PASAA Journal
Volume 72, January – June 2026, 47–69

A Competence-Based, Technology-Mediated Speaking Intervention in Vietnamese EFL Classrooms: A Comparative Study of Collaborative and Expert-Novice Dyads

Tran Anh Vu[a, b, *] and Kwanjira Chatpunnarangsee[b]

[a] Thu Dau Mot University, Ho Chi Minh, Vietnam
[b] Thammasat University, Bangkok, Thailand
**Corresponding author: trananhvu@tdmu.edu.vn*

Article information	
Abstract	This study investigates speaking development during participation in a speaking instructional model combining Competence-Based Language Teaching (CBLT), dyadic interaction, and mobile social networking site (Zalo)-based speaking practice for secondary EFL learners in Vietnam, where students have few opportunities for authentic oral communication despite years of studying English. Two classes of seventy-two Grade 8 learners (36 in each class) participated in a 12-week program. Both classes followed the same CBLT and Zalo-based speaking curriculum but worked in different dyads: collaborative pairs or expert–novice pairs with English-major university buddies. Data included pre- and post-speaking tests, focus group interviews, and weekly Zalo speaking recordings. Quantitative results from pre- and post-tests scores suggested that the model was associated with significant improvements in grammar, vocabulary, pronunciation, and interactive communication for both dyads, with expert–novice dyads showing larger gains in accuracy-related areas. Qualitative data, analyzed via thematic analysis, showed that the three components operated as an interdependent ecology: CBLT provided the pedagogical structure, dyadic interaction functioned as the mechanism for linguistic development, and Zalo served as the low-pressure delivery environment. These findings offer promising evidence regarding the practicality of this instructional model for supporting speaking development in resource-limited contexts.
Keywords	expert-novice dyad, collaborative dyad, CBLT, mobile social networking site, speaking performance
APA citation:	Tran, A. V., & Chatpunnarangsee, K. (2026). A competence-based, technology-mediated speaking intervention in Vietnamese EFL classrooms: A comparative study of collaborative and expert–novice dyads. *PASAA Journal, 72*, 47–69.

E-ISSN: 2287-0024

1. Introduction

EFL learners in countries where English is used as a foreign language, particularly in Vietnam, have limited opportunities to use spoken English in everyday situations; as a result, improving their speaking skills remains a significant challenge (McDonough, 2004; Yashima, 2002). Despite many years studying English, many EFL learners still struggle to maintain basic conversations, which is mainly because of limited opportunities to practice, the exam-focused teaching approach, and the widely used grammar-translation method (Le, 2011; Nguyen, 2013; Sun & Yang, 2015). As a result, speaking anxiety and a lack of confidence among Vietnamese EFL learners (Bui & Nguyen, 2016; Sundkvist & Nguyen, 2020) contribute to the nation's consistently poor standing in global English proficiency rankings (Education First, 2024).

This situation is worsened by current local assessment and testing practices. While recent governmental reforms promoted communicative competence and CEFR alignment (Thao & Mai, 2022), classroom practices still largely focused on testing because of institutional constraints, and teachers' perceptions (Phan, 2021; Nguyen, 2022). When students are under pressure to do well on high-stakes exams, they tend to focus less on speaking but to pay more attention to reading and vocabulary (Van, 2010; Nguyen, 2013). For this reason, it is observed that authentic EFL speaking interaction remains insufficiently addressed in many educational settings. Despite teachers' theoretical support for the communicative teaching approach, they still focus on the skills that are tested in real exams (Ha et al., 2021).

To address this issue, Competence-Based Language Teaching (CBLT) provides a promising approach for EFL teaching. Instead of focusing on the knowledge of language, CBLT emphasizes what learners can do with the language in meaningful situations (Richards & Rodgers, 2014). Specific CBLT-based activities encourage communication with clear goals, reflection on students' learning, and an increase in their autonomy over time. Previous studies claimed that when CBLT-based activities were supported with structured pedagogical assistance, learners' proficiency and confidence in speaking could be greatly improved (Dooly & Sadler, 2016; Eltanskaya et al., 2019; Pambudi et al., 2024). However, the time for CBLT activities in the classroom alone is usually insufficient to provide enough speaking practice for EFL learners.

