

Construction and Practice of an AI-Enabled Stratified College English Teaching Model in Application-Oriented Undergraduate Universities

Xiaoting Li^{1*}

¹ School of Foreign Languages, Guangzhou Institute of Science and Technology, Guangzhou, China

*Correspondence: Xiaoting Li

Received: 26 July 2026; Accepted: 19 August 2026; Published: 2 September 2026

Abstract:

Artificial intelligence (AI) technology provides new possibilities for addressing real-world difficulties in stratified college English instruction at application-oriented undergraduate universities, including reliance on a single stratification basis, inadequate differentiation, and the lack of dynamic adjustment. Based on an analysis of the theoretical foundations of stratified teaching and its implementation challenges, and on a review of research progress in AI-enabled foreign language teaching, this study constructs an “AI-enabled stratified college English teaching model for application-oriented undergraduate universities.” Guided by four dimensions (stratification basis, teaching system, AI enablement, and dynamic evaluation), the study designs the corresponding implementation pathway. With the help of intelligent platforms such as PiGaiWang, the teaching practice takes A-level and B-level classes as pilot programs and verifies the effectiveness of the model using platform behavioral data and two teaching cases—argumentative essay writing and Vlog production. The findings indicate that the proposed model has a positive effect on students’ learning engagement and vocabulary learning outcomes across different stratification. AI also plays an effective supportive role in generating writing feedback and facilitating integrated output. This study offers an operable implementation pathway and process-based evidence for application-oriented undergraduate universities to carry out stratified teaching using AI technology.

Keywords:

Artificial Intelligence, Stratified Teaching, Application-oriented Undergraduate University, College English, Data-driven

1. Introduction

The Outline of Plan for Building China into a Leading Country in Education (2024–2035) proposes “exploring effective approaches for digitally enabled large-scale stratified education tailored to individual differences and innovative teaching,” and also calls for “deepening the development of teacher teams empowered by artificial intelligence. [1]” *The Guidelines for College English Teaching* (2020 edition) likewise points out that AI technology has become an important means in foreign

language education and teaching, advocating that universities implement blended teaching models so that students can develop toward individualized learning [2]. These policy directions indicate that using AI to empower college English teaching has entered the stage of practical promotion.

As application-oriented undergraduate universities primarily serve local economic development, they cultivate application-oriented talents that meet the needs of local industries (Yuan et al., 2025) [3]. Under the context of digital transformation, for such universities, language competence has gone beyond the traditional attribute of being merely a language tool, becoming a core supporting element for new-quality talents in cross-cultural collaboration and digital application abilities. Therefore, the college English curriculum in these universities should place greater emphasis on the application of language in real-world contexts. However, with the expansion of student enrollment, the differences in students' English proficiency are substantial, and traditional teaching modes make it difficult to meet the learning needs of students at different stratification, leading to a long-standing contradiction between "not being sufficiently challenged" and "being unable to cope."

Although stratified teaching is widely recognized, practical implementation still faces challenges. From the level of instructional implementation, Yuan Jun'e (2026) [4] identifies technical difficulties such as a single basis for setting objectives, difficulties in matching resources, low classroom implementation efficiency, and untimely evaluation feedback. Sun Qigang (2021) [5], from the perspective of institutional design, reveals that in some schools, stratified teaching raises concerns about educational equity due to unsound class assignment standards and non-transparent procedures. Therefore, it can be seen that the reasonable implementation of stratified teaching needs to satisfy two prerequisites simultaneously: stratification standards should be scientific and transparent, so as to avoid turning into mere selection tools; instructional implementation should be supported by data to ensure that students in all stratification receive appropriate support. Whether and how AI can respond to the above-mentioned difficulties in practice still requires further verification.

Based on intelligent teaching platforms such as Chaoxing, VocabGo program and PiGaiWang, this study constructs an AI-enabled stratified college English teaching model for application-oriented undergraduate universities. The study designs the implementation pathway around four dimensions—stratification basis, teaching system, AI enablement, and dynamic evaluation—and analyzes and tests it based on teaching practice in A-level classes and B-level classes. The study intends to answer the following research questions: (1) How can an AI-enabled stratified teaching model be constructed for application-oriented undergraduate universities? (2) How can AI platforms support learning-student profile diagnosis, differentiated teaching, and dynamic evaluation in teaching practice? (3) What effects does this model have on students' learning engagement and learning outcomes across different stratification?

2. AI Enablement and Stratified College English Teaching: Research Status and Problems

The rationality of stratified teaching has been widely recognized. Vygotsky's theory of the Zone of Proximal Development and Krashen's Input Hypothesis respectively provide theoretical grounds for differentiated instructional design based on students' existing levels—from the perspective of cognitive developmental space and

comprehensible input (Vygotsky, 1982; Krashen, 1978) [6, 7]. However, in college English teaching practice, the effectiveness of stratified teaching has not been fully realized.

