An Adaptive Video Based Learning for English as Second Language Learning

V. Chithra Devi

Associate professor, Department of English, PSNA College of Engineering and Technology, Dindigul, India.

Mail id: chithrapsnavijay@gmail.com[1]

HOW TO CITE:

V. Chithra Devi (2026). An Adaptive Video Based Learning for English as Second Language Learning. International Journal of Special Education, 41(7s), 233-251.

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors. Open access publication underthe terms and conditions of the Creative CommonsAttribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

ABSTRACT:

The growing use of digital media in English as a Second Language (ESL) teaching and learning requires teaching models that no longer focus on passive video presentation to learners but rather on responsive and interactive learning platforms. This paper introduces a Video-Based Interactive and Adaptive Learning Framework of ESL (VIA-ELF), which personalizes the process of learning video based on the profiling of the learners, interactive video based assessment, and difficulty conscious content recommendation in the context of widely available video streaming tools. The offered methodology takes the form of a mixed-methods approach that necessitates the needs analysis of the learners, the diagnostic proficiency testing, the interactive video design, the in-video and post-video formative assessment, the feedback on the perceived difficulty provided by the learners, and the update of the profile through the iterative approach. Adaptive recommendation mechanism bases its recommendations on objective performance measures, as well as subjective notions of difficulty, to bring the difficulty of the video in line with the dynamic competencies of the learners in language. The framework was tested with ESL students by pre- and post-testing the listening, speaking, reading and writing comprehension, evaluation and analysis of engagement and progression through the adaptive learning cycles. The effectiveness of the suggested adaptive video-based method is proven by the results of the experiment where all language abilities are enhanced, the number of learners involved is high, and the process of reaching the right level of difficulty is effective. These results suggest that lightweight, scalable video-based adaptive systems can be used to improve ESL learning, especially in learning settings with limited resources.

Keywords: English as a Second language, Video streaming applications, Digital language learning, personalized video recommendation, adaptive learning.

INTRODUCTION

English is the language that is taught to non-native speakers, which is especially important when it comes to expanding opportunities in the non-native English speaking regions, especially to the students who may want to continue their education or pursue a career in a foreign country [1]. Even though traditional classroom settings are effective to teach students, they do not allow the teacher to provide guided speaking practice due to the high number of students per teacher [2]. As a result, learners have to be frequently asked to engage in the speaking practice independently, which leads to the emergence of both written and oral English proficiency [3]. Nevertheless, lack of practice in

speaking might lead to good performance in writing and restrict oral speech [4].

It means that in the contemporary educational environment students can improve and expand the learning provided by teachers through independent development of their abilities, which on the one hand leads to better educational standards [5]. Due to the shortage of human resources to conduct individual instruction in large learning classes, the institutions usually offer English pre-recorded instructional videos as an auxiliary means of learning [6]. Consequently, there is a growing need of innovative and scalable instructional models that put the learner in the center with the application of technology to facilitate language learning. Video-based learning has been shown to be a useful tool in ESL instruction as it is the combination of visual, auditory, and contextual information that reinforces and aids in long-term retention of the information. Such content when being used over video streaming platform results in convenience, wide access and scalability in various learning environments. Nevertheless, most of the old models of video-based learning are passive and do not allow active engagement, personalised learning trajectories, and engagement with learning.

Meanwhile, AI-driven systems like speech recognition systems, chatbots, virtual tutors, and language learning apps have taken up as useful resources in language education [7]. These applications use advanced algorithms and natural language processing to provide personalized feedback, interactive exercises and real time practice environments [8]. It is worth noting that the speech recognition systems are the most popular means of advancing AI-aided language learning, as they allow the learner to train speaking, examine the speech produced, and receive feedback on the essential areas of speech production, including pronunciation, intonation, and fluency, promptly [9].

Despite the fact that video-based learning is now a popular activity due to its ease of use and the presence of multimedia capability, the majority of ESL instructional videos remain passive and one-directional delivery systems as opposed to interactive learning experiences. This form of passive interaction is not usually effective in watching to individual differences in learners, allowing continuous motivation, and providing sufficient support to balanced development of listening, speaking, reading and writing abilities. In addition, most of the current video-based ESL interventions do not provide the option of real-time interaction, adaptive sequencing of content, and systematic feedback. Consequently, educators are less able to observe engagement behaviors and learning outcomes of the learners and, therefore, it is not easy to identify gaps in learning, provide learners with necessary support in time, or tailor instruction according to students progress based on performance. The restrictions are particularly noticeable where the scarcity of educational set-ups, where instructional support and infrastructural difficulties further exert constraints on learning disparities. Thus, an urgent need is a systematic, interactive video-based teaching model that will transform video streaming platforms into active data environments to learn ESL.

An incisive analysis of the available literature [] indicates that there are various gaps in the current ESL teaching and technology-enhanced learning studies. First, despite the fact that numerous studies have focused on the effects of multimedia and video based learning in ESL settings, there is comparatively less research on interactive video based learning which will involve the direct involvement of learners in video streaming media. In the majority of previous research, videos are presented as additional or inactive learning tools but not the main tool of adaptive and interactive language acquisition. Recent adaptive learning solutions are usually based on external learning management systems or mobile apps, which overlooks the potential of video streaming information to be analyzed. The importance of teacher-assisted learning analytics in video-based teaching on ESL is not properly investigated. The body of research ignores the role of understanding that occurs between learners and videos that can inform decisions about instruction, provide timely feedback, or help teachers intervene by targeting the learner in the real classroom settings.

Moreover, there is a significant lack of studies that consider the implementation and the testing of interactive video-based ESL systems in rural and semi-urban schools, where the limitation caused by infrastructure and a variety of learner profiles are introduced as unique challenges.

The existing studies [10-15] utilize mostly the static instructional designs, providing the same video to all learners regardless of the proficiency and learning difficulty. These blanket strategies are unable to meet the diversity of the learner and may lead to cognitive overload or inadequate instructional difficulty. Whereas adaptive learning strategies have been investigated in ESL instructions, most of them are computationally intensive and require sophisticated AI models or mobile application environments, which cannot be deployed in rural or resource-restricted environments. Lightweight and scalable adaptive frameworks that can work with common video streaming platforms are yet to be researched. The current adaptation interventions are mainly based on objective performance indicators, including the test scores, and refuse to consider the perceived difficulty and affective feedbacks by the learners. This curtails responsiveness to cognitive load, confidence and subjective learning experience of learners which is important in effective language acquisition. Furthermore, it is characterized by the absence of empirically tested and integrated models that involve the integration of learner profiling, interactive video design, adaptive content sequencing, and continuous assessment especially with college level ESL learners in rural and semi-urban settings.