At this point, technology becomes important. Mobile social networking sites such as Facebook, Skype, and Vietnam's Zalo create communication platforms that extend beyond the classroom walls. Lin (2012) and Sun and Yang (2015) found that using these platforms for practicing speaking helped learners feel less anxious, or stressed, and more motivated. Additionally, according to sociocultural theories of second language acquisition, these platforms encourage communication and peer support (Ohta, 2001; Swain, 2000).

While mobile platforms create the space for interaction, the way learners actually use that space matters. Dyadic interaction provides one concrete mechanism through which these benefits can emerge. Working in pairs, whether with the same-level peers or in expert–novice pairs, is particularly effective for co-constructing meaning and providing feedback. For example, according to Storch (2002), Chan et al. (2019), and Qureshi et al. (2023), both pairing types can promote learning, though they do so through different forms of scaffolding.

However, despite promising findings, most previous studies have examined CBLT, dyadic interaction, and mobile learning in complete isolation from one another. There is limited empirical research examining what occurs when these elements are intentionally combined into

an integrated instructional model. Additionally, there is limited evidence regarding the differences between collaborative interactions and expert-novice interactions in supporting EFL learners' speaking development.

To address this gap, the present study investigates:

1) To what extent does Vietnamese EFL learners' speaking performance change during participation in an integrated CBLT, dyadic, and Zalo-mediated speaking program?
2) How do different dyad types (collaborative vs. expert–novice) relate to learners' gains in grammar & vocabulary, pronunciation, and interactive communication?
3) How do the EFL learners perceive the new instructional model for developing their English speaking skills?

2. Literature Review

2.1 Sociocultural Perspectives on Peer Interaction

The present study is grounded in Vygotsky's (1978) Sociocultural Theory, particularly the concept of Zone of Proximal Development (ZPD). Within this framework, learning is commonly viewed as a socially mediated process, which is facilitated through scaffolding interaction with more capable peers. The process of scaffolding (Wood et al., 1976) offers assistance that enables learners to complete tasks beyond their current abilities, fostering self-reliance over time. Donato (1994) and Lantolf (2000) extended this perspective to peer-mediated learning, showing that scaffolding could develop through collaborative dialogue rather than through teacher-led instruction only.

Swain (2000) claimed that collaborative dialogue allowed learners to discuss meaning, negotiate language forms, and engage in language reflection, which was important for language development. Based on this idea, empirical research on dyadic interaction has identified specific patterns that support effective L2 production. For example, Storch (2002, 2004) suggested that both collaborative and expert-novice interactions helped learners work through problem-solving and respond to feedback. Later studies also showed that mixed-proficiency dyads led to more meaning negotiation and greater opportunities for speaking practice (Watanabe & Swain, 2007; Tian & Jiang, 2021).

Although strong evidence shows that peer scaffolding supports L2 development, several important gaps in the research remain. Specifically, most research concentrates on university students, investigates only short-term assignments, or takes place in in-person environments. Few have investigated peer scaffolding in lower-secondary classrooms, Asian exam-centric contexts, or mobile-assisted learning environments. Furthermore, there are not many studies that have previously integrated interactional scaffolding with organized cycles of specific tasks, such as CBLT activities. The present study is designed to bridge these gaps by looking closely at collaborative dyad and expert-novice dyad working within a competence-based speaking model, which is mediated through the Zalo application.

2.2 Competence-Based Language Teaching (CBLT)

CBLT focuses on building useful communication skills by using defined competencies, structured learning sequences, and performance-based assessment (Richards & Rodgers, 2014). It promotes progression through tasks that are oriented toward specific goals, promotes practice that is repeated, and provides feedback which collectively assists students to become

independent and master their specific subjects (DeKeyser, 2007). CBLT is quite different from conventional methodologies such as grammar-translation or presentation–practice–production because learning is embedded in contextualized settings that are highly task-oriented, where the accuracy and communicative appropriateness of the student are evaluated simultaneously.

