

*International Journal of Learning, Teaching and Educational Research*  
 Vol. 25, No. 2, pp. 217-242, February 2026  
<https://doi.org/10.26803/ijlter.25.2.11>  
 Received Nov 01, 2025; Revised Jan 27, 2026; Accepted Jan 28, 2026

## Enhancing Conversational Competence with Short Simulation-Based Clips in a Thai University First-Year EFL Course

Nattana Boontong<sup>ID</sup>, Atitaya Pakladapongsin<sup>ID</sup>, Jiratchaya Metee<sup>ID</sup>,

Pimmada Japang<sup>ID</sup>

Rajamangala University of Technology Srivijaya  
 Songkhla, Thailand

Thapanee Khemanuwong<sup>ID</sup>

King Mongkut's Institute of Technology Ladkrabang  
 Bangkok, Thailand

**Abstract.** This study has examined whether short, simulation-based animated clips strengthen listening and speaking skills in a Thai university first-year EFL course. Using a convergent mixed-methods design, the study employed a one-group pretest-posttest quantitative design complemented by qualitative interviews to examine a clip-supported instructional intervention. One intact cohort ( $n = 34$ ) completed pre- and post-assessments of listening and speaking during the intervention. The quantitative results showed significant gains in both listening and speaking performance ( $p < .001$ ), with large, paired effect sizes (listening  $d = 1.22$ , speaking  $d = 1.30$ ). To complement these findings, semi-structured interviews with a select number of students and the course instructor were analyzed thematically to explore the learning processes and instructional feasibility. The qualitative findings indicated that dual-channel input, familiar everyday scenarios, and brief post-clip micro-drills (60–90 seconds) supported comprehension, the noticing of language forms, and greater speaking engagement and confidence. The integration of quantitative and qualitative findings informed a practical, instructor-led model that converts multimodal input into immediate speaking practice. The results suggest a scalable, low-preparation approach for enhancing listening-to-speaking development in EFL contexts.

**Keywords:** simulation-based clips; multimodal input; conversational competence; Thai EFL; role-play transfer

---

\*Corresponding author: Nattana Boontong; [nattana.b@rmutsu.ac.th](mailto:nattana.b@rmutsu.ac.th)

## 1. Introduction

Digital technologies are standard in Thai universities, yet the classroom issue is specific: many first-year EFL students find conversational listening and speaking difficult and are reluctant to speak. Prior Thai studies note limited opportunities to use English outside the classroom and anxiety about pronunciation or errors, which can reduce participation and willingness to speak (Akkakoson, 2016; Qin & Poopatwiboon, 2023). At the same time, AI-supported tools and multimedia can broaden access to comprehensible input and low-stakes practice (Kukulska-Hulme & Viberg, 2018; Qiao & Zhao, 2023), while the role of English as a lingua franca (ELF) keeps the academic and career stakes high for undergraduates (Crystal, 2003).

In everyday practice, many first-years encounter English mainly in scheduled lessons, so opportunities for repeated, low-pressure speaking are scarce. When they do arise, concerns about pronunciation and the risk of making mistakes often dampen participation (Akkakoson, 2016; Qin & Poopatwiboon, 2023). Consequently, there is a need for a classroom routine that increases exposure and practice without adding extra class time. Here, AI-supported multimedia is helpful because it enables brief, reusable input cycles that align with the lesson aims, feed short drills, and lead naturally into pair work (Kukulska-Hulme & Viberg, 2018; Qiao & Zhao, 2023).

Many lecturers already integrate technology into routine lessons, but short, structured tasks built around brief animation or video, followed by in-class practice, are especially workable. With university EFL learners, unit-linked task sequences have been associated with gains in fluency, vocabulary, pronunciation, and confidence (Menggo et al., 2022; Wahyuni & Pratiwi, 2021). In practical terms, this looks like a tight pre-while-post cycle: a 60-90-second clip to set context, a focused drill to surface target patterns, and pair work that requires those patterns in a short exchange. Keeping the clips brief reduces the extraneous load and leaves time for speaking, which is where confidence tends to build.

Simple scaffolds such as key phrases on the board, one example turn, and a visible timer also make the speaking step clear without adding materials. At entry level, students often avoid formal speaking because it feels risky. Low-stakes or semi-informal exchanges, including online warm-ups, can lower that barrier and help them try out the language (Akkakoson, 2016; Pusey & Nanni, 2022). Used this way, light digital practice can warm the room, so the same patterns transfer into face-to-face talking during the lesson (Zadorozhnyy et al., 2025). What courses need are materials that allow repeated, low-anxiety practice in familiar settings with AI-supported options when they genuinely help that fit inside the existing period without extra sessions.

This study addresses that need by embedding short, animated clips that model everyday conversational skills (greeting and introductions, small talk, apologizing, telephoning, and asking/giving directions) into regular lessons. Each clip is followed by a brief focus on the target patterns and a short pair task, so the timetable stays intact. The design follows Mayer's (2009) multimedia

learning principles because aligning verbal and visual channels and reducing the extraneous load helps first-year undergraduate students focus on the message rather than the effects. It also draws on Krashen's (1982) affective filter because low-pressure practice and quick successes encourage hesitant students to speak. In other words, the routine is built to keep the cognitive load manageable and anxiety low while still giving students more chances to talk within the same period. The practical goal is simple: support speaking and listening with classroom-ready materials that fit existing periods and equipment, not extra sessions or special labs.

Despite growing evidence for the effectiveness of multimedia in EFL instruction, research in the Thai context has paid limited attention to simulation-based animated clips for developing undergraduates' English conversational skills. Existing studies have largely focused on general video use, providing little insight into simulation-oriented environments that support repeatable interaction and systematic teacher mediation. Because pedagogical uptake depends on both learning outcomes and instructional feasibility, lecturer implementation in authentic classroom settings remains underexplored. To address these limitations, the present study targets three gaps in the literature: (1) an empirical gap regarding integrated listening–speaking development through simulation-based animated clips that has been rarely examined in Thai EFL research; (2) a methodological gap, given the limited documentation of classroom-embedded, low-preparation interventions; and (3) a practical gap, concerning lecturer feasibility and curricular alignment within regular university EFL courses.

Against this backdrop, the study investigates a classroom-ready set of short, simulation-based animated clips designed to enhance conversational competence among first-year EFL undergraduates at a southern Thai university. Specifically, it examines whether: (1) exposure to the clips leads to gains in listening and speaking from pre-test to post-test; (2) students perceive improvements in motivation, confidence, and engagement; and (3) the course lecturer judges the approach to be feasible, instructionally effective, and well aligned with the existing curriculum. By foregrounding both learning outcomes and implementation conditions, the study contributes a replicable, instructor-led classroom routine that translates the multimodal input from short simulation-based clips into immediate speaking performance, while explicitly specifying the feasibility parameters for routine EFL instruction.

### **1.1 Research Questions**

**RQ1:** Is there a significant difference in students' listening and speaking performance from pre-test to post-test after exposure to simulation-based animated clips?

**RQ2:** How do students perceive the usefulness of simulation-based animated clips for learning English conversation?

**RQ3:** How do the clips influence students' motivation, confidence, and engagement in speaking English?

**RQ4:** How does the course lecturer perceive the feasibility, classroom integration, and effectiveness of using simulation-based animated clips?

## 2. Literature Review

This section synthesizes relevant literature and outlines the theoretical framework that informs the design and interpretation of the study.