The research suggests a systematic teaching and learning model that places interactive video broadcast on streaming media as the fundamental instrument of teaching and learning ESL. Compared to the conventional blended or content-intensive methods, the framework puts more emphasis on interaction with learners, use of contextual language, and multimodal support in videos. It proposes a comprehensive system of profiling learners involving diagnostic tests, in-video interaction data, worksheet results, and learner perceived difficulty. With this integration, there is the ability to conduct data-driven instructional planning and individualized learning progression. The suggested methodology will include the use of an automatic video suggestion strategy to operate in alignment of video difficulty levels and perceived learning difficulties of the learners with their assessed proficiency. This helps in having proper cognitive load and long-term involvement among different skills of the learners. The research integrates both formative and reflective assessment into the video-based learning environment, which allows tracking the level of understanding, interaction, and mastering of learners in real time without interfering with teaching.

This paper intends to overcome these constraints by proposing a Video-Centric Interactive dynamic Learning Framework (VC-IDLF) that uses interactive video streaming technologies to facilitate adaptive, interactive, and pedagogically sound ESL teaching so that it provides valuable input to ESL pedagogy and educational technology research. The proposed framework integrates the activities of profiling the learner, building interactive video, dynamic learning paths, group video-based exercises and teacher-prompted analytics to create a comprehensive ESL learning ecosystem. Using the video-based interactions, the framework aims at increasing the accessibility and ensuring high pedagogical rigor and instructional effectiveness. This method is particularly applicable to rural and semi-urban educational settings, in which the limitations associated with both infrastructure and the scarce teaching materials tend to hinder the achievement of ESL education. The research will eliminate inequality in access to language learning and develop inclusive, scalable, and engaging ESL education through the application of video streaming services that are widely accessible and data-driven teaching methodologies. In general, the framework promotes the development of technology-enhanced language pedagogy and provides a viable and flexible model of implementing interactive video-based learning in a variety of ESL classroom contexts.

The main aims of the research are as follows:

- The purpose of the study is to plan and introduce a video-based interactive ESL teaching model that would increase student interaction and language acquisition in college-level learners.
- To create a learner profiling method with the help of quantitative tests and qualitative feedback to reveal the needs and learning obstacles related to specific skills.
- To implement an automatic video recommendation system that would adjust the instructional material depending on student performance and subjective challenge.
- To assess the efficacy of embedded video based assessment to track the learning progress and to guide adaptive instruction.
- To examine how the proposed framework affects the listening, speaking, reading, and writing abilities among learners.
- To determine the viability and scalability of the given approach in rural and resource limited educational settings.

1. LITERATURE SURVEY

As given in [10], the successful integration of technology in classroom instruction largely depends on teachers' roles, beliefs, and pedagogical practices, which significantly influence how technological tools are implemented. In technology-enhanced language learning environments, teachers must anticipate potential challenges that may arise during course delivery, particularly when incorporating video-based materials. In [11], the authors emphasized that English language teachers are responsible for selecting video content that aligns with instructional objectives and learner needs. This requires careful planning and informed decision-making to ensure pedagogical relevance. However, the abundance of readily available online video resources presents a significant challenge for educators in the advanced technology era. Berk [12] notes that a wide variety of video formats can be utilized in classroom settings, further increasing the complexity of content selection. Consequently, teachers must invest additional effort in filtering and curating appropriate video materials that effectively support teaching and learning objectives while addressing learners' diverse needs.

Media in teaching and learning has become a mandatory need to the teaching profession especially in the distance and technology mediated learning. Instructional media in the English language teaching context influences the motivation aspect of learners and improve their proficiency of language skills in EFL classrooms [13]. Learning materials in the form of media are also associated with learner autonomy because they are available at their convenience, both in time and place, which allows students to learn on their own, outside the classroom limits [14]. Furthermore, the authors in [15] have highlighted that instructional media help achieve the engaging and appealing learning environment which are beneficial to the interest and involvement of students in the learning process. All these studies point out the essentiality of the media in English language teaching, which can be of value to both teachers and learners because it can enhance teaching efficiency, understanding, and active participation of the learners.

2. METHODOLOGY

The research design used in this study is a systematic mixed-method research design that could design, implement, and assess a Video-Based Interactive and Adaptive ESL Learning Framework (VIA-ELF). The methodology is logically structured into a series of steps, such as analyzing the needs of learners and profiling, designing the instruction based on the video-centric pedagogy, developing the interactive video content, piloting with the help of the video streaming platforms, and multi-level assessment and evaluation. All the stages are tightly interconnected to assure the consistency between the learner traits, pedagogical goals, and adaptive video-based instructional techniques. The proposed methodology allows sustained personalization of video content by combining both quantitative performance metrics and qualitative feedback provided by learners, which makes the proposed methodology effective in promoting interaction and skill development, as

well as effectiveness in instruction in ESL learning settings.

2.1. Participant Selection

The participants of this work were ESL students who were selected among rural higher educational institutions of the selected state. The sample population consisted of undergraduate students aged between 17 and 19 years, which is the undergraduate entry level (first year degree programs), one of the crucial phases in acquisition of academic and professional English communication skills. This group was chosen because the need to acquire English skills in post-secondary education, job marketability and e-learning facilities is increasing.

The students that underwent the needs analysis phase were 120 undergraduate students. A stratified random sampling method was used in order to achieve demographic diversity and minimum sampling bias on the basis of the socio-economic background, gender, academic stream, and baseline English proficiency levels. The parental occupation, household income category, and availability of educational resources were used to identify the socio-economic classification. Based on the diagnostic pre-test result, English proficiency was classified into beginner, intermediate and low-advanced levels.

Majority of the participants indicated that they had little exposure to English outside of formal classroom instruction and little chances of real-time conversational practice. Even though most of them had the accessibility to smartphones and internet connections, the video streaming applications were used very informally and with an entertainment purpose, which is why it was necessary to organize and pedagogically oriented interactive video-based learning. The present context justified the use of video-based instruction as a replicable and available ESL intervention among rural college students. The same was done to ensure gender representation to facilitate inclusive analysis and the sample was selected among different academic disciplines to represent different learning needs and contexts of language use. The demographic distribution depicts the actual situation in rural college ESL classrooms hence increasing the ecological validity of the study. The details of selected participants with demographic information are provided in Table 1.

