

Effects of Using Microsoft Flip on Undergraduate Students' Speaking Skills

Santhoshkanth Kamaraj^{1*} / Angcharin Thongpan² / Sunchai Hamcumpai³

^{1*}English Language Teaching Graduate Program, Faculty of Liberal Arts and Science, Roi Et Rajabhat University, Roi Et, Thailand

E-mail: santhoshkanth88@reru.ac.th

²English Language Teaching Graduate Program, Faculty of Education and Human Development, Roi Et Rajabhat University, Roi Et, Thailand

E-mail: angcharin@reru.ac.th

³English Language Teaching Graduate Program, Faculty of Liberal Arts and Science, Roi Et Rajabhat University, Roi Et, Thailand

E-mail: shamcumpai@gmail.com

Received: January 24, 2026

Revised: February 24, 2026

Accepted: March 20, 2026

Abstract

A persistent disparity exists between EFL learners' theoretical linguistic knowledge and their practical oral proficiency—a phenomenon often described as 'mute English.' This study addresses this critical gap by examining the efficacy of integrating Microsoft Flip, an asynchronous video discussion platform, into a structured teaching-speaking cycle model based on Goh and Burns' (2012). The research aimed to (1) compare undergraduate students' speaking skill scores before and after the intervention and (2) investigate students' perceptions of the platform's utility. Using a pre-experimental, one-group pre-test/post-test design, the study employed cluster random sampling to recruit 53 first-year nursing undergraduate students at a public university in Thailand. Over an eight-week intervention, participants engaged in asynchronous speaking tasks designed to foster planning, rehearsal, and reflection across eight thematic modules. Data were analyzed using descriptive statistics and a paired-samples t-test. Quantitative analysis revealed a statistically significant improvement in speaking proficiency, with mean scores rising from 10.23 to 16.56 ($t(52) = 30.28, p < .001$), representing a large effect size (Cohen's $d = 4.16$). All sub-skills showed growth, with vocabulary exhibiting substantial gains. Furthermore, quantitative data from questionnaires indicated positive student perceptions ($M = 3.96$), particularly regarding the platform's ability to expand vocabulary ($M = 4.17$) and encourage regular practice. These findings suggest that combining asynchronous technology with teaching-speaking cycle instruction effectively ameliorates speaking deficits, offering a scalable pedagogical model for EFL educators seeking to enhance learner autonomy and oral competence in non-native contexts.

Keywords: microsoft flip, speaking skills, teaching speaking framework, EFL, undergraduate students

Introduction

Among the four macro-skills of language acquisition, speaking is widely recognized as the most fundamental component of communication and the primary mode of oral interaction (Woodrow, 2006; Richards, 2008). It is a productive skill that enables learners to express ideas, thoughts, and emotions effectively in academic, professional, and social settings (Sudarmo, 2021). Mastery of speaking significantly boosts learner confidence, promotes active participation in discussions, and enhances employment opportunities (Gillis, 2013; Copland et al., 2024). Conversely, a lack of oral proficiency can lead to professional stagnation and exclusion from global discourse.

In Thailand, English is officially recognized and mandated throughout the national curriculum. Despite this policy emphasis, a persistent gap remains between students' theoretical understanding of the language and their practical ability to use it, especially in spoken contexts (Nithideechaiwarachok et al., 2022). This phenomenon, often termed "mute English," is characterized by learners who may excel on grammar tests but struggle to articulate simple thoughts in real-time conversation. Many Thai students find it challenging to communicate effectively due to limited opportunities to practice outside the classroom (Shen & Chiu, 2019), fundamental difficulties with English grammar (Hamilton et al., 2024), and limited lexical resources (Noom-ura, 2013).

This proficiency gap creates a significant disadvantage, hindering both individual job prospects and the country's overall economic competitiveness (Siddoo et al., 2019; Thadphoothon, 2017). The problem is particularly severe for undergraduates in rural areas, where limited exposure to English speakers creates a profound disconnect between classroom knowledge and real-world application. Furthermore, the pressure to perform often leads to high levels of communication apprehension, inhibiting students' Willingness to Communicate (WTC) and perpetuating a cycle of avoidance and low proficiency.

Despite the proliferation of research highlighting the benefits of technology in EFL classrooms (Gilakjani, 2017; Parvin & Salam, 2015) and the specific affordances of Microsoft Flip within the Thai context (Pornwasanying, 2023; Robillos, 2023), a critical gap remains regarding its pedagogical integration. Existing literature predominantly isolates the tool's impact on affective factors such as engagement and anxiety, largely overlooking its efficacy when operationalized within a structured instructional cycle. Specifically, there is a paucity of empirical inquiry into how the asynchronous capabilities of Microsoft Flip—such as planning, recording, and self-monitoring—synergize with the systematic stages of Goh and Burns' (2012) Teaching Speaking Cycle. This study seeks to bridge this gap by moving beyond the tool's inherent features to examine how its strategic integration with a robust framework facilitates the holistic development of speaking skills for Thai EFL learners.

This study addresses that gap by examining not only whether the tool is effective but also how it functions when combined with a framework for developing strategic skills. By integrating the asynchronous capabilities of Microsoft Flip, which allow planning, recording, and re-recording, with a pedagogical cycle that explicitly teaches learners to focus, plan, and reflect, this research proposes a comprehensive solution to the speaking challenges faced by Thai EFL learners.