Yao Zhongping (2022) [8] notes that stratified teaching currently faces threefold dilemmas: first, the stratification basis is mainly based on entrance-exam scores, lacking systematic consideration of students' learning process behavioral data; second, after stratification, the differences in teaching objectives and content are not obvious, failing to truly implement differentiated teaching; third, stratification outcomes become fixed, with no mechanism for dynamic adjustment, making it difficult to respond to the continuous changes in students' abilities. These limitations often lead stratified teaching to become merely a formality in practice, making it difficult to genuinely address students' learning needs under differentiation.

These dilemmas are particularly prominent in application-oriented undergraduate universities. Constrained by resource shortages, such universities show a clear gap from first-class universities in building digital resource systems, making it difficult to sufficiently meet the needs of teachers and students for high-quality foreign language education resources (Feng et al., 2025) [9]. Moreover, their cultivation systems are also restricted by traditional disciplinary barriers, institutional inertia in major setting, and rigid cultivation model, which makes it difficult to achieve precise and dynamic responsiveness to industry needs (Li et al., 2026) [10]. In terms of stratification, related research still mainly relies on experience-based summaries, with insufficient systematic effectiveness evaluation and model refinement. Foreign language teachers also face structural constraints such as misalignment between the evolution of technologies and the construction of teachers' competencies (Dai & Li, 2026) [11]. As a result, AI-enabled stratified teaching is both necessary and particularly challenging to implement in this type of institution.

The rapid development of AI technology offers new possibilities for addressing the above dilemmas. The rapid advancement of AI technology offers new possibilities for addressing the aforementioned challenges. In a systematic review of AI applications in EFL (English as a foreign language), Alshumaimer and Alshememry (2023) [12] point out that AI-powered technologies have distinctive advantages in synthesizing large volumes of learners' academic data and modeling individualized learning pathways. Their core value lies in enabling dynamic adaptation of instructional content and providing immediate feedback through data-driven approaches, thereby offering technical support for differentiated instruction. In the "diagnosis–configuration–feedback–adjustment" cycle that stratified teaching emphasizes most, AI technology has demonstrated potential functions that align with instructional needs. However, Sun Chengzhi and Liu Wenyu (2025) [13] argue that current research on AI-enabled foreign language teaching mainly focus on functional implementation of AI tools, but fail to deeply integrate AI into the theoretical framework of foreign language teaching and into authentic problem contexts.

Overall, most of existing studies provide theoretical support and technical references for this study, but there remain clear research gaps. Specifically, current research either focuses on demonstrating the functions of AI tools or advocating the stratified teaching; it lacks systematic integration of the two, and lacks pathway design covering the entire process of stratified teaching. In addition, many studies evaluate effects mainly using examination scores, with insufficient attention to

process indicators such as learning engagement and behavioral characteristics, which makes it difficult to reveal the mechanism of how stratified teaching works.

Therefore, this study attempts to use college English courses in application-oriented undergraduate universities as the context, construct an AI-enabled stratified teaching model, and examine its implementation effectiveness through platform behavioral data and case analyses. It is expected to provide reference for educational reform practices in similar institutions.

3. Construction of an AI-Enabled Stratified College English Teaching Model for Application-Oriented Universities

To address the challenges in stratified college English teaching at application-oriented universities—such as an overly singular stratification basis, the lack of dynamic adjustments, insufficient individualized support, and the difficulty of homogenous instruction to meet differentiated learning needs—this study relies on AI-enabled intelligent teaching tools. By integrating the cultivation characteristics of application-oriented universities—career orientation, application orientation, and ability-based stratification—the study reconstructs the stratified college English teaching model from four dimensions: stratification standards, instructional system, technological enablement, and dynamic evaluation.

3.1. Core Principles for Model Construction

The construction of an AI-enabled stratified college English teaching model for application-oriented universities emphasizes the embedding of technology into instructional mechanisms, rather than a simple overlay of technology on top of existing teaching. Based on the career-oriented cultivation goals and the differences in students' learning situations, this study establishes four core principles: precision, applicability, dynamism, and effectiveness. These principles guide the design of the stratification standards, instructional system, technological enablement, and evaluation mechanism, respectively, and run through the entire process of model implementation.

Table 1. Core Principles of the AI-Enabled Stratified Teaching Model and Improvement Points.

Principle	Core Meaning	Key Improvement Points
Precision	Student grouping and teaching level matching supported by multi-source learning data	To reduce errors caused by relying on a single score for placement
Applicability	Teaching “objectives-content-assessment” are oriented toward workplace English application	To strengthen language practice and contextual transfer
Dynamism	Stratification criteria are iteratively updated based on learning performance	To ensure ongoing suitability and alignment of learning levels
Effectiveness	AI handles data analytics and organizes feedback, enabling teachers to focus on crucial guidance	To alleviate teachers' workload for high-intensity grading and feedback

The model developed in this study consists of “implemented” and “proposed for improvement.” Specifically, the stratified design of the instructional system (Section 3.3), the overall architecture enabled by AI technology (Section 3.4), and the dynamic assessment mechanism (Section 3.5) have been implemented in this round of pilot. In contrast, the construction of multidimensional learner profiles as the basis for stratification (Section 3.2), due to constraints related to the pilot period and data

accumulation, remains at the stage of conceptual discussion and will require further research to validate in the future.