Table 1 Participant demographic characteristics

Parameter	**Description**
Age group	17-19 years
Educational level	Undergraduate degree
Institution type	Rural public institutions
Academic streams	Arts – 38%, Science – 34%, Commerce – 28%
Background language	Non-native English speakers
English proficiency levels	Beginner – 40% Intermediate – 37% Low advanced – 23%
Gender distribution	Male – 51% Female – 49%
Economic and social background	Low income – 44% Low middle income – 36% Middle income – 20%
Prior English exposure	Mainly classroom based
Access to digital devices	Smartphones – 82% Shared access – 12% Limited access – 6%

2.2. Qualitative and Quantitative Data Collection

The needs analysis was done through a quantitative and qualitative approach. Pre-tests were done using diagnostic tests to determine the level of pre-knowledge of listening, speaking, reading and writing ability. Simultaneously, the qualitative research techniques of semi-structured interviews and focus group discussions were used to obtain the in-depth information about the perceptions, learning experience, and difficulties of students in learning English. The observation in the classroom also helped in identifying patterns of engagement, levels of attention, and interaction behavior in the presence of the current ESL instruction.

Quantitative data collection

Standardized diagnostic pre-tests in form of quantitative data were collected before the intervention. These tests were meant to gauge the proficiency of the students at the basic level in the four main skills of language i.e. listening, speaking, reading and writing. The listening section evaluated the understanding of spoken academic and conversational English by use of audio-based tests, whereas the speaking section evaluated pronunciation, fluency and simple communicative proficiency by use of pretested oral prompts. Reading ability was measured by the comprehension passages according to the levels of undergraduate reading, and writing ability was measured by using the criteria of coherence, grammar, use of vocabulary, and the structure of a sentence in short written answers. The diagnostic tests were objective and gave numerical results on the level of proficiency of learners and helped classify them into the categories of beginner, intermediate, and low-advanced. These findings used as a quantitative reference to determine skill-specific shortages and were used to customize instructions of content, speed, and difficulty within the interactive video modules.

Qualitative data collection

Semi-structured interviews and focus group discussions (FGDs) were used to gather qualitative data to supplement the quantitative results. These approaches were used to investigate the subjective experiences of the learners, their perceptions towards the learning of the English language, motivational variables, and the difficulties they face both in classroom learning and self-directed learning. The semi structured interview format ensured that there was flexibility in probing individual responses and that the probing is also consistent with the participants. The interactive and reflective peer-to-peer sessions of the focus groups allowed identifying the common challenges, learning preferences, and attitudes toward video-based instruction. Moreover, non-participant observation in the classroom was performed in ordinary ESL sessions to investigate the patterns of engagement of the students, their attention span, level of participation, and their interactions. The indicators that were systematically recorded using an observation checklist included: responsiveness of learners, frequency of interaction, collaboration with peers and dependence on passive learning strategies by learners. These observations gave some contextual understanding of what is happening in the current instructional practices, and what the traditional ESL instructional practices are not able to achieve, especially in terms of active learning among the learners.

The combination of test results, interview outcomes, focus group themes and the findings of observation was able to address both the cognitive, behavioral as well as the affective aspects of the ESL learning process of each of the learners. These profiles were important inputs in designing interactive video content and influencing the decisions regarding the complexity of language, the level of instructional support, support of the subtitles, the real life contextualisation and interactivity incorporated in the content. In addition, the triangulated analysis helped to develop the adaptive instructional practices that were adjusted to the needs identified in the learners, so that the suggested video-based ESL framework was learner-focused, context-sensitive, and pedagogically based. This logical data-based strategy reinforced the methodological rigor of the study and increased the efficacy and applicability of the suggested instructional intervention.

2.3. Development of Interactive Video Content

The interactive video content was organized in a sequential manner with a set of the language learning objectives that included listening, speaking, reading, and writing. The different video parts were each associated with the particular skill outcomes and also contained incorporated interactive features like comprehension questions, pronunciation prompt, and vocabulary highlights. An authenticity was achieved through contextualizing language use through real-life communicative situations. Instructional videos have served as the main content delivery, communication, and interaction channel. The

design was based on constructivist learning theory, where it is important to note that the learners actively create linguistic knowledge by interacting with meaningful audiovisual information and not through passive viewing. In this regard, every video lesson was created to place language learning in real life situations that are applicable to the university level ESL students like discussions in classroom and real life scenarios.

Multimedia learning theory principles were applied in order to maximize the pedagogical effectiveness of the video-based instruction. The videos were combined with a synchronized audio narration, contextual images, textual indication on screen, and subtitles in order to facilitate dual-channel processing and better understanding. Cognitive scaffolds were deliberately employed to support the listening comprehension, vocabulary and reading skills especially among lower proficient learners through the use of subtitles to support the learning process. The visual signalling methods, including the highlighted keys and the situational images were included so that it would direct the attention of the learners to the most important linguistic features, but would not impose extra load on their cognitive processes.

Instead of the external gamification mechanisms, the interaction between the learner was encouraged by means of video interactions embedded within the context. There were also interactive prompts, pause and respond segments, reflective questions, and prediction tasks that were strategically placed in the videos to promote continued thinking. This interaction based methodology made sure that engagement occurred through valuable language application and contextual interpretation, which is the same as the video based nature of the recommended instructional framework.

2.4. Video integrated supplementary activities

To extend learning beyond video viewing while maintaining instructional coherence, supplementary activities were designed as direct extensions of the video lessons. These activities reinforced key language concepts introduced in the videos and promoted active application of listening, speaking, reading, and writing skills.

Video-linked interactive worksheets were developed to support post-viewing practice. These worksheets included vocabulary reinforcement tasks, sentence construction exercises, guided writing prompts, and comprehension questions derived explicitly from video scenarios. By revisiting video content through structured activities, learners were able to consolidate understanding and transfer language knowledge to new contexts.

2.5. Automatic Video Recommendation

To support the suggested learner profiling and data-driven instructional design, the methodology incorporates an automatic mechanism of video recommendation in order to customize the learning pathways in the video-based ESL system. This mechanism has the purpose of generating dynamically relevant video difficulty based on the changing language proficiency and perceived learning challenges of learners, to improve the relevance and effectiveness of instruction. The proficiency levels of the learners were constantly updated with the help of the data received during diagnostic pre-tests, in-video formative assessment, and post-video practice tasks covering the skills of listening, speaking, reading, and writing. According to the aggregated performance scores, the learners were divided into basic, intermediate, and advanced levels. Parallel to this, the linguistic complexity, speech rate, vocabulary density, sentence constructions, and cognitive challenges of the embedded tasks were used to pre-classify instructional videos into low, medium, and high difficulty levels. Such a dual classification brought about a systematic mapping between the video difficulty and the learning ability of a learner.

As an adaptive parameter used to complement the objective performance measures, learner-perceived difficulty was included. Learners were asked to rate the perceived level of difficulty on a short self-assessment scale (easy, moderate, difficult) after every video session. The performance metrics for video recommendation can be predicted by considering learner-perceived difficulty level. The

difficulty level has been rated on a scale as easy, moderate and difficult on the end of every video. This feedback measure reflected affective and cognitive loading factors that cannot be entirely reflected using assessment scores. The recommendation logic logically analyzed the performance results and the perceived difficulty to decide the video recommendations in the future. Students who were performing well with an easy perceived difficulty were gradually suggested more difficult videos with less scaffolding, faster narration, and more complicated language structures. Moderately performing learners or those that indicated moderate difficulty were given similar difficulty videos to ensure that they solidify their skills. Students with poor performance and high perceived difficulty were directed to less challenging videos with more support tools including subtitles, slower speech, and visual aids. This two criteria adaptation allowed challenge adequate progression and reduction of cognitive load.