Research Objectives and Questions

This study seeks to answer the following research questions:

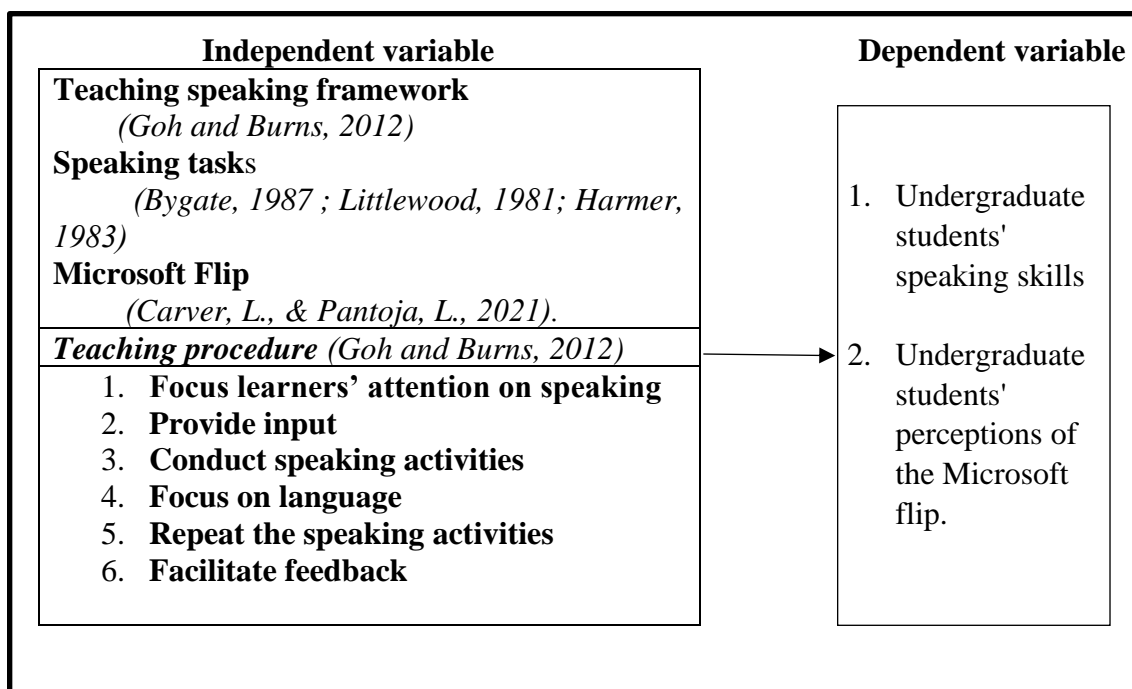
1. Does Microsoft Flip affect the speaking skills scores of undergraduate students?
2. What are the undergraduate students' perceptions of Microsoft Flip?

Conceptual Framework

This research aims to compare undergraduate students' speaking skill scores before and after using Microsoft Flip, and to investigate their perceptions of the tool. The following is the conceptual framework of this study.

Figure 1

Conceptual Framework of the Study



Literature Review

1. Theories in Second Language Acquisition (SLA)

To understand the mechanisms through which Microsoft Flip enhances speaking skills, this study is grounded in several key Second Language Acquisition (SLA) theories that prioritize interaction, output, and cognitive processing. First, the *Interactionist Theory*, primarily associated with Lev Vygotsky (1978), emphasizes the social nature of learning and the concept of the Zone of Proximal Development (ZPD). According to this theory, learners achieve higher levels of proficiency when more knowledgeable others scaffold them in a social context. In the specific context of this research, the Microsoft Flip platform functions as a technological scaffold. Although asynchronous, it creates a social space in which peer interaction and teacher feedback help learners navigate their ZPD, bridging the gap between their current ability and their potential for development through guided practice.

Complementing the foundational role of Krashen's (1982) *Comprehensible Input Hypothesis*, which asserts that learners acquire language best when exposed to input that is slightly beyond their current proficiency level ($i+1$). This theory underpins the "input" stage of the instructional framework used in this study, ensuring that, before students are asked to produce language, they are provided with the necessary lexical and grammatical resources.

While the Input Hypothesis is vital, this study acknowledges that the social aspect is the *Output Hypothesis* proposed by Swain (1985), which argues that exposure to comprehensible input alone is insufficient for language acquisition. Swain posits that learners must produce "pushed output"-language that is stretched to meet communicative goals-to acquire the target language truly. Producing language forces learners to move from semantic processing, where they merely understand meaning, to syntactic processing, where they must construct grammar. The intervention in this study operationalizes this hypothesis through Microsoft Flip's video recording feature. The act of recording compels students to articulate their thoughts, thereby forcing them to notice gaps in their interlanguage and to test hypotheses about language structure in ways that passive listening does not.

Finally, the research draws upon *Task-Based Language Teaching (TBLT)*, as articulated by Ellis (2003). TBLT prioritizes the use of language as a tool to achieve non-linguistic outcomes, such as solving a problem or telling a story. By integrating specific speaking tasks within the Microsoft Flip environment, this study shifts the pedagogical focus from rote memorization of forms to meaningful, authentic communication.

2. Speaking Skills: Definition and Components

Oral communication has long been recognized as a fundamental component of second language acquisition, with speaking skills often regarded as the most difficult of the four macro-skills to master (Bygate, 1987; Goh & Burns, 2012). Speaking is a dynamic process of constructing meaning that involves producing, receiving, and processing information (Brown, 1994; Burns & Joyce, 1997). It is an interactive, two-way process that requires the speaker to manage multiple linguistic and cognitive demands simultaneously. To operationalize speaking skills for this study, five core components are evaluated. First, fluency is the ability to speak smoothly and coherently without undue hesitation (Brown, 1994). Second, accuracy (grammar) refers to the correct use of sentence structure and word order to convey precise meaning (Thornbury, 1999). Third, vocabulary encompasses the range of lexical items a speaker can retrieve and use effectively (Richards & Renandya, 2002). Fourth, pronunciation involves producing clear, intelligible speech sounds (Harris, 1974). Finally, comprehension is the ability to understand the interlocutor and respond appropriately, facilitating interaction (Harris, 1974). In conclusion, the components of speaking-vocabulary, grammar, pronunciation, fluency, and comprehension-are interconnected and together define oral proficiency in a second language. Each element plays a specific yet related role in enabling effective communication: vocabulary enhances expressiveness, grammar maintains structural accuracy, pronunciation ensures clarity, fluency supports smooth speech, and comprehension enables meaningful interaction. For the purpose of this study, comprehension was explicitly evaluated through a face-to-face speaking test in which students were required to process and respond to live questions from an