3.2. Reshaping the Basis for Stratification: From Single Scores to the Construction Path of Multidimensional Learner Profiles

The basis for stratification determines the appropriateness of the proposed model. Traditionally, college English streaming relies mainly on a one-time placement test score to form a static grouping, which has several limitations, including a relatively narrow range of criteria, the solidification of instructional levels, and difficulty in promptly adjusting to fluctuations in students' learning differences. Against this backdrop, this study proposes an approach to optimize the stratification basis: by leveraging the multidimensional data provided by an AI platform, learners' profiles can be constructed at the stages of "foundational competence—learning process—application potential." This would strengthen the stratification diagnosis by shifting the rationale from "single-test scores" toward "multidimensional evidence."

It should be noted that the multidimensional profile construction and inference mechanism described in this section are intended as a conceptual discussion for improvement; they have not been implemented in the current reform pilot, nor have their effectiveness been validated.

If, in subsequent phases of reform implementation, the multidimensional evidence approach is further operationalized, pilot students may be categorized—based on comprehensive diagnosis—into two types of class structures: "A level" and "B level", and class assignments can then be made accordingly. These two class types keep consistent teaching hours and credit requirements to ensure comparability of the training plans. The specific profile dimensions include: (1) Examination-oriented foundational competence: stage test results as the core. This dimension characterizes students' levels in vocabulary, grammar, and reading. (2) Comprehensive learning behaviors: drawing on platform learning data. This dimension records learning process information such as in-class performance, assignment quality, learning duration, and error distribution. (3) Potential for professional application: by aligning with the needs of students' intended professional directions and their performance in classroom application-based tasks, this dimension evaluates students' language practice capability and their propensity for contextual transfer.

Table 2. Stratified Learner Profile Characteristics and Stratified Objectives.

Class Type	Learner Profile Characteristics	Core Objectives
Level-A Class	Students who scored above 100 in the Gaokao English exam have a solid foundation, strong learning ability, and relatively outstanding listening and speaking ability.	To improve workplace English and cross-cultural / critical-thinking expression, and promote the generation of application tasks
Level-B Class	Students who scored below 100 in the Gaokao English exam have a relatively weaker foundation. Their learning methods and self-regulation need to be developed, and there is a more concentrated demand for improvement in listening and speaking skills.	To strengthen foundational language knowledge, correct pronunciation and learning methods, build basic listening/speaking competence and learning confidence, and lay scaffolding for high-order task abilities in subsequent stages.

3.3. Stratified Design of the Teaching System: Differentiated Allocation of Goals—Content—Methods—Tasks

The stratification criteria clarify students' differences across three dimensions: examination-oriented foundational competence, learning-process behavioral patterns, and potential for professional application. Accordingly, the elements of the teaching system should be configured differently at the design. This approach should both avoid the common problem of traditional stratified teaching—where students are divided into tiers but receive largely similar instruction—and ensure that teaching goals, teaching content, teaching methods, and after-class tasks form an implementable support chain.

Guided by the cultivation principle of application orientation in an application-oriented undergraduate university, this study develops an integrated stratified teaching system across four dimensions—teaching goals, teaching content, teaching methods, and after-class tasks. The system is intended to match the differing learning starting points and development needs of students in different class types, and to provide observable instructional evidence for subsequent multidimensional profile diagnosis.

Table 3. *An Integrated Stratified Teaching System for Two Student Groups.*

Dimension	A-level	B-level
Teaching goals	To enhance professional application potential, strengthen language use and integrated output (oral and written) in workplace contexts while accommodating extended training in cross-cultural communication and critical expression	To focus on test-oriented foundational competence and learning accessibility and consolidate basic language knowledge and learning strategies; to enhance listening, speaking, reading, and writing abilities through scaffolding, thereby laying the groundwork for the subsequent development of higher-order task capabilities.
Teaching content	Using the textbook as a structural framework, supplement it with industry-themed professional cases; conducting a “development-oriented” extension and transfer of key concepts and expressions.	With the textbook as the main thread, addressing foundational weaknesses: streamlining difficult points while supplementing transitional vocabulary, basic sentence patterns, and foundational listening materials
Teaching methods	Project-based instruction, seminar-style classrooms, and task-driven independent inquiry; with the teacher intervening as a facilitator, providing evaluative criteria and process-oriented feedback for high-level output.	Online-and-offline integration, concise instruction with ample practice, and scenario-based simulation; with the teacher lowering task thresholds.
After-class tasks	Assigning comprehensive extended outputs such as professional English research reports, English Vlog, and industry-themed presentations, thereby enhancing applied competence and integrated productive skills.	Assigning foundational consolidation tasks such as vocabulary repetition, sentence-pattern imitation, and short-dialogue reading, thereby solidifying test-oriented competence and establishing effective learning habits.

