# Practical Application and Effect Evaluation of an AI-Empowered Training Model for Enhancing Language and Writing Abilities

**Jin Shi***

*Gansu Finance and Trade Vocational College, Lanzhou 730000, China*
**Corresponding author: 40585577@qq.com*

***Abstract:*** *Constant interaction, feedback-based correction and task matching are essential for improving language and writing ability. The existing training models have problems such as poor refinement of feedback, fixed learning paths and difficulty providing personalized content. Artificial intelligence provides technical help to reform the language and writing training, but merely stacking tools cannot upgrade the training model. In this paper, we construct an AI-enabled training model for improving language and writing ability according to cognitive linguistics and adaptive learning theory from three aspects: theoretical foundation; operating mechanism; and effect evaluation and iterative optimization. We clarify the cognitive characteristics of learners' language processing and constraints of technological intervention, analyse the implementation path of multimodal corpus supply, adaptive learning paths, immediate feedback and human–machine collaboration, establish a multidimensional evaluation and iterative closed loop, choose 42 higher vocational students who majored in finance and business to carry out an 8-week small-sample pilot practice. Our research shows that our model has achieved internal unity among objectives, tasks, content, feedback and learning paths, however, there are some drawbacks (e.g., insufficient AI evaluation of pragmatic appropriateness in workplace scenarios), which can provide a model reference for AI-enabled language and writing training in vocational colleges without facing the problem of technological idling or content generalization.*

***Keywords:*** *AI-enabled; language and writing ability; training model; multimodal corpus; adaptive learning; human–machine collaboration; effect evaluation; language training in vocational education*

## Introduction

The language and writing ability of vocational talents belongs to comprehensive literacy (phonology, vocabulary, syntax, discourse, pragmatics, expressive strategies), converting declarative language knowledge into procedural expressive ability needs frequent language retrieval, combination and adjustment; Uniform pacing traditional training cannot meet individual differences in feedback focused on learning outcomes which leads to mismatched training content and training objectives. Integrating generative AI into diagnosis, content generation, feedback intervention, path adjustment and so on gives opportunities for training reform, but its precondition for technology empowerment is conforming to the laws of language cognition, matching training objectives, leaving room for teacher's pedagogical judgment.

Most of the research uses AI tools and reformation of local classroom; but we should further systemically construct the whole training loop's model construction. We clarify our study on the boundary of human-machine division of labor, improve the feedback precision and path adaptability, give a full reference for designing training programs and reduce technological idling. Improvement of language and writing ability requires continuous interaction and refinement of correction: it is impossible to develop long-term capacity through just accumulation of resources or single score feedback. So, we construct a training model framework and do empirical testing combined with a pilot that can provide an interpretable basis for language and writing training in vocational colleges.

## 1. Theoretical Foundations of Embedding Intelligent Technology in Language and Writing Ability Training

### *1.1 Cognitive Characteristics of Language and Writing Ability Development and Conditions for Technological Intervention*

There are so many levels in ability for language and writing like phonological identification, lexical retrieval, syntactic assembly, discourse cohesion, pragmatic appropriateness, expressive strategies etc., there is a limit because of working memory capacity, distribution of attentional resources and rate of long-term memory retrieval. Declarative knowledge needs to transform into procedural ability by repetition of retrieval, combination and adjustment; Input must be understandable, Output has practical meaning, Feedback shows correctable problem points. Metacognitive ability: Learner recognizes his/her expression deviation, Judges the origin of problems, Chooses modification method which determines that the improvement of language ability depends on continuous interaction not one time theoretical explanation[1].

The combination of AI technology with different cognitive links in language processing (speech recognition to capture pronunciation and fluency) ,semantic analysis to find out lexical and syntactic errors,generative models to generate rewriting and contextual dialogues,adaptive systems dynamically adjust task difficulty and feedback rhythm. The technological intervention shall satisfy these constraints:control extra cognitive load without consuming learners' energy for operating the system;match feedback timing and granularity according to task types so as not to interrupt expression due to being too dense or missing a correction window due to being too sparse;keep training difficulty at the edge of learners' ability but balance between completability and moderate challenge;relying on learners' basic self-regulation ability transform external feedback into internal monitoring.It is supposed to be embedded in the core cognitive processes of teaching rather than detached from teaching logic create an independent learning process.

### *1.2 Key Elements and Collaborative Structure of the AI-Enabled Training Model*

There are six basic components including learner model, language resource repository, diagnostic and generative engine, interactive interface, feedback mechanism, and trainer regulation. Learner model tracks the language foundation, task performance, error type, and progress change of learners which can be used as the base for task matching; Language resource repository stores multimodal corpus of different genres, contexts, and difficulty level; Diagnostic and generative engine detects the state of learners' abilities and generate adaptive exercises, examples, and revision suggestions; Interactive interface conducts various language tasks in listening, speaking, reading, writing, and translation; Feedback mechanism converts the diagnostic result into actionable revision clues; Trainer regulation performs goal setting, contextual interpretation, strategy guidance, and learning maintenance. These should interlock and coordinate each other within the same training cycle rather than being juxtapositioned.