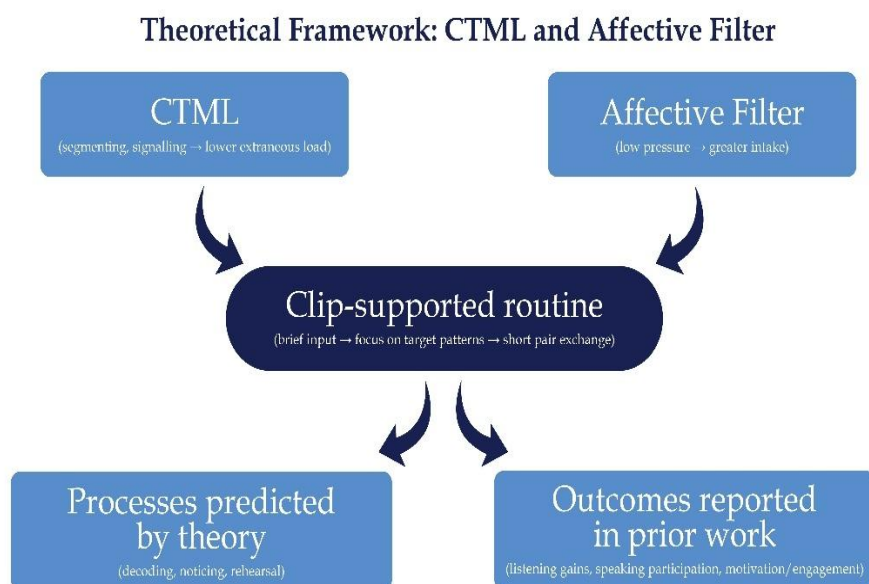
### 2.1 Theoretical Framework

Based on Cognitive Theory of Multimedia Learning (CTML) (Mayer, 2009) and the Affective Filter Hypothesis (Krashen, 1982), this study assumes two things: pairing visuals with audio and signaling key points in a logical order reduces the cognitive load, and affective factors (motivation, confidence, anxiety) determine how much input learners can use. When the filter is low, more language is taken in. Accordingly, concise, student-centered multimedia activities (e.g., animated clips with a quick exchange pair) tend to model the target language and lower anxiety more effectively than teacher-guided controlled practice. This emphasis is consistent with CTML's segmenting and signaling principles, with evidence that low-pressure tasks support uptake (Mayer, 2009; Krashen, 1982).

In the Thai context, many EFL classrooms emphasize teacher-centered, memorization-focused methods; students often experience more anxiety, less engagement, and fewer authentic opportunities to use English (Yuh & Kaewurai, 2021). Learner-centered alternatives trend differently: short, well-scaffolded video, games, or simulation tasks have raised speaking performance, participation, and confidence in university EFL settings (Angelini, 2019; Franciosi et al., 2016; Lee, 2023; Pittayanantakul & Phusawisot, 2024).

Beyond general video use, simulations point to specific benefits. Simple simulation tasks can yield better delayed vocabulary retention than non-game practice and higher engagement (Franciosi et al., 2016). Classroom simulations integrated into lessons have improved speaking-related outcomes (Angelini, 2019), and broader reviews report positive effects for simulation-supported foreign language learning (Zvarych et al., 2023). These patterns align with CTML's dual-channel account and support Krashen's view that low-pressure practice aids learning (Mayer, 2009; Krashen, 1982).

As multimedia and simulation platforms have become more accessible, language teachers are increasingly using these tools to create interactive, immersive activities. Commercial platforms can provide repeatable, context-rich scenarios paired with clear English audio, enabling brief input cycles that lead directly into pair work. Since classroom adoption depends on both student outcomes and practical feasibility, the lecturer's perspective on integration is therefore a key part of this framework. Figure 1 illustrates the theoretical framework guiding this study, integrating the Cognitive Theory of Multimedia Learning (CTML) with the Affective Filter Hypothesis to explain how multimodal input and affective conditions jointly support listening-to-speaking development.



**Figure 1: Theoretical Framework: CTML and Affective Filter**

## 2.2 Animated Video Clips in Language Learning

Animated video clips are now widely used in English classes because they offer dynamic, multimodal input that mimics real interactions. In a classroom study, Wahyuni and Pratiwi (2021) noted that lessons supported by animation improved speaking skills (particularly pronunciation, vocabulary, fluency, and sentence construction) and also increased engagement and motivation by making meaning easier to watch and hear. Similarly, Menggo et al. (2022) discovered that undergraduates who completed video-based tasks outperformed those taught with traditional methods on oral assessments and reported increased confidence. The use of video also promoted independence and more active participation. Complementary evidence from university EFL settings shows that short, well-scaffolded simulation/video tasks can strengthen vocabulary retention and encourage oral participation (Angelini, 2019; Franciosi et al., 2016).

These findings align with Mayer's (2009) idea that learning is more effective when students receive information through both visuals and sound, with Krashen's (1982) view being that lower anxiety supports language learning. Rather than dense text, animated materials give contextualized input in a visually engaging, low-pressure setting, which improves understanding and communication in English. Related game-based interventions have also reported gains in speaking (Pittayanantakul & Phusawisot, 2024). Overall, the literature suggests that brief, targeted clips paired with explicit signaling and immediate practice support the development of conversational skills in ways that conventional text-heavy tasks often do not (Angelini, 2019; Menggo et al., 2022; Wahyuni & Pratiwi, 2021).

Many Thai EFL students have limited exposure to English outside class and often experience speaking anxiety. Prior studies report that undergraduates struggle with confidence and have limited speaking vocabulary (Akkakoson, 2016; Qin &

Poopatwiboon, 2023). Animated media can help by creating structured, low-anxiety opportunities for meaningful spoken practice (Franciosi et al., 2016; Menggo et al., 2022; Pittayanantakul & Phusawisot, 2024; Wahyuni & Pratiwi, 2021). However, most research focuses on younger learners or overall fluency, with fewer studies on conversational competence, involving integrating listening and speaking through turn-taking, among first-year Thai university students (Angelini, 2019; Menggo et al., 2022; Wahyuni & Pratiwi, 2021; Zvarych et al., 2023). This gap motivates the present study's emphasis on short clip sequences embedded in regular lessons to support listening-speaking integration and confidence.

### **2.3 Simulation-based machinima clips for EFL conversation**

A commercial life-simulation platform was used not for gameplay but as a machinima tool to produce short, teacher-controlled clips. Scenes were scripted and screen-captured into 60–90-second videos with precise turns and gestures, then dubbed in English. This workflow affords control over the context, turn length, and camera angle, enabling tight alignment to unit targets while managing extraneous load and lowering pressure before speaking. Evidence from language education contexts suggests that teacher-guided or student-created machinima can support communication outcomes and oral skills development, providing a practical bridge from multimodal input to performance (Ochoa Alpala & Ortíz García, 2018; Pellas & Christopoulos, 2022).

Overall, the literature offers a strong theoretical and practical case for the use of animated clips in EFL. Building on Mayer's Cognitive Theory of Multimedia Learning and Krashen's Affective Filter Hypothesis, prior work shows that dual-channel input and supportive emotional climates aid acquisition. Evidence from international and Thai contexts also points to gains in speaking, confidence, and classroom engagement through simulation-based or animated lessons (Akkakoson, 2016; Angelini, 2019; Franciosi et al., 2016; Lee, 2023; Menggo et al., 2022; Pittayanantakul & Phusawisot, 2024; Qin & Poopatwiboon, 2023; Wahyuni & Pratiwi, 2021; Zvarych et al., 2023). What remains less explored is how short classroom-embedded clips build integrated listening-speaking skills, conversational competence, among Thai undergraduates. This study addresses that gap by using short teacher-controlled machinima clips embedded within regular lessons to provide contextualized, low-anxiety opportunities for interactive English communication.

## **3. Methodology**

To address the research questions while capturing both outcomes and experiences, a mixed-methods approach was adopted, as detailed in the following subsection.