The stratified teaching system maintains internal consistency across its four dimensions in terms of goal orientation, content support, and task continuity. Specifically, the B-level primarily provides scaffolding for examination-oriented foundational competence and correction of learning-process deviations, whereas the A-level strengthens task-based output and transfer-oriented improvement of professional application potential. Meanwhile, as the multidimensional profiling mechanism is gradually introduced in the pilot stage, the stratified configurations can

form a stage-wise feedback loop with the learner-state review reports, thereby offering evidence for the dynamic adjustment of both A-level and B-level instruction.

3.4. Implementation Pathways for AI-Enabled Empowerment: From Difficulty Diagnosis to Functional Adaptation

Practically, stratified college English teaching often faces three common challenges: unclear and overly subjective stratification criteria, difficulty in sustaining differentiated instruction, and limited dynamic adjustment across tiers. The value of integrating AI lies in functional adaptation: mapping each of the above pain points to corresponding technical supports that can be implemented, thereby transforming stratified teaching from a concept into an operational model.

This study relies on intelligent platforms such as Chaoxing, VocabGo program and PiGaiWang to establish an implementation pathway of “diagnosis—configuration—support—review.”

First, AI-supported learning analytics enable objective stratification. These platforms aggregate placement diagnostics, stage assessments, and assignment process data to produce stage-wise learner-state representations, thereby providing auditable evidence for stratified decision-making and reducing bias caused by experience-based judgment.

Second, AI-enabled question-bank configuration supports differentiated practice. Based on tier identification and error distributions, the question bank automatically generates exercises and tests with varying difficulty levels and focus areas, facilitating differentiated implementation of “one common goal, multiple tiered tasks” and reducing teachers’ workload in repetitive test construction.

Third, AI-driven resource recommendation supports teaching tailored to individual differences. In the post-class stage, adapted resources (e.g., micro-lectures, reading materials, vocabulary exercises) are recommended according to learning trajectories, enabling relatively personalized learning pathways within the framework of whole-class instruction. This helps ensure that support more closely matches learners’ needs.

Fourth, AI-based data review supports dynamic adjustments. The platforms continuously collect performance indicators from the learning process; at stage milestones, learner-state review outputs are generated to provide evidence for whether tiers should be maintained or adjusted.

It should be emphasized that technology and teachers form a complementary closed loop, jointly sustaining the operation of the stratified teaching system. The integration of AI does not weaken the central role of teachers; instead, it reinforces their irreplaceable functions in metacognitive regulation, critical thinking cultivation, and ethical judgment (Vogt & Flindt, 2023) [14].

3.5. Dynamic Evaluation Mechanism: Data-Driven Effect Assessment and Iterative Feedback

Sustained and effective operation of stratified teaching requires a dynamic evaluation mechanism aligned with the stratification basis and AI-enabled functional adaptation described above. This study follows the principles of “precision—application—dynamics—effectiveness,” constructing a “process—output—feedback” evaluation chain that runs in parallel with instructional implementation. This shifts

evaluation from a “terminal” step to a driving engine embedded throughout the entire teaching process.

(1) Process-oriented evaluation: learning analytics as the core

AI platforms continuously collect learning-behavior data (e.g., task completion, accuracy, learning duration, interaction frequency, and error-pattern distributions). At each unit, the system generates stage-wise learning-state briefs and delivers them to teachers for diagnosing issues and adjusting subsequent teaching. The same outputs are also provided to students to help them identify weaknesses and clarify improvement directions, thereby enabling real-time in-course correction.

(2) Output-oriented evaluation: validating learning through language application

Beyond stage assessments, this study establishes output-oriented evaluation tasks centered on whether learners can transform language into demonstrable communicative competence. These tasks are designed to match the competency orientations of the two stratification:

Level A emphasizes the potential for disciplinary or professional application and the transfer of competence. More integrative, task-driven output tasks are assigned. A representative output task for Level-A is producing a Vlog, which guides students to integrate language knowledge and expressive strategies in authentic contexts. Through theme elaboration, content organization, stance expression, and contextual interaction, students demonstrate language use and cross-context transfer capability.

Level-B focuses on consolidating basic competence and correcting the learning process. Application-oriented tasks directly aligned with learning objectives are used to verify whether students achieve “basic competence–understanding–usefulness.”

The results of these output-oriented evaluations serve as key evidence for subsequent stratified adjustments. They are used not only to determine the extent to which stage goals are achieved, but also to provide support for optimizing instructional strategies in the next teaching cycle.