There are several layers for our collaboration (data, algorithm, pedagogy, interaction, feedback). Collecting learning behaviors, infer ability state by algorithm, plan task sequence in the pedagogical layer, complete language input and output in the interaction layer, optimize future learning path in the feedback layer. Division of labor between humans and machines: AI takes over repetitive tasks like high-frequency diagnosis, batch feedback, content generation, difficulty adjustment while trainers take charge of contextual judgment, clarification of expressive intentions, guidance of learning strategies, and maintenance of learning participation. Collaboration is not merely a stack of tools but a dynamic division of labor that forms a closed loop "state identification—task matching—language production—feedback correction—model updating". Operation of the training model depends upon the coupling of many elements, if any one link becomes disconnected, then it will be a scattered set of exercises[2].

### *1.3 Matching Logic Between Ability Objectives and Intelligent Content Generation*

The macro level ability objectives should be broken down into sub-tasks for execution: phonological norms focus on sound discrimination and repetition training; vocabulary objectives focus on collocation, substitution and contextual selection; discourse objectives focus on summarizing, expanding and textual cohesion training; pragmatics objectives focus on role simulation and contextual judgment tasks. Not only does the AI generate texts/exercises but also difficulty parameters, topic scope, modality forms and feedback dimensions too. Vague objectives result in AI producing lots of

content with no relationship to the needs of ability cultivation.

The consistency of objectives, tasks, content, and feedback is central to the matching logic: tasks trigger target abilities; content allows task completion; feedback targets deviations in abilities. Dynamic adjustment of difficulty avoids fear-driven avoidance due to excessive difficulty or lack of promotion of new language combinations due to excessive difficulty. Modalities for training correspond with attribute of tasks (speaking training focuses on pronunciation and immediate interaction whereas writing training focuses on discourse structure and revision traces); topics and registers of generated corpus have to suit a training scenario so that learners can achieve knowledge transfer. Constraints like corpus quality, difficulty range, feedback granularity and value orientation are set based on correspondence of objectives and content instead of blindly increasing resource quantity.

## 2. Operating Mechanism and Support Paths of the AI-Enabled Training Model

### *2.1 Multimodal Corpus-Driven Language Input and Expression Training*

Multimodal corpus combines text, speech, images, video and interactive dialogue into a language learning field in an authentic context. Authenticity prevents memorizing language forms only while enhancing practical use of language. Convert modalities (transcribe key points from listening materials or orally retell after watching videos or rewrite viewpoints after reading texts) constantly connect language form, meaning and context. AI does corpus difficulty grading, collocation extraction, rewriting and expansion, context completion and generate variants in the same semantic field. Supplying corpus is not about quantity it serves learners' language processing. It's necessary for controlling cognitive load and keeping topic continuity, repeating in the same semantic field helps learners consolidate vocabulary network and syntactic chunks[3].

Bridging of input-output by expression training. The focus of speaking training includes fluency, pronunciation accuracy, and immediate organization ability while writing training involves discourse structure, cohesive devices, and revision traces. Scaffolding (collocation prompts, sentence pattern variants, logical prompts) from AI gradually withdraws so that learners don't rely too much on system output and copy-and-confirm them; multimodal corpus's main advantages lie in promoting learners' meaning conversion among various modalities and establishing a form – meaning – context association. Discourse schemata, meaning negotiation, and output-driven learning mechanisms achieve specific tasks: corpus, generative engine, and interactive interface need cooperation for one unified ability objective otherwise the training effect will tend to fragment.

### *2.2 Adaptive Learning Paths and Immediate Feedback Regulation*

The adaptive learning paths are dependent upon the learner model. Besides answer correctness/incorrectness, there are other dimensions (error types, response time, revision traces, expressive complexity, task completion methods) contained in the learner model. Finer diagnosis makes adjustment path sufficient. Dynamically adjust task difficulty/type/sequence/repetition frequency so that there will be a branch where practice lies within learners' zone of proximal development. Adjusting the path does not mean labeling all the learners with fixed abilities but constantly updating our judgment about learning conditions based on learning performance. Adjusting the path needs to be stable since too many switches add up cognitive load and cause learners losing their sense of direction.

Feedback needs to be controlled on time/type/granularity: Direct error correction for objective formal errors; Use prompts/recasts for expressive hesitation; Metacognitive prompts for strategic deviations to help learners find problems by themselves. Make content actionable – no more just scores or vague comments. Differentiate feedback density (we don't want to interrupt thinking with high frequency feedback), sometimes it's better to wait for feedback and summarize feedback. Cyclic regulation function of formative assessment. Feedback starts a new round of learning; Learners need to understand external prompts and internalize them into self-monitoring. A higher level of self-regulation will benefit more from adaptive paths, and feedback strategies should align with task objectives to avoid deviation from the main line of learning[4].