### **3.1 Research Design**

Following Creswell and Plano Clark (2018), this study employed a convergent mixed-methods design in which quantitative and qualitative strands were implemented concurrently over a four-week period and given equal priority. Quantitatively, the study followed a one-group pretest–posttest design, while

qualitative interviews were collected concurrently and integrated through a convergent mixed-methods framework. This approach enabled the capture of measurable learning outcomes alongside the participants' classroom experiences within the same instructional cycle. Quantitative measures tracked changes in listening and speaking performance, while the qualitative data explained how and why the changes occurred, especially in relation to learner motivation, confidence, and the feasibility of the intervention. The study was conducted within a single intact class under institutional scheduling constraints, making a quasi-experimental design impractical. Withholding instructional materials from a comparison group would also have been pedagogically inappropriate. Integrating the quantitative and qualitative strands thus enabled the triangulation of findings and provided a richer, more comprehensive interpretation of the intervention's effects.

Quantitatively, listening was measured using 15 multiple-choice items reflecting everyday conversational English, and speaking was assessed through brief role-play tasks rated on an analytic rubric (speaking was collected for the matched subsample). The two strands were analyzed separately and then merged through side-by-side comparison and a joint narrative to identify convergence, complementarity, and any divergence in explaining observed changes.

The instructional clips consisted of six short simulation-based machinima illustrating familiar exchanges (e.g., apologizing, giving directions, small talk). Each clip was dubbed in English (AI text-to-speech or local recordings) and embedded in weekly lessons to scaffold key expressions, vocabulary, and sentence patterns while keeping production classroom-ready and straightforward. Clips were aligned with the unit objectives to preserve lesson flow and authenticity.

Students from the tested cohort and the course instructor participated in semi-structured interviews within the same four-week window, focused on mechanisms (dual-channel support and micro-drill), motivation and confidence, feasibility, classroom fit, and perceived effectiveness. The interview findings were then integrated with the test results to interpret how and why changes occurred.

### **3.2 Participants**

The participants were first-year undergraduates enrolled in the English for Communication program at a public university in southern Thailand who were taking a compulsory English Conversation course during the first semester of 2025. The entry requirements and course content were standardized across the course sections, ensuring comparable instructional exposure. However, the students' initial English proficiency levels were not assumed to be equivalent. The total student population for the course was 65.

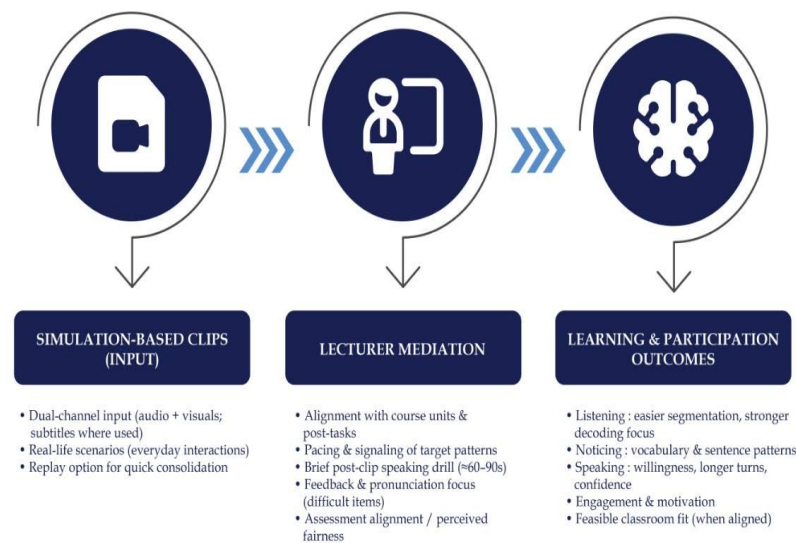
To account for potential attrition, 40 students were initially recruited. The priori power analysis of the paired-samples *t*-test ( $\alpha = .05$ , power = .80, expected effect size  $d = 0.50$ ) indicated that a minimum of 34 participants was required for the planned analyses. Proportionate stratified random sampling was employed, with

the course sections serving as strata to ensure proportional representation across timetable groups, followed by simple random selection within each stratum. Completely matched pre- and post-test data was obtained from 34 students, who constituted the final analytic sample and met the statistical power requirements for within-subject comparisons.

Although Slovin's formula (Sevilla et al., 1992) yields a similar figure, the power analysis was used to guide sample planning because it was more directly aligned with the statistical procedures employed. Because of standard attendance patterns and scheduling, complete speaking-task data was collected from 34 students. Students performed the speaking task in pairs, but each person received an individual score. From this participant pool, six students were selected for interviews using purposive sampling to capture the variation in speaking performance (high, medium, low), confidence, and participation, ensuring representation across the observed performance range while maintaining analytical depth. The course instructor taught the class, used simulation-based clips during regular lessons, served as one of two independent raters, and wrote a brief reflection on feasibility, classroom integration, and perceived effectiveness.

The instructor provided informed consent, did not recruit any of the participants, and all reporting used anonymized IDs because the instructor was also a rater. Potential expectancy bias is acknowledged and mitigated through independent double-rating and ICC checks, with the remaining risk noted in the Limitations section. Participant selection and the instructional context were aligned with the conceptual framework, which conceptualizes simulation-based clips as mediated input within regular classroom instruction and highlights learner engagement, interaction, and instructor facilitation as key mechanisms in listening-to-speaking development. Figure 2 illustrates this hybrid, clip-supported, instructor-led model.

**Conceptual Framework of Simulation-Based Clips and  
Conversation Development in a Thai University First-Year EFL Course:  
A Convergent Mixed-Methods Study**  
(Hybrid, clip-supported & lecturer-led model)



Frameworks: Mayer (2005) Cognitive Theory of Multimedia Learning; Krashen (1982) Affective Filter.  
Evidence aligned with Wahyuni & Pratiwi (2021); Menggo, Basir, & Halum (2022); Franciosi et al. (2016); Tonkonoh (2023); Thai EFL studies.

**Figure 2: Conceptual framework of simulation-based clips and conversation development in a Thai University First-Year EFL Course (hybrid, clip-supported, and lecturer-led model)**

### 3.3 Research Instruments

Three instruments were implemented: a listening comprehension test, a speaking task, and semi-structured interviews. Each instrument was aligned with six simulation-based animated units to maintain close correspondence between instruction and assessment. Conversation themes (greetings and introductions, apologizing, telephoning, and giving directions) were taken from the university's first-year English Conversation curriculum. The conversational themes were selected based on three criteria: (1) alignment with the official first-year English Conversation syllabus, (2) frequency and functional relevance in everyday university and service encounters, and (3) suitability for brief, role-play-based assessment. Topics such as greetings, apologizing, telephoning, giving directions, expressions of gratitude, and agreement or disagreement were prioritized because they require turn-taking, pragmatic appropriateness, and formulaic language, which are core components of conversational competence at the A2-B1 level.

More abstract or extended discourse types (e.g., debate or academic discussion) were excluded to maintain task authenticity, reduce the cognitive load, and ensure feasibility within short classroom speaking tasks. Classroom clips were produced as instructional machinima using a commercial life-simulation platform (The Sims 4; Electronic Arts/Maxis) to stage repeatable everyday scenes. Sequences were scripted, screen-captured, and assembled in a desktop editor, with subtitles added only when they supported the noticing of target phrases.

This workflow allowed control of the context, turn length, and camera angle, keeping clips short (60–90 seconds) and reusable across sections while minimizing extraneous load. Clips were produced as instructional machinima using a commercial life-simulation platform; the platform is named in the acknowledgments for transparency, and standard educational-use attribution is included.