(3) Feedback and iteration: closed-loop adjustment via review reports

The platform generates stage-wise learning-state review reports to support longitudinal tracking and horizontal comparison, producing two types of improvement recommendations: evidence for tier adjustment (supporting advancement or downgrading) and suggestions for optimizing instructional strategies (targeting weak knowledge points or skill dimensions). After teachers review the reports, they make the final decisions and implement the adjustments in the next instructional cycle, thereby forming an iterative loop of “evaluation–feedback–improvement–re-evaluation.”

Overall, process-oriented evaluation focus on “how well learners are learning,” output-oriented evaluation aims at “can learners use it,” and “feedback–iteration” pays attention to “how to optimize next.” Together, these components integrate the stratification basis, instructional design, and AI functional adaptation described above, supporting continuous refinement of stratified teaching.

4. Teaching Practice and Case Analysis

The study is conducted based on routine instructional scenarios. Data analysis primarily relies on cross-class comparisons between the two instructional classes and longitudinal trends across learning tasks.

4.1. Overview of Teaching Implementation

This study implements teaching practice in the first semester of the 2024–2025 academic year within the college English course. The teaching period lasts 16 weeks and involves two instructional classes. Based on students' entrance scores, learners are assigned to the Level-A Class (including 43 students) and the Level-B Class (including 48 students). The two classes are kept consistent in terms of instructional hours and credit allocation to ensure the comparability of the training programs.

Instructional platforms and AI-supported tools include Chaoxing (for daily assignments, attendance, and discussions), Vocabgo program (for unit vocabulary assessments), PigaiWang (for writing review and grading), and Doubao, among others.

4.2. Analysis of Learning Engagement and Behavioral Performance

4.2.1. Comparative Analysis of Assignment Completion Rates Across Classes

Assignment completion reflect students' learning engagement and willingness to complete learning tasks. The semester-wise changes in assignment completion rates on the Chaoxing platform for the two instructional classes are as follows:

Level-A Class: 100% (Weeks 1–4) → 100% (Weeks 5–10) → 100% (Weeks 11–16);

Level-B Class: 93% (Weeks 1–4) → 96% (Weeks 5–10) → 100% (Weeks 11–16)

Overall, both classes maintained high assignment completion rates. The Level-B Class showed a gradual increase, indicating that after stratified teaching, the task difficulty progressively matched students' ability levels, thereby strengthening their motivation to complete tasks. Level-A Class remained stable, suggesting that the expanded tasks did not negatively affect high-level students' learning engagement.

4.2.2. Changes in Correctness Rates on VocabGo program

Four unit vocabulary assessments are conducted via VocabGo program. The changes in correctness rates for each instructional class can be seen from Table 4.

Table 4. Changes in Correctness Rates for the Two Classes

Class Type	Unit 1	Unit 2	Unit 3	Unit 4	Average
Level-A Class	93.20%	84.40%	89.00%	82.60%	87.30%
Level-B Class	81.30%	74.50%	81.40%	78.20%	78.90%
Gap	11.90%	9.90%	7.60%	4.40%	8.40%

The Level-A Class achieved an average correctness rate of 87.3%, whereas the Level-B Class scored 78.9%, resulting in a difference of 8.4 percent. In terms of trends, both classes exhibited fluctuations characterized by a decline in Unit 2, a rebound in Unit 3, and a slight drop in Unit 4. This suggests that the difficulty curve of the vocabulary assessments affected students at both levels in a similar manner.

Notably, the correctness-rate gap between the two classes narrowed from 11.9% in Unit 1 to 4.4% in Unit 4. The primary reason is that the Level-A Class experienced a

larger decline in later units, rather than the Level-B Class catching up. This finding indicates that the vocabulary tasks for the Level-A Class increased in difficulty more rapidly in the later units, and some students had not fully adapted. By contrast, the task difficulty for the Level-B Class remained relatively well matched to students' ability levels.

Overall, this highlights the need for finer design of the difficulty progression slope in stratified teaching, to prevent the expanded tasks assigned to higher-level students from exceeding the range of their current ability.

4.2.3. Comparative Analysis of Average Answering Time on VocabGo program

Answering time can serve as an auxiliary indicator for students' perceived task difficulty and learning engagement. The average times per student for each instructional class in the unit-based vocabulary tests are as follows:

Table 5. Average Answering Time for Vocabulary Tests

Class Type	Unit 1	Unit 2	Unit 3	Unit 4
Level-A Class	4 min	5 min	6 min	6 min
Level-B Class	4 min	5 min	3 min	6 min

Considering the unit correctness data (Level-A Class: 93.2%, 84.4%, 89.0%, 82.6%; Level-B Class: 81.3%, 74.5%, 81.4%, 78.2%), the two classes demonstrate distinct learning-behavior patterns in the relationship between response time and correctness rate.