Recent studies show that CBLT works effectively across different situations. Research indicates that CBLT facilitates self-regulated learning, purposeful communication, and enhanced speaking performance, encompassing improvements in fluency, accuracy, and confidence (Dooly & Sadler, 2016; Eltanskaya et al., 2019; Pambudi et al., 2024). It has also been shown that combining CBLT with CLIL can improve communication skills and lower anxiety (López-Medina & Casado, 2024). Digital adaptations of CBLT also show that students are more interested and learn useful skills (Adeyeye et al., 2022; Boahin & Hofman, 2014; Saint, 2021).

Despite these numerous positive findings, there has been limited research focusing on how CBLT works when it is used with peer scaffolding or through mobile-supported interaction. There are also few empirical studies regarding the enhancement of CBLT tasks through dyadic collaboration, particularly expert-novice pairings, and the extension of these tasks beyond the classroom via mobile platforms in secondary school environments. This study directly addresses the existing gap.

2.3 Mobile Social Networking for Language Learning

Mobile-Assisted Language Learning (MALL) uses ubiquitous mobile devices to help with flexible, multimodal content in L2 practice. Platforms such as Zalo, WhatsApp, LINE, and Telegram are all social networking sites that let students talk to each other synchronously and asynchronously, which is less stressful for the students. These kinds of platforms make it easier to speak and provide students more freedom and chances to practice and receive feedback (Akkara et al., 2020; Han & Keskin, 2016; Shih et al., 2015).

Mobile tools not only make things easier, but they also help people learn specific skills. For instance, microblogging has been shown to greatly assist with pronunciation awareness and self-correction, while Telegram voice messages support sustained listening and pronunciation practice, and make complex language features easier for learners to understand (Fouz-González, 2017; Xodabande, 2017). Collaborative mobile tasks on sites like Facebook and Skype have been shown to improve fluency, interactional competence, and student-centered communication (Yen et al., 2015; Avci & Adiguzel, 2017).

Despite the growing body of MALL research, many of these studies focus on university students or use short-term intervention. While human-tool interaction has received significant attention, less focus has been given to organized instructional models, and the investigation into the use of mobile devices for enhancing scaffolding interaction and CBLT activities remains significantly limited. Besides, there is limited research on Zalo, Vietnam's most popular messaging app, though it has the potential to help lower-secondary students practice speaking.

This study aims to fill in these gaps by introducing a new instructional model, in which CBLT provides structured, outcome-based tasks; dyadic interaction offers scaffolded support; and Zalo extends practice opportunities beyond the classroom time.

2.4 Conceptual Framework

Rather than viewing Competence-Based Language Teaching (CBLT), dyadic interaction, and Zalo as separate interventions, this study conceptualizes them as mutually reinforcing layers within a single sociocultural learning ecology. Drawing on Sociocultural Theory (Vygotsky, 1978), the model assumes that learning occurs through mediated activity within the Zone of Proximal Development (ZPD). In this ecology, dyadic scaffolding functions as the primary mechanism of development, while CBLT structures learning activity and Zalo provides the mediational environment that sustains interaction over time (see Figure 1).

2.4.1 Dyadic interaction as the developmental mechanism

The center of the model is dyadic scaffolding, grounded in Sociocultural Theory (Vygotsky, 1978; Lantolf, 2000), which suggests that students are able to co-construct meaning via discussion and negotiation in pairs (see Figure 1). Research claimed that interactional moves, such as asking for clarification, reformulating, and encouraging learners to produce language, could lead to their noticing and immediate responses to feedback, which helped improve their control of language forms and pronunciation (McDonough, 2004; Storch, 2002; Watanabe & Swain, 2007).

In expert–novice pairings, more proficient partners can support learners' performance in the zone of proximal development while maintaining learner autonomy (Ohta, 2001; Swain, 2000). Accordingly, the framework assumes the pathway: scaffolding → noticing / modified output → uptake → accuracy gains.

2.4.2 The CBLT task cycle as pedagogical structure

The Competence-Based Language Teaching (CBLT) task cycle forms the next layer around the central dyadic interaction (see Figure 1). Rather than focusing on isolated language forms, CBLT emphasizes what learners can do with language in meaningful situations (Richards & Rodgers, 2014). Specifically, structured cycles involving goal setting, guided practice, rehearsal, and self-assessment, provide performance criteria and a sense of responsibility.