2.6. Embedded Assessment Strategy for Adaptive Decision-Making

The evaluation was integrated in the whole process of video-based learning in order to facilitate the regular profiling and updating of recommendations. The assessments are performed with in-video questions like small quizzes and pause and respond prompts. The skill of applying the knowledge learnt outside the video setting was assessed through post-video testing, including video-linked worksheets and guided writing activities. Moreover, self-reflection and difficulty-rating scales were used to obtain self-reported information on the learning experience and confidence of learners. Regular cumulative evaluations allowed longitudinal evaluation of the progress of learners at different levels of difficulty in videos.

Evaluation was integrated in the entire course of video based learning to facilitate constant profiling and recommendation revisions. Short quizzes and pause and respond questions in the form of in-video formative assessments assessed real-time understanding and interest. The transfer of knowledge outside the video context between the learners was assessed using post-video tests like video-linked worksheets and guided writing assignments. Moreover, difficulty-rating measures and self-reflection measures was used to capture the subjective experience of learning in learners and their level of confidence. Regular cumulative evaluations made it possible to do longitudinal analysis on the progression of learners through the various videos of different levels of difficulty.

The video based learning process had an integrated evaluation procedure to help the continuous profiling and recommendation of the changes. Real-time comprehension and engagement were determined by using formative assessment via in-video tests, short quizzes, and pause and respond questions. The ability of the learners to transfer the knowledge to the video setting was measured using post-video tests such as video-linked worksheets and directed writing tasks. The subjective learning experience of the learners and their confidence were also explained by the self-reflection and difficulty-rating measures. The longitudinal study of the progression through the various levels of video difficulty of learners could be conducted by cumulative evaluation at specific intervals.

The results of all the assessments and feedback of the learners were integrated in the learners profiling module and this allowed the learners to be updated continuously about the profiles. The data analysis part took the form of these multi-source inputs to facilitate automatic video recommendations decision-making and measure learning progression patterns. The proposed methodology, which combines objective performance feedback with perceived difficulty feedback, would create a responsive and scalable and learner-centred video-based ESL framework, which can be specifically applied to college level students in resource limited educational environments.

3. IMPLEMENTATION

Immediately the instructional cycle starts, the learner profile is set up through diagnostic pre-test scores in the areas of listening, speaking, reading and writing skills. Such scores determine the first

level of proficiency (basic, intermediate, or advanced) of the learner and are initial inputs of adaptive recommendation module. Assessment of learners forms the bottom level of the adaptive framework. The diagnostic pre-tests and post video tests measure the proficiency of the learners in the four main ESL skills- listening, speaking, reading, and writing. Each of the skills has the same level of contribution to the overall proficiency score, which allows achieving a proportional reflection of language skills of learners which is depicted in Table 2.

Table 2 Core language skills assessed

Skill assessed	Assessment type	Score range (%)	Weight (%)	Description
Listening	Audio comprehension test	0-100	25	Understanding ability
Speaking	Pronunciation and oral response		25	Accuracy, fluency and clarity in spoken responses
Reading	Passage comprehension test		25	Understanding ability
Writing	Short paragraph writing		25	Grammar, vocabulary usage, coherent writing

For evaluating English proficiency in college students learning ESL, the fundamental keys like listening, speaking, reading and writing have been selected. Listening skill is evaluated by use of audio comprehension test, this skill evaluates the ability of the learners to understand spoken English in realistic situations. Effective interaction with interactive video content requires a high level of listening skills because interference with auditory processing is the basis of involvement in a further procedure based on in-video cues and mimicking dialogues. Speaking skill is tested through the examinations of pronunciation and oral response, speaking proficiency describes the fluency, clarity and accuracy of the learners in expressing the language. Such an evaluation will help establish that the system can track the growth of oral skills, which is essential to perform interactive role-plays and practice pronunciation, which will be incorporated into the videos. The reading skill is measured by completing tests based on passage comprehension to determine the capability of the learners to comprehend written text and draw inferences. Good reading proficiency helps the learners to follow the subtitles, comprehend vocabulary in videos, and complete comprehension worksheet tasks. Writing is evaluated using short paragraph writing which is used to measure grammar skills, vocabulary use, coherence skills and composition skills. Writing assessment is an addition to the reading and speaking sections, which allows determining whether the learners can convert the knowledge gained during the comprehension and vocabulary development into meaning.

All skills were given equal weights (25 each) in calculating the composite proficiency score, which indicates the purpose of the framework to create a balance between all four domains on a learner profile. This composite score is used as the starting point of video difficulty assignment and adaptation in the automatic recommendation mechanism. The framework ensures that the strengths and weaknesses of the learners in the skills are captured in a holistic way to ensure that the adaptive system can deliver skill-based video content that is individualized and can also direct the learners through the right learning channels depending on the challenges encountered.

3.1. Video difficulty assessment

The learners are typically divided into three categories as per the composite score. The categories include basic, intermediate and advanced as given in Table 3. The learner proficiency level is evaluated and updated dynamically when the learners interact with the content.

Table 3 Video assessment based on learner score

Learner proficiency level	Composite score	Assigned video difficulty level	Video characteristics
Basic	<60	Low	Slow, simple vocabulary, small sentences
Intermediate	60-79	Medium	Moderate, contextual vocabulary, compound sentences
Advanced	≥80	High	Natural/fast speech, complex sentence forms, domain-rich vocabulary

This table standardizes the preliminary mapping of learner competence and video challenge in the proposed video-based ESL system. The framework enables one to use video features to match the perceived proficiency level of a learner to initiate him/her with either a too challenging or too easy video. This systematic task provides a consistent baseline on the basis of which further adaptive suggestions based on post video performance and perceived difficulty feedbacks are made. The process of mapping learner proficiency to video assessment is depicted in Figure 1.