For the Level-A Class, answering time increased gradually (4–5–6–6 minutes), which generally aligns with fluctuations in correctness rate (93.2% → 84.4% → 89.0% → 82.6%). Specifically, correctness rate in Unit 2 decreased by 8.8 percent while answering time increased by 1 minute, suggesting that students began adjusting their engagement strategies but had not fully adapted. In Unit 3, correctness rate rebounded to 89.0 percent with time increasing to 6 minutes, indicating that greater engagement effectively supported learning outcomes. In Unit 4, answering time remained at 6 minutes while correctness rate declined slightly, implying that the difficulty further increased and students had reached a phase-specific ceiling in their engagement.

For the Level-B Class, answering times were 4, 5, 3, and 6 minutes, respectively, and showed no clear positive association with correctness rates. In Unit 2, correctness fell to 74.5 percent while response time was 5 minutes; in Unit 3, correctness rose to 81.4 percent while response time was only 3 minutes. This suggests that students' performance at this level was mainly determined by knowledge mastery rather than the length of time spent. Students who could answer did so quickly and accurately, whereas those who could not answer continued to spend a long time without achieving better results—an indicator of a typical learning pattern for the Level-B Class.

Overall comparison shows that in Unit 4, both classes experienced a decline in correctness rate accompanied by increased or unchanged response time, indicating that rising vocabulary difficulty in later units posed challenges to both levels. The key difference between the two classes is that the Level-A Class maintained relatively higher correctness rates through sustained increases in engagement, whereas increased engagement in the Level-B Class did not transform effectively into improved correctness. This points to the presence of “ineffective time-on-task” in the Level-B

Class when facing high-difficulty tasks. This difference further supports the necessity of stratified teaching: students at different levels exhibit distinct learning-behavior patterns, requiring separate alignment of task difficulty design and guidance on learning strategies.

4.3. Case One: AI-Assisted Argumentative Writing

4.3.1. Teaching Objectives and Differences Across Stratified Levels

This case centers on argumentative writing as the core task, with objectives focusing on the structurality and readability of English expression, while implementing differentiated requirements across two class types.

For the Level-A Class, the writing task emphasizes comprehensive output ability and higher-order expression quality, including clear viewpoints, coherent argumentation, and richer and more appropriate language use, while for the Level-B Class, the writing task focuses on accurate expression supported by foundational scaffolding, including more stable sentence structures, fewer grammatical errors, clearer cohesion, and expression outcomes that are more readily understandable when listened to or read.

4.3.2. Implementation Process: Human-AI collaborative Writing

This case follows the implementation pathway of task-driven instruction → AI initial evaluation → teacher verification → second-draft revision → consolidation through feedback and instruction:

Step 1: The teacher releases the argumentative writing task through the learning platform and specifies stratified requirements. In particular, it clearly delineates the differences between the two class tracks in terms of the depth of argumentation, linguistic complexity, and revision priorities. Students then complete the first draft in accordance with the requirements and submit it.

Step 2: After submission, students receive automated assessment and feedback from AI. The AI provides targeted comments on grammatical accuracy, spelling, punctuation, and overall readability of the expressions, while also offering supplementary support for vocabulary consolidation. Students review the feedback, identify and locate errors, and record the relevant issues for subsequent revision.

Step 3: The teacher conducts a review and calibration of the AI-generated feedback. Specifically, the teacher filters the AI suggestions, determines the “priority areas for revision,” and transforms them into operational and task-relevant revision instructions, thereby enhancing the feasibility and instructional alignment of the feedback.

Step 4: Students submit a revised second draft based on the teacher-calibrated feedback. The revision requirements emphasize adjustments to rhetorical and linguistic strategies rather than mere word substitution. Students submit the second draft via the platform and retain version-tracking evidence to demonstrate improvements in both structure and language from the first to the second draft.

Step 5: In class, the teacher provides commentary and conducts instruction centered on common problems identified across students’ drafts. Through guided feedback and targeted resource provision, the teacher facilitates the transfer of strategies to subsequent writing tasks, thereby consolidating learning outcomes.

4.3.3. Stratification Manifestation: Same Prompt with Different Revision Thresholds

The same argumentative writing task is differentiated between the two class types primarily through “evaluation focus and revision thresholds.” For Level A, greater emphasis is placed on argumentation structure and overall expression: teacher review prioritizes clarity of thesis, completeness of argumentation within paragraphs, and the extent to which cohesive devices serve the progression of reasoning; AI-generated feedback is employed to pinpoint language accuracy issues. For Level B, greater emphasis is placed on foundational attainment and readability: teacher review prioritizes high-frequency grammatical and syntactic errors, clarity of cohesive expressions, and whether the writing meets basic standards; AI-generated feedback is used for rapid error correction and exemplar reinforcement.

4.3.4. Outcome Summary

The evidence for this case is drawn primarily from comparisons between students’ first and second drafts, platform revision records, and observational feedback from classroom sessions.

