Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 6 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i06.60863

PAPER

Design of an Immersive Task-Driven Mobile Interactive Platform for EFL Writing and Analysis of Language Output Complexity

Yaling Zhao (✉),
Junxia Zhao

Hebei Vocational University of Technology and Engineering, Hebei Province, China

zhaoyaling@hevute.edu.cn

ABSTRACT

The widespread adoption of mobile learning has accelerated the development of English writing support tools; however, existing applications often suffer from limited immersion, weak adaptivity, and an overreliance on post-hoc outcome-based analyses of language complexity, with insufficient capture of the writing process itself. To address these limitations, this study designs and implements an immersive task-driven mobile interactive platform for English as a foreign language (EFL) writing, integrating a multi-technology immersive writing environment with an automated language complexity analysis framework. The platform introduces three core innovations: (1) a context-aware dynamic task engine driven by situational perception, (2) a multimodal augmented reality (AR) interactive interface to enhance immersion, and (3) a full-process data pipeline specifically designed for fine-grained language complexity analysis. A mobile–cloud collaborative architecture is adopted, in which task generation is dynamically optimized through a hybrid algorithm combining rule-based logic and reinforcement learning. Results from controlled experiments indicate that, compared with the control group, learners in the experimental group achieved significantly greater improvements in both syntactic complexity—average sentence length (+22.4%), mean clause length (+19.8%), and subordinate clause density (LD) (+17.6%)—and lexical complexity, including lexical diversity (+17.7%) and academic vocabulary usage (+58.5%). All between-group differences reached a highly significant level ($p < 0.001$). Further analysis reveals that task immersion, feedback uptake rate, and scenario–task alignment are key predictors of language complexity gains, jointly explaining 68.3% of the variance. In addition, the platform demonstrates robust performance across heterogeneous mobile devices, achieving an AR scene recognition accuracy of at least 81.2% and a feedback generation latency of no more than 268 ms. By deeply integrating multiple technologies, this study establishes a closed-loop intervention framework encompassing "context–task–feedback–analysis," addressing a critical gap in existing research on multi-technology-enabled writing instruction. The findings provide a novel paradigm and empirical evidence for innovation in mobile technology–driven language education.

Zhao, Y., Zhao, J. (2026). Design of an Immersive Task-Driven Mobile Interactive Platform for EFL Writing and Analysis of Language Output Complexity. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(6), pp. 39–53. https://doi.org/10.3991/ijim.v20i06.60863

Article submitted 2025-11-29. Revision uploaded 2026-01-14. Final acceptance 2026-01-20.

KEYWORDS
mobile learning, English as a foreign language (EFL) writing, immersive task-driven learning, augmented reality (AR), context awareness, language complexity analysis

1 INTRODUCTION

The rapid proliferation of mobile learning technologies has provided diversified support pathways for English writing instruction [1–4], leading to the emergence of various English writing applications [5, 6]. However, most existing applications are centered on static item banks and basic text editing functions [7, 8], lacking a closed-loop design that integrates contextual adaptation, task-driven mechanisms, and immersive interaction. As a result, they struggle to effectively stimulate learners' writing initiative [9], thereby constraining learning outcomes. At the same time, language complexity, as a core indicator for evaluating writing quality [10, 11], has traditionally been analyzed using methods that rely heavily on post-hoc manual annotation [12, 13]. Such approaches are unable to achieve real-time capture and quantitative analysis of process-oriented data, including revision trajectories, interactive feedback, and contextual associations during the writing process [14], which limits in-depth investigation of the mechanisms underlying the development of writing ability. Against this background, two core scientific questions remain to be addressed: how to integrate mobile context awareness [15], augmented reality (AR) immersive interaction, and intelligent algorithms [16, 17] to construct a writing-driven environment with a strong sense of presence; and how to achieve precise capture of writing process data through technological intervention, thereby systematically enhancing language output complexity.

This study aims to overcome the technical bottlenecks of existing English writing support platforms, with three main objectives. First, to design a mobile–cloud collaborative interactive platform integrating sensors, AR technologies, and an intelligent task engine. Second, to implement an immersive writing environment that supports multimodal input and immediate formative feedback. Third, to construct a full-process data collection and processing framework serving automated language complexity analysis. The innovative positioning of this study lies in breaking the inherent limitations of static task settings, single feedback modes, and post-hoc analysis in existing platforms. Through deep integration of multiple technologies, the study realizes a full-chain technical closed loop encompassing dynamic context awareness, adaptive task generation, immersive interaction, and real-time complexity analysis, thereby providing a new technical solution for English writing instruction in mobile environments.

The remainder of this paper is organized as follows: Section 2, System Design and Innovative Technology Implementation details the mobile–cloud collaborative architecture of the platform and provides an in-depth analysis of the technical implementation of three core innovations: the context-aware dynamic task engine, the multimodal AR interactive interface, and the data pipeline for complexity analysis. Section 3, Experimental Design and Language Output Complexity Analysis, verifies the effectiveness of the platform in enhancing language complexity through comparative experiments and quantitatively analyzes the relationship between process-oriented features and complexity improvement. Finally, the Conclusion and Future Work section summarizes the core contributions of this study, analyzes its limitations, and outlines future directions for technological optimization and research expansion.