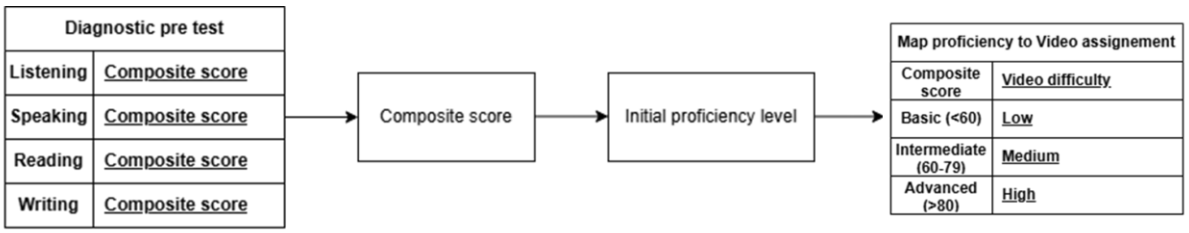

Figure 1 Initial video assignment based on learner proficiency

After every instructional cycle, post video assessment and perceived difficulty are evaluated to revision the profile of the learner. This is to make sure that profiling does not only capture objective improvement in performance, but also subjective learning experiences, which are vital in successful adaptation. Adaptive video recommendation module is based on the updated profile of the learner to identify the next video in the learning sequence that would be most suitable. Pre-classification of videos according to linguistic complexity, the speech rate, the vocabulary density, and cognitive load of tasks inside the video is done into low, medium, and high categories. The logic of recommendation incorporates two important inputs including assessment-based performance-based indicators and perceived difficulty feedback experienced by learners at video completion. Based on these inputs the video recommendation has been done as provided in Table 4.

Table 4 Video recommendation based on learning perception and performance

Performance	Difficulty of perception	Difficulty level of Video
High	Low	High difficulty videos for progression
Moderate	Moderate	Reinforce learning
Low	High	Enhanced instruction support

The dynamism of the framework to the changing needs of learners is facilitated by the continuous feedback given by the execution of the steps of assessment, profiling and recommendation. Video interactions produce new performance and perception data that improves the learner profiles and affects further suggestions. The proposed methodology will enable one to personalize, scale, and learn the ESL process, within a video-based learning environment, through a tight integration of assessment results with learner profiling and selection of adaptable videos, which is why it can be applied to both resource-limited and regular education environments.

3.2. Video Delivery via Streaming Platform

After the profiling of learners and the introduction of initial video assignment, learning content is presented using a video streaming-based learning platform that facilitates interactive and adaptive attributes. The learners are assigned to an interactive video that is based on their desired level of difficulty to make the instructional process consistent with their current proficiency profile. The videos presented have interactive features of learning, such as subtitles, pause-and-answer cues, embedded understanding questions, and guided pronunciation. These characteristics are dynamically set depending on the set video difficulty. The videos typified as low-difficulty level generally poses full-subtitle support and recurrent communication to encourage comprehension, whereas the videos typified as higher-difficulty videos have limited subtitle support to encourage independent learning. The streaming platform allows the platform to streamline across all devices and logs the data of interaction by learners in the form of the length of viewing, frequency of pauses, quiz responses, and the use of subtitles. These interaction records can be used as live signs of involvement and understanding, to be used in later evaluation and updates in learner profiles. With a proposed framework, which uses a streaming based delivery, scalability, accessibility, and pedagogical flexibility will be achieved giving it applicability in both resource constrained and mainstream educational settings. This action defines the essential interaction of instruction wherein learners will access the video-based ESL content and produce performance information to enable adjustive learning movement.

3.3. Video linked post assessment

The interactive video learning session is conducted through the use of embedded formative assessment of the learners. They are short comprehension checks in such forms as multiple-choice or true/false questions involving the newly covered material, Prediction questions where the learner has to imagine what is coming next in the video or how something will be used in the future, Interactive prompts when the learners have to repeat a phrase by speaking it out loud, complete a sentence, or the correct vocabulary words in the video context. The online learning platform captures comprehensive performance data of every learner as indicated in Table 5.

Table 5 Video embedded assessment metrics

Metric	Description
Response accuracy	Accuracy of answers in quizzes and prediction tasks
Completion time	Time taken to answer each prompt while watching video
Interaction frequency	Number of interactions with clicks, pause-and-response prompts
Engagement patterns	Subtitles, replays handling and navigation activities

Such metrics are considered real time in order to assess understanding, involvement and mastery. Its results are directly inputted into the learner profiling module, which is used to update the learner with their current level of proficiency and input into the adaptive video recommendation engine.

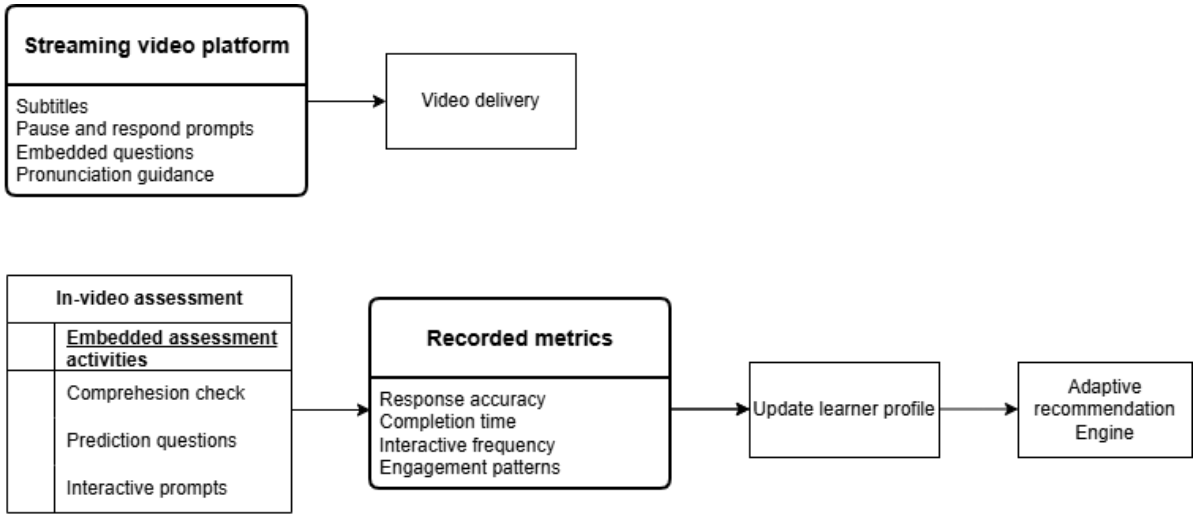

Figure 2 Video streaming and In-video assessment

3.4. Post-video assessment mapping

After the interactive video lessons, the learners get to engage in video-linked post-viewing tests, which aim at cementing the knowledge and solidifying the language. These tests are Video-Linked worksheets, Guided writing tasks, Sentence development tasks. Video-Linked worksheets contain exercises that are immediately connected with video material, including vocabulary training, sentence structure, comprehension questions and grammar tasks with contexts. Guided writing tasks are short paragraphs or story completions or descriptive writing requests based on the scenarios shown in the videos. Sentence development tasks are focused tasks in which learners practice making syntactically and semantically accurate sentences with newly acquired vocabulary and expressions in the video. The post-video assessment mapping has been provided in Table 6. Some of the performance measures that are obtained at this phase are accuracy, use of vocabulary, grammatical correctness, written response coherence, time of task completion, and the levels of engagement. Such indicators give an idea about the capacity of learners to translate video based input to active language production. The results of the post video tests are also dual purpose, they will reinforce the learning process via the structured practice and will also provide the necessary performance information to make the necessary changes to the learner profile. Together with the in-video formative assessment outcomes and perceived difficulty feedback, the information will underpin adaptive video recommendations and provide progressive and learner-focused instruction in ESL. The metrics measured from these evaluations are described in Table 7.