The listening test consisted of 15 multiple-choice items, each pairing a brief dialogue with four response options, aligned with one of the six instructional units. The dialogues mirrored the communicative functions and scenarios later assessed in the speaking role-plays, ensuring thematic and pragmatic alignment between the receptive and productive tasks. Given the short instructional window and the focus on receptive growth rather than item learning, identical items were retained to maximize construct alignment, with practice effects addressed as a limitation. Audio materials were generated using TTSMaker, primarily employing British and American accents, with a limited number of additional international accents to reflect real-world EFL/ELF listening contexts.

The materials were dited in Audacity for clarity, volume normalization, and consistent timing. The inclusion of multiple accents was intended to assess communicative robustness rather than accent-specific familiarity. However, the narrower accent range used in the instructional clips is acknowledged as a limitation and discussed in relation to score variability and transfer (Sections 5 and 9). Synthetic voices ensured uniform pacing and controlled accent variation across the items. The test was administered via Google Forms in a supervised computer lab, with students using personal headphones, limited to a single playback per item. The same 15-item test was used for both administrations to ensure that any score changes reflected gains in receptive ability rather than practice effects.

Speaking performance was determined through brief role-play prompts aligned to the six units. Two speaking prompts were used to reflect progression across the course units rather than to establish causal sequencing. The post-test prompt required a broader range of pragmatic and interactional resources than the pre-test, allowing for the comparison of performance across time while remaining aligned with the instructional content. Improvements were therefore interpreted as developmental change within the course context rather than as direct effects of a specific task prompt.

The pre-test covered basic conversational skills (meeting a new friend, answering simple questions, self-introduction), whereas the post-test required more complex exchanges (giving directions, making an appropriate apology, conducting a short phone call). All performances were recorded and rated using a rubric adapted from Brown (2004) across the five criteria of fluency, vocabulary use, pronunciation, grammatical accuracy, and task completion, each rated on a five-point scale. Although the students worked in pairs, each student received an individual score. Two raters (the researcher and the course instructor) scored independently; discrepancies of more than 2 points within a criterion were

resolved by consensus. Inter-rater reliability for the total speaking scores was estimated with two-way random-effects and absolute-agreement ICC (2,2).

Six students were purposively selected from the speaking sample to represent varied abilities, and levels of confidence and participation. A researcher developed a semi-structured guide following established interview design guidance, covering perceived usefulness for listening and speaking, willingness and motivation to speak, confidence during participation, engagement, and suggested improvements. The course lecturer also provided a brief reflective interview on feasibility, classroom fit, and perceived effectiveness in regular teaching, including preparation time, syllabus alignment, and observed effects on listening and speaking confidence. Content validity was confirmed via expert review using the item-objective congruence (IOC) method, with a threshold of 0.67 or higher. All listening items met or exceeded the threshold; both speaking prompts recorded 1.00, and all five rubric criteria likewise recorded 1.00. The student interview protocol met the threshold after minor clarity edits.

### 3.4 Data Collection and Analysis

Instruction and testing occurred during regular class periods. Students completed the 15-item listening test twice in the computer lab using Google Forms. Each session lasted about 30 minutes. Speaking performances were collected twice through pair role-plays aligned with course themes; each student received an individual score from two independent raters. Although three or more raters may strengthen consensus in some designs, two independent raters were used in this classroom-based study due to feasibility constraints. Reliability was addressed through analytic rubrics, independent scoring, the resolution of large discrepancies by consensus, and the reporting of inter-rater agreement (ICC). This approach aligns with established practices in classroom speaking assessment and allows for the transparent evaluation of scoring consistency. Discrepancies of more than 2 points within a criterion were resolved by consensus, and the total score used for analysis was the average of the two raters.

For the quantitative analysis, descriptive statistics were computed, and pre- to post-test differences were evaluated using paired-samples *t*-tests (two-tailed,  $\alpha = .05$ ) and 95 percent confidence intervals. Effect sizes were reported as Cohen's *d* for the paired data, using the mean difference divided by the standard deviation of the differences. The normality of the pre-post difference scores was evaluated using the Shapiro-Wilk test. When the diagnostics suggested violations, Wilcoxon signed-rank tests were reported as robust checks. Inter-rater agreement on the total speaking scores was assessed using a two-way random-effect, absolute-agreement ICC (2,2); agreement was excellent at pre-test, 0.865 with a 95 percent confidence interval of 0.772 to 0.919, and fair at post-test, 0.449 with a 95 percent confidence interval of 0.205 to 0.620. Category-level kappa was explored but is not reported due to the analytic scale design.

The qualitative data was analyzed following Braun and Clarke's six-phase thematic approach. A small codebook was developed and applied across the corpus; two researchers independently coded the transcripts, checked the coding consistency on a 20% random sample, and resolved disagreements through

discussion. Thematic refinement produced five principal themes—perceived usefulness, listening support, speaking confidence, motivation and engagement, and feasibility and integration with negative and atypical cases examined during analysis. Credibility was supported through verbatim quotations, checks of thematic consistency, and triangulation between student and lecturer perspectives. During interpretation, qualitative themes were considered alongside observed changes in listening and speaking and then integrated with the quantitative findings using a side-by-side convergent approach. Data saturation was reached after the sixth interview, when no new themes emerged.

### 3.5 Ethical Considerations

Ethical approval (SCPHYLRB-2568/591) was granted by the host institution prior to the data collection. Participation was voluntary, and written informed consent was obtained from all students and the course lecturer. Students were informed that participation or non-participation would not affect their course grades. All data was anonymized using coded identifiers, and all recordings were stored securely and accessed only by the research team.

## 4. Research Findings

To address the four research questions, the findings are presented in two strands: the quantitative analyses of pre- and post-test listening and speaking performance (RQ1), followed by the qualitative results from the student and lecturer interviews examining perceived usefulness, motivation, confidence, engagement, and instructional feasibility (RQ2–RQ4).

### 4.1 Quantitative Findings

Table 1 shows that listening performance significantly improved after instruction with the simulation-based animated clips. For the matched cohort ( $n = 34$ ), the mean scores rose from 10.41 (SD = 2.46) at pre-test to 12.59 (SD = 1.81) at post-test (out of 15). The average gain was +2.18 points (95% CI [1.55, 2.80]), and a paired-samples  $t$ -test confirmed that this improvement was statistically significant,  $t(33) = 7.12$ ,  $p < .001$ . The effect size was Cohen's  $d_p = 1.22$ , showing a meaningful educational change. The normality of the difference scores was supported (Shapiro–Wilk  $W = 0.944$ ,  $p = .082$ ).

**Table 1: Listening Pre-Post (paired; full score = 15)**

n (pairs)	Pre M (SD)	Post M (SD)	$\Delta M$ (Post–Pre)	95% CI for $\Delta M$	t(df)	p	Cohen's $d_p$	Normality (W, p)
34	10.41 (2.46)	12.59 (1.81)	2.18	[1.55, 2.80]	7.12 (33)	< .001	1.22	0.944, .082

Note.  $\Delta M$  = mean difference; Cohen's  $d_p$  = paired-samples effect size; Shapiro–Wilk  $W$  tests normality of difference scores ( $p > .05$  indicates assumption not violated).

Table 2 shows that speaking performance clearly improved after the clips. For the matched cohort ( $n = 34$ ), the mean scores (max = 25) increased from 15.22 (SD = 3.13) at pre-test to 20.13 (SD = 1.94) at post-test—an average gain of +4.91 points (95% CI [3.60, 6.23]). A paired-samples  $t$ -test confirmed a significant improvement,  $t(33) = 7.59$ ,  $p < .001$ , with a large effect size (Cohen's  $d_p = 1.30$ ).