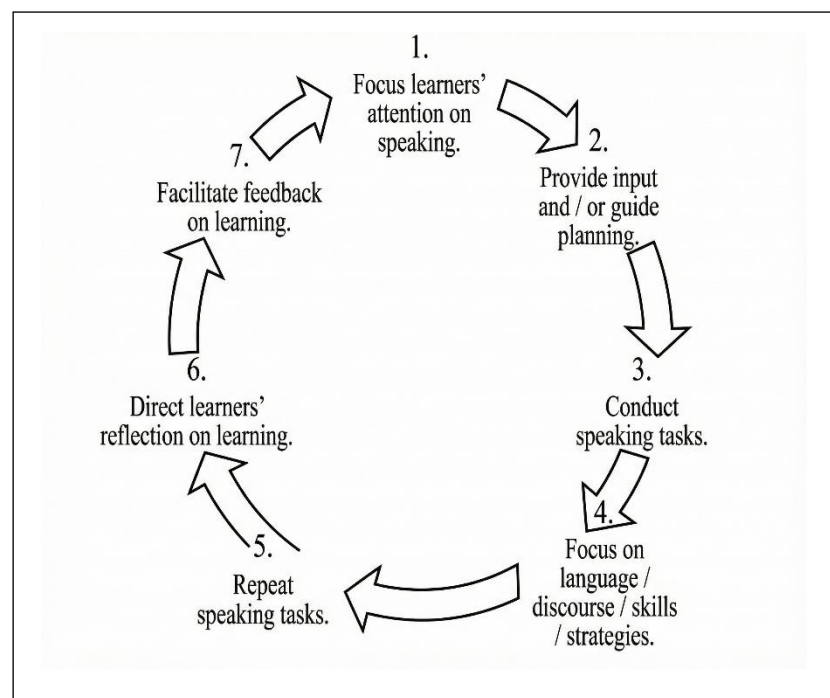
interlocutor, ensuring that their ability to interpret and respond to input was measured alongside their production skills. Addressing these components holistically is essential for developing skilled and confident speakers.

3. The Teaching-Speaking Framework (Goh & Burns, 2012)

Traditional methods such as Audio-Lingualism (drills) or PPP (Present-Practice-Produce) often fail to develop the strategies needed for autonomous speaking. To address this, Goh and Burns (2012) proposed a holistic seven-stage cycle that integrates cognitive and metacognitive instruction. The cycle begins by focusing learners' attention to raise awareness of the task and required skills, followed by providing input to supply necessary vocabulary and grammar. The third stage involves conducting speaking tasks in which students engage in the actual communicative activity, after which instruction shifts to focus on language, offering detailed guidance on linguistic form. A crucial component of this framework is the fifth stage, repeating speaking tasks, which develops fluency and accuracy through iteration. The cycle concludes with facilitating feedback to provide constructive critique and reflecting on learning to encourage self-assessment and planning for future improvement. This study adapts strategically, incorporating Microsoft Flip into a modified six-stage teaching model to facilitate the "Repeat" and "Reflect" stages. The platform's re-recording feature allows students to rehearse and self-monitor before submitting their final output.

Figure 2

Teaching Speaking Framework Cycle



Goh and Burns (2012: 151–152)

4. Speaking tasks

The historical trajectory of speaking tasks reveals a clear progression from isolated methodologies to integrated pedagogical models. Early scholarship by Harmer (1983) established a critical distinction between accuracy-focused practice and fluency-driven activities, such as role-plays. Bygate (1995) expanded this scope by applying a psycholinguistic framework, positing that tasks such as information-gap tasks are essential for developing the cognitive automaticity required for spontaneous speech. This emphasis on authentic interaction was consolidated within Task-Based Language Teaching (TBLT), with Nunan (2003) defining tasks by their non-linguistic outcomes and Ellis (2003) proposing a three-phase structure to enable "focus-on-form".+1

These theoretical developments are effectively synthesized in Goh and Burns' (2012) seven-stage teaching cycle. Rather than displacing earlier methodologies, their framework organizes them: fluency activities are implemented in "Stage 3," accuracy is addressed reactively in "Stage 4," and the critical role of task repetition is formalized in "Stage 5." Consequently, the Goh and Burns model serves as a robust structure that unifies fluency and accuracy instruction within a cohesive lesson cycle. Anchoring these tasks within specific stages of a teaching speaking framework ensures that speaking tasks are not isolated events but are instead aligned to scaffold the learner's progression from cognitive planning to autonomous production.

5. Web-Based Instruction and Microsoft Flip

Web-Based Instruction (WBI) uses internet technologies to create meaningful learning environments (Khan, 1997). Microsoft Flip (formerly Flipgrid) is a WBI tool designed for asynchronous video discussion. It offers a social media-like interface with video filters, emojis, and "likes," which appeal to digital-native students. Key features include the "Discovery Library" for resources, "Mixtapes" for compiling videos, and the "Immersive Reader" for accessibility. Research indicates that Flipgrid reduces anxiety by allowing students to record in private settings (Miller, 2021). It promotes "pushed output" and increases willingness to communicate (Carver, 2021; Chaisiri, 2023). Previous studies in EFL contexts (Pornwasanying, 2023; Robillos, 2023; Damayanti & Citraningrum, 2021) have consistently found improvements in fluency and confidence. However, the asynchronous nature limits real-time negotiation of meaning, a key component of interactional competence. Additionally, technical barriers and the pressure of being recorded can induce performance anxiety in some learners (Rinfret, 2023; Hammett, 2021).