Table 6 Post Video assessment mapping

Assessment type	**Skill assessed**	**Difficulty level**
Video-linked worksheets	Listening, reading and vocabulary	Low/Medium/High
Guided writing tasks	Writing, grammar and vocabulary	Medium/High
Sentence development tasks	Writing, grammar and syntax	Low/Medium/High
Contextual vocabulary	Vocabulary and reading	Low/Medium

workouts		
Comprehension and recall prompts	Reading, Listening and understanding	Low/Medium/High

Table 7 Post video assessment metric

Metric	Description
Accuracy	Correctness of answers given in worksheets; answers from writing tasks
Vocabulary usage	Appropriate and diverse use of words from video content
Grammar & Syntax	Sentence structure and grammar accuracy
Coherence & Fluency	Flow and readability in given responses
Completion time	Time for completing each task
Engagement	Interaction rate while providing worksheet responses

3.5. Learner reported difficulty feedback

As an addition to objective measures of performance, the intended methodology uses the learner-perceived difficulty feedback as a subjective measure of cognitive load and affective reaction in the process of video-based learning. After each interactive video lesson and post-video tests, the learners report short self-assessments as to whether they found the learning experience easy, moderate or difficult on a scale of three; easy, moderate, difficult. This feedback will provide the level of confidence, perceived understanding and mental effort used by learners and which could not be effectively solely measured in terms of quantitative scores. The combination of perceived difficulty and assessment results makes the framework more comprehensible in terms of learner development and difficulties. The gathered ratings of the difficulty are documented with assessment metrics and added into the learner profiling module. This allows the adaptive recommendation mechanism to discriminate between learners who achieve well but have high cognitive load and those who perform poorly following lack of adequate proficiency. Therefore, perceived difficulty feedback is essential in modifying the video recommendations, balancing the instructional complexity and controlling the extent of the scaffolding that is given in the next learning cycles. The following Figure 3 represents video-linked post assessment metrics and profile updating flows.

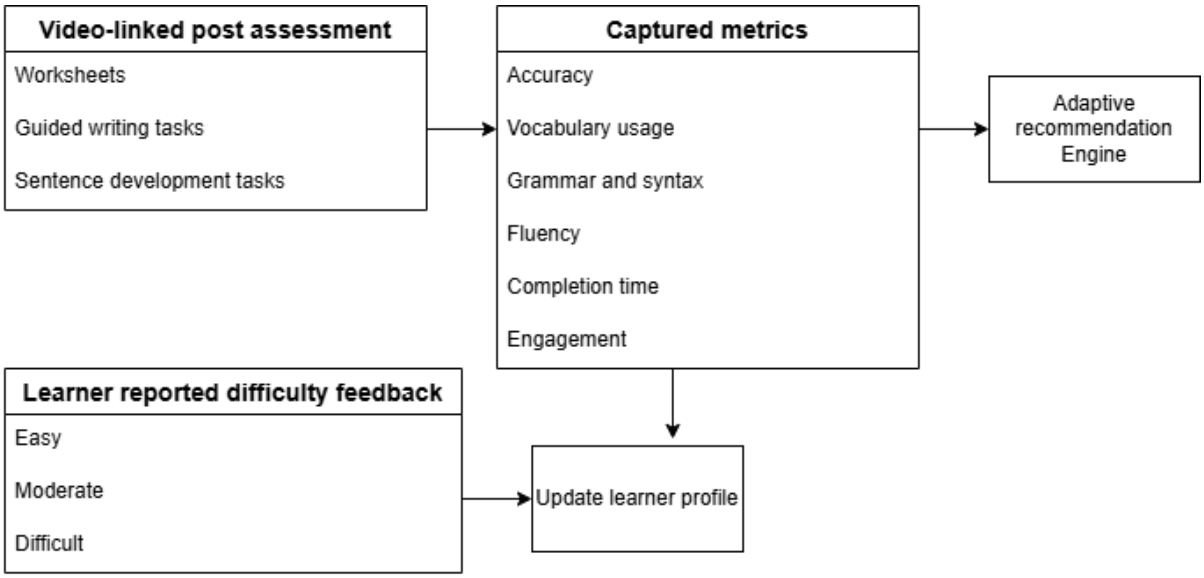

Figure 3 Video-linked post assessment and learner difficulty feedback

3.6. Profile Update and recommendation decision logic

After the in-video and post-video tests have been completed, the data of the learners are compiled to refresh the profile of a learner dynamically. It is a process that combines both quantitative performance outcomes such as the results of the listening, speaking, reading, and writing tests and the qualitative information on the perceived difficulty of the learner gathered after the completion of the video. The data accumulated would indicate the objective performance development and the subjective difficulties of the learning process, which allows a complete description of the developing competence of each learner. The most important indicators, including improvement in skills, accuracy in the response, time to complete the task, frequency of engagement, and perceived cognitive load are continually analyzed and included in the learner profile. The framework enables the informed adaptive decision-making in the next instructional cycles by having an updated and data-driven profile of the learners. The revision of the profile is an input of the logic of the decision of the recommendation that guarantees a correspondence of the video difficulty level and scaffolding strategies with the current competence and learning experience of the learners.

According to the dynamically revised profile of the learner, the proposed framework relies on an adaptive decision logic of recommendation to decide on the next instructional action to follow. The mechanism of decision is a combined assessment of the learner proficiency development based on quantitative assessment results and the assessment of the learner perceived difficulty as indicators of cognitive and affective learning situations. The logic of recommendations works on the basis of preset rules of decisions. Students with high performance and low perceived difficulty level in the evaluated language competencies that claim the low difficulty will be advised to watch higher-level videos. Such videos have higher linguistic levels of complexity, increased speech rate and less instructional scaffolds to facilitate higher language acquisition. Students having moderate performance and moderate learning difficulty are given with videos with difficulty level same as previous ones. This holds continuity in instruction and gives the learner time to distil the knowledge of repeated exposure and practice. Those learners who face low performance and high perceived difficulty are placed in lower-difficulty videos with more enhanced scaffolding. Such videos use simplified vocabularies, talk slower, use longer subtitles, and other aids to minimize thinking to help people grasp the message.