Although the Shapiro–Wilk test indicated that the difference scores were not perfectly normal ( $W = 0.915$ ,  $p = .012$ ), a nonparametric check (Wilcoxon signed-rank) reached the same conclusion ( $T = 0$ ,  $p = 7.8 \times 10^{-7}$ ). In short, the students' speaking scores increased substantially, and the result is robust to assumption checks.

**Table 2: Speaking Pre-Post (paired; full score = 25)**

n (pairs)	Pre M (SD)	Post M (SD)	$\Delta M$ (Post - Pre)	95% CI for $\Delta M$	$t(df)$	$p$	Cohen's $d_p$	Normality (W, p)
34	15.22 (3.13)	20.13 (1.94)	4.91	[3.60, 6.23]	7.59 (33)	< .001	1.30	0.915, 0.012

Note.  $\Delta M$  = mean difference; Cohen's  $d_p$  = paired-samples effect size; Shapiro–Wilk  $W$  tests normality of difference scores ( $p > .05$  indicates assumption not violated).

## 4.2 Qualitative Findings

Based on the qualitative findings, the thematic analysis of six student interviews (S1–S6) and one lecturer interview (L) identified three main areas of experience with the short simulation-based clips: (1) perceived learning mechanisms for listening and speaking, (2) classroom engagement and participation, and (3) conditions for effective implementation and improvement.

### 4.2.1 Perceived Learning Mechanisms (Listening & Speaking)

#### 4.2.1.1 Multimodal Support for Comprehension

All six students reported that combining audio with visuals made the spoken input easier to follow and segment, while the gestures and on-screen context helped them infer meaning, and replaying strengthened their understanding.

One student said,

*"We both listen and see the characters' gestures... it's close to real life"* (S2).

Another added,

*"The subtitles helped... the language is usable in real, everyday situations"* (S6).

The lecturer confirmed this pattern:

*"Comprehension was easier because of the visual cues"* (L).

#### 4.2.1.2 Real-Life Relevance and Transfer to Role-Play

Five students described the clips as close to everyday life, which in turn helped them speak more naturally during pair work and role-plays.

One student explained,

*"The clips are very close to real situations... when I meet the same situation, I remember how to respond"* (S1).

Another was concise:

*"The situations are like everyday life"* (S4)

The lecturer similarly noted,

*"It almost felt like the conversations were from real-life situations" (L).*

#### 4.2.1.3 Language Uptake: Vocabulary, Patterns, and Formulaic Sequences

Four students reported noticing the target vocabulary and sentence patterns during viewing and immediately reused them afterward. For example,

*"I also learned vocabulary and sentence patterns" (S1)*

*"The clips help with new vocabulary and sentences" (S3).*

Supporting this, the lecturer said,

*"The animation helped students notice and use new vocabulary and sentence patterns" (L).*

#### 4.2.2 Classroom Engagement and Participation

##### 4.2.2.1 Confidence and Willingness to Speak

According to four students, the clips increased their willingness to speak because the visual and audio models reduced anxiety and enabled them to sustain longer turns during pair work. One student stated,

*"I feel more confident speaking because I can rewatch... I remember how to respond" (S1).*

Similarly,

*"Seeing the characters move... builds confidence" (S2).*

Furthermore, the lecturer agreed that the clips

*"Gave students the confidence to participate, especially in role-plays" (L).*

##### 4.2.2.2 Engagement and Motivation

Five students reported greater engagement when the text and audio were presented together rather than as they are in single-mode lessons. Short narrative hooks kept students focused and motivated them to try speaking. For example,

*"The clips made me more interested and focused" (S3), and*

*"They made me more interested in learning English... I felt like trying to speak" (S6).*

Consistently, the lecturer observed,

*"Using the animated clips made my lesson more engaging than usual" (L).*

#### 4.2.3 Conditions for Implementation and Improvement

##### 4.2.3.1 Feasibility and Alignment (Lessons and Assessment)

Two students highlighted that integration was simple when the clip topics matched the course units; replay supported lesson flow, and the aligned post-tasks felt fair. One informed,

*"Being able to replay helps... I know which situation the conversation belongs to" (S5).*

Another recommended pairing with the textbook to

*"learn while watching" (S6).*

Meanwhile, the lecturer explained that

*"lesson topics were aligned with the clips, so synchronization was easy,"*  
and confirmed that the clips

*“fairly prepared students for the post-listening and speaking tasks” (L).*

#### 4.2.3.2 Suggested Improvements (Actionable Refinements)

Five students suggested practical adjustments to improve learning: clearer, louder audio; greater accent variety; slightly longer clips with clearer mini-stories; and brief pronunciation or drill segments. For example, one student said,

*“Add a wider range of accents” (S2),* and another suggested,

*“Make the audio louder... use clips alongside regular teaching” (S3).*

To complement these ideas, the lecturer recommended keeping the clips concise while adding

*“difficulty levels – easy, medium, hard” (L).*

In short, the interviews show that when lessons used short scenario clips, the students said the speech was easier to track and remember. Because the scenes feel like everyday life, students can lift the expressions straight into pair work and role-plays. Learners noted that they learned the target language fast and, compared with single-mode lessons, they felt more confident and kept the floor for longer. One put it: *“The clips made me more interested and focused” (S3).* Delivery was easiest when the clip topics lined up with the unit aims and set post-task; both the students and lecturer viewed the marking as fair. The proposed improvements were pragmatic – to make the audio clearer and louder, to increase the variety of accents, to extend the story a bit, and to add a brief pronunciation and drill segment. The lecturer also suggested tiered levels (easy, medium, hard). With these changes and consistent teacher guidance, the clips can strengthen both listening and speaking.

Overall, the quantitative gains in listening (Table 1) and speaking (Table 2) mirror the interview themes: the students noted a clearer decoding of target phrases, higher confidence and willingness to speak, and steady engagement with a routine of short clips followed by micro-drill and pair work. The lecturer similarly highlighted feasible integration and perceived effectiveness. Two caveats for the Discussion are the lower post-test inter-rater agreement ( $ICC(2,2) = 0.449$ ) and recurring remarks about accent variety, which may account for the variability in individual speaking gains.

## 5. Discussion

Building on prior reports of the learning gains, this study extends the literature by specifying a practical classroom pathway: dual-channel input support; a brief 60–90-second micro-drill that consolidates forms; the students then applying those forms in near-transfer role-play; and clear classroom conditions that enable the routine in everyday teaching. This study pursued three goals: (1) to determine whether short simulation-based clips enhance listening and speaking skills; (2) to investigate the students’ perspectives on motivation, confidence, and engagement; and (3) to assess the lecturer’s views on feasibility, classroom fit, and effectiveness in routine teaching. Using a convergent mixed-methods approach, the quantitative and qualitative findings are interpreted together through a conceptual model where the dual-channel input supports noticing and

confidence, which then translates into role-play, with lecturer mediation (including alignment, pacing, a brief post-clip drill, and feedback), linking the input to the outcomes.

Because the study used a convergent mixed-methods design, the quantitative results (Tables 1–2) and interview themes have been interpreted side by side to highlight convergence and note any meaningful divergences. Students repeatedly pointed to dual-channel support (audio with visuals) and real-life relevance as reasons they could follow, remember, and reuse the target language in pair work. The lecturer emphasized alignment with the course units, a brief post-clip consolidation, and careful pacing as practical choices that preserved lesson flow and made the post-tasks feel fair.

