback may become fragmented or inconsistent across instructors and over time.

The AI-Assisted Learning Framework provided a structured method for organizing recurring learner patterns, monitoring changes over time, and supporting more consistent instructional follow-up. By clustering recurring forms and identifying frequency patterns, instructors were better able to focus attention on persistent weaknesses rather than isolated errors. AI functioned as a diagnostic support tool that assisted instructors in identifying patterns, preparing targeted materials, and organizing feedback, while instructional decisions and final validation remained instructor-mediated throughout the process.

Although instructor workload increased during the initial setup phase because of prompt development, calibration, and validation of AI outputs, the framework later reduced repetitive manual tracking and improved continuity across rotating teaching teams. Early investment in diagnostic organization was therefore partially offset by improved efficiency and consistency of feedback.

Reflection and Future Directions

Several limitations should be acknowledged. First, the implementation relied primarily on written production and transcribed speech because of institutional AI tool policies. Future research incorporating real-time oral analysis and prosodic tracking may provide additional insight into spontaneous speaking development. Second, although several learners demonstrated substantial gains, not all participants reached target proficiency levels during the 19-week cycle. Future studies may benefit from delayed post-tests, controlled production tasks, or comparison groups to examine the durability of observed improvements more precisely. Finally, the small sample size (n = 5) and absence of a control group limit generalizability. This study should therefore be understood as exploratory classroom-based action research demonstrating feasibility and instructional potential rather than definitive evidence of causal impact.

CONCLUSION

This action research proposed an AI-Assisted Learning Framework designed to support learner development and autonomy among intermediate and advanced Chinese learners in an intensive instructional setting. Over the 19-week instructional cycle, several participants demonstrated measurable proficiency gains, while recurring grammatical, lexical, and discourse-level patterns declined across multiple task types and modalities. The primary contribution of this study is not the claim that AI independently improves proficiency. Rather, the study demonstrates how AI-assisted diagnostic processes may support identification, reinforcement, review, and monitoring of recurring learner patterns within intensive instruction. Rather than replacing instructor expertise, the framework supported greater visibility of learner patterns and more consistent coordination within a team-taught environment.

Author

Hanwei Tan, PhD, is an Associate Professor at the Defense Language Institute Foreign Language Center (DLIFLC). He earned his PhD in Comparative Literature from Indiana University and has taught and administered language programs in both the US and China, including serving as Chair of the Chinese Department in the School of Resident Education at DLIFLC and Associate Dean of Foreign Languages at Hebei Normal University. His research interests include diagnostic assessment, learner autonomy, AI-assisted language learning, visual poetry, and comparative literary studies. Dr. Tan is the author of *Chinese Visual Poetry: The Path of a Picturesque Literature* (2018) and *A Course in American Literature* (1995), and has published internationally on language teaching, assessment, intercultural learning, literary criticism, and visual poetics.

Ethics and Consent

This action research followed DLIFLC and Department of War guidelines for minimal-risk educational studies. Students were informed of the project's purpose, the types of data collected (test scores, classroom artifacts, AI logs, journals), and their right to opt out at any time without academic consequence. All data were anonymized using pseudonyms, and identifying information was removed from excerpts and tables. AI tools were used only for instructional diagnostics, not for high-stakes evaluation. Data were stored on secure, password-protected institutional systems and accessed only by the instructor-researcher. All procedures adhered to DLIFLC data-security and confidentiality requirements.

References

Amini Harsin, N. (2026). Leveraging AI-driven ChatGPT for pedagogical personalization. *Dialog on Language Instruction, 36*(1), 33–49. https://www.dliflc.edu/ojs/dialog/article/view/44/38

Benson, P. (2011). *Teaching and researching autonomy in language learning* (2nd ed.). Routledge.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching, 5*(1–4), 161–170.

DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). Teachers College, Columbia University. (Original work published 1885).

Han, Z.-H. (2004). *Fossilization in adult second language acquisition*. Multilingual Matters.

Holec, H. (1981). *Autonomy and foreign language learning*. Pergamon Press.

Little, D. (1991). *Learner autonomy 1: definitions, issues and problems*. Authentik.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.

Schmidt, R. (1990). The role of consciousness in second language learning. A*pplied Linguistics, 11*(2), 129–158.

Selinker, L. (1972). Interlanguage. I*nternational Review of Applied Linguistics in Language Teaching, 10*(3), 209–231.

Slamet, J. (2024). Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discover Artificial Intelligence, 4*, Article 46. https://doi.org/10.1007/s44163-024-00143-2

Wallace, L., & Lima, E. (2025). Using ChatGPT to elevate pronunciation pedagogy with the CTPSC prompting framework. *Proceedings of the 15th pronunciation in second language learning and teaching conference.* Iowa State University Digital Press. https://doi.org/10.31274/psllt.18671