Copyright © the author(s)
Jurnal Onoma: Pendidikan, Bahasa dan Sastra, **Vol. 12, No. 4, 2026**

Developing and Piloting an ADDIE-Based Microlearning EduKit for Mandarin Grammar in SIPEJAR

Octi Rjeky Mardasari[1]
Galuh Bunga Usadani[2]
Amira Eza Febrian Putri[3]
Dhevy Olivia Firdaus[4]
Farhan Dwi Saputra[5]
Sherly Aulia[6]
[123456] Department of Mandarin Language Education, Universitas Negeri Malang, Indonesia

Corresponding author: octi.rjeky@um.ac.id

Abstract

University students learning Mandarin as a foreign language often report that they understand grammatical rules better than they can apply them, while supplementary online resources are generally not integrated with the formal curriculum. This development study aimed to design, produce, and pilot an interactive microlearning EduKit for Mandarin grammar integrated into SIPEJAR, the institutional Learning Management System (LMS). Following the ADDIE model, the article reports the analysis, design, development, and implementation stages, focusing on the development process rather than on effectiveness testing. Data were collected through a learning-needs questionnaire (n = 100), review by one material expert and one media expert, a pretest-posttest trial (n = 32 paired participants), and a post-use response questionnaire (n = 31). In the needs survey, students' mean self-ratings on a 5-point scale were 3.84 for understanding the grammar taught in class and 3.61 for applying it in speaking and writing, and 92% reported needing additional learning media. Six grammar units were designed and produced, each comprising a micro video, an infographic, and interactive exercises, and all were embedded in SIPEJAR. Expert review led to revisions in grammatical accuracy and naturalness, particularly for passive sentences, and in audiovisual presentation. In the pilot of the existential-sentence (存现句 cúnxiàn jù) unit, the mean score rose from 76.72% on the pretest to 93.44% on the posttest, although 17 of the 32 students reached the maximum posttest score; the mean response rating was 4.45 out of 5. Six units were produced, but learner use was assessed only for the existential-sentence unit; the results support further testing rather than establish effectiveness.

Keywords: *microlearning EduKit, Mandarin grammar, ADDIE model, SIPEJAR, learning media implementation*

Introduction

Advances in educational technology have driven the transformation of higher education toward blended learning, a model that integrates face-to-face activities with synchronous and asynchronous online learning within a coherent pedagogical design encompassing content delivery, learning interaction, and assessment (Korson, 2023). Blended learning is designed to enhance learning flexibility and student engagement and to promote student-centered learning, while emphasizing a shift from content-delivery-oriented instruction toward activity-based learning (McKenzie et al., 2022; Müller & Mildenberger, 2021; Rayens & Ellis, 2018). At Universitas Negeri Malang, these principles

are implemented through SIPEJAR, the university's official learning management system (LMS), which enables the integration of learning content, activities, and evaluation.

In Mandarin language learning, however, grammar remains a persistent challenge, since command of sentence structures such as the aspect particles 着(zhe), 了(le), and 过(guò), the 把字句(bǎzìjù) and 被字句(bèizìjù) constructions, and the functional distinctions among 的(de), 得(dé), and 地(dì) is a prerequisite for accurate production. Students often struggle to apply these structures consistently, largely because many grammatical rules have no direct equivalent in their first language (Chang et al., 2017). This difficulty is both conceptual and procedural, in line with the distinction between explicit and implicit knowledge in second language acquisition (Ellis, 2005): students may be able to explain the rule of a sentence pattern yet be unable to use it accurately when producing their own sentences. In this article, understanding a rule refers to recognizing or explaining a correct form, whereas applying it refers to producing the form in a sentence. These involve different performances, ranging from recognizing a correct sentence and completing a gap to producing an original spoken or written sentence. The needs survey reported below captures students' self-ratings of both abilities, whereas the pilot test measures the recognition of correct forms in sentence contexts through multiple-choice items rather than production. Cognitive Load Theory provides a design rationale for this study: because working memory capacity is limited, material whose complexity exceeds processing capacity can impair learning performance (Chen, 2017; Sweller et al., 2019), and dense, unsegmented instruction delivered in a second language further increases the risk of such overload (Shadiev & Huang, 2020). This rationale motivates the segmentation of grammar content into small units; cognitive load itself was not measured in this study.

These limitations lead students to seek supplementary learning resources on social media platforms such as YouTube, TikTok, and Instagram. Such content is presented in engaging short-video formats, and platforms such as TikTok have shown potential as microlearning environments in higher education (Conde-Caballero et al., 2024). However, these resources are not designed around a specific course curriculum; they are therefore not mapped to the learning outcomes of a course such as Grammar II and do not necessarily undergo expert review. Because no systematic audit of these resources was conducted in this study, this point is offered as a contextual rationale for the development rather than as an empirical finding. Previous research on Mandarin learning in Indonesia has likewise focused largely on vocabulary media or on identifying learning difficulties in specific skills, such as an illustrated mini-dictionary (Sholikah & Hidayat, 2022), augmented reality media for vocabulary (Paramitha et al., 2022), and an analysis of reading difficulties in the HSK proficiency test (Irawati & Anggraeni, 2018), as well as audiovisual media for learning Mandarin characters developed with the ADDIE model (Muyassaroh et al., 2024), whereas the development of media that specifically target grammar and are integrated into an institutional LMS remains relatively limited. A recent bibliometric study shows that microlearning research has grown substantially worldwide, with an annual growth rate of 19.37%, yet most studies remain general and technically oriented (Monib et al., 2025). This situation highlights the gap between the abundance of general digital content and the need for grammar instruction that is structured, measurable, and aligned with the Semester Learning Plan (Rencana Pembelajaran Semester, RPS). It also underscores the need to develop interactive microlearning that is integrated into the formal learning system and reviewed before being implemented with students.