Several studies showed that participation in CBLT activities could contribute to the enhancement of students' confidence, competence, and reflective awareness (Dooly & Sadler, 2016; Eltanskaya et al., 2019; DeKeyser, 2007; Pambudi et al., 2024). Accordingly, the expected pathway is: goal clarity + self-monitoring → regulated performance → more appropriate and effective interaction.

2.4.3 Zalo as a mediational mobile environment

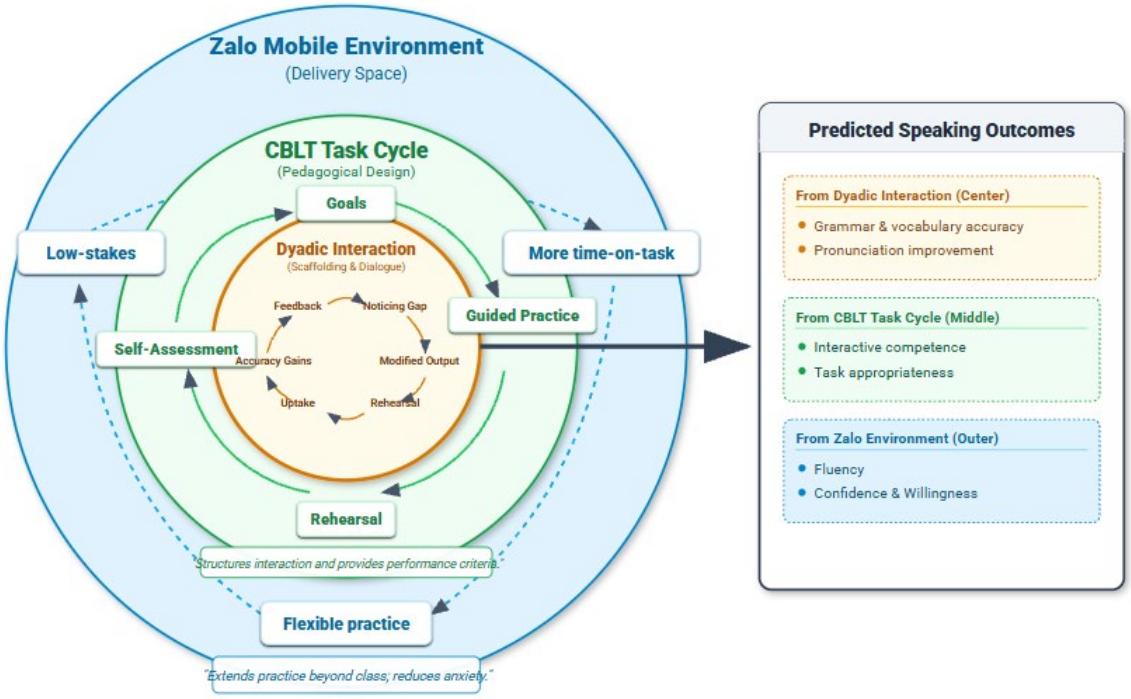
The outer layer of the ecology is the Zalo mobile environment, which mediates practice beyond the classroom (see Figure 1). Previous studies have shown that mobile platforms can reduce anxiety, increase time on task, and provide flexible environments for low-stakes practice (Lin, 2012; Mao, 2014; Sun & Yang, 2015; Han & Keskin, 2016). Since students are permitted to speak privately and asynchronously during practice periods, they usually experience greater comfort when attempting new skills, practicing new pronunciations, and perfecting their overall speech production. This aligns with sociocultural perspectives which suggest that tools can influence the development of involvement throughout time (Peterson,

2012; Ohta, 2001). Thus, the model anticipates the pathway: mobile affordances → reduced anxiety + sustained practice → fluency and willingness to speak.

In general, these theoretically grounded pathways provide a basis for examining not only whether the model works, but also how its components contribute to learners' speaking development.

Figure 1

Conceptual framework: Competence-Dyadic-Zalo-based speaking practice instruction



3. Methodology

3.1 Research Design

This study used a mixed-methods design. Quantitative data were obtained via pre- and post-speaking tests, while qualitative data were collected from semi-structured focus group interviews and students' Zalo speaking recordings. This research design enabled triangulation of test-based performance, observations, and learners' perspectives.