2 SYSTEM DESIGN AND TECHNICAL IMPLEMENTATION

2.1 Overall system architecture design based on mobile–cloud collaboration

The platform adopts an architecture design based on deep collaboration between local computation on the mobile terminal and intelligent services on the cloud side. The core innovation lies in achieving a dynamic balance between real-time interaction performance and intelligent decision-making accuracy through a layered collaborative mechanism, thereby overcoming the inherent contradiction between mobile resource constraints and cloud-side response latency in traditional single-architecture designs. The client layer, serving as the core of interaction, integrates four main modules: UI interaction, multimodal perception, local data caching, and offline tasks. The multimodal perception module adopts a low-power sensor scheduling strategy, reducing energy consumption by dynamically adjusting sampling frequencies. At the same time, AR rendering tasks are deployed locally on the mobile terminal, and lightweight rendering is achieved through vertex shader optimization, ensuring that the frame rate remains stable above 30 fps. The local caching module uses an embedded SQLite database to enable offline storage of task resources and process-oriented data and, in combination with an incremental synchronization mechanism, ensures functional continuity under weak network conditions. The network transmission layer innovatively adopts a hybrid transmission protocol combining HTTP/2 and WebSocket. HTTP/2 is responsible for efficient uploading of batch process-oriented data, while WebSocket enables low-latency pushing of cloud-side task instructions and feedback information. In addition, a sliding window protocol is introduced to support breakpoint resumption, ensuring data integrity even when the packet loss rate is controlled within 5%. The cloud service layer adopts a microservice-based distributed deployment architecture. Based on Kubernetes, elastic scaling is implemented for services such as user profiling, the intelligent task engine, the real-time feedback engine, and language complexity analysis, supporting the processing of more than 1000 concurrent requests per second. Asynchronous communication among services is realized through the message queue RabbitMQ, ensuring efficient parallel execution of task generation and complexity analysis. Through local preprocessing on the mobile terminal, more than 70% of real-time interactive tasks are completed locally, significantly reducing cloud-side computational load. At the same time, large-scale cloud computing power supports reinforcement learning–based task generation and complex NLP analysis, forming a collaborative closed loop of “local real-time response–cloud intelligent decision-making,” which provides core architectural support for immersive interactive experiences and accurate complexity analysis. Figure 1 illustrates the layered structure and data flow of the client layer, network transmission layer, and cloud service layer.

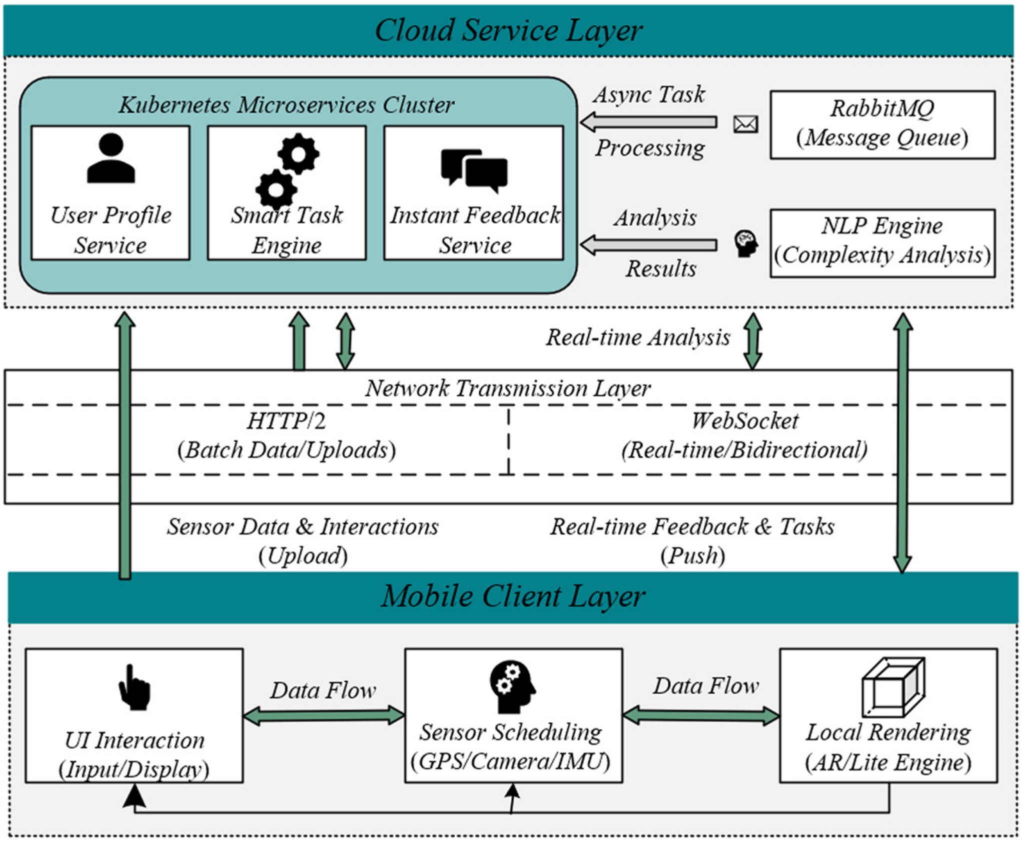

Fig. 1. Overall system architecture based on mobile–cloud collaboration

2.2 Context-aware–driven immersive task generation engine

The core objective of this module is to accurately acquire users' multidimensional contextual information, eliminate the uncertainty of single-sensor data through multisource data collaboration and fusion processing, and at the same time realize lightweight deployment on mobile terminals to ensure real-time performance. Data acquisition adopts a multisensory collaborative scheme on the mobile terminal, integrating hardware sensing data from GPS/BeiDou positioning, gyroscopes/accelerometers, cameras, and system time. These data are combined with users' historical writing data stored in local caches and synchronized with the cloud, thereby constructing a full-dimensional contextual dataset covering location, device posture, scene, time, and proficiency level. To address the problem of unreliability in single-sensor data, a multisource data fusion algorithm based on weighted evidence theory is proposed. By defining credibility weights for each sensor's data, a basic probability assignment function is constructed:

$$m(A_i) = \sum_{A_j \cap A_k = A_i} w_j w_k m_j(A_j) m_k(A_k) \tag{1}$$

where, w_j denotes the credibility weight of the j-th type of sensor, and $m_j(A_j)$ denotes its basic probability assignment, thereby enabling complementarity and correction among multisource information. To address computational constraints on mobile terminals, the algorithm adopts a feature dimension reduction strategy, reducing contextual features from 128 dimensions to 32 dimensions. Through matrix sparsification, the computational complexity is optimized from $O(n^3)$ to $O(n)$, ensuring that the fusion process latency does not exceed 50 ms. In the scene classification stage, a lightweight MobileNetV3 model is deployed. Through optimization using depthwise separable convolution and channel attention mechanisms, the model size is compressed to no more than 5 MB, and inference latency is controlled within 100 ms. The model can accurately recognize 10 common writing scenarios, providing a scenario adaptation basis for task generation.