Unlike previous studies, which have tended to focus on vocabulary media or on identifying learning difficulties in general, this study offers an approach to developing a grammar-focused microlearning EduKit that is integrated directly into the institutional LMS and designed according to course learning outcomes. The novelty of this study lies in three main aspects: (1) the development of a microlearning EduKit specifically intended for Mandarin grammar learning and implemented in SIPEJAR, the official LMS of Universitas Negeri Malang; (2) content design based on the RPS and on survey data on student needs, in contrast to generic and unstructured social media content; and (3) a staged quality assurance process consisting of material- and media-expert review, followed by an initial implementation trial with students. Table 1 illustrates this design contribution by tracing how one Grammar II learning outcome was translated into a video segment, a practice task, and feedback in SIPEJAR. Accordingly, this article aims to describe the development of an interactive microlearning EduKit for Mandarin grammar learning using the ADDIE model, covering needs analysis, content design, development (production and review), and an initial implementation trial on one learning unit. The pilot question addressed here is whether the 存现句 (cúnxiànjù) (existential sentences) unit could be implemented as intended through SIPEJAR and how students responded to it; whether the EduKit improves learning relative to regular instruction is a separate question left for future effectiveness testing.

Method

This study employed a *Research and Development* (R&D) approach using the ADDIE model (*Analysis, Design, Development, Implementation, Evaluation*), a procedural model that emphasizes a gradual, staged cycle of designing, producing, and trialing instructional products (Branch, 2009). This article reports four stages of the model, namely analysis, design, development, and implementation. Evaluation was not treated as a separate final stage; instead, formative evaluation was embedded in the reported stages through expert review of the design (design stage), expert review of the produced media (development stage), and a one-unit pilot with students (implementation stage). Summative evaluation of the complete EduKit, including effectiveness testing across all six units, was deferred to a later study. Accordingly, the quantitative data are interpreted descriptively as part of formative evaluation (Tessmer, 2013), consistent with the nature of development research, which is oriented toward producing valid and practical products (Richey & Klein, 2007; Sugiyono, 2019).

The analysis stage was conducted by administering a learning-needs questionnaire to 100 students of the Mandarin Language Education Study Program, complemented by a review of the Semester Learning Plan (RPS) and the learning outcomes of the Grammar II course, as well as a mapping of the use of SIPEJAR as the institutional LMS. The questionnaire covered conceptual understanding, practical application, the level and sources of learning difficulties, the need for supplementary media, and preferences for learning methods and media types. All 300 students of the study program in the 2022–2025 cohorts were invited, regardless of whether they had already taken Grammar II, and 100 (33.3%) completed the questionnaire on 5 May 2026 (by student ID: 46 from the 2025 cohort, 37 from 2024, 10 from 2023, and 7 from 2022). The needs analysis therefore reflects the views of students both before and after taking Grammar II. The two key items were “I understand the grammar material taught in class” and “I can use Mandarin grammar in conversation or writing”; both were rated on the same 5-point scale (1 = strongly disagree, 5 = strongly agree), so their means are

comparable as self-ratings. The questions on the most difficult aspects, causes of difficulty, helpful methods, and needed media allowed more than one answer. The 32 pilot participants were third-semester students of the 2025 cohort in one Grammar II class (offering A); 18 of the 31 who completed the pilot response questionnaire had also responded to the needs survey four months earlier, before they took Grammar II.

The design stage focused on translating the results of the needs analysis into a concrete EduKit blueprint, including the selection of six core grammar topics, the development of micro video scripts, the design of infographics, the preparation of interactive exercises, and the formulation of pretest-posttest instruments. The six topics were taken from the materials listed in the Grammar II Semester Learning Plan (RPS). They were therefore determined by curricular coverage rather than ranked by survey responses, as the needs survey asked about grammar difficulties in general rather than about individual topics; the survey instead informed the format and components of each unit. The design was reviewed by one material expert, a Mandarin teacher at SMAK Cor Jesu Malang who holds a bachelor's degree in Mandarin Language Education from Universitas Negeri Malang and has five years of experience teaching Mandarin, using a qualitative review sheet consisting of notes and revision recommendations; the suggested changes recorded on the sheet were followed up through revisions. This review served as an *expert review* in formative evaluation (Tessmer, 2013); because it involved a single reviewer, it is reported as one expert's judgment rather than as agreement among experts.