These outcomes are consistent with the Cognitive Theory of Multimedia Learning, which explains the comprehension benefits of well-sequenced visual–auditory input (Mayer, 2009), and the Affective Filter Hypothesis, which helps account for the increased confidence and willingness to speak (Krashen, 1982). Prior studies likewise report speaking gains and higher participation from animated and simulation tasks (Angelini, 2019; Menggo et al., 2022; Wahyuni & Pratiwi, 2021), as well as vocabulary and engagement improvements due to game-enhanced practice (Franciosi et al., 2016; Lee, 2023). The subsections that follow discuss how the clips supported listening and speaking (5.1), how noticing became used (5.2), how participation changed (5.3), and the conditions that made the clips work in class (5.4), before outlining the study’s contribution (5.5).

### **5.1 How the clips supported listening and speaking**

Both listening and speaking improved significantly, and assumption checks supported the reliability of these gains. With the clips, learners consistently described dual-channel support—audio with gestures and on-screen context—that made the speech easier to segment and follow. Afterwards, they reused the target expressions in role-plays. Because the scenarios felt close to everyday life, the students reported longer turns, more volunteering, and increased confidence, indicating that comprehension translated into production. This listen-to-speak pathway aligns with multimedia learning (Mayer, 2009) and a lower anxiety climate (Krashen, 1982). It also matches reports that animation/simulation activities enhance decoding, pronunciation focus, engagement, and oral performance (Angelini, 2019; Menggo et al., 2022; Wahyuni & Pratiwi, 2021; Zvarych et al., 2023) and that simple simulations improve vocabulary retention and motivation (Franciosi et al., 2016; Lee, 2023; Pittayanantakul & Phusawisot, 2024). In this setting, the gains appear tied to design features (multimodal scaffolding plus a short drill) rather than high production values.

### **5.2 Noticing to use: vocabulary, patterns, and transfer**

The interview evidence shows a quick shift from noticing to use. Short simulation-based clips served as compact exemplars: the students picked up useful expressions and sentence patterns during viewing and applied them immediately in pair work. Familiar scenarios reduced uncertainty about when and how to speak, lowering hesitation and supporting longer turns. Prior simulation/video studies report similar effects—lasting vocabulary benefits and increased oral

participation (Angelini, 2019; Franciosi et al., 2016; Lee, 2023; Zvarych et al., 2023; Pittayanantakul & Phusawisot, 2024). In this setting, brief context-rich clips helped learners move from input to output within a single lesson, consistent with CTML (Mayer, 2009) and a low-pressure climate (Krashen, 1982).

### **5.3 Participation, confidence, and willingness to speak**

Students associated the clip-based sequence—short clip, brief micro-drill, then pair work—with lower anxiety, a greater willingness to speak, more frequent volunteering, and longer turns. A visible model and clear target forms reduced the perceived risk, while the brief drill helped shift forms from recognition to production without disrupting lesson flow. This pattern aligns with the affective filter account (Krashen, 1982) and with reports that animation or game-like tasks raise engagement and confidence in EFL contexts, including Thailand (Pittayanantakul & Phusawisot, 2024; Wahyuni & Pratiwi, 2021; Zvarych et al., 2023). It also contrasts with evidence that teacher-centered routines can heighten anxiety and restrict authentic use (Akkakoson, 2016; Yuh & Kaewurai, 2021). For learners with limited out-of-class exposure and lower starting confidence (Qin & Poopatwiboon, 2023), the sequence provided manageable speaking opportunities. The participation trend is also consistent with meta-analytic evidence that the informal digital learning of English is associated with a higher willingness to communicate (Zadorozhnyy et al., 2025).

### **5.4 Conditions for effective implementation (feasibility, fit, effectiveness)**

Implementation was straightforward when the clip topics aligned with the course units and post-tasks, supporting the perception of fair assessment. Preserving lesson flow required trimming some materials but engagement remained high. The participants recommended clearer and louder audio, moderate accent variety, slightly longer mini-stories to clarify pragmatic intent, and a brief pronunciation or micro-drill pause before pair work—this is consistent with research on multimodal clarity and structured, low-pressure practice (Mayer, 2009; Menggo et al., 2022; Wahyuni & Pratiwi, 2021). A minor risk concerns accent transfer: the instructional voices were relatively uniform, whereas the post-test mixed accents. This mismatch may help explain the variability in individual speaking gains (Qin & Poopatwiboon, 2023). Overall, the balance of evidence favors design fundamentals—dual-channel input, a short post-clip drill, and explicit task alignment—over production effects.

### **5.5 New contributions**

This study offers a classroom-level account of how short simulation-based animated clips function as targeted examples in a Thai university conversation class. Rather than treating video as a generic add-on, the findings point to a repeatable process: (1) dual-channel input makes spoken forms easier to understand, (2) familiar real-life scenarios anchor meaning, (3) learners adopt patterns and immediately reuse them, and (4) a short micro-drill (60–90 seconds) reinforces the shift from form to use. The result is a hybrid, clip-supported, instructor-led model in which clips initiate noticing and confidence, while the teacher ensures alignment, feedback, and spontaneous talk. This model aligns with CTML (Mayer, 2009) and the affective filter (Krashen, 1982), and it echoes international evidence that animation/simulation activities can improve

speaking, engagement, and vocabulary (Angelini, 2019; Franciosi et al., 2016; Lee, 2023; Menggo et al., 2022; Wahyuni & Pratiwi, 2021; Zvarych et al., 2023; Pittayanantakul & Phusawisot, 2024). It also addresses Thai-specific needs for low-anxiety speaking practice (Akkakoson, 2016; Qin & Poopatwiboon, 2023; Yuh & Kaewurai, 2021).

## 6. Conclusion

This study indicates that short simulation-based animated clips can improve first-year EFL learners' listening and speaking when implemented in a teacher-led format. Multimodal input (audio with visuals) made speech easier to follow and remember, supported the noticing of vocabulary and patterns, and—together with brief, teacher-guided post-clip activities—led to immediate role-plays, longer turns, and greater willingness to speak. Both students and the lecturer judged the approach to be practical when the clip topics matched the unit aims and post-tasks, and they suggested modest refinements: clearer or louder audio, a broader but still manageable accent range, slightly longer mini stories to clarify pragmatic intent, and a brief pronunciation focus.

Overall, the findings support a hybrid, clip-supported, instructor-led model. The clips supply dual-channel input that lowers processing demands and affective barriers. The move from understanding to production depends on teacher mediation, alongside choosing clip-unit alignments, regulating the viewing pace, and running focused micro-drills, so then the modeled language becomes usable in conversation. This extends the CTML–Affective Filter framing where the dual-channel input, coupled with brief, teacher-supported practice, facilitates near-immediate transfer to role-play without reducing interaction time.

For the course design and delivery, the guidance is straightforward: align the clips tightly to syllabus units and post-tasks; include a 60–90-second post-clip micro-drill; ensure clear, adequately loud audio with moderate accent variety; provide easy, medium, and hard bands to accommodate proficiency differences; and add a brief pronunciation moment for challenging items. Under these conditions, technology supports attention, clarity, and willingness to speak. At the same time, the teacher sustains accuracy, fairness, and spontaneity, the conditions under which animated clips most reliably strengthen conversational skills.

## 7. Implications for Practice and Policy: From Multimodal Input to Usable Conversation

In practice, the model is straightforward: short simulation-based clips provide multimodal scaffolding, and lecturer mediation turns that become practical conversation. A light rhythm keeps the prep minimal and discussion focused: highlight one or two target patterns and the communicative move (1–2 minutes), view the clip for gist and briefly replay the key exchange, run a 60–90-second micro-drill (short pronunciation focus plus one choral or paired cue), move into a 3–6-minute role-play in a near-transfer scenario, and close with a quick debrief (one note on accuracy, one on appropriacy, one student example). Clips work best when tied closely to unit goals and post-tasks, with a small pattern log (target phrase plus learner variant) to make recycling visible across weeks. Because the

listening test includes multiple accents, transfer should improve if the clip bank adds a broader range of accents (for example, one or two extra accents per unit). Pair key lines with contrasting playback (Accent A followed by Accent B) and include 30–45 seconds of shadowing on the most challenging lines. These are small steps that raise robustness without overload. Keep the audio clear and at a comfortable loudness.