Research Methodology

This study used a quantitative research approach to assess improvements in Nursing undergraduate students' speaking skills through Microsoft Flip, which supported their speaking. Quantitative data were collected from speaking tests administered before and after the Microsoft Flip intervention, as well as from a questionnaire given after the intervention to gauge students' perceptions.

Population and Sample

The population consisted of 110 first-year undergraduate students enrolled in the "English for Communication" course (GED 1101) at Roi Et Rajabhat University during the first semester of 2025. The sample was selected using cluster random sampling, with one intact classroom serving as the sampling unit. The final sample comprised 53 first-year nursing undergraduate students.

Demographically, the participants were aged 18-20 years and predominantly from rural backgrounds in Thailand. Prior to the intervention, their English proficiency was assessed at the A2 level (Basic Users) on the Common European Framework of Reference for Languages (CEFR), a level notably below the B2 target set by the Ministry of Higher Education, Science, Research, and Innovation (MHESI).

Research Design

This study employed a one-group pre-test/post-test pre-experimental design, a design commonly used in educational settings to evaluate changes in a group following an intervention (Campbell & Stanley, 1963). The design is represented as:

$$O1 \rightarrow X \rightarrow O2$$

O1 is the pre-test, X is the 8-week Microsoft Flip intervention, and O2 is the post-test. While this design has limitations regarding internal validity, it is practical for exploratory research within intact classroom settings.

Research Instruments

To gather comprehensive data, two primary research instruments were meticulously developed and validated for this study.

1. Speaking Test

To assess speaking skills in alignment with the first research objective, a standardized speaking test was constructed based on the theoretical frameworks of Bailey (2005) and Bachman and Palmer (1996). This assessment was structured into three distinct segments designed to elicit different aspects of speaking skills: a one-minute self-introduction to gauge basic information exchange; an interview regarding familiar topics such as family and hobbies lasting four to five minutes; and a structured monologue based on a specific prompt card, covering themes like their life experiences, memories, and opinions about the community, services, and cultures, which were designed according to the lessons, which required five to ten minutes of sustained speech. Performance was evaluated using an analytic scoring rubric that rated vocabulary, grammar, fluency, pronunciation, and comprehension on a six-band scale. The test's content validity was verified by three experts, yielding an Index of Item-Objective Congruence (IOC) between 0.67 and 1.00. In contrast, inter-rater reliability between the researcher and an external expert proved excellent, with an Intraclass Correlation Coefficient (ICC) of 0.98.

2. Perception Questionnaire

Complementing the proficiency test, a perception questionnaire was tailored to address the second research objective by assessing students' attitudes toward the Microsoft Flip platform. Adapted from the works of Stepp-Greany (2002) and Soureshjani and Riahipour (2012), the tool was divided into two sections: demographic data collection and a thirteen-item assessment of student perceptions. The perception items used a five-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree". They were categorized into three dimensions: content and lesson design, the benefits of using Microsoft Flip, and students' perceptions of usage. The instrument demonstrated high internal consistency during pilot testing, achieving a Cronbach's Alpha coefficient of 0.85, indicating high reliability.

Research Procedure

The study was conducted over 10 weeks, divided into two phases.

Phase 1: Preparation The researcher analyzed the course syllabus (GED 1101) and mapped it to CEFR standards. A needs analysis survey (N=15) was conducted to identify the 14 topics in language learning offered by the Council of Europe (2022) that were most relevant to the English for Communication course. The top eight topics selected by descriptive statistics of need analysis, as chose were: Business, Friends, Future, Education, Stories, People, Culture, and Services. The eight-week lessons were designed with an integrated teaching-speaking framework using the Microsoft Flip intervention.

Phase 2: Implementation (The Intervention) The intervention followed the adapted 7-stage Goh and Burns framework, integrated with Microsoft Flip tasks.

Week 1: Orientation to Microsoft Flip and administration of the **Pre-test**.

Weeks 2-9: Implementation of the 8 instructional units. Each unit (120 minutes). The 120-minute duration refers to the onsite (face-to-face) instructional sessions held once per week. During these sessions, the researcher facilitated Steps 1 (Focus), 2 (Input/Planning), and 4 (Language Focus). The online components-Steps 3 (Conducting Speaking Tasks), 5 (Repetition), and 6 (Feedback)-were designed as asynchronous tasks completed by students outside of class hours via the Microsoft Flip platform. This structure allowed class time to be dedicated to scaffolding and linguistic preparation, while speaking practice and peer interaction occurred independently.

1. Focus Attention (On-site): Warm-up videos and quizzes to engage interest.

2. Provide Input (On-site): Explicit teaching of vocabulary (e.g., "Entrepreneur," "Stakeholder" for the Business unit) and grammar structures (e.g., Future tense with "will/going to").

3. Conduct Speaking Tasks (Online/Asynchronous): Students recorded videos on Microsoft Flip. Tasks included:

Describing: Describing personal information (Unit 1).

Sharing: Sharing opinions on idols (Unit 2).

Presentation: Presenting local business ideas (Unit 3).

Information Gap: Comparing education systems (Unit 4)

Problem Solving: Planning future goals (Unit 5).

Spot the Difference: Discussing cultural customs (Unit6).

Storytelling: Narrating folk tales (Unit 7).

Role-play: Simulating service encounters (Unit 8).