The development stage involved the actual production of content, including the creation of infographics, the production of micro videos, and the preparation of interactive exercises, followed by review by one media expert using an instructional media review sheet. The media expert first reviewed the first two videos (把字句 and 存现句); the resulting recommendations were applied to the remaining four videos, which were then checked by the same media expert. This second check raised no additional recommendations, and all content was then uploaded to SIPEJAR.

The implementation stage was conducted as a pilot trial of one microlearning unit, 存现句 (existential sentences). This unit was selected because it was the topic scheduled to begin at the third meeting of the Grammar II course, which fell within the research period. It represents the other units in structure, following the same sequence of micro video, infographic, exercises, quiz, and reflection in SIPEJAR, but not necessarily in grammatical complexity, as its video was the shortest of the six. This trial was part of the formative evaluation of the product rather than an effectiveness test with a comparison group. Students took a *pretest* before accessing the EduKit and a *posttest* after a period of independent use through SIPEJAR. The test consisted of 20 multiple-choice items on the use of the 存现句 pattern in sentence contexts. Each correct answer scored one point (maximum score 20), and the pretest and posttest used parallel items. Because the test included no speaking or free-writing task, it measured the recognition of correct forms rather than production, and claims about application in this article are limited to performance on these items. Pretest and posttest data were presented descriptively (means, medians, standard deviations, minimum–maximum scores, and the direction of individual score changes) to identify trends in score change as part of the evaluation of product implementation. After the trial period, students also completed a response questionnaire on a 1–5 Likert scale (strongly disagree to strongly agree) with 18 items covering five dimensions: material and micro video, infographics, interactive exercises and quizzes, ease of access via SIPEJAR, and overall usefulness and satisfaction,

complemented by three open-ended questions on the most useful components, technical problems, and suggestions for improvement.

Data from all stages were analyzed descriptively: qualitatively for the review data and open-ended responses, and quantitatively, in the form of means and percentages, for the needs questionnaire, the pretest-posttest scores, and the response questionnaire scores. Responses to the three open-ended questions were coded by the components or problems they mentioned; a response could name more than one component, and the number of students in each category was counted. All 31 respondents answered every item, so there were no missing responses. Expert comments were followed up through revisions to scripts and content, and the revised materials were returned to the reviewers, who confirmed that their recommendations had been addressed (reviewer confirmation of revisions). This procedure supports the quality of the revisions but is not evidence that the tests measured the intended abilities.

Results

This section reports the results of the four development stages, namely analysis, design, development, and implementation, in the order in which they were carried out; the findings of each stage informed the next. The interpretation of these results in relation to the theoretical framework and previous studies is presented in the Discussion section.

Analysis Stage

The analysis stage aimed to identify students' needs in Mandarin grammar learning as the basis for designing the microlearning EduKit. Students' mean self-rating for the item "I understand the grammar material taught in class" was 3.84 on a 5-point scale (SD = 0.81; strongly agree 21, agree 46, neutral 30, disagree 2, strongly disagree 1). Their mean self-rating for "I can use Mandarin grammar in conversation or writing" was 3.61 (SD = 0.89; strongly agree 16, agree 39, neutral 36, disagree 8, strongly disagree 1). The mean self-rating for application was thus 0.23 points lower than that for understanding, and responses shifted from agree toward neutral.

For the item "I have difficulty understanding Mandarin grammar" (mean = 3.37, SD = 0.90), the most frequent response was *neutral* (52), followed by *agree* (24) and *strongly agree* (13), while 11 respondents disagreed. Respondents could select more than one option for the most difficult aspect of Grammar II (77 selected two or more): remembering grammatical patterns (70/100), using grammar in conversation (54), understanding sentence structures (45), using grammar in writing (39), and understanding theoretical explanations (25). Multiple selection was also allowed for causes of difficulty (62 selected two or more). The most frequently selected cause was insufficient practice (74/100), followed by a weak grasp of prerequisite material (57), material perceived as too difficult (31), a lack of examples (19), and lecturer explanations perceived as too fast (9). Conversely, the item "The exercises given help me understand the material" yielded a mean of 4.11 (SD = 0.84; 79 agreed or strongly agreed). These responses informed the inclusion of interactive exercises and quizzes in every unit, whereas the micro video length, infographic layout, and game-based quiz format reflected the developers' judgment.

The need for supplementary learning media was also high: 91 of the 99 students who answered this item (92%) reported needing additional learning media beyond classroom materials. The media most often selected were more example sentences (65), instructional videos (61), Mandarin learning applications (54), additional practice exercises (51), and additional books (22), and open-ended answers described desired

characteristics including interactivity, animation, everyday-life contexts, and interactive quizzes. Regarding methods, students most frequently chose direct explanation by the lecturer (78) and practice exercises (65) as the most helpful, followed by conversation practice (57), instructional videos (45), and group discussion (26). Together, these results served as the main reference for determining the content structure, the sequence of sentence patterns, and the types of exercises developed in the design stage.