This adaptive logic is rule-based, which means that the progress of instruction is based on the level of challenges to avoid cognitive overload and the disengagement of learners but support the development of skills. The results of the recommendation are directly given into the subsequent cycle of learning and allow tailor-made and adaptable video-based teaching of ESL. In Figure 4, flow recommendation decision to next assignment is depicted.

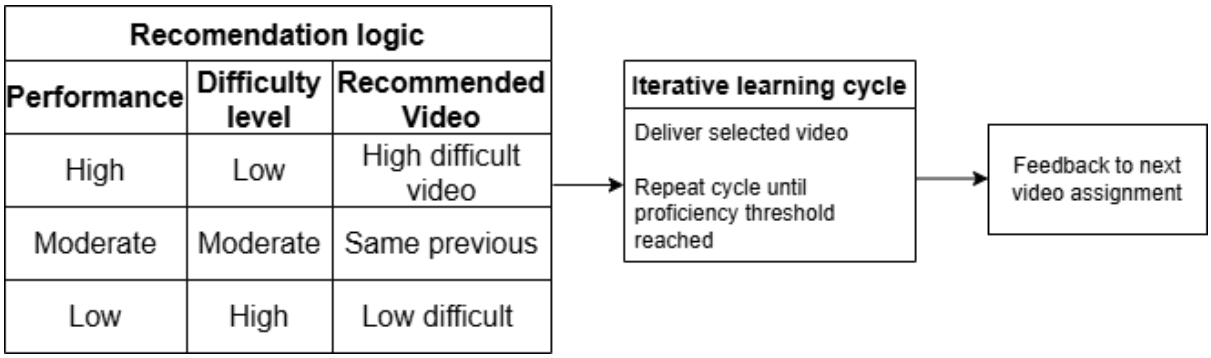

Figure 4 Recommendation decision to next assignment

The adaptive framework is a cyclical process of constant iteration whereby every learner advances on the basis of their respective levels of proficiency and levels of learning. After the recommendation logic, the video is shown through the streaming platform with the relevant interactive functions and the level of difficulty. In every iteration, the learners will undergo in-video formative assessments, post-video skill tasks and submit perceived difficulty feedback. These data are subsequently summed up into the learner profile and used to inform the next video recommendation. This process is repeated until accomplishment of the pre-determined learning outcomes or proficiency levels in listening, speaking, reading, and writing skills are achieved among the learners.

4. Results and analysis

The given adaptive video-based ESL framework was tested with the sample of 45 ESL learners of college level in the rural and semi-urban settings. The experiment involved the effects of interactive video lessons on post video assessment and adaptive recommendations on language learning in terms of listening, speaking, reading and writing.

4.1. Learner interaction

The learner engagement and perceived difficulty were also compared to determine the effectiveness of the adaptive personalization mechanism. In-video interaction measures, such as the rate of quiz responses, the rate of pause and respond prompt participation, and the rate of video completion were used to monitor learner engagement. At the level of aggregate scores of interaction, the learners were grouped into four engagement levels.

Table 8 Distribution of learner engagement levels

Engagement level	Percentage of learners
Highly engaged	62%
Moderately engaged	28%
Neutral	8%
Disengaged	2%

The findings show that the rate of learners participation is very high with 62 percent of the learners rated to be highly engaged with a consistent interaction with embedded quizzes and prompt-based activities. This is an indication that the interactive video elements are effective in maintaining the attention of learners and encouraging them to participate actively instead of passively consuming the content. A further 28 per cent of the learners showed moderate engagement levels which implied partial interaction which, nonetheless, had a positive impact on the learning outcomes. Very few learners were reported to be neutral (8%) or disengaged (2%), and it was found that the adaptive video-based instructional design was effective. Notably, the engagement measures

were analyzed together with the feedback of perceived difficulty in learners, which allowed the adaptive algorithm to differentiate between not engaging with the content because it was inappropriate and because of actual mental stress. The integrated understanding led to better decisions on video sequencing which meant the learners would be provided with instructional content based on their level of performance as well as their engagement patterns. Altogether, the results affirm that interactive video systems are important in improving the engagement of learners and facilitating adaptive decisions on recommendations in the proposed framework.

4.2. Language skill improvement

Diagnostic pre-tests, post-video tests, and cumulative skill development in the process of several adaptive video cycles was used to measure learner performance. Table 9 summarizes the average scores of the various skills of the language prior to and after the intervention.

Table 9 Language skill assessment

Skill	Pre-test average (%)	Post-test average (%)	Improvement (%)
Listening	**55**	**73**	**18**
Speaking	**51**	**72**	**21**
Reading	**59**	**77**	**18**
Writing	**56**	**74**	**18**

Observed results illustrate a statistically significant and constant enhancement in all the four language skills which validates the efficacy of the suggested adaptive video learning structure. The greatest improvements were on the listening and speaking skills with improvements of 18% and 21% respectively. This can be explained by real life context videos, controlled speech speed and interactive audio-visual prompts that essentially involve the learners in comprehension and oral response activities. The gains in reading and writing were also significant with improvement of 18% each. These gains underscore the effectiveness of video-linked worksheets and guided writing exercises which prompted the learners to apply the knowledge gained in video content to the structured reading comprehension and sentence making exercises. These reinforcement systems guaranteed the deeper processing other than passive video viewing. In addition, the equalizing enhancement of receptive (listening, reading) and productive (speaking, writing) skills points out that the adaptive recommendation mechanism successfully adjusted the video challenge and scaffolding to the level of the learner proficiency. This alignment minimized cognitive overload, but had enough challenge, which sustained skill development across the iterative learning cycles. Comprehensively, the results confirm that the proposed framework can provide quantifiable and comprehensive language skill improvement due to adaptive and interaction-based video-based learning.

4.3. Effectiveness in adaptive recommendation

The adaptive recommendation mechanism was tested in terms of its effectiveness in the progression of the learners through the video levels of difficulty depending on the objective performance measures and the subjective feedback on the difficulty. Table X provides the preliminary and final levels of difficulty of the videos in the various categories of learners.