The tasks were categorized into two distinct types based on the interactional demands:

A. *Monologic Tasks* (Units 1, 2, 3, 5, and 7): These tasks required students to produce sustained, individual speech on a given prompt. Examples include Describing personal information (Unit 1), Presenting business ideas (Unit 3), and Storytelling (Unit 7). These tasks were designed to foster fluency, complexity, and vocabulary retrieval by allowing students to plan and record their speech without the immediate pressure of real-time turn-taking.

B. *Dialogic/Interactive Tasks* (Units 4, 6, and 8): These tasks simulated social interaction and required students to engage with peer input. Examples include Information Gap activities (Unit 4) and Role-plays (Unit 8). In these units, students were required to function as interlocutors by viewing peer videos and recording appropriate responses, thereby targeting comprehension and interactional competence.

The progression of these tasks was designed with specific task design variables in mind: moving from familiar to abstract topics (e.g., from 'Friends' to 'Cultural Customs') and from structured to open-ended responses, ensuring a scaffolded development of oral proficiency.

4. Focus on Language (On-site): Reviewing common errors and teaching communication strategies (e.g., circumlocution, stalling).

5. Repeat Task (Online): Students re-recorded their videos after receiving feedback, utilizing the "rehearsal" strategy.

6. Facilitate Feedback (Online): Peer feedback and teacher evaluation were provided via the platform's comment section.

Week 10: Administration of the **Post-test** and the **a questionnaire about the perception**.

Research Findings

Effect on Speaking Skills (Research Objective 1)

The first objective was to compare undergraduate students' speaking skill scores before and after the Microsoft Flip intervention. The results of the paired-samples t-test are presented in Table 1

Table 1

Overall Scores from the Pre- and Post-speaking Tests (n=53)

Test	N	Full Score	Mean (M)	S.D.	Mean Diff	t-value	Sig. (2-tailed)
Pre-test	53	30	10.23	3.33			
Post-test	53	30	16.56	3.64	6.33	30.28	.001*

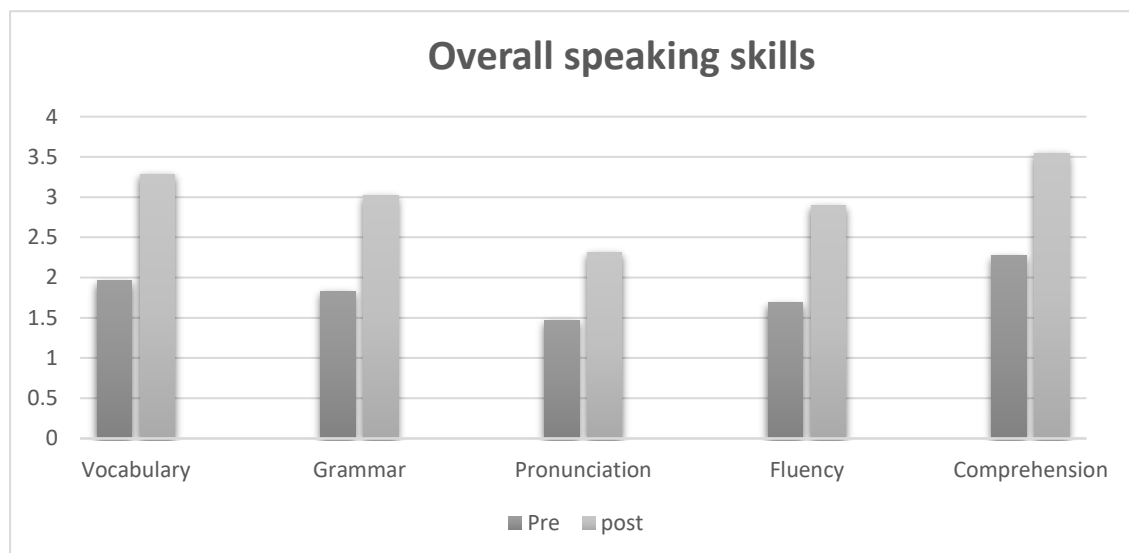
Note: Significance level $p < .05$

As shown in Table 1, the mean score for speaking skills increased significantly from the pre-test ($M = 10.23$, $SD = 3.33$) to the post-test ($M = 16.56$, $SD = 3.64$). The paired-samples t-test analysis yielded ($t(52) = 30.28$, $p < .001$), a p-value of .000, which is less than the significance level of 0.05. This indicates a statistically significant improvement in the students' English-speaking proficiency following the intervention. The mean difference of 6.33 points represents a substantial performance gain. Additionally, Cohen's d of 4.16 indicates that Microsoft Flip had a considerable effect on the students' speaking abilities.

Sub-skill Analysis: Further analysis examined improvements across the five specific components of speaking.

Figure 3

Comparison of Pre- and Post-test Speaking Skills in Each Sub-skills



The results for the five sub-skills are presented below, organized by the magnitude of the mean score increase. Vocabulary showed the highest nominal growth across all components, rising from a pre-test mean of 1.96 to a post-test mean of 3.28. This represents a gain of 1.32 points (67.31%), indicating that the asynchronous planning time allowed students to effectively retrieve and use diverse lexical items. Comprehension followed with a gain of 1.28 points (56.02%), maintaining its position as the strongest overall skill with a post-test mean of 3.55. Fluency showed the most dramatic relative improvement; the point gain of 1.21 marked a substantial 71.51% increase over the baseline score of 1.69, indicating that the video-recording tasks significantly reduced hesitation. Grammar showed a consistent upward trend, with a gain of 1.19 points (64.95%). Finally, Pronunciation recorded the smallest point gain of 0.84, rising from 1.47 to 2.31. While this still reflects a 57.06% improvement, the lower magnitude compared to other skills suggests that phonological accuracy may require more intensive synchronous feedback to match the gains seen in fluency and vocabulary.