Design Stage

The design stage focused on translating the needs analysis results into a concrete EduKit design, including the development of micro video scripts, the design of infographics, the preparation of interactive exercises, and the formulation of *pretest-posttest* instruments. The design covered six core grammar topics taken from the Grammar II RPS: 把字句 (the 把 construction), 存现句 (existential sentences), 比较方式 (comparative sentences), 被动句 (passive sentences), 联合复句 (coordinate complex sentences), and 偏正复句 (subordinate complex sentences). Each topic was allocated two meetings, covering 12 of the 16 meetings planned for one semester. At the time of writing, the infographics for all six topics had been fully produced, while the micro video scripts and exercise designs had been completed as production blueprints for the next stage.

Each microlearning unit was designed with consistent components: a micro video explaining the sentence pattern, an infographic summarizing the pattern, contextual examples, interactive exercises, and a formative quiz, giving students a complete learning cycle from concept presentation to independent self-testing. The infographics were designed with a consistent visual approach, presenting the sentence pattern, its grammatical function, usage notes, and bilingual example sentences (Mandarin and Indonesian), so that they can serve both as a companion to the micro videos and as a *quick reference* for independent study outside class hours. Table 1 shows one complete learning sequence from the 存现句 unit, illustrating how a course learning outcome was translated into a video segment, a learner prompt, an expected response, and feedback after an incorrect answer.

Table 1. An example learning sequence in the 存现句 unit

Component	Example from the 存现句 unit
Course learning outcome (CPMK) and sub-outcome	CPMK: Students can analyze and distinguish types of sentences in Mandarin and apply this understanding in communicative contexts without reference aids. Sub-CPMK: Students can correctly write 存现句 sentences in Mandarin without consulting a book or dictionary.
Video segment	Segment ending at about 05:26 (video length 9:54). Pattern: Location + verb + 着 + person/object, e.g., 窗边还坐着几个学生。 ('A few more students are sitting by the window.')
Learner prompt (interactive exercise, item 3)	房间里（　）十几位客人。 ('In the room ... a dozen or so guests.') Options: A. 站着 B. 坐着 C. 放着 D. 来了
Expected response	B. 坐着, giving 房间里坐着十几位客人。 ('A dozen or so guests are sitting in the room.')
Feedback after an incorrect answer (e.g., C. 放着)	Guests (客人) are people. To describe the presence of people in a room as a static state, 坐着 (sitting) is the most common choice. Option C (放着) is incorrect because 放 is used only for inanimate objects.

Vol. 12, No. 4, 2026
ISSN 2443-3667(print) 2715-4564 (online)

As part of quality assurance before full production, all material designs were validated by a material expert, a Mandarin teacher at SMAK Cor Jesu Malang. The validation yielded several important revision notes, particularly for 被动句 (passive sentences), which has several markers (被, 叫, and 让) that appear similar on the surface but differ in their rules of use. The first note concerned the completeness of a passive example in its dialogue context. The reviewer pointed to the particle 了 or a resultative complement after the verb, which makes a 被 sentence sound complete; for example, "你差点儿被老师批评" was revised to "你差点儿被老师批评了". In the dialogue, one student tells a classmate who has arrived late, after his phone was stolen, that he had nearly been reprimanded during the teacher's roll call; in this context, the reviewer judged that the sentence sounded unfinished without 了, which commonly closes the 差点儿…了 construction. This is the reviewer's judgment about natural wording in this example rather than a grammatical requirement for every passive sentence. The second, more fundamental note identified several dialogue examples using 叫 and 让 that are, in fact, 兼语句 (pivotal sentences) rather than true 被动句; these examples were therefore replaced so that they consistently represent the correct 被动句 structure. The third note concerned naturalness: grammatical but unidiomatic sentences such as "我的手机被我忘在宿舍了" were identified, and the more natural 把字句 pattern "我把手机忘在宿舍了" was recommended instead. All feedback was thoroughly addressed through revisions of scripts and example sentences before the 被动句 material was produced as final video and infographic content.

Development Stage

The development stage involved the actual production of EduKit content, including the creation of infographics, the production of micro videos, the preparation of interactive exercises, and media-expert validation before the content was uploaded to SIPEJAR. At the time of writing, all infographics and micro videos for the six topics had been produced, uploaded to the research team's official YouTube channel (Figure 1), and fully embedded in SIPEJAR for student use.