Table 10 Adaptive video difficulty progression

Learner level	Initial video difficulty	Final video difficulty	Average difficulty
Basic	Low	Medium	Moderate
Intermediate	Medium	High	Moderate
Advanced	High	High	Low

Table 11 Correlation between difficulty and post-video performance

Metric	Correlation coefficient
Listening performance vs difficulty	-0.62
Speaking performance vs difficulty	-0.59
Reading performance vs difficulty	-0.55
Writing performance vs difficulty	-0.53

The negative correlation that is observed between all subjects of the four language skills shows that there is a linear inverse relationship between the perceived difficulty and the performance of the learner. An increased level of self-reported difficulty was broadly linked to low levels of post-video evaluation scores and the highest correlation levels between the two were during listening and speaking skills. This tendency is connected with the cognitive needs of audio processing and oral production, which are offered by ESL learning. These results confirm the incorporation of perceived difficulty to the learner as an essential input to the adaptive recommendation logic. Although objective performance scores describe quantifiable performance, subjective difficulty feedbacks offer complementary information on the cognitive load and affective conditions of the learners. The proposed system will help to avoid cognitive overload by allowing timely interventions based on scaffolds (simplified narration, improved subtitles, or repeated exposure) to advance the learning process. All in all, the design decision of using performance-based and perception-based adaptation as a way of making video recommendations more responsive and learner-centered is supported empirically by the correlation analysis.

A cumulative study of skill learning through repeated cycles showed that the learning process was on a gradual improvement curve with learners attaining predetermined level of proficiency in an average of 6-8 adaptive cycles. This iterative process ensured that there were no cognitive overloads among low-performing learners and gave the advanced learners enough challenge. The framework made it possible to have individualized learning journey within a diverse learner population. Adaptive scaffolding enhanced confidence and use of engagement especially among learners whose initial level of proficiency is at the basic level. The dual-criteria adaptive approach was confirmed to be effective in integrating perceived difficulty feedback with assessment scores to improve decision-making when making video recommendations. From the results it is observed that video-based interactive learning is very effective in enhancing ESL skills in the field of listening, speaking, reading and writing. Fine-tuning of content difficulty and keeping it interesting are essential as a result of learner engagement and subjective feedback. The results prove the originality and efficiency of the offered methodology, especially within the resource-restricted educational setting, and offer a solid empirical basis to the further studies in the adaptive, video-based teaching of ESL.

5. Conclusion

This paper has proposed an adaptive, video based ESL learning model, which combines with learner profiling, interactive video presentation, continuous assessment and flow based recommendation logic that facilitates individualized language learning. In contrast to fixed, universal teaching strategies, the set methodology allows on-the-fly adjustment of video challenge according to the changing skill levels and perceived learning complexity of the learners to provide personalized learning channels in a video streaming framework that is scalable. The outcomes of the experiments proved that all four language skills, including listening, speaking, reading, and writing, were improved significantly thanks to the interactive video functions and video-related after-test activities that facilitate the active use of learning and knowledge transfer. Introduction of learner-perceived difficulty feedback was quite useful, as correlation analysis indicated that there was a strong inverse correlation between performance and perceived

difficulty. This observation justifies the fact that subjective cognitive load has to be taken into account in conjunction with objective test results in the adaptive ESL systems. The adaptive recommendation mechanism was effective in bringing the learners to the proper level of challenge, reducing the cognitive load, but not excessive, instructional challenge. Students advanced gradually through levels of video difficulty and the measures of engagement showed that the students were highly participative, which proved the efficiency of interactions embedded in the video. Moreover, the cyclic training system allowed updating profiles and sequencing responsive content, which will guarantee the ability to develop skills in the long-term. In general, the suggested framework provides a lean, practical, and learner-centred approach to teaching video-based ESL and works best with college students in rural and semi-urban settings with limited access to sophisticated learning facilities. Using both qualitative and quantitative learning information and taking advantage of the widely available video streaming tools, the proposed work adds to a scalable and empirically-tested method of learning a (language) adaptation. Future labor can build on the framework to include such features as speech analytics, fine-grained affective sensing, and long-term retention analysis in order to increase the level of personalization and learning efficiency.

References

Azimova, Dilobarkhon, and Dilyorjon Solidjonov. "Learning English language as a second language with augmented reality." Qo ‘Qon Universiteti Xabarnomasi 1 (2023): 112-115.

Gerard, Libby, Marcia C. Linn, and Marlen Holtmann. "A comparison of responsive and general guidance to promote learning in an online science dialog." *Education Sciences* 14.12 (2024): 1383.

Al-Jarf, Reima. "Online vocabulary tasks for engaging and motivating EFL college students in distance learning during the pandemic and post-pandemic." *International Journal of English Language Studies (IJELS)* 4.1 (2022): 14-24.

Phan, Thi Minh Uyen, Thi Thuy Hang Nguyen, and Thi Xuan Dao Phan. "Factors affecting oral presentation skills of English majored sophomores at Tay Do university, Vietnam." *European Journal of English Language Teaching* 7.2 (2022).

Lakkala, Suvi, et al. "Teachers supporting students in collaborative ways—An analysis of collaborative work creating supportive learning environments for every student in a school: Cases from Austria, Finland, Lithuania, and Poland." *Sustainability* 13.5 (2021): 2804.

Yasmin, Tahira, Mazhar Hussain Bhadroo, and Naeem Ullah. "Investigating the Use of Pre-Recorded Lecture Videos in Flipped Classroom Technologies on Students' Learning at Secondary Level." *Journal of Asian Development Studies* 13.2 (2024): 1734-1744.

AbuSahyon, Alia Sa’ad Eldin, et al. "AI-driven technology and chatbots as tools for enhancing English language learning in the context of second language acquisition: A review study." *International Journal of Membrane Science and Technology* 10.1 (2023): 1209-1223.

Hussain, Syed Muddasir, Sumra Mussarat Jabeen Satti, and Zahra Khan. "AI-Powered Personalized Learning: Advancing Language Education in the Digital Era." *Journal of Social Signs Review* 2.4 (2024): 730-740.

Burns, Anne, and Barbara Seidlhofer. "Speaking and pronunciation." *An introduction to applied linguistics*. Routledge, 2019. 240-258.

Abukhattala, Ibrahim. "The use of technology in language classrooms in Libya." *International Journal of Social Science and Humanity* 6.4 (2016): 262-267.

Bell, Lynn, and Glen Bull. "Digital video and teaching." *Contemporary issues in technology and teacher education* 10.1 (2010): 1-6.

Berk, Ronald A. "Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom." *International Journal of Technology in Teaching & Learning* 5.1 (2009).

Simamora, Marta Widiawitasari Br, and Lulud Oktaviani. "What is your favorite movie?: a strategy of English Education students to improve english vocabulary." *J. English Lang. Teach. Learn* 1.2 (2020): 44-49.

Aminatun, Dyah, and Lulud Oktaviani. "Memrise: Promoting students' autonomous learning skill through language learning application." *Metathesis: Journal of English Language, Literature, and Teaching* 3.2 (2019): 214-223.

Oktaviani, Lulud, and Fatimah Mulya Sari. "REDUCING SOPHOMORE STUDENTS' DILEMA IN CREATING AN APPEALING TEACHING MEDIUM THROUGH SLIDESGO USAGE." *Jurnal IKA PGSD (Ikatan Alumni PGSD) UNARS* 8.2 (2020): 342-349.