### *2.3 Human–Machine Collaborative Training Organization and Process Management*

Role division in human-machine collaboration training: AI takes charge of high-frequency diagnosis, content generation, data tracking, message push, group suggestions; Trainers take charge of goal setting, contextual explanations, strategy guidance, and maintaining learning participation. The learning analytics dashboard shows process data, but the right to interpret data belongs to teachers; algorithm output is only a probabilistic trend which can't be used for final pedagogical judgment. Distributed cognition theory: Machines take charge of the learning rhythm, teachers take charge of the pedagogical direction; Collaborative effect depends on smooth interfaces, transparent information, and clear rights and responsibilities.

Different foundation of languages, preferences for tasks and available times among students; if we follow one pace, some students will lag while others are repeating at lower levels. We can dynamically group by ability but we should take into account both system scores and interaction performance and expressive needs when grouping. Process management is flexible: tasks can be preset and we allow dynamic adjustment during implementation. Output from AI programs provide suggestions which trainers choose, protecting the space of learner's autonomy. Eliminate role overreach: do not let our system over-intervene in teaching or teachers give up all judgment authority. Incorporate diagnosis, generation, feedback and management into a unified closed loop, retain traceable process evidence for later effect evaluation.

## 3. Effect Evaluation and Iterative Optimization of the Training Model

### *3.1 Multidimensional Evaluation Indicators for the Improvement of Language and Writing Ability*

Language and writing ability assessments should focus on not only vocabulary size and grammar accuracy but also include phonological identification, pronunciation accuracy, lexical richness, collocation appropriateness, syntactic complexity, discourse cohesion, pragmatics appropriateness, expression strategy and fluency matching with different tasks in listening, speaking, reading, writing, and translation. The differences between four levels: knowledge, skills, communication, and self-regulation. Foundation integrates knowledge by skills using communicative tasks that guarantee the continuous development of ability via self-regulation[5].

Evaluation indicators should be observable, traceable and interpretable. Interaction logs, revision traces, speech features, response times and semantic similarity are process evidence; work products and comprehensive tests are outcome evidence. Combination formative assessment and summative assessment including criterion- and norm-referenced assessment. Emphasis placed on reliability and validity but behavior like click counts or page dwell time do not always map onto ability level. Weights vary by task type – speaking emphasizes immediacy and fluency whereas writing emphasizes discourse structure and revision quality. Consider fairness in evaluations too – we cannot use one set of standards to just measure learners with different foundations and different modality preferences.

### *3.2 Comprehensive Judgment Based on Process Data and Outcome Performance*

Process data track learner's behavior in learning, revising, and responding to feedback; Outcome data tracks final products -- each one alone is limited (focusing only on outcomes fails to show how an ability develops while focusing only on processes easily conflate activity level with ability improvement) we need to combine behavioral logs, work products, and trainer observation into a complete evidence chain using time-series analysis comparing learners' revision strategies, expressive complexity, and feedback response patterns across different stages to find patterns of ability change.

We accept that data has noise, missingness, and bias: we can't capture learners' offline thinking or offline language practice; more clicks don't always mean deep learning, long dwell doesn't mean high engagement. Trends from algorithms just as reference instead of replacing teachers' professional judgment. Trainers interpret data based upon task context, learner's background and expressive intention. Adjusting dynamic evaluation criteria: emphasizing participation and basic form at an early stage, task completion and strategy use at a middle stage, expressing stable and fitting with context at a late stage. The quantity of data does not ensure accuracy, but it matters if data explains a change in language ability. Our research conclusion is very cautious, correlation does not necessarily mean causality[6].

### *3.3 Model Adjustment and Optimization Strategies Driven by Evaluation Feedback*

Evaluation results should be returned to the whole process of the model for iteration upgrade. Feedback on evaluations influences each link like goal setting, content generation, task arrangement, feedback strategy, path adjustment, and human–machine division of labor. If some abilities are still weak after a long time period, then adjust objective weights and task types; if there's general comprehension obstacle in the corpus, then optimize difficulty parameters and contextual support; if the feedback is largely ignored by learners, then adjust feedback timing, granularity, and expression methods. Adjustments include short-cycle fine-tune and long-cycle restructuring: fine tune adjusts difficulty, topics, task sequence, and prompt intensity, whereas restructuring adjusts ability objectives, collaborative structure, evaluation indicators, and resource organization. Don't have excessive fine tuning undermining stability or delayed restructuring causing problems to get entrenched.