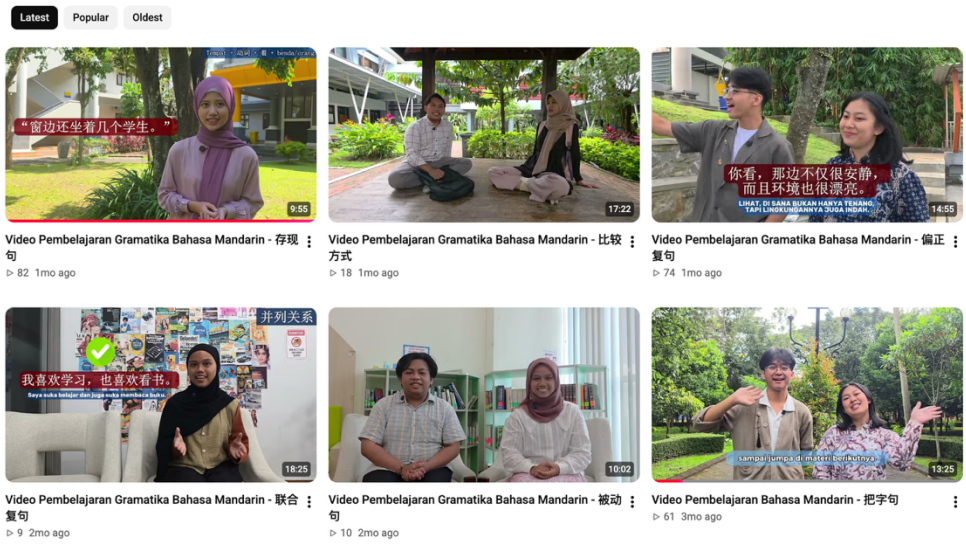

Figure 1. Instructional videos uploaded to the researchers' YouTube channel

An important finding in the development stage concerns the duration of the micro videos, which in practice ranged from 9 to 18 minutes per video, longer than the 5–8 minutes initially planned in the research proposal. The longest videos were 联合复句 (liánhé fùjù) (18:25) and 比较方式 (bǐjiào fāngshì) (17:22), and the shortest was 存现句 (9:54). These differences correspond to the complexity of each sentence pattern: topics with more sub-patterns or rule exceptions (such as 联合复句 (liánhé fùjù) and 比较方式 (bǐjiào fāngshì)) required longer explanations and more varied examples to be fully understood by students. In SIPEJAR, each video is watched as a whole, and each sentence pattern forms a segment introduced by a section marker and a pattern label. For example, the 存现句 (cúnxiàn jù) video includes a segment, ending at about 05:26 of the 9:54 video, on the pattern Location + verb + 着 (zhe) + person/object, illustrated with 窗边还坐着几个学生 (chuāng biān hái zuò zhe jǐ gè xuéshēng) ('A few more students are sitting by the window'), whose objective is to show how 着 (zhe) marks a continuing state of a person or object at a location. Practice follows the whole video, through five comprehension-check questions and then the interactive exercises in SIPEJAR.

The media expert first reviewed the first two videos (把字句 (bǎ zì jù) and 存现句 (cúnxiàn jù)), yielding three main recommendations. First, the Mandarin subtitles should be larger than the Indonesian subtitles, since the learning target is the Mandarin language itself and its level of difficulty is relatively high for students. Second, each transition between explanations of sentence patterns should include a section marker to give students a pause to process the information just presented before moving on to the next pattern. Third, the explanation of each sentence pattern should be labeled with a subtopic marker, such as "Pattern 1", using animated text that appears at intervals, so that students receive clear structural cues about their position in the flow of the video. All recommendations were addressed before the videos were embedded in SIPEJAR, and the resulting format guidelines, such as subtitle size, section markers, and the animation style of sentence-pattern labels, became the standard presentation reference for the remaining four videos, which were subsequently checked by the same media expert; this check raised no additional recommendations.

With all six micro videos and infographics produced and embedded in SIPEJAR, planned for use in 12 of the 16 meetings of the Grammar II course, the development stage of the EduKit can be considered complete, allowing the study to proceed to the implementation stage with students. In each SIPEJAR unit, students go through a sequence of learning steps: watching the micro video, answering five comprehension-check questions, studying the infographic, completing interactive exercises with feedback, taking a game-based formative quiz, and writing a reflection (Figure 2). At the reporting date, all of these components were available in each of the six units. Planned coverage across 12 meetings differs from observed use: only the 存现句 (cúnxiànjù) unit was used by students during the two-week pilot reported below.