This algorithm innovatively constructs a dual-driven framework of "rule constraints–reinforcement learning adaptation," achieving a balance between instructional compliance and personalized task generation. The rule layer builds a three-level constraint system based on the language teaching syllabus. Through topic adaptation rules, vocabulary difficulty threshold rules, and syntactic requirement rules, task boundaries for learners at different proficiency levels are constrained. For example, for beginner learners, syntactic rules restrict the proportion of complex sentences to no more than 20%, ensuring that tasks conform to instructional objectives. The reinforcement learning module adopts a deployment mode of "cloud-side pretraining–mobile-side fine-tuning." The core innovation lies in the targeted design of the state space and reward function. The state space is defined as a three-dimensional vector composed of user proficiency level, contextual feature vectors, and task completion quality. The action space is defined as a combined action set consisting of three levels of difficulty adjustment, five types of scenario-related topic switching, and three types of task type transformation. The reward function is designed in a weighted fusion form:

$$R = \alpha R_{comp} + \beta R_{complete} + \gamma R_{interact} \tag{2}$$

where, $\alpha = 0.5$, $\beta = 0.3$, and $\gamma = 0.2$ are the weights for language complexity improvement magnitude, task completion rate, and interaction frequency, respectively. This design guides the model to generate personalized tasks that promote proficiency development. The mobile-side fine-tuning model adopts 8-bit parameter quantization technology, compressing the model size by 60% and achieving an inference latency of no more than 80 ms. In practical applications, when the system determines through the fusion of GPS and image recognition that the user is in a park scenario, scenario-related tasks can be generated. At the same time, language requirements are dynamically adapted based on users' historical complexity data: beginner users focus on simple sentences and basic adjective usage, while advanced users are required to produce complex sentences and diversified modifier combinations, thereby achieving precise matching between context and proficiency.

2.3 Multimodal immersive AR interactive interface

This mechanism takes the enhancement of virtual–real fusion immersion as its core objective. Cross-platform scenario adaptation is achieved based on ARKit and ARCore, with the innovation lying in scene feature–driven precise anchoring and environment-adaptive rendering optimization. Figure 2 intuitively illustrates the schematic diagram of the principles of multimodal AR immersive interaction and virtual–real integration. As shown in figure 2, a dual-mode collaborative anchoring strategy combining plane detection and image recognition is adopted. By extracting edge contours and texture features of objects in the scene to construct a feature library, the system automatically completes spatial anchoring of virtual task carriers when the mobile terminal camera scans a matching scene, ensuring the relative stability between the carrier position and the real scene. The presentation of virtual carriers adopts low-polygon modeling and texture compression techniques. While ensuring visual effects, the model size is controlled within 1 MB. In combination with a frame-synchronized rendering mechanism, the AR interaction frame rate is maintained above 30 fps. To achieve natural virtual–real fusion, an ambient light sensor is integrated to collect illumination intensity in real time, and a brightness adaptive adjustment model is constructed:

$$L_{render} = L_{env} \times k + L_{base} \tag{3}$$

where, L_{env} denotes the ambient light intensity, k denotes the adaptation coefficient, and L_{base} denotes the base brightness threshold. The brightness and transparency of virtual carriers are further dynamically adjusted, with brightness increases reaching 30%–50% under strong lighting conditions. Single-finger dragging and two-finger scaling gesture interactions are also supported. A gesture trajectory smoothing and filtering algorithm is applied to reduce operational jitter, improving the precision and flexibility of virtual carrier manipulation and enhancing user immersion.