3.2 Setting and Participants

The study was conducted over a 16-week semester in 2023 at an urban public secondary school in Binh Duong Province, Vietnam. The school was selected because of its representative demographics and students' familiarity with mobile technology.

Two intact Grade 8 classes (N = 72) were recruited through convenience sampling. Students' average age was 14.5. After a KET-aligned speaking pretest, the two classes were randomly assigned to one of two dyadic interaction types:

- Collaborative dyads (n = 36): students practiced with same-level classmates.
- Expert–novice dyads (n = 36): students practiced with trained senior English majors ("English buddies") from a public university in Binh Duong province, Vietnam.

In the latter condition, the pairing represented cross-age, near-peer mentoring rather than simple peer interaction. English buddies (n = 36; aged 20–22) were selected from approximately 500 senior English majors based on: (1) GPA ≥ 7.0/10.0, (2) willingness to participate, and (3) a B1-or-higher rating on the Cambridge Speak & Improve test. These English buddies were given guidance on how to provide supportive scaffolding, including techniques such as prompting, asking for clarification, and offering constructive feedback.

In this study, dyadic interaction was not considered an independent variable but rather two forms of learning support within the instructional environment. The collaborative dyad reflected mutual interaction between learners of comparable proficiency, while the expert–novice dyad operated as a structured form of near-peer scaffolding within a shared pedagogical and technological setting.

Because the expert–novice dyad involved older and more proficient partners who had received prior training, it differed from the collaborative dyad in more than just proficiency pairing. These differences may have introduced further influences, including differences in feedback quality and frequency, mentoring approaches, and interactional leadership patterns.

Accordingly, the outcomes are interpreted not as the result of dyad type alone, but rather as reflecting differences in the intensity and frequency of scaffolding, the immediacy of feedback, and the interactional support structures within the integrated instructional context. Process indicators reported in the Results section (such as examples of corrective feedback, instances of prompted elaboration, and relative talk time) are used to illustrate how this scaffolding operated in practice.

Both classes were taught by the same English teacher, who has seven years of experience using technology in education, to ensure instructional consistency. Ethical approval was secured, and informed consent was obtained from students and parents.

3.3 Instructional Procedure

Following one week of pretesting and orientation, students completed a 12-week CBLT-based speaking program (two hours per week) across six Grade 8 topics: family and friends, school, pets, clothing, restaurants, and shopping. Instruction followed a 10-step CBLT cycle, which involved goal-setting, activation of background knowledge, targeted vocabulary learning, watching/listening to model dialogues, comprehension checks, vocabulary application, language form application, pronunciation practice, modified role plays, and self-assessment.

To ensure systematic and replicable implementation, the speaking intervention followed a two-week online CBLT cycle delivered entirely through Zalo. Each topic was divided into ten sequenced steps that moved learners from comprehension to controlled practice and finally to “pushed output” and self-reflection. Activities were completed in dyads, which included pairs meeting synchronously via Zalo voice or video calls and submitting weekly recordings. Table 1 summarizes the timing, content focus, and sample activities for one instructional topic (Family and Friends), which illustrates how the 10-step sequence was operationalized.

Table 1

Two-Week Online CBLT Cycle Delivered via Zalo (Example Topic: Family and Friends)

Week	Step	Time	Content Focus	Activities (Dyadic, via Zalo)
Week 1	1.Learning targets	10 min	Clarify communicative objectives	Partners read aloud and discuss objectives; identify what they expect to practice.
Week 1	2.Activate background knowledge	10 min	Personal experiences with family & friends	Partners ask/answer guiding questions about families and friendships.
Week 1	3.Key vocabulary & expressions	10 min	Vocabulary and useful phrases for the topic	Practice target words/phrases; create short sample sentences related to the topic.
Week 1	4.Model conversation	15 min	Listening/viewing for meaning	Watch/listen to sample dialogue; discuss context and main ideas of the model dialogue.
Week 1	5.Comprehension check	15 min	Understanding details and main ideas of the model dialogue	Take turns asking and answering prepared comprehension questions about the model dialogue.
Week 2	6.Vocabulary in use	10 min	Applying vocabulary in sentences	Create brief exchanges using target words; partners respond and extend.
Week 2	7.Language forms	15 min	Grammar structures in context	Practice short sentences focusing on forms (e.g.,