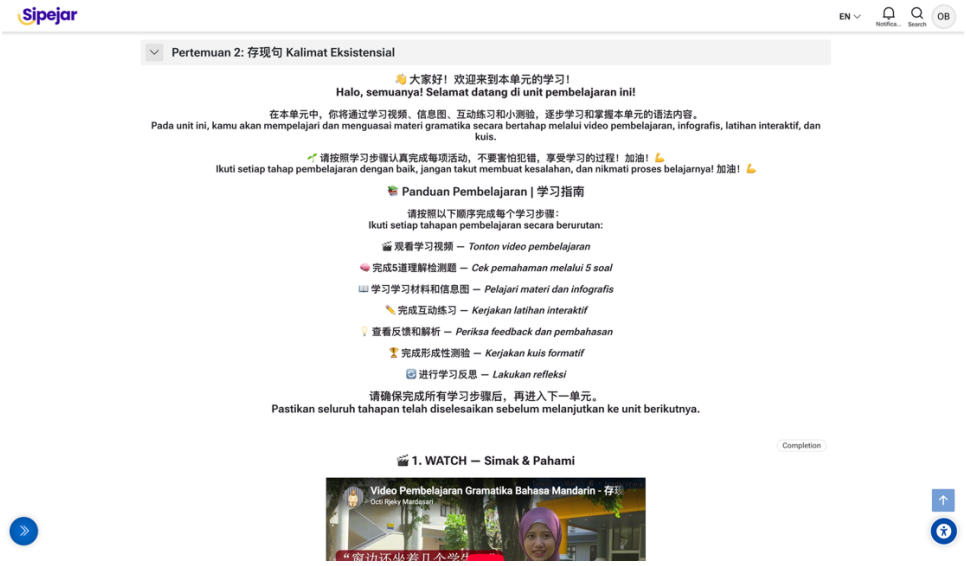

Figure 2. Learning interface of an EduKit unit in SIPEJAR

Implementation Stage

The implementation stage was conducted as an initial trial of the EduKit with students, as part of formative evaluation in the ADDIE cycle (Branch, 2009; Tessmer, 2013). The trial focused on one microlearning unit, 存现句 (cúnxiànjù) (existential sentences), as an initial step before extending it to the remaining five units. It aimed to describe the implementability of the EduKit after students accessed it independently through SIPEJAR and to gather student responses to the developed product.

The 存现句 (cúnxiànjù) unit was used for two weeks as an independent learning resource through SIPEJAR, without lecturer assistance. The Grammar II course was taught by another lecturer, while the researchers acted as the EduKit developers and administered the *pretest*, *posttest*, and questionnaire. The *pretest* was given before the 存现句 (cúnxiànjù) material was introduced, and the *posttest* was given two weeks later. The four-week research period was organized as follows: Week 1: briefing and pretest; Weeks 2–3: independent use of the 存现句 (cúnxiànjù) unit in SIPEJAR; Week 4: posttest and response questionnaire. During the two weeks of use, the regular instructor did not teach 存现句 (cúnxiànjù) in class but covered other material. Students' independent study outside the EduKit, however, was not monitored.

The *pretest* results showed that, of the 32 respondents, the mean initial score was 15.34 out of 20 (76.72%), with a minimum of 45% and several students already achieving a perfect score of 100%. Descriptively, students' initial performance was fairly good for the basic pattern Place + 有(yǒu) + person/object (above 90% correct on several items) but remained weak in the contextual use of 着(zhe), with the lowest percentage of correct answers, only 40.63%, on an item requiring direct application within a sentence structure. After the period of EduKit use, the *posttest* results for the same respondents showed a mean score of 18.69 out of 20 (93.44%), with a minimum of 70% and a median of 95% (pretest median: 75%); the standard deviation decreased from 2.66 to 1.77 points. Seventeen of the 32 students achieved a perfect score. This mean difference of 16.72 percentage points (3.34 raw-score points) is consistent with the item-level analysis, including the item on the use of 着(zhe), which showed the largest change, from 40.63% to 90.63%, and the item requiring students to distinguish the functions of 有(yǒu), 着(zhe), and 了(le), which changed from 62.50% to 93.75%. Individually, of the 32 students who took both the *pretest* and the *posttest*, 27 improved their scores, 3 obtained the same score, and 2 obtained lower scores. These data are presented descriptively as part of the evaluation of product implementation in a single trial group, not as a test of an effectiveness hypothesis.

In addition to the *pretest-posttest* data, the response questionnaire was completed by 31 students after using the 存现句(cúnxiànjù) unit. The overall mean of 4.45 is the mean of all 18 closed items across the 31 respondents (558 responses); all respondents answered every item. Responses were positive across all dimensions measured. The infographic dimension obtained the highest mean (4.56), followed by interactive exercises and quizzes (4.46), ease of access via SIPEJAR (4.44), material and micro video (4.43), and overall usefulness and satisfaction (4.38). Within the material and micro video dimension, the item on subtitle clarity received the highest score (4.58). In the open-ended responses, 13 of the 31 students named the micro video as the most useful component (10 named it alone), mainly because it aided understanding through spoken explanation and everyday examples. Eleven named the interactive exercises or the quiz because they offered direct practice, 5 named the infographic, 6 considered all components equally useful, and 1 did not name a component. Eight students reported technical problems: unstable connections or lag (4), difficulty locating the course or materials in SIPEJAR (2), and difficulty with the game-based quiz, whose game mechanics sometimes ended a round before they could answer (2). No student reported being unable to access the unit.

Discussion

The needs analysis indicates that students feel less confident applying Mandarin grammar than understanding it. Because both measures are self-ratings, the 0.23-point difference does not establish a measured gap between grammatical knowledge and productive performance; it shows where students locate their own difficulty. Students attributed this difficulty mainly to limited opportunities for structured and repeated practice, and they rated the exercises they had received as helpful. At the same time, direct explanation by the lecturer remained their most preferred method. This suggests that students do not want digital media to replace the lecturer but rather to complement and strengthen face-to-face learning, in line with the principle of blended learning, which positions technology as an integral part of the learning experience without replacing pedagogical interaction between lecturers and students (Korson, 2023).