AI-assisted scoring enhances the timeliness and precision of feedback, allowing students to receive automated diagnostic feedback at the linguistic level immediately after submitting their first drafts. However, as Godwin-Jones (2024) points out, automated writing evaluation tools, while effective at identifying surface-level language errors, struggle to make accurate judgments regarding the argumentative structure or originality of ideas [15]. This limitation explains the irreplaceability of the teacher review stage in this case—AI undertakes linguistic diagnosis, while the teacher guides the refinement of argumentation and expression strategies.

The second-draft revision mechanism provides institutional support for translating feedback into tangible learning actions. Students revise their work based on dual feedback from AI-based initial evaluation and teacher review. The differences between the first and second drafts constitute visible evidence of learning, reflecting students’ actual improvement in linguistic accuracy and discourse organization. Classroom debriefing further transforms individual cases into shared experience, guiding students to transfer revision strategies to subsequent writing tasks.

4.4. Case Two: English Vlog Project for Level-A

4.4.1. Task Design and Evaluation Criteria

In alignment with the application-oriented cultivation orientation, the Level-A class is assigned an English Vlog project as a comprehensive output task. The Vlog requires not only linguistic accuracy but also the successful execution of content planning, video presentation, and interactive communication. The evaluation of the Vlog consists of the following eight criteria:

The video content demonstrates originality and creativity, with no plagiarism or extensive reliance on others’ work;

The video editing is smooth, with natural transitions, appropriate use of transition effects, and reasonable application of special effects and background music;

The voiceover is clear, with accurate pronunciation and rhythm, moderate pacing, and coherent delivery without excessive pausing;

Bilingual subtitles are complete and accurately match the content;

The video length falls between 90 and 120 seconds and plays properly across different devices;

The video title is visually appealing and relevant to the content, and the cover design is appropriate;

The Vlog concludes with an effective summary or narrative that engages the audience and incorporates interactive elements (e.g., questions or prompts for comments);

The content avoids discriminatory language or inappropriate behavior, ensuring it is wholesome and appropriate.

4.4.2. Implementation Process: A Collaborative Mechanism of AI Assistance and Teacher Oversight

The Vlog project emphasizes “AI empowerment without replacement” throughout its implementation.

Script preparation and language polishing: Students may use the learning platform and AI tools for sentence-level refinement, expression substitution, and language organization; the teacher provides guidance on script structure and expression to ensure that students complete the task through “content creation” rather than simple generation and pasting.

Recording and delivery practice: Students engage in trial readings and multiple recordings, with a focus on pacing, pausing, and coherent delivery; the teacher provides in-class demonstrations and corrective feedback.

Editing and subtitle refinement: Students complete editing, transitions, subtitles, and title/cover design according to task requirements; the teacher exercises critical quality control over readability and accuracy.

Submission and debriefing: After submission, the teacher provides formative feedback based on the rubric, and selects representative works for in-class presentation and experience sharing.

4.4.3. Stratification Manifestation: an Evidence for Level-A Comprehensive Output

The Vlog project is highly consistent with the cultivation objectives of the Level A class: it requires students to integrate language knowledge into authentic contextual expression, and to accomplish the transfer from “comprehension to expression to interaction” through video narration. Compared with purely writing-based tasks, the Vlog more directly demonstrates students’ attainment in comprehensive expression, pronunciation and intonation control, information organization, and communicative strategies.

4.4.4. Outcome Summary

The evidence for this case is drawn primarily from student works (content and presentation quality), classroom presentation observations, and rubric-based attainment records. Practice indicates that the Vlog project enhances student engagement and expressive motivation; meanwhile, the explicit rubric makes evaluation more transparent, enabling students to make revisions and improvements based on feedback.

To sum up, this chapter, through data trend analysis and the presentation of two teaching cases, provides preliminary validation of the implementation effectiveness of the AI-empowered stratified teaching model constructed in Chapter 3.

In terms of learning engagement and vocabulary learning performance, the assignment completion rates and vocabulary accuracy rates of both classes generally show an upward or stable trend, with the accuracy gap between Level B and Level A gradually narrowing—a preliminary indication that stratified teaching has had a positive effect on vocabulary learning among Level-B students. Answering time data reveal that engagement fluctuations among Level-A students broadly correspond to changes in accuracy rates, whereas performance in Level B is predominantly determined by the level of knowledge mastery. This distinction corroborates the design rationale articulated in Section 3.5—namely, that productive assessment prioritizes the verification of language application attainment: students at different proficiency levels exhibit distinct behavioral patterns in learning, necessitating differentiated task difficulty and evaluation approaches.

The two teaching cases further substantiate the model. The argumentative writing case exemplifies the principle of “AI empowerment without replacement” set forth in Section 3.4—AI undertakes initial evaluation and feedback generation, while the teacher is responsible for review and calibration of key points, with the two roles complementing each other; the two-draft revision mechanism transforms feedback into observable learning actions. The Vlog case, meanwhile, embodies the productive assessment outlined in Section 3.5—using authentic expressive tasks as the vehicle to integrate language knowledge into demonstrable comprehensive output. Together, the two cases constitute an evidence chain extending from “process improvement” to “authentic expression,” enabling the dynamic assessment mechanism to move from conceptual framework to actionable implementation.