				present simple, pronouns).
Week 2	8.Pronunciation	15 min	Word stress & connected speech	Drill key words/phrases; repeat lines from model conversation with feedback.
Week 2	9.Modified role-play	15 min	Pushed output & turn-taking	Adapt the model to create an original dialogue about their families; record via Zalo.
Week 2	10.Self-assessment	5 min	Reflection and goal-setting	Complete checklist; note strengths, difficulties, and next-week targets.

To ensure fidelity across dyads, three strategies were used.

1. Structured implementation guides

All dyads received a step-by-step task guide for each topic. The guide included:

- the learning target,
- prompts for discussion and clarification,
- sample vocabulary and sentence frames,
- role-play instructions,
- the self-assessment checklist.

This guide ensured that both collaborative and expert–novice dyads followed the same pedagogical structure, regardless of individual differences.

2. Monitoring of online sessions

To verify that the 10-step cycle was followed:

- Students recorded details of each Zalo session, including the date, partner, duration, and mode (audio or video),
- a random sample of 25% of recordings was reviewed by the researcher using a brief observation checklist (e.g., "role-play completed," "pronunciation step attempted," "feedback exchanged").

Across weeks, logs indicated that dyads typically completed 50–60 minutes per week, and sampled recordings showed clear evidence of Steps 7–9 (practice, pronunciation work, and modified role-plays), suggesting high adherence.

3. Consistent facilitator support

Before each topic, the teacher clarified expectations and demonstrated how to move through the steps in Zalo. Specifically, students in collaborative dyad (low proficiency-low proficiency pairs) needed to (1) make shared choices, (2) repeat and expand on each other's utterances, or (3) request for clarification from the other (Storch, 2002, 2004; Walls, 2018). On the other hand, students in expert-novice dyad (high proficiency-low proficiency pairs) needed to (1) ask the novice to confirm concepts, (2) verify for understanding on the novice's side, (3) correct the novice's mistakes and (4) translate the novice's contributions from L1 into English (Storch, 2002, 2004; Walls, 2018). Importantly, no additional grammar instruction beyond the planned materials was provided, which helped isolate the procedural structure of the instructional model.

Taken together, these strategies improved implementation consistency and increased confidence that the observed outcomes reflected the intended instructional ecology rather than uncontrolled variation in instruction.

3.4 Instruments

3.4.1 Speaking tests

Pre/post speaking tests used in this study were based on the Cambridge Assessment English A2 Key Speaking test, which includes two parts: part 1 (3-4 minutes) – students take turns to listen to and individually answer the examiner's question related to their personal information such as name, age, home, family, etc.; and part 2 (5-6 minutes) – students work in pairs to talk about the pictures related to familiar topics (e.g., hobbies). In this part, they have to choose what they like or don't like and give reasons for their choices. To assess students' speaking performance, the A2 speaking rubric developed by Cambridge Assessment English was used. This rubric helps to assess three aspects of students' speaking: grammar and vocabulary, pronunciation, and interactive communication. Scores for all three aspects range from 0 to 5, which measures students' grammar & vocabulary, pronunciation, and interactive communication.

All assessments were administered in a quiet room by the classroom teacher following Cambridge timing guidance. Two Cambridge-trained KET-speaking examiners were invited to score all performances independently, achieving inter-rater reliability above 0.85. Both examiners were blinded to the time point and students' dyads to avoid confirmation bias.

3.4.2 Focus Group Interviews

The second instrument used in this study was the focus group interview. Eight interview questions were validated by three experts in English teaching using the Item Objective Congruence method. These interview questions explored students' experiences with the integrated model combining dyadic interactions, CBLT activities, and Zalo-based speaking practice. Students were then invited to describe their overall experiences and emotional responses to using this model to improve their speaking skills, as well as to recount specific moments when they found the integration particularly helpful or unhelpful. They were also asked to discuss the model's pros and cons and how they would describe it to others. Finally, the interviews examined whether students would recommend this integrated model to other English learners and give any additional comments. All interviews were done in Vietnamese, then were transcribed and translated for analysis.