Student Perceptions (Research Objective 2)

The second objective investigated the students' perceptions of using Microsoft Flip. The data from the 5-point Likert scale questionnaire (where 5 = Strongly Agree) revealed generally positive attitudes. The overall mean score for perception was 4.19 (SD = 0.94), interpreted as a "High" level of positive perception. The results are categorized into three key dimensions:

Content and Lesson Design (M = 4.19). Students responded positively to the lesson structure. The highest-rated item in this category was "The instructions in the lesson are clear" (M = 4.40), indicating that the combination of on-site instruction and online tasks was easy to follow. Students also agreed that "The lesson has clearly defined learning objectives" (M = 4.30). The lowest-rated item in this section was "The exercises in the lesson are interesting" (M = 3.91). While still positive, this suggests that task design could be further diversified to maintain high engagement over 8 weeks.

The Benefits of Using Microsoft Flip (M = 4.36). This was the highest-rated dimension, reflecting strong student endorsement of the tool's utility. The highest-rated item in the entire survey was "Microsoft Flip can be further implemented in real classroom settings" (M = 4.45). This indicates a strong acceptance of the technology and a desire for its continued use. Students also firmly believed that 'Using Microsoft Flip in Enhancing English Speaking Skills' was effective (M = 4.36).

Student Perception of Usage (M = 4.06). This dimension explored the affective and autonomous aspects of learning. Students reported that "The lesson promotes independent learning" (M = 4.30). This is a critical finding, suggesting that the platform's asynchronous nature successfully shifted the locus of control to the learner. However, the item 'The lesson caters to students' varying learning abilities' received the lowest rating in the survey (M = 3.70, SD = 1.15). This lower mean score suggests that, while the intervention was viewed positively overall, participants perceived the platform as less effective at accommodating individual learning differences than its other benefits. This finding highlights a distinction between the tool's general utility and its perceived adaptability, indicating that differentiation remains a critical area for pedagogical refinement.

Discussion

Enhancement of speaking skills

Research Question 1: Does Microsoft Flip affect the speaking skills scores of undergraduate students?

The quantitative data reveal a robust and substantial gain in undergraduate speaking proficiency following the Microsoft Flip intervention. The overall mean score rose from 10.23 to 16.56 ($t(52) = 30.28, p < .000$), with a large effect size (Cohen's $d = 4.16$). These findings corroborate recent computer-assisted language learning (CALL) research in the Thai EFL context, specifically aligning with Pornwasanying (2023), who observed similar advancements in learner confidence and speaking performance, and with Robillos (2023), whose work confirmed that technology-mediated metacognitive tasks significantly improve delivery and clarity. However, the distinct hierarchy of gains observed in the present study—where vocabulary and comprehension outpaced pronunciation—necessitates a multi-theoretical interpretation that draws on Input, Output, ZPD, and Task-Based Learning perspectives. The most marked improvement was observed in Vocabulary, which exhibited the highest

nominal gain of 1.32 points ($t(52) = 30.28, p < .000$). This finding is best understood through Swain's (1985) Output Hypothesis. The platform's asynchronous nature facilitated a "pushed output" environment in which learners, upon reviewing their recordings, engaged in self-monitoring and hypothesis testing. This recursive process of recording and re-recording enabled the retrieval and integration of more sophisticated lexical items than are typically available in spontaneous speech, a phenomenon also noted by Robillos (2023) regarding the benefits of planned asynchronous tasks. Furthermore, Comprehension remained the strongest component overall ($M = 3.55$), showing a significant increase ($t(52) = 30.28, p < .000$). This development supports Vygotsky's Zone of Proximal Development (ZPD), as the peer feedback stage (Stage 6) required students to critically evaluate interlocutor output, thereby reinforcing receptive skills through social scaffolding. Concurrently, the gains in Fluency (+1.21 points) and Grammar (+1.19 points) reflect the efficacy of Task-Based Language Teaching (TBLT) principles. By engaging in meaning-focused tasks such as role-plays and business presentations, students were driven by non-linguistic goals, which, combined with the Input Hypothesis ($i+1$) operationalized in the planning stages, reduced cognitive load and hesitation. Conversely, Pronunciation demonstrated the lowest gain (+0.84 points). While statistically significant, this disparity suggests a pedagogical divergence: while asynchronous video tools effectively cultivate lexical complexity and fluency through planning and repetition, phonological accuracy may require immediate, synchronous corrective feedback to achieve comparable development. Thus, while the intervention fostered holistic growth, the results indicate that asynchronous platforms are particularly well-suited for developing the cognitive and lexical dimensions of speaking, whereas phonological refinement may require supplementary synchronous instruction.

Research Question 2: What are undergraduate students' perceptions of Microsoft Flip?

The quantitative findings from the questionnaire indicate that students perceive Microsoft Flip as a highly beneficial tool for language learning. The high mean score for the domain "The Benefits of Using Microsoft Flip" ($M = 4.36$) indicates strong student endorsement of the platform's utility in enhancing their speaking abilities. This positive reception is further supported by high ratings for "Independent Learning" ($M = 4.30$), corroborating the Interactionist Theory. Although the interaction was asynchronous, the presence of a "digital audience" (peers and teacher) provided the necessary social pressure to perform, while the ability to edit provided psychological safety to take risks.

This aligns with Chaisiri's (2023) findings on Willingness to Communicate (WTC). Thai students, often culturally reticent to speak in public due to "face" concerns, found the video recording's private nature liberating. They could present their "best self," which boosted confidence. However, the finding regarding differentiation ($M = 3.70$) serves as a critical caveat. As noted by Rinfret (2023), technology is not a panacea; the lower score suggests that the platform's implementation did not fully account for varying levels of digital literacy or anxiety. For students with lower proficiency, the rigid requirement to produce video content-without sufficient scaffolding or alternative modes such as audio-only submissions-likely turned the task into a source of stress rather than a learning opportunity. This indicates that, while the tool is effective, future implementation strategies must include tiered technical support

and flexible participation options to prevent anxiety from inhibiting learning among vulnerable students.