For implementation, maintain a simple clip bank tagged by topic, runtime, target patterns, and level band. If the sound or video fails, printed frames and a brief excerpt from the transcript can sustain the lesson. Use a single standard educational-use attribution across all materials. At the program level, clarify whether the clips count toward assessment, include a brief, non-stigmatizing pronunciation moment, and uphold strict privacy (no recording without consent). Monitor class energy with three indicators – volunteers per lesson, average pair-turn length, and one accuracy note – and lighten the load or adjust pacing if these dip. In short, the clip sparks the language and the interaction refines it.

## **8. Limitations and Directions for Future Research**

This single-site study used one intact cohort with a single lecturer, so the findings reflect a specific institutional context and teaching style. The one-group pre–post design without a control condition limits causal inference and leaves room for novelty or expectancy effects. The listening test reused the same 15 items at the pre- and post-test, which may introduce minor practical effects despite the four-week interval.

Speaking scores were based on human ratings; the inter-rater agreement was excellent at pre-test but only fair at post-test ( $ICC(2,2) = 0.449$ ), so residual measurement error cannot be ruled out. The interviews involved six students and one lecturer, offering depth but a limited range of perspectives. The instructional materials draw on a specific set of short simulation-based clips (topics, pacing, accent mix), which may not generalize to other media, classrooms, or hardware configurations. In addition, the instruction featured a narrower accent range than the multi-accent listening assessment, which may have contributed to variability in individual gains.

Future work should expand the scope and strengthen inference by using multi-site samples and randomized or quasi-experimental comparisons that match time-on-task. To curb practice effects, use parallel forms or an item bank and include delayed post-tests to examine retention and transfer. Improve measurement with rater training, anchor samples, periodic reliability checks, and, where feasible, many-facet Rasch or generalizability studies to separate the person, task, and rater effects. Align the instructional and assessment accents more closely or test accent transfer explicitly with planned contrasts. Add process indicators – pattern logs, volunteers per lesson, average pair-turn length, brief confidence scales – and analyze outcomes with mixed-effects models that account for student and task clustering. Finally, compare practical design choices (clip length, mini-story clarity, number and length of micro-drills, difficulty bands) and conduct cost-prep analyses to guide the scalable adoption in Thai and similar EFL contexts.

## 9. Acknowledgments

Appreciation is extended to the participating students, the course lecturer, and the host institution for their support of the classroom study. During the manuscript preparation, language editing tools (Grammarly) and ChatGPT (OpenAI) were used to suggest wording and improve clarity/grammar. No AI tools were used to generate, alter, or analyze data; all content reflects the authors' own work and intellectual contributions. A commercial life-simulation platform (The Sims 4; Electronic Arts/Maxis) was used solely to create short noncommercial instructional clips for educational purposes in accordance with fair-use principles.

## 10. References

- Akkakoson, S. (2016). Speaking anxiety in English conversation classrooms among Thai university students. *Procedia - Social and Behavioral Sciences*, 232, 214–220. <https://doi.org/10.1016/j.sbspro.2016.10.051>
- Angelini, M. L., & García-Carbonell, A. (2019). Developing English speaking skills through simulation-based instruction. *Teaching English with Technology*, 19(2), 3–20. <https://files.eric.ed.gov/fulltext/EJ1215381.pdf>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Crystal, D. (2003). *English as a global language (2nd ed.)*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511486999>
- Franciosi, S. J., Yagi, J., Tomoshige, Y., & Ye, S. (2016). The effect of a simple simulation game on long-term vocabulary retention. *CALICO Journal*, 33(3), 355–379. <https://doi.org/10.1558/cj.v33i2.26063>
- Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, 72(12), 2954–2965. <https://doi.org/10.1111/jan.13031>
- Krashen, S. D. (1982). Principles and practice in second language acquisition. *Pergamon*. [https://www.sdkrashen.com/content/books/principles\\_and\\_practice.pdf](https://www.sdkrashen.com/content/books/principles_and_practice.pdf)
- Kukulka-Hulme, A., & Viberg, O. (2018). Mobile collaborative language learning: State of the art. *British Journal of Educational Technology*, 49(2), 207–218. <https://doi.org/10.1111/bjet.12580>
- Lee, S. M. (2023). Factors affecting incidental L2 vocabulary acquisition and retention in a game-enhanced learning environment. *Language Teaching Research*, 27(6), 1124–1149. <https://doi.org/10.1177/13621688221142739>
- Mayer, R. E. (2009). *Multimedia learning (2nd ed.)*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>
- Menggo, S., Basir, N., & Halum, M. (2022). Video-based tasks in strengthening speaking skills of EFL college students. *Indonesian Journal of EFL and Linguistics*, 7(2), 269–288. <https://doi.org/10.21462/ijefl.2022.7.2.6>
- Ochoa Alpala, C. A., & Ortíz García, W. R. (2018). Creating machinima (3D) and real-life videos in an ESP classroom. Profile: *Issues in Teachers' Professional Development*, 20(1), 13–27. <https://files.eric.ed.gov/fulltext/EJ1165938.pdf>
- Pellas, N., & Christopoulos, A. (2022). The effects of machinima on communication skills in students with developmental dyslexia. *Education Sciences*, 12(10), 684. <https://doi.org/10.3390/educsci12100684>

- Pittayanantakul, N., & Phusawisot, K. (2024). The effect of board games on speaking skills of Thai EFL students. *Journal of Education and Learning*, 13(6), 226–235. <https://doi.org/10.5539/jel.v13n6p226>
- Pusey, K., & Nanni, A. (2022). Thai University Students' Attitudes Toward Online Language Learning in a Time of Crisis. *Sage Open*, 12(2). <https://doi.org/10.1177/21582440221091712>
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in a Chinese EFL context. *Frontiers in Psychology*, 14, 1255594. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Qin, X., & Poopatwiboon, S. (2023). Thai EFL undergraduates' oral communication anxiety and its predictors. *English Language Teaching*, 16(11), 70–83. <https://doi.org/10.5539/elt.v16n11p70>
- Sevilla, C. G., Ochave, J. A., Punsalan, T. G., Regala, B. P., & Uriarte, G. G. (1992). *Research methods*. Quezon City, Philippines: Rex Book Store.
- Songkhro, J., & Thammasitboon, S. (2022). Animated videos via Google Sites: Learning achievement and satisfaction among Thai undergraduates. *Education Quarterly Reviews*, 5(2), 371–381. <https://doi.org/10.31014/aior.1993.05.02.499>
- Wahyuni, L. G. E., & Pratiwi, N. P. M. (2021). Enhancing EFL students' speaking competency through animated learning video. In *Proceedings of the 4th International Conference on Interdisciplinary Research on Education, Arts, and Design (ICIRAD 2021)* (pp. 176–181). Atlantis Press. <https://doi.org/10.2991/assehr.k.211222.029>
- Yuh, A. H., & Kaewurai, W. (2021). Teacher-centred practices in Thai classrooms: Issues and implications. *The Golden Teak: Humanity and Social Science Journal*, 27(2), 91–107. <https://so01.tci-thaijo.org/index.php/GoldenTeak/article/view/247632>
- Zadorozhnyy, A., & Lee, J. S. (2025). Informal Digital Learning of English and willingness to communicate in a second language: self-efficacy beliefs as a mediator. *Computer Assisted Language Learning*, 38(4), 669–689. <https://doi.org/10.1080/09588221.2023.2215279>
- Zvarych, I., Tonkonoh, I., Bopko, I., Melnychuk, S., Mehela, K., & Shyrmova, T. (2023). The effectiveness of using simulation in learning a foreign language. *Forum for Linguistic Studies*, 5(3), Article 1916. <https://doi.org/10.59400/FLS.v5i3.1916>