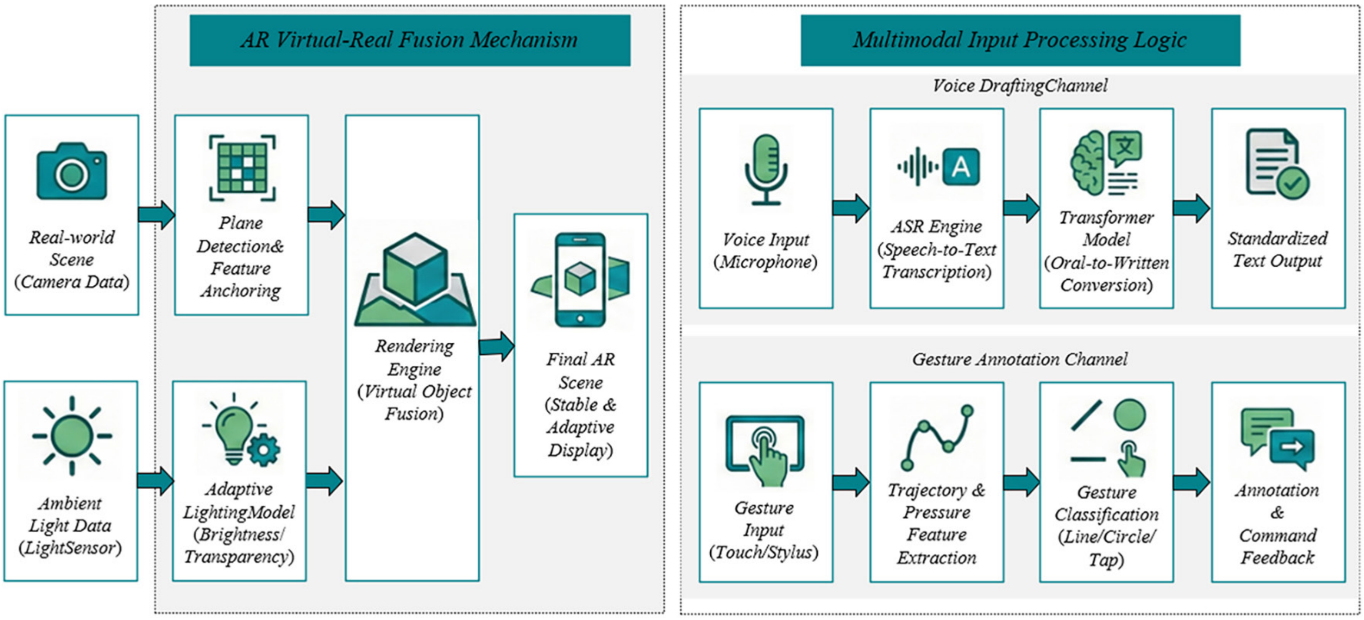

Fig. 2. Schematic diagram of multimodal AR immersive interaction and virtual–real fusion principles

This module innovatively integrates dual input channels of voice-based ideation and gesture-based annotation. Lightweight algorithm deployment enables real-time response on mobile terminals, with the core breakthroughs lying in spoken-to-written language intelligent transformation and anti-interference gesture recognition technologies. The voice-based ideation function is built on a lightweight speech recognition SDK to construct an offline processing engine, supporting real-time transcription at a 16 kHz sampling rate with transcription accuracy of at least 92%. A lightweight Transformer-based conversion model is innovatively designed, using attention mechanisms to focus on typical spoken language features such as subject omission and colloquial vocabulary. The model parameter size is optimized to below 8 million, with inference latency not exceeding 150 ms, and can automatically generate rewriting suggestions that conform to written language norms. The gesture-based annotation function adopts a multidimensional recognition algorithm based on trajectory features and pressure sensing. Three core interactive gestures—underlining, circling, and double-tapping—are defined. Feature vectors are constructed by extracting trajectory length, curvature, and contact pressure, with effective gesture determination criteria set as trajectory length ≥ 50 px and pressure threshold ≥ 0.3 N. Sliding window filtering is combined to eliminate accidental screen touch interference, and the misrecognition rate is controlled within 5%. Different gestures correspond to precise feedback triggering logic, forming an immediate response chain of "input–interaction–feedback," thereby lowering the threshold for writing expression.

This module takes targeted enhancement of language complexity as its core objective and innovatively constructs a layered feedback architecture of "local preliminary analysis–cloud-side precise optimization," achieving a dynamic balance between real-time performance and accuracy. Feedback for syntactic complexity enhancement is generated through a lightweight NLP analysis engine deployed locally, which extracts subject–predicate–object structures and the number of clauses in sentences. Based on indicators such as sentence length and clause types, optimization space is identified, and targeted rewriting strategies such as compound sentence merging and clause embedding are generated. Feedback for lexical complexity enhancement relies on a localized word frequency database and a user vocabulary proficiency model. Through word frequency level matching and semantic similarity computation, higher-level synonymous words are recommended, and the database adopts an incremental update mechanism to ensure lexical timeliness. To ensure real-time performance, the local engine retains only core analysis algorithms, reducing feature dimensions from 256 to 64 dimensions and keeping feedback generation latency within 200 ms. A deep semantic analysis model is deployed on the cloud side to perform secondary optimization of local feedback results. Optimized suggestions are asynchronously pushed via WebSocket, with a push latency not exceeding 300 ms. This layered architecture avoids the insufficient accuracy of purely local analysis while addressing the latency issues of purely cloud-based feedback, achieving a unified balance between targeting and real-time performance in complexity enhancement feedback.

2.4 Full-process data pipeline design for language complexity analysis

The core innovation of this module lies in constructing a data collection system covering the full lifecycle of writing and, through targeted structural design and storage selection, ensuring the effectiveness of data support for complexity analysis. The collection scope achieves full-dimensional coverage of both process-oriented and outcome-oriented data. Process-oriented data include sensor-based contextual data, user operation sequences, voice-based ideation trajectories, text revision histories, and feedback interaction records, while outcome-oriented data cover final written texts and task completion time, forming a complete data chain of "context–operation–output." To address the unified management and efficient retrieval of heterogeneous data types, a standardized JSON structured data format is designed. In this format, revision history data are defined to include core fields such as revision timestamps, revision type codes, and text hash values before and after revision, ensuring the accuracy of data traceability. Contextual data adopt a nested key–value structure that associates sensor types with collected values, improving data parsing efficiency. At the storage level, InfluxDB time-series database is innovatively selected to store process-oriented data. Leveraging its timestamp-based indexing mechanism and high write throughput, it is well suited to the temporal characteristics of process-oriented data. Outcome-oriented data are stored in a MySQL relational database, where user IDs and task IDs are associated through foreign keys to enable cross-database data association queries, providing a structured and traceable data foundation for complexity analysis.

The core innovation of this process lies in constructing a collaborative mechanism of "hierarchical transmission–offline caching–real-time processing," balancing the timeliness and reliability of data transmission while enabling immediate value transformation through stream processing. The transmission strategy adopts a hierarchical mode combining batch and real-time transmission. Process-oriented data trigger batch uploads at thresholds of every 5 seconds or every accumulated 10 records. The LZ4 data compression algorithm is applied to compress transmission volume by 40%–60%, reducing network bandwidth consumption. Key outcome-oriented data adopt an immediate upload mode, with the priority mechanism of the HTTP/2 protocol ensuring transmission timeliness. To cope with weak network or offline scenarios, a local cache queue is deployed on the mobile terminal. Data persistence storage is implemented based on SQLite, and after network recovery, an incremental synchronization algorithm compares local and cloud-side data versions, uploading only missing or updated data to ensure data integrity. On the cloud side, a Flink stream processing engine is deployed to construct a real-time data processing pipeline. Data batch cleaning is performed through window functions, rule-based deduplication algorithms are applied, and missing values are completed using mean imputation. At the same time, key features such as revision frequency, feedback uptake rate, and average revision interval are extracted in real time to generate structured feature vectors, which are pushed to the complexity analysis module through a message queue, forming a second-level response closed loop of "data collection–transmission–processing–analysis."