Conclusion

This study concludes that the integration of Microsoft Flip within a structured, adapted teaching-speaking cycle is a highly effective pedagogical strategy for the Thai EFL context. It successfully addresses the 'mute English' problem by providing a scaffolded, low-anxiety environment that compels active language production. The intervention yielded a statistically significant improvement in overall speaking proficiency ($t(52) = 30.28, p < .000$), with consistent gains observed across all five sub-skills, particularly in vocabulary acquisition. Furthermore, findings from the perception questionnaire substantiate the tool's broader educational value; students reported high levels of agreement that the platform fostered learner autonomy ($M = 4.30$) and equipped them with skills applicable to real-world career contexts ($M = 4.06$).

Recommendations

Pedagogical Implications

1. *Strategy is Key*: Technology should not be used in isolation. The success of this study depended on the *framework* (Goh & Burns) rather than just the *app*. Teachers must explicitly teach strategies like planning and self-monitoring.

2. *The "Rehearsal"*: Educators should legitimize "rehearsal" as a learning activity. Allowing students to re-record assignments is not "cheating"; it is a valuable learning process of error correction.

3. *Hybrid Models*: The study suggests a blended approach is best. "Input" and "Language Focus" (grammar/pronunciation correction) may be best delivered on-site, while "Production" and "Repetition" are effectively offloaded to the asynchronous platform.

Limitations

- *Design*: First, the study employed a one-group pre-test/post-test design, which inherently limits the internal validity of the findings. Without a control group, it is difficult to attribute the observed improvements in speaking skills exclusively to the Microsoft Flip intervention. Threats to validity, such as maturation effects (students naturally improving over the eight-week period) or history effects (external exposure to English outside the classroom), cannot be entirely ruled out. Consequently, while the significant gains in test scores ($t(52) = 30.28, p < .000$) strongly suggest the intervention's efficacy, causal claims regarding the specific impact of the technology versus general instruction must be interpreted with caution the one-group pre-test/post-test design limits the ability to attribute causality strictly to the intervention, as maturation effects cannot be entirely ruled out.

- *Sample*: The study was limited to nursing students at a single university, limiting generalizability to other disciplines or proficiency levels.

- *Duration*: The 8-week period was sufficient for short-term gains, but long-term retention of these skills remains untested.

- *One size fits all*: Additionally, while the quantitative results indicated a lower perception of the platform's ability to cater to varying learning abilities ($M=3.70$), this study did not explicitly measure variables such as digital literacy levels

or communication apprehension in relation to this specific item. Therefore, it is possible that a 'one-size-fits-all' approach may not fully address the diverse needs of mixed-ability classes, particularly for students with higher anxiety, but further research is required to confirm this correlation

Recommendation For Future Studies

1. *Experimental Design*: Future studies should employ a quasi-experimental design with a control group (receiving traditional face-to-face speaking practice) to compare the efficacy of the asynchronous video-mediated framework against conventional methods. Rather than attempting to isolate the technology itself, this design would evaluate the comparative effectiveness of the technology-enhanced pedagogical model, acknowledging that the tool's impact is intrinsically linked to its instructional implementation.

2. *Longitudinal Analysis*: Research tracking students over a full academic year would provide insights into the sustainability of the motivation and skill gains.

3. *Task Design Variables*: While this study employed a scaffolded task design that integrated multiple variables-specifically, task complexity (moving from concrete to abstract topics) and interactional mode (alternating between monologic and dialogic tasks)-the specific impact of each variable on speaking performance remains aggregated. Future research should therefore aim to isolate these task design variables to determine their individual efficacy. For instance, experimental studies could compare the effects of monologic tasks (e.g., storytelling, presentations) versus dialogic tasks (e.g., information gaps, role-plays) on distinct speaking components such as fluency and accuracy. Additionally, investigating the role of topic familiarity could reveal how moving from personal to academic subjects within asynchronous video platforms specifically influences learner anxiety and willingness to communicate.

4. *Differentiation Strategies*: Research is needed into how to adapt WBI platforms for students with learning disabilities or severe communication apprehension, ensuring that technology bridges gaps rather than widening them.

References

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Bailey, K. M. (2005). *Practical English language teaching: Speaking*. McGraw-Hill.
- Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy*. Prentice Hall Regents.
- Burns, A., & Joyce, H. (1997). *Focus on speaking*. National Centre for English Language Teaching and Research.
- Bygate, M. (1995). *Speaking, language teaching, a scheme for teacher education*. (Eds. C.N. Candlin & H. G. Widdowson), Oxford University Press.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.
- Carver, L., & Pantoja, L. (2021). Visual and auditory production. In L. Carver & H. S. Atkins (eds.). *Leading in the Digital Environment: Being a Change Agent* (pg. 97-110). Lanham, MD: Rowman & Littlefield.