Build up a loop: "evaluation — interpretation — decision-making — adjustment — re-evaluation", with humans having the decision-making power (algorithms give suggestions; trainers decide; learners interpret feedback), improve data governance on clarification of data sources, collection scope, and use boundaries, improve trainers' literacy in learning analytics and language assessment and disclose evaluation logic and feedback principles to learners, optimize models toward consistent objectives, tasks, content, feedback, and paths instead of chasing an increase of indicators quickly, accumulate interpretable and traceable improvement evidence so that we can achieve long-term improvement of language and writing ability.

## 4. Pilot Application Practice of the Model

### *4.1 Pilot Design*

Pilot: The pilot study involved 42 (experimental group: 21, control group: 21) first-year business and trade students from Gansu Finance and Trade Vocational College. No significant differences existed between pre-tests of the two groups regarding their proficiency. Experimental group used this AI-enabled training model while control group still adopted a traditional lecture plus after-class exercise model. Duration of the pilot is 8 weeks with 2 class hours of offline teaching per week and online tasks.

To explore the impact of our model on Chinese character writing, oral expression and business practical writing we collected feedbacks from teachers and students to identify shortcomings. The tools are as follows: Language Module of the National Smart Education Platform; AI-based Mandarin self-testing system; unified test papers; student questionnaire; teacher interview outline.

### *4.2 Pilot Implementation Process*

Pre-test diagnosis in week 1. AI creates learning profiles, personalizes tasks based on business/trade workplace scenarios. Tiered implementation from weeks 2-7: Teachers teach key/difficult points offline (experimental), AI pushes training with immediate feedback after class; teachers retrieve data to review/tutor individually. In the control group unified instruction is given, manual grading of assignments are done without any personalized push. Summative assessment in week 8 along with questionnaires/interviews.

### *4.3 Analysis of Preliminary Pilot Effects*

Quantitative results: The experimental group has a higher overall average score than that of the control group (especially in speaking and practical writing); Participation in after-class exercises increased; Wait time for error correction was decreased with AI feedback.

Qualitative feedback: Students appreciated personalized training and immediate feedback but noted that AI lacked in judging pragmatically appropriate for business context. Teachers felt relieved of repetitive grading tasks though there were some omissions with AI-generated materials which required manual review.

### *4.4 Reflections on the Pilot*

Small sample, short term pilot has limited generalisability. The lack of pragmatic evaluation on workplace scenarios is one of our main limitations; we need to collaborate with industry experts to build a workplace corpus and improve AI evaluation standards for the future.

## Conclusion

The essence of our AI-based training model to improve language and writing ability lies in maintaining the inner consistency between objectives, tasks, content, feedback and learning paths: nothing can be upgraded by technology alone. In future research, we need to further develop semantic alignment and difficulty grading of multimodal corpus, enhance the stability and interpretability of adaptive paths, optimize the timing, types and granularity of feedback, explore the boundary of rights and responsibilities in human-machine collaboration; we should establish a comprehensive evaluation system combining both process data and outcome performance, build an evaluation-driven iterative closed loop. We pay attention to learners' individual difference, modality preference and long-term retention of abilities, strengthen research on cross-task transfer and language use in authentic scenarios. The optimization of the model should not only achieve short-term scores, but also coordinate the operation of all teaching links, accumulate traceable improvement evidence, provide robust support for cultivating language and writing ability in vocational colleges.

## Fund Projects

- Research on the Construction and Practice of a "Three-Dimensional and Four-Stage" Language and Writing Ability Improvement Training Model Empowered by AI (Project Number: GSYJ-Y2025172)

## References

*[1] Min Shangchao, et al. A Longitudinal Study on the Effects of GenAI-Enabled Personalized Adaptive English Learning. Foreign Language World, 2025(5): 42–49.*
*[2] Tong Ling. Research on Language Learning Assessment from the Perspective of Artificial Intelligence. Chinese Character Culture, 2025(23): 172–174.*
*[3] Xie Jingwei, Shi Wei. Construction of a Digital Human System for Teaching Chinese as a Foreign Language Based on Large Language Models. Journal of China Multimedia and Network Teaching (Mid-Month), 2025(10): 21–24.*
*[4] Liu Hua, Chen Kaiyi. From Expression to Scheduling: Prompt-Driven Human–Machine Collaboration and Reunderstanding of Language Ability. Journal of Social Science of Hunan Normal University, 2025, 54(5): 138–147.*
*[5] Wu Jianhao, Zhou Wanting, Cao Chao. An Empirical Study on Generative Artificial Intelligence Technology Empowering Oral Teaching. China Educational Technology, 2024(4): 105–111.*
*[6] Chen Hua, Cao Yating, Ma Dongmei. Construction of a Multimodal Database of Academic English Speaking for Chinese University Students. Technology Enhanced Foreign Language Education, 2022(3): 45–51+109.*