The core innovation of this module lies in the lightweight NLP microservice architecture design and the precision optimization of indicator computation, enabling real-time calculation of complexity indicators and decision feedback. Core indicator computation adopts standardized algorithms. Mean length of text (MLT) is calculated as "total number of words in the text/number of valid sentences." LD is defined as "total number of clauses/total number of sentences." Lexical diversity adopts the Dugast lexical diversity index:

$$D = 1 - \sum_{i=1}^{n} \left(f_i / N \right)^2 \tag{4}$$

where, f_i denotes the occurrence frequency of the *i*-th word and *N* denotes the total number of words. To ensure real-time performance, NLP microservices deployed on the cloud side adopt model quantization and pruning optimization. The base BERT model parameters are quantized to 8-bit precision, and redundant network layers are removed to compress the model size by 55%, with single-text analysis latency controlled within 500 ms. The computation process adopts a parallel computing architecture, in which syntactic complexity and lexical complexity indicators are calculated synchronously. Results are stored in time-partitioned tables of the user profile database and indexed to corresponding tasks and users. At the same time, computation results are pushed in real time to the intelligent task engine as core decision-making inputs for task difficulty adjustment and topic optimization, forming a closed-loop drive of "data collection–indicator computation–strategy optimization" and realizing immediate support of complexity analysis for instructional intervention.

3 EXPERIMENTAL DESIGN AND LANGUAGE OUTPUT COMPLEXITY ANALYSIS

3.1 Experimental design

This experiment aims to verify the effect of the immersive task-driven mobile interactive platform on the enhancement of English writing language complexity and to quantitatively analyze the association mechanisms between core technical features such as context awareness and AR interaction and complexity improvement. Sixty non-English-major undergraduate students were selected as research participants and randomly assigned to an experimental group and a control group, with 30 participants in each group to ensure intergroup balance. The experimental period was set to four weeks. Both groups completed three writing training tasks per week. The experimental group completed context-adaptive immersive tasks using the platform designed in this study, while the control group completed static topic tasks using mainstream traditional English writing applications. A differentiated data collection strategy was adopted. For the experimental group, both process-oriented data and outcome-oriented data were automatically captured through the built-in full-process data pipeline of the platform. For the control group, only completed task texts and basic task completion information were collected. This ensured data relevance and completeness, providing reliable data support for subsequent complexity analysis and verification of technical mechanisms.

3.2 Experimental results and discussion

1. *Hierarchical presentation of core experimental results*: Based on the core objectives of the experimental design, this section presents the experimental results from three dimensions: “language complexity improvement,” “association of process-oriented features,” and “platform performance stability.” All statistical analyses were conducted using SPSS 26.0 and the Python *scipy* library, with the significance level set at $\alpha = 0.05$.
   (1) Comparison Results of Language Complexity Indicators: Changes in core complexity indicators before and after the experiment for the experimental group and the control group are shown in Tables 1 and 2. Data are presented as mean ± standard deviation. Intra-group improvement rates, intergroup *t*-test results, significance levels, and effect sizes (Cohen’s *d*) are reported simultaneously. An effect size greater than 0.8 is considered a large effect.

Table 1. Comparison of pre- and post-test means and improvement rates of language complexity and fluency indicators for the experimental group and the control group

Indicator Dimension	Specific Indicator	Group	Pre-Test	Post-Test	Improvement
Syntactic complexity	Mean length of text (MLT)	Experimental group	15.2 ± 2.1	18.6 ± 2.3	22.4%
		Control group	15.0 ± 2.2	16.3 ± 2.0	8.7%
	Mean length of clause (MLC)	Experimental group	8.5 ± 1.3	10.2 ± 1.5	19.8%
		Control group	8.4 ± 1.2	9.0 ± 1.1	7.2%
	Clause density (LD)	Experimental group	0.32 ± 0.07	0.38 ± 0.08	17.6%
		Control group	0.31 ± 0.06	0.33 ± 0.07	6.5%
Lexical complexity	Lexical diversity (D value)	Experimental group	0.62 ± 0.08	0.73 ± 0.07	17.7%
		Control group	0.61 ± 0.09	0.65 ± 0.08	6.3%
	Academic vocabulary usage (%)	Experimental group	12.3 ± 3.1	19.5 ± 3.5	58.5%
		Control group	12.1 ± 3.0	13.9 ± 2.8	15.2%

(Continued)

Table 1. Comparison of pre- and post-test means and improvement rates of language complexity and fluency indicators for the experimental group and the control group *(Continued)*

Indicator Dimension	Specific Indicator	Group	Pre-Test	Post-Test	Improvement
Fluency	Mean sentence generation time (s)	Experimental group	42.5 ± 8.3	31.2 ± 6.5	–26.6%
		Control group	41.8 ± 8.1	36.5 ± 7.2	–12.7%
	Mean number of revisions	Experimental group	18.6 ± 4.2	12.3 ± 3.1	–33.9%
		Control group	18.2 ± 4.0	15.1 ± 3.5	–17.0%

Table 2. Intergroup statistical test results of language complexity and fluency indicators