3.4.3 Supplementary Performance Data (Speaking Recordings)

The third instrument was students' weekly Zalo recordings, which were submitted to the researchers weekly, and served as performance evidence. Despite not being scored, these recordings gave a clearer picture of learners' fluency development, pronunciation awareness, self-correction, and interactional patterns, which helped triangulate the findings.

3.5 Data Collection Procedures

The study lasted 16 weeks with two intact Grade 8 classes. After informed consent, all students completed a Week 1 pre-speaking test. Classes were then randomly assigned to collaborative or expert–novice dyads. In Week 2, both groups received training on how to do CBLT tasks for a sample topic (e.g., family and friends), dyadic interactions, and Zalo-based speaking practice. After that, from Weeks 3–14, both groups followed the same curriculum and 10-step CBLT activities, completing weekly Zalo speaking tasks and submitting recordings as evidence of practice. Each topic was practiced over two weeks.

In Week 15, students completed the post-speaking test, followed by focus group interviews in Week 16. Zalo recordings produced during the program were stored (with consent) and used as supplementary qualitative data to support observed performance changes and reported learning experiences.

3.6 Data Analysis

3.6.1 Quantitative Analysis

SPSS 26 was used for statistical analysis. Descriptive statistics, paired-samples t-tests and one-way ANOVA were conducted by the research team to assess the observed gains in speaking performance both within the specific dyads and across the distinct dyads. Fundamental statistical assumption requirements, which included normality and homogeneity of variance, were checked before the official testing of the stated hypothesis began. Effect sizes, including Cohen's d and eta squared (η^2), were calculated to determine the extent of participants' learning gains.

3.6.2 Qualitative Analysis

The qualitative data obtained from the interview transcripts were then analysed using reflexive thematic coding (Braun & Clarke, 2021). Reflexive thematic analysis was chosen because it emphasizes researcher interpretation and the development of meaningful patterns, rather than relying on mechanical code counting. The data were analyzed in units of complete ideas (sentences or short exchanges expressing a single meaning), and each segment could be coded under multiple themes when needed. Analysis proceeded through six phases. First, the researcher read the transcripts several times to become familiar with the data and to note initial ideas. Second, initial codes were generated inductively to capture key features of motivation, engagement, and learning processes. Third, similar codes were grouped into early themes to show broader patterns in the data. Fourth, the themes were checked against the dataset to ensure they were consistent and clearly distinct from one another. Fifth, themes were refined, named, and defined with attention to their conceptual links to sociocultural theory and competence-based pedagogy. Finally, illustrative excerpts were selected to represent thematic claims.

In this present study, the initial coding process was done by the first author, who had prior training in qualitative research methods. To enhance the reliability of the analysis, the

second author, who is an experienced qualitative researcher, independently examined a subset of the coded transcripts (approximately 30%), which was considered sufficient to reflect differences between dyad types and stages of instruction. Then, the two researchers discussed their coding differences through a negotiated agreement process (Campbell et al., 2013) to clarify interpretations and refine themes rather than focus on statistical reliability. Throughout the analysis, the researchers kept reflexive notes to record key decisions and reflect on their assumptions about technology use, peer interaction, and EFL teaching.

3.6.3 Use of Supplementary Speaking Recordings for Triangulation

A subset of the collected speaking recordings was reviewed by the research team to triangulate and support the key themes which were generated from the interview data. These speaking recordings were analyzed to find examples of self-correction, pronunciation monitoring, effective vocabulary use, and interactional engagement among participants.

To ensure systematic integration into the analytical framework, speaking recordings were selected at different time points. They were selected from weeks 3, 8, and 14 to capture the early, middle, and later stages of the process. At each time point, recordings from both dyad types were randomly selected.

The recordings were then analyzed in line with a predefined set of guidelines to examine examples of self-correction, feedback exchanges, the duration of uninterrupted speech, and the use of new vocabulary from the CBLT tasks.

These notes were not used to generate scores or to independently identify new themes. Instead, they were used to triangulate the findings from the interview data. When discrepancies emerged, the interview interpretations were re-examined to