## Appendix 1

Listening blueprint (sample items, 1 page)

No.	Listening Comprehension (Pre-Post Test)
1	<p><b>Audio Dialogue:</b>            Man: Where are you going?            Woman: I'm heading to the cafeteria. Do you want to come?            Question: Where is the woman going?</p> <ul style="list-style-type: none"> <li>• A. To the bookstore</li> <li>• B. To the cafeteria</li> <li>• C. To the library</li> <li>• D. To the classroom</li> </ul>
2	<p><b>Audio Dialogue:</b>            Woman: Hi, I'm Jane.            Man: Nice to meet you, Jane. I'm Tom.            Question: What are the speakers doing?</p> <ul style="list-style-type: none"> <li>• A. Saying goodbye</li> <li>• B. Introducing themselves</li> <li>• C. Talking about plans</li> <li>• D. Giving directions</li> </ul>
3	<p><b>Audio Dialogue:</b>            Man: Let's hang out after class.            Woman: That sounds fun! Where should we go?            Question: What does the woman mean?</p> <ul style="list-style-type: none"> <li>• A. She agrees to hang out.</li> <li>• B. She is too busy.</li> <li>• C. She doesn't want to go.</li> <li>• D. She is angry.</li> </ul>

## Appendix 2

Speaking Task	
1	<p><b>Task Pre-Test Speaking Task</b></p> <p><b>Prompt:</b> Imagine you're meeting a new exchange student on their first day at university. Introduce yourself and help them feel welcome. Then, ask and answer basic questions to get to know each other. You can also offer help if they seem confused about anything.</p> <p><b>Themes Covered:</b> Greeting and Introduction (Clip 2), Understanding Questions (Clip 1), Socializing (Clip 3)</p> <p><b>Preparation Time:</b> 1 minute <b>Target Time:</b> 1.5–2 minutes</p>
2	<p><b>Task Post-Test Speaking Task</b></p> <p><b>Prompt:</b> Your foreign friend is visiting your city. They're lost and call you for help. Give them directions. Later, they tell you they accidentally offended someone and feel bad. Give advice or apologize for them. Finally, invite them to a group hangout with your friends.</p> <p><b>Themes Covered:</b> Giving Directions (Clip 6), Apologizing (Clip 4), Telephoning (Clip 5), Socializing (Clip 3 revisited)</p> <p><b>Preparation Time:</b> 1 minute <b>Target Time:</b> 1.5–2 minutes</p>

### Appendix 3

Speaking performance was assessed with an analytic rubric during brief conversation tasks prompted by short, simulation-based animated clips. Each criterion was rated on a five-point scale by two independent raters.

Criteria	5 - Excellent	4 - Good	3 - Satisfactory	2 - Needs Improvement	1 - Poor
Task Completion	Fully and appropriately responds to the prompt with clear, relevant, and detailed information.	Adequately responds to the prompt with mostly relevant and clear information.	Responds to the prompt but with some missing or unclear information.	Partially responds to the prompt with limited or somewhat off-topic content.	Fails to respond appropriately; response is minimal or off-topic.
Fluency	Speaks smoothly with natural pacing and minimal hesitation.	Generally fluent with occasional hesitation or self-correction.	Noticeable pauses or hesitation but meaning is mostly clear.	Frequent pauses or disruptions that interfere with communication.	Speech is fragmented or difficult to understand due to fluency issues.
Pronunciation	Clear and natural pronunciation with minimal mispronunciations.	Mostly clear with few mispronunciations that do not hinder understanding.	Some mispronunciations that occasionally interfere with understanding.	Frequent mispronunciations that make understanding difficult.	Pronunciation severely hinders communication.
Grammar	Consistently uses accurate grammar with complex sentence structures.	Generally accurate grammar with minor errors.	Some grammatical errors that do not significantly affect meaning.	Frequent grammatical errors that hinder understanding.	Major grammar issues that severely impede communication.
Vocabulary Use	Wide range of vocabulary used accurately and appropriately.	Good range of vocabulary with occasional imprecision.	Limited vocabulary but meaning is generally conveyed.	Very basic vocabulary; lacks variety or appropriateness.	Vocabulary is insufficient and prevents communication.

### Appendix 4

No.	Students' Interview
1	<p><b>What do you think about using the short simulation-based clips in your conversation class? Did the clips help your listening and speaking? If yes, how?</b></p> <p><b>Probes:</b></p> <ul style="list-style-type: none"> <li>- Can you give an example of something you learned from a specific clip?</li> <li>- Did the conversations feel similar to real-life situations?</li> <li>- Was it easier to understand the English because of the animation?</li> <li>- Did you find it helpful for learning new vocabulary or sentence patterns?</li> </ul>
2	<p><b>What did you enjoy or find most useful about using the short simulation-based clips in class?</b></p> <p><b>Probes:</b></p> <ul style="list-style-type: none"> <li>- Was it more fun or interesting than regular lessons?</li> <li>- Did the characters' voices or actions help you understand the meaning?</li> <li>- Did the clips make you feel more confident to speak?</li> <li>- Which part did you like the most – listening, repeating, or role-playing?</li> </ul>
3	<p><b>3a. Do you have any suggestions for improving this type of English learning?</b></p> <p><b>3b. How did using the short simulation-based clips affect your motivation to learn English?</b></p> <p><b>Probes:</b></p> <ul style="list-style-type: none"> <li>- Would you prefer longer or shorter clips?</li> <li>- Do you want more speaking practice after watching?</li> <li>- Should the video topics be more challenging or easier?</li> <li>- What kind of situations would you like to see next time?</li> <li>- Did it make learning more enjoyable? Did you want to participate more?</li> </ul>

## Appendix 5

Simulation-Based English Conversat...



# RUTS


RAJAMANGALA  
UNIVERSITY  
OF TECHNOLOGY  
SRIVIJAYA

Enhancing Conversational Competence with Short Simulation-Based Clips in a  
Thai University First-Year EFL Course: A Convergent Mixed-Methods Study


Home Unit v Pre-Test Post-Test




Simulation-Based English Conversat...




UNIT 3  
Socializing



UNIT 4  
Apologizing




UNIT 5  
Telephoning



UNIT 6  
Asking and Giving direction

Speaking English is challenging for many Thai undergraduates, particularly in real-life situations that demand confidence and fluency. This project integrates short, simulation-based animated clips to model everyday conversations (e.g., greetings, apologizing, telephoning). The clips provide multimodal input—visuals, voice, and situational context—that makes spoken English easier to perceive and process. Grounded in simulation-based and game-enhanced learning, the approach is designed to build speaking fluency, strengthen listening comprehension, and increase willingness to speak through authentic, classroom-embedded practice aligned to routine communicative needs.

CONTACT



Downloaded from: [Mitsuru Fukuhara, Shih-Wei Hsiao, and Pimwadee Jaing](#)

Home Unit v Pre-Test Post-Test

### Unit 6

Asking and Giving direction



#### Description

This clip focuses on functional language for asking and giving directions. Students will hear expressions like "Excuse me, how do I get to...?", "Go straight ahead," and "Turn left at the corner." The scenes reflect real-life navigation scenarios such as finding a location or helping someone who is lost. The goal is to support students in building confidence using English in public settings where giving directions is necessary.

Previous

Home