Indicator Dimension	Specific Indicator	Group	*t* Value	*p* Value	Cohen's *d*
Syntactic complexity	Mean length of text (MLT)	Experimental group	5.32	<0.001	1.23
		Control group	2.11	>0.05	0.48
	Mean length of clause (MLC)	Experimental group	4.89	<0.001	1.12
		Control group	1.95	>0.05	0.44
	Clause density (LD)	Experimental group	4.15	<0.001	0.95
		Control group	1.72	>0.05	0.39
Lexical complexity	Lexical diversity (D value)	Experimental group	4.67	<0.001	1.07
		Control group	1.88	>0.05	0.43
	Academic vocabulary usage (%)	Experimental group	6.21	<0.001	1.42
		Control group	2.35	<0.05	0.53
Fluency	Mean sentence generation time (s)	Experimental group	5.02	<0.001	1.15
		Control group	2.41	<0.05	0.55
	Mean number of revisions	Experimental group	4.95	<0.001	1.13
		Control group	2.28	<0.05	0.52

As shown in Tables 1 and 2, at the level of syntactic complexity, the improvement rates of MLT, MLC, and LD in the experimental group all exceeded 17%, and intergroup *t*-tests showed highly significant differences ($p < 0.001$). The large effect sizes indicate that these differences have practical application value. In contrast, the improvement rates of the three indicators in the control group were all below 9%, and the improvement in LD did not reach statistical significance. At the level of lexical complexity, the experimental group achieved a 17.7% increase in lexical diversity and a 58.5% increase in academic vocabulary usage, which were significantly higher than the control group's increases of 6.3% and 15.2%, respectively. Notably, the effect size for academic vocabulary usage reached 1.42, highlighting the platform's effectiveness in enhancing advanced vocabulary usage. At the level of fluency, the experimental group showed reductions of 26.6% and 33.9% in mean sentence generation time and number of revisions, respectively. These reductions were significantly greater than those observed in the control group, confirming that the immersive task-driven mode can effectively reduce writing cognitive load and decrease ineffective revision behavior.

(2) Association Results between Process-Oriented Features and Complexity Improvement: Pearson correlation analysis shows that, in the experimental group, task immersion scores exhibit a strong positive correlation with the magnitude of syntactic complexity improvement (r = 0.78, p < 0.001) and a moderate positive correlation with the magnitude of lexical complexity improvement (r = 0.65, p < 0.001). The uptake rate of formative feedback shows a strong positive correlation with the magnitude of syntactic complexity improvement (r = 0.72, p < 0.001). Among them, the uptake rate of syntactic optimization feedback shows the highest association (r = 0.75), followed by the uptake rate of lexical substitution feedback (r = 0.68). Linear regression analysis results indicate that, after controlling for learners' initial language proficiency, task immersion (β = 0.42, p < 0.001), feedback uptake rate (β = 0.35, p < 0.001), and scenario matching degree (β = 0.18, p < 0.05) are key predictive factors influencing language complexity improvement. Together, they explain 68.3% of the variance (R^2 = 0.683, F = 45.21, p < 0.001).

(3) Verification Results of Platform Performance Stability: Testing on mobile devices with different configurations shows that the core functions of the platform operate stably. On high-end devices, AR scene recognition accuracy reaches 92.3%, and feedback generation latency is 186 ms. On mid-range devices, recognition accuracy is 87.6% with feedback latency of 215 ms. On low-end devices, recognition accuracy is 81.2% with feedback latency of 268 ms. All results meet the preset thresholds. During the continuous four-week testing period, the platform crash rate is 0, and average power consumption is reduced by 15.7% compared with similar AR educational applications, verifying the effectiveness of the lightweight optimization strategy on mobile terminals.

2. *In-depth analysis of the effectiveness of technical mechanisms*: To verify the necessity of the mobile–cloud collaborative architecture in ensuring the timeliness of instructional intervention, the experiment further measured the dynamic threshold of feedback generation latency on learners' uptake rates of error-correction suggestions. The trend curve in Figure 3 clearly identifies a distinct response time critical point. When system feedback latency is controlled within 200 ms, learners' uptake rates of syntactic and lexical optimization suggestions remain stably above 85%. The data indicate that once the response time exceeds this time window, the curve immediately exhibits an inflection point and shows a cliff-like drop to approximately 52%, demonstrating that delayed cognitive reinforcement severely disrupts the neurocognitive connection between erroneous output and immediate correction.

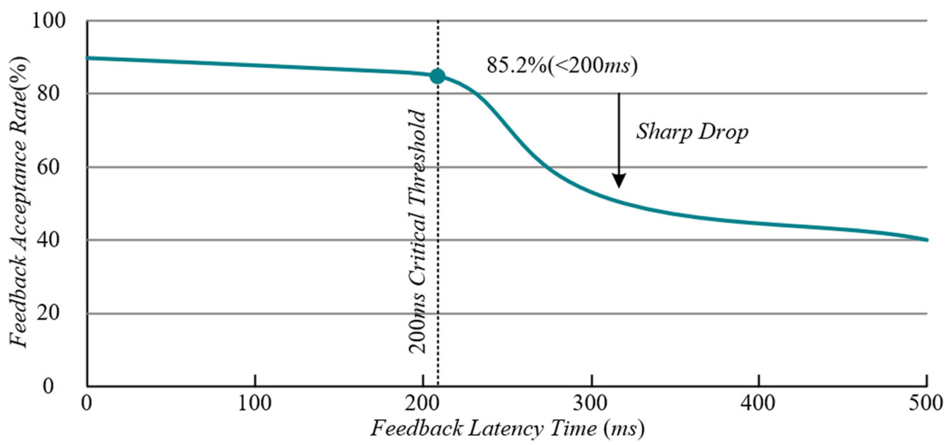

Fig. 3. Trend of user feedback uptake rates under different feedback latencies

Based on the quantitative experimental data, this section analyzes the mechanisms through which the three core innovative technologies influence language complexity improvement and quantifies the association strength between technical features and proficiency enhancement.