- Chaisiri, P. (2023). Promoting Thai EFL learners' willingness to communicate in the virtual classroom through technology-mediated oral tasks. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 97-120.
- Copland, F., Garton, S., & Barnett, C. (2024). *Global practices in teaching English to young learners: Ten years on*. British Council.
<https://doi.org/10.57884/JHCP-DS26>
- Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.
- Damayanti, I. L., & Citraningrum, E. (2021, April). Flipgrid: A pathway to enhance students' speaking performance. In Thirteenth Conference on Applied Linguistics (CONAPLIN 2020) (pp. 90-95). Atlantis Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Council of Europe. (2022). The importance of plurilingual and intercultural education for democratic culture, Recommendation CM/Rec(2022)1 and explanatory memorandum. [Document/Report number or series, if available].
https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://ealta.eu/documents/resources/Threshold-Level_CUP.pdf&ved=2ahUKEwi9tfvHh8mQAxUDZvUHHXQ-CHMQFnoECBoQAAQ&usg=AOvVaw3SJQmVbA-guKWjeC8i6iWM
- Gilakjani, A. (2017). A review of the literature on the integration of technology into the learning and teaching of English language skills. *International Journal of English Linguistics*, 7(5), 95-106. <https://doi.org/10.5539/ijel.v7n5p95>
- Gillis, G. (2013). *The importance of speaking skills*.
<http://www.geraldgillis.com/importance-speaking-skills>
- Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. Cambridge University Press.
- Graddol, D. (2006). *Why global English may mean the end of "English as a Foreign Language"*. British Council.
- Hamilton, S., et al. (2024). Grammar challenges in the Thai EFL context. *Asian EFL Journal*, 28(1), 45-62.
- Hammett, D. (2021). Utilizing flipgrid for speaking activities: A small-scale university-level EFL study. *Theory and Practice in Language Teaching and Learning*, 3(2), 51-64. <https://doi.org/10.29140/tl.v3n2.509>
- Harmer, J. (1983) *The Practice of English Language Teaching*. Longman.
- Harris, D. P. (1974). *Testing English as a second language*. McGraw-Hill.
- Hung, S.-T., & Huang, D. (2021). Exploring EFL learners' willingness to communicate in multimodal videoconferences. *International Journal of Computer-Assisted Language Learning and Teaching*, 11(4), 79-99.
<https://doi.org/10.4018/IJCALLT.2021100106>
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.
- Khan, B. H. (1997). *Web-based instruction*. Educational Technology Publications.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.
- Littlewood, W. (1981). *Communicative language teaching*. Cambridge University Press.

- Miller, J. S. (2021). FlipGrid [Review of the software Flipgrid]. *The French Review*, 95(1), 286.
- Nithideechaiwarachok, B., Maneekanon, O., & Bubphada, T. (2022). Exploring English language proficiency, English language problems, and English needs among first-year undergraduate students. *International Journal of Learning, Teaching and Educational Research*, 21(12), 273-290.
<https://doi.org/10.26803/ijlter.21.12.15>
- Noom-ura, S. (2013). English-teaching problems in Thailand and Thai teachers' professional development needs. *English Language Teaching*, 6(11), 139-147.
<https://doi.org/10.5539/elt.v6n11p139>
- Parvin, R. H., & Salam, S. F. (2015). The effectiveness of using technology in English language classrooms in government primary schools in Bangladesh. *FIRE: Forum for International Research in Education*, 2(1), 47-59.
<https://doi.org/10.18275/fire201502011049>
- Pornwasanying, K. (2023). Using Flipgrid to enhance students' speaking performance in listening-speaking course. *Journal of Humanities and Social Sciences for Sustainable Development*, 6(2), 114-127.
- Richards, J. C. (2008). Second Language Teacher Education Today. *RELC Journal*, 39, 158-177. <http://dx.doi.org/10.1177/0033688208092182>
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511667190>
- Rinfret, S. R., & Forster, E. (2023). FlipGrid: Re-imagining online discussion forums for the public administration classroom. *Journal of Public Affairs Education*, 29(1), 92-106. <https://doi.org/10.1080/15236803.2022.2063632>
- Robillos, R. J. (2023). Improving students' speaking performance and communication engagement through technology-mediated pedagogical approach. *International Journal of Instruction*, 16(1), 551-572.
<https://doi.org/10.29333/iji.2023.16131a>
- Shen, M., & Chiu, T. (2019). *EFL Learners' English speaking difficulties and strategy use*. Education and Linguistics Research.
- Siddoo, V., Sawattawee, J., Janchai, W., & Thinnukool, O. (2019). An Exploratory Study of Digital Workforce Competency in Thailand. *Heliyon*, 5(5), 1723.
<https://doi.org/10.1016/j.heliyon.2019.e01723>
- Soureshjani, K. H., & Riahipour, P. (2012). Demotivating factors on English speaking skill: A study of EFL language learners and teachers' attitudes. *World Applied Sciences Journal*, 17(3), 327-339.
- Stepp-Greany, J. (2002). Student perceptions on language learning in a technological environment: Implications for the new millennium. *Language Learning & Technology*, 6(1), 165-180.
- Sudarmo, S. (2021). The importance of speaking in English as a foreign language between skillful and thoughtful competencies: Studying sociolinguistics perspectives. *Linguistics and Culture Review*, 5(S1), 113-124.
<https://doi.org/10.21744/lingcure.v5nS1.1321>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House.

- Thadphoothon, J. (2017). Thailand 4.0 and English language teaching: Impact and implications. *International Journal of Languages, Literature and Linguistics*, 3(1), 1-5.
- Thornbury, S. (1999). *How to teach grammar*. Pearson Education.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. In M. Cole, V. John-Steiner, S. Scribner and E. Souberman (Eds.), Cambridge, MA: Harvard University Press.
- Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal*, 37, 308-328. <http://dx.doi.org/10.1177/0033688206071315>

Author

Santhoshkanth Kamaraj is a lecturer in the Business English Program at the Faculty of Liberal Arts and Sciences, Roi Et Rajabhat University. He holds a Bachelor's in Computer Science Engineering from Annamalai University, India. He graduated with a diploma in English Language Teaching from the Philippine Christian University, Bangkok Campus. He is pursuing a Master of Arts in English Language Teaching at Roi Et Rajabhat University. His research interests focus on English for Specific Purposes, Teaching English as a Second or Foreign Language, Technology-Aided Teaching, and innovative materials and learning technologies in English Language Teaching.