Cognitive Load Theory (Sweller et al., 2019) provides the rationale for how the EduKit responds to these needs: because working memory capacity is limited, presenting dense grammar content without segmentation or practice support risks exceeding learners' processing capacity. Shadiev and Huang (2020) showed that learning in a second language can increase students' cognitive load, particularly without adequate segmentation of material. Because cognitive load was not measured in this study, the survey results are not attributed to working-memory overload; rather, the theory guided the decision to present material in small, focused units, in accordance with the principle of bite-sized learning proposed by Alias and Razak (2025) and supported by recent evidence that mobile-based microlearning improves engagement, knowledge retention, and learning outcomes in higher education (Al-Zahrani, 2024).

The same rationale explains two decisions made during development. The micro videos became longer than the 5–8 minutes initially planned, but this adjustment remains consistent with the spirit of microlearning as bite-sized learning (Alias & Razak, 2025), provided that the material is segmented by sentence pattern and accompanied by pauses for information processing. The media expert's recommendation to insert section markers between sentence patterns follows the same logic: it is consistent with the segmenting principle of multimedia learning (Mayer, 2021) and Cognitive Load Theory (Sweller et al., 2019), which emphasize the importance of processing pauses when presenting highly complex material.

The material expert's feedback shows that expert review checks not only the accuracy of grammatical rules but also the appropriateness of language use in real communicative practice, an aspect that, according to Chang et al. (2017), often poses a particular challenge for foreign learners acquiring Mandarin structures that have no direct equivalent in their first language. The involvement of an external material expert from secondary education also offered a relevant additional perspective, given that teaching Mandarin grammar to beginner and intermediate learners at senior high school often involves challenges like those faced by first-year university students, namely the tendency to produce grammatical but unnatural sentences because of direct translation from first-language patterns. Overall, the EduKit underwent a staged quality assurance process, from needs analysis based on survey data, through review of linguistic substance by the material expert, to review of media presentation by the media expert, before the content was distributed to students. This supports the argument that the development of foreign-language grammar learning media cannot rely solely on the designer's assumptions but should be supported by data on student needs and by expert review of both substance and presentation before its implementation is trialed with students.

The implementation findings are consistent with the needs analysis in showing that students view structured supplementary media as a complement to, rather than a replacement for, face-to-face learning (Korson, 2023). Within the pilot, the lowest pretest accuracy was on an item requiring the contextual use of 着 (zhe) (40.63%), which also showed the largest gain (to 90.63%). This identifies a difficult item in the pilot test and should not be equated with the difficult topics reported in the needs survey. The segmented presentation of 着 (zhe) in the micro video, with pauses between sections as recommended by the media expert, is consistent with the segmenting principle of Cognitive Load Theory (Sweller et al., 2019) and is a plausible explanation for this gain. However, practice with the exercises, familiarity from repeated testing, or independent study outside the EduKit could also have contributed; 存现句 (cúnxiànjù) was not taught in class during the pilot.

Several features of the pilot data limit how far these results can be interpreted. Because more than half of the students (17 of 32) reached the maximum posttest score, the test had limited room to distinguish higher levels of performance, and the posttest mean likely understates differences among stronger students. The high rating for subtitle clarity (4.58) indicates students' satisfaction with the final subtitles, but because no version without the resized subtitles was tested, it does not isolate the effect of the resizing. The reported difficulty with the game-based quiz, whose mechanics sometimes ended a round before students could answer, indicates that the quiz format should be reviewed before the EduKit is extended to the other units. With these qualifications, the positive initial trial results for one microlearning unit, together with the production of all six units in SIPEJAR, show that the four ADDIE stages reported here were carried out from needs analysis to an initial implementation trial, while summative evaluation remains to be conducted.

Conclusion

This development study produced six interactive microlearning units for Mandarin grammar in SIPEJAR, one of which was piloted with students. The analysis stage showed a difference between students' mean self-ratings for understanding (3.84) and applying (3.61) grammar, and 92% of respondents reported a need for supplementary media, which informed the design of the six units. In the development stage, all infographics and micro videos for the six topics were produced and embedded in SIPEJAR. Review by one material expert and one media expert yielded revisions concerning the accuracy and naturalness of grammatical examples and the visual and audio presentation, and the revised video format was applied to all six videos. In the pilot of the 存现句 (cúnxiànjù) unit, the mean score rose from 76.72% on the pretest to 93.44% on the posttest (n = 32), and the mean response rating was 4.45 on a 5-point scale (n = 31). These results support further testing of the EduKit but do not establish that it closed a gap in productive grammar use.

This study has several limitations. The implementation trial covered only one of the six units and involved a single group without a comparison group, so the reported score changes should be interpreted descriptively and cannot be generalized as evidence of effectiveness. More than half of the pilot students (17 of 32) reached the maximum posttest score, which limited the test's ability to distinguish higher levels of performance. In addition, the test consisted of multiple-choice items and did not directly assess spoken or written production, although the related sub-learning outcome requires students to write 存现句 (cúnxiànjù) sentences. Students' independent study outside the EduKit during the pilot was also not monitored. Future studies are recommended to extend the implementation to all six units, to test the effectiveness of the EduKit using a quasi-experimental design with a control group, and to include novel speaking or writing prompts and a delayed posttest to examine transfer and retention.