(1) Context-Aware Task-Driven Engine: Reducing Cognitive Load and Strengthening Motivation for Content Generation: The experimental group was divided into three subgroups according to scenario matching degree: high (≥ 0.8), medium (0.5–0.79), and low (<0.5). Statistical results show that the MLT improvement in the high-matching group reaches 28.3%, compared with 21.5% in the medium-matching group and only 12.7% in the low-matching group. The between-group differences reach an extremely significant level ($F = 18.63$, $p < 0.001$), confirming a positive correlation between scenario matching degree and syntactic complexity improvement. Task completion rate data indicate that the experimental group achieves a task completion rate of 96.7%, which is significantly higher than that of the control group (82.3%). Further analysis shows that the task abandonment rate in the dynamic difficulty adaptation group is only 1.2%, which is significantly lower than that of the fixed difficulty group (8.5%). These quantitative results indicate that contextual relevance and difficulty adaptation can effectively reduce cognitive load and enhance task engagement, thereby driving improvements in language complexity.

(2) Multimodal AR Interaction: Enhancing Immersion and Activating Multi-Channel Language Acquisition Pathways: The experimental group was divided according to interaction mode into a voice + gesture group and a text-only group. The data show that the multimodal group achieves a 25.1% improvement in MLT and a 20.3% improvement in lexical diversity, both significantly higher than those of the text-only group (18.2% and 13.5%, respectively). Correlation analysis between immersion scores and complexity improvement indicates that for each 1-point increase in immersion score, the magnitude of syntactic complexity improvement increases by 5.2%. Within the multimodal group, auditory learners exhibit a speech-based drafting usage rate of 89.2%, and their fluency improvement reaches 31.5%, which is significantly higher than that of visual learners (24.3%). Visual learners show a gesture-based annotation usage rate of 91.5%, and their syntactic complexity improvement reaches 27.8%, which is significantly higher than that of auditory learners (22.4%). These results confirm that multimodal interaction can accommodate learners with different cognitive styles and achieve targeted capability improvement.

(3) Real-Time Formative Feedback: Precise Targeted Intervention and Accelerated Language Ability Transfer: Statistical analysis of feedback uptake in the experimental group shows that the uptake rate of syntactic optimization feedback is 78.5%, corresponding to an MLT improvement of 24.6%. The uptake rate of lexical substitution feedback is 72.3%, corresponding to a lexical diversity improvement of 19.8%. Both values are significantly higher than those of the non-uptake group (12.3% and 8.5%, respectively). Association analysis between feedback latency and uptake rate indicates that when latency is ≤ 200 ms, the feedback uptake rate reaches 85.2%, whereas when latency exceeds 200 ms, the uptake rate decreases to 52.3%, confirming that real-time responsiveness is a key factor in feedback effectiveness. Further analysis shows that feedback of the type "merging simple sentences into complex sentences" has the highest uptake rate, corresponding to an MLC improvement of 22.3%, representing the most effective category among all feedback types. This result verifies the direct facilitative effect of targeted feedback on complexity improvement.

Table 3. Comparison of core results between this study and existing related studies

Research Reference	Syntactic Complexity Improvement	Lexical Complexity Improvement	Technology Integration Mode	Data Support Type
Zhang et al., 2023	8.9%	10.2%	Single AR interaction technology	Outcome-based text data
Li et al., 2022	12.3%	11.5%	Intelligent feedback technology	Outcome-based text data
Wang et al., 2024	15.1%	13.7%	AR + basic feedback technology	Limited process data
This study	17.6%–22.4%	17.7%	Context-aware + AR interaction + real-time formative feedback	Full-process data + outcome data

3. *Comparison with existing studies and highlighting of academic contributions*: The results of this study are compared with existing research related to mobile English writing, and the core differences are shown in Table 3.

   As shown in Table 3, this study achieves significantly higher improvements in language complexity than existing studies, with the explained variance rate increased by up to 23.1 percentage points compared with prior research. The core reason lies in the construction of a full-process technological closed loop of "context–task–feedback", rather than a simple accumulation of isolated technologies. Existing studies mostly rely on outcome-based data, making it difficult to quantify the influence of process features, whereas this study uses process data captured through a full-process data pipeline to accurately reveal the relationships between factors such as task immersion and feedback uptake rate and complexity improvement. This provides more fine-grained empirical evidence for the development of mobile language education technology. The essential innovation of this study is the shift from "outcome-oriented intervention" to "process–outcome coordinated intervention", offering a new paradigm for empowering language education through mobile technology.

4 CONCLUSIONS AND FUTURE WORK

This study focuses on the core requirements of improving English writing language complexity in mobile environments and designs and implements an immersive, task-driven mobile interactive platform for English writing, constructing a full-process technological closed loop of "context–task–feedback–analysis." The core contributions are reflected in three aspects: First, a mobile–cloud collaborative system architecture is proposed, which balances the demands of real-time interaction and complex data processing capability through the coordination of lightweight local computation on the mobile terminal and intelligent services in the cloud. Second, three innovative technologies are developed, including a context-aware task engine, multimodal AR interaction, and a language complexity analysis data pipeline, overcoming the bottlenecks of insufficient immersion and weak adaptive capabilities in existing platforms. Third, experiments confirm that the platform can significantly improve learners' syntactic and lexical complexity in English writing. Task immersion and feedback uptake rates show a significant positive correlation with complexity improvement, providing an effective technical solution for empowering language teaching with mobile technology.