5. Conclusions

This study, set within the context of college English courses at application-oriented undergraduate universities, constructs an AI-enabled stratified teaching model and designs an implementation pathway encompassing four dimensions: stratification basis, instructional system, AI empowerment, and dynamic assessment. Through teaching practice and data analysis in both Level-A and Level-B classes, the study preliminarily verifies the positive effects of this model in enhancing student learning engagement and vocabulary learning outcomes. The two cases—writing feedback and Vlog production—further demonstrate that AI plays a practicable role in supporting differentiated instruction and cultivating comprehensive expressive competence.

Nevertheless, this study has the following limitations. First, the stratification basis remains primarily reliant on entrance exam scores; the multi-dimensional learner profiling enabled by AI has not been fully implemented in this pilot, and it awaits further validation in subsequent research. Second, the study spans only one semester with a limited sample size, warranting caution in the generalization of findings. Third, behavioral data (e.g., answering time) may be subject to student disengagement or perfunctory responses; future research could employ manual spot-checking and classroom observation for cross-validation.

Future research may proceed in the following directions: introducing a multi-dimensional profiling mechanism to extend the stratification basis from entrance scores to process-oriented data; extending the experimental duration to track the long-

term effects of stratified teaching; and exploring the optimal division of labor between AI and teachers, seeking a balance between technological support and humanistic care. As AI technologies continue to evolve, stratified teaching is poised to benefit from more robust technical support for greater precision and dynamism.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

Funding

This work was supported by Guangzhou Institute of Science and Technology, grant number 2024XB080.

Acknowledgments

The author would like to acknowledge Guangzhou Institute of Science and Technology for funding the institutional research project that served as the basis for this study.

References

- [1] Outline of Plan for Building China into a Leading Country in Education (2024–2035). 2025. Available online: https://www.gov.cn/zhengce/202501/content_6999913.htm (accessed on 15 August 2026).
- [2] Ministry of Education, College Foreign Language Teaching Advisory Board. College English Teaching Guide (2020 Edition); Higher Education Press: Beijing, China, 2020; pp. 32; 9787040552942.
- [3] Yuan, W.; Hu, Z.X.; Bao, Z.M. Building an effective mechanism for the adjustment of academic disciplines and program offerings through systematic thinking. *China Higher Education*, 2025, 8, 9-15.
- [4] Yuan, J.E. Exploring AI-empowered stratified foreign language teaching practice at the university level. *Journal of Beijing Union University*, 2026, 1, 17-23.
- [5] Sun, Q.G. True “stratified teaching” better aligns with the principle of educational equity. *People’s Education*, 2021, 5, 78-80.
- [6] Vygotsky, L.S. *Mind in Society: The Development of Higher Psychological Processes*; MIT Press: Cambridge, MA, USA, 1978.
- [7] Krashen, S.D. *Principles and Practice in Second Language Acquisition*; Pergamon Press: New York, NY, USA, 1982.
- [8] Yao, Z.P. On the practice of implicit stratified teaching mode of college English. *Journal of Jiujiang Vocational & Technical College*, 2022, 2, 41-44.
- [9] Feng, Z.G.; Tong, Y.Z.; Xiao, Y.P.; Zhao, F.F. Practical pathways for cultivating new-quality foreign language talents in application-oriented universities in the digital education era. *China Metallurgical Education*, 2025, 6, 25-29.

- [10] Li, H.Y.; Zhou, L.; Zhang, J. Research on the construction of the talent supply and demand matching mechanism in application-oriented universities. *Jiangsu Higher Education*, 2026, 7, 70-76.
- [11] Dai, S.W.; Li, X. A grounded theory study on the development of AI-TPACK among foreign language teachers in application-oriented universities. *Journal of Hubei Engineering University*, 2026, 4, 79-85.
- [12] Alshumaimer, Y.A.; Alshememry, A.K. The extent of AI applications in EFL learning and teaching. *IEEE Transactions on Learning Technologies*, 2023, 3, 653-663.
- [13] Sun, C.Z.; Liu, W.Y. Artificial intelligence in foreign language teaching: applications and prospects. *Foreign Languages and Their Teaching*, 2025, 5, 77-87.
- [14] Vogt, K.; Flindt, N. Artificial intelligence and the future of language teacher education: A critical review of AI tools for use in the foreign language classroom. In *Innovating the Future of Teacher Education*; Hohaus, P., Heeren, J.-F., Eds.; Brill: Leiden, The Netherlands, 2023; pp 179-199.
- [15] Godwin-Jones, R. Distributed agency in second language learning and teaching through generative AI. *Language Learning & Technology*, 2024, 2, 4-31.



© 2026 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)