Acknowledgment

The authors would like to thank Universitas Negeri Malang for the funding and research facilities provided, as well as for the valuable feedback that helped improve this article.

References

Alias, N. F., & Razak, R. A. (2025). Revolutionizing learning in the digital age: A systematic literature review of microlearning strategies. *Interactive Learning Environments*, *33*(1), 1–21. https://doi.org/10.1080/10494820.2024.2331638

Al-Zahrani, A. (2024). Enhancing postgraduate students' learning outcomes through Flipped Mobile-Based Microlearning. *Research in Learning Technology*, *32*. https://doi.org/10.25304/rlt.v32.3110

Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer US. https://doi.org/10.1007/978-0-387-09506-6

Chang, J.-W., Lee, M.-C., Su, C.-Y., & Wang, T.-I. (2017). Effects of using self-explanation on a web-based Chinese sentence-learning system. *Computer Assisted Language Learning*, *30*(1–2), 44–63. https://doi.org/10.1080/09588221.2016.1241806

Chen, S. (2017). The construct of cognitive load in interpreting and its measurement. *Perspectives*, *25*(4), 640–657. https://doi.org/10.1080/0907676X.2016.1278026

Conde-Caballero, D., Castillo-Sarmiento, C. A., Ballesteros-Yánez, I., Rivero-Jiménez, B., & Mariano-Juárez, L. (2024). Microlearning through TikTok in Higher Education. An evaluation of uses and potentials. *Education and Information Technologies*, *29*(2), 2365–2385. https://doi.org/10.1007/s10639-023-11904-4

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*(2), 141–172. https://doi.org/10.1017/S0272263105050096

Irawati, R. P., & Anggraeni, A. (2018). Analisis kesulitan mahasiswa semester VI angkatan 2013 Prodi Pendidikan Bahasa Mandarin Universitas Negeri Semarang dalam memahami teks 阅读 pada 汉语水平考试 (HSK) level IV. *Longda Xiaokan: Journal of Mandarin Learning and Teaching*, *1*(2). https://doi.org/10.15294/longdaxiaokan.v1i2.12689

Korson, C. (2023). A place-based approach to blended learning. *Journal of Geography in Higher Education*, *47*(4), 569–588. https://doi.org/10.1080/03098265.2022.2122032

Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.

McKenzie, S., Hains-Wesson, R., Bangay, S., & Bowtell, G. (2022). A team-teaching approach for blended learning: An experiment. *Studies in Higher Education*, *47*(4), 860–874. https://doi.org/10.1080/03075079.2020.1817887

Monib, W. K., Qazi, A., & Apong, R. A. (2025). Mapping microlearning development and trends across diverse contexts: A bibliometric analysis (2007–2023). *Interactive Learning Environments*, *33*(3), 1865–1910. https://doi.org/10.1080/10494820.2024.2402556

Müller, C., & Mildenberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, *34*, 100394. https://doi.org/10.1016/j.edurev.2021.100394

Muyassaroh, L. U., Ventivani, A., Mardasari, O. R., Marah, Z. I. N. E., Setyawan, R., & Ermawati, V. (2024). Audiovisual media innovation: Hanyu Shuzi learning solution. *Bahasa Dan Seni: Jurnal Bahasa, Sastra, Seni, Dan Pengajarannya*, *52*(1), 30–42. https://doi.org/10.17977/um015v52i12024p30

Paramitha, A., Afifah, L., & Muyassaroh, L. U. (2022). Pengembangan Media Yuxihan Berbasis Augmented Reality untuk Pembelajaran Kosakata Bahasa Mandarin Siswa Kelas XI SMAN 5 Malang. *Journal of Language Literature and Arts*, *2*(1), 67–76. https://doi.org/10.17977/um064v2i12022p67-76

Rayens, W., & Ellis, A. (2018). Creating a Student-Centered Learning Environment Online. *Journal of Statistics Education*, *26*(2), 92–102. https://doi.org/10.1080/10691898.2018.1475205

Richey, R. C., & Klein, J. D. (2007). *Design and Development Research: Methods, Strategies, and Issues*. Routledge.

Shadiev, R., & Huang, Y. M. (2020). Investigating student attention, meditation, cognitive load, and satisfaction during lectures in a foreign language supported by speech-enabled language translation. *Computer Assisted Language Learning*, *33*(3), 301–326. https://doi.org/10.1080/09588221.2018.1559863

Sholikah, E. P., & Hidayat, E. (2022). Pengembangan Media Kamus Mini Bahasa Mandarin Bergambar sebagai Penunjang Pembelajaran Kosakata Bahasa Mandarin Siswa Kelas XI SMA Negeri 6 Malang. *Journal of Language Literature and Arts*, *2*(9), 1328–1338. https://doi.org/10.17977/um064v2i92022p1328-1338

Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif dan R&D*. Alfabeta.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tessmer, M. (2013). *Planning and Conducting Formative Evaluations*. Routledge. https://doi.org/10.4324/9780203061978