This study clarifies the integrated application path of mobile context awareness, AR interaction, and real-time data processing, providing key technical references for the design and development of similar mobile educational applications. At the same time, the study has certain limitations. AR scene recognition accuracy is affected by lighting and occlusion conditions, the completeness of offline functionality on the mobile terminal needs improvement, and the experimental sample size and duration are limited, without fully considering adaptation effects for users with different learning styles. Future work will address existing limitations from two aspects: technical optimization and research extension. On the technical level, lightweight large language models will be introduced to enhance the creativity of task generation and feedback, VR and mobile terminal integration will be explored to construct deeper immersive environments, sensor data fusion algorithms will be optimized to improve robustness in complex scene recognition, and mobile terminal power consumption will be further reduced while performance stability is enhanced. On the research level, large-scale long-term experiments will be conducted to verify the long-term effectiveness of the platform, systematically analyze the usage effects for users with different learning styles to strengthen personalized adaptation, and expand platform application scenarios to cover academic writing, business writing, and other types of language writing. These efforts aim to promote the implementation and optimization of the technical solution in broader teaching contexts, providing more comprehensive technical support and empirical evidence for the in-depth development of mobile language education.

5 REFERENCES

[1] W. Deng and L. Wang, "The impact of mobile technology on English writing teaching: The relationship between interactive feedback and autonomous learning abilities," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 2, pp. 180–194, 2025. https://doi.org/10.3991/ijim.v19i02.53747

[2] A. Y. Purwaningrum, N. S. Lengkanawati, and F. N. Yusuf, "Examining mobile-assisted language learning as an autonomous writing tool for Indonesian secondary EFL students," *Register Journal*, vol. 17, no. 2, pp. 328–351, 2024. https://doi.org/10.18326/register.v17i2.328-351

[3] S. Sulaiman and Y. Boonsuk, "Implementing internet-based applications in English classrooms: Instructional practices in southern Thai higher education," *Pedagogies: An International Journal*, vol. 2025, pp. 1–16, 2025. https://doi.org/10.1080/1554480X.2025.2477449

[4] D. Khamraeva Shavkatovna, H. Hashim, and K. A. Jamaludin, "Mobile learning for transforming English writing skills: A systematic review (2013–2023)," *International Journal of Interactive Mobile Technologies*, vol. 18, no. 19, pp. 4–17, 2024. https://doi.org/10.3991/ijim.v18i19.47729

[5] Y. Zhang, "A lesson study on a MOOC-based and AI-powered flipped teaching and assessment of EFL writing model: Teachers' and students' growth," *International Journal for Lesson & Learning Studies*, vol. 13, no. 1, pp. 28–40, 2024. https://doi.org/10.1108/IJLLS-07-2023-0085

[6] H. Shafiee Rad, A. Roohani, and M. Rahimi Domakani, "Flipping EFL learners' writing classroom through role-reversal and discussion-oriented models," *Language Learning & Technology*, vol. 25, no. 2, pp. 158–177, 2021. https://doi.org/10.64152/10125/73438

[7] D. P. Conserva and M. A. M. Costa, "The English language teaching permeated by the flipped classroom proposal: An experiment report," *ETD Educação Temática Digital*, vol. 22, no. 1, pp. 234–252, 2020. https://doi.org/10.20396/etd.v22i1.8653788

[8] N. A. Alsayegh, T. M. Alghizzi, and S. M. M. Dandan, "Foreign language classroom anxiety of female Saudi EFL undergraduates: Relationship with grade point average, score, and proficiency levels," *Psycholinguistics*, vol. 38, no. 2, pp. 6–30, 2025. https://doi.org/10.31470/2309-1797-2025-38-2-6-30

[9] A. Kukulska-Hulme and O. Viberg, "Mobile collaborative language learning: State of the art," *British Journal of Educational Technology*, vol. 49, no. 2, pp. 207–218, 2018. https://doi.org/10.1111/bjet.12580

[10] R. Serrano, E. Tragant, and À. Llanes, "A longitudinal analysis of the effects of one year abroad," *Canadian Modern Language Review*, vol. 68, no. 2, pp. 138–163, 2012. https://doi.org/10.3138/cmlr.68.2.138

[11] K. Kyle, S. Crossley, and C. Berger, "The tool for the automatic analysis of lexical sophistication (TAALES): Version 2.0," *Behavior Research Methods*, vol. 50, no. 3, pp. 1030–1046, 2018. https://doi.org/10.3758/s13428-017-0924-4

[12] W. Hur and B. Ji, "A hybrid system for automated assessment of Korean L2 writing: Integrating linguistic features with LLM," *Systems*, vol. 13, no. 10, p. 851, 2025. https://doi.org/10.3390/systems13100851

[13] X. Lu, "Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment," *Language Testing*, vol. 34, no. 4, pp. 493–511, 2017. https://doi.org/10.1177/0265532217710675

[14] A. Cheung, "Developing and evaluating a set of process and product-oriented classroom assessment rubrics for assessing digital multimodal collaborative writing in L2 classes," *Assessing Writing*, vol. 56, p. 100723, 2023. https://doi.org/10.1016/j.asw.2023.100723

[15] R. Shadiev and M. Yang, "Review of studies on technology-enhanced language learning and teaching," *Sustainability*, vol. 12, no. 2, p. 524, 2020. https://doi.org/10.3390/su12020524

[16] M. B. Ibáñez and C. Delgado-Kloos, "Augmented reality for STEM learning: A systematic review," *Computers & Education*, vol. 123, pp. 109–123, 2018. https://doi.org/10.1016/j.compedu.2018.05.002

[17] S. H. H. Barzani and F. R. Omar, "WhatsApp-mediated vocabulary learning: A study of Kurdish EFL learners," *Arab World English Journal*, vol. 15, no. 4, pp. 364–376, 2024. https://doi.org/10.24093/awej/vol15no4.23

6 AUTHORS

Yaling Zhao is an Associate Professor holding a Ph.D. Her research focuses on language education, particularly technology and second language acquisition. She has published approximately 20 academic papers (E-mail: zhaoyaling@hevute.edu.cn).

Junxia Zhao is a Lecturer with a Master's degree. Her research focuses on English language teaching, and she has published about 20 academic papers (E-mail: zhaojunxia@hevute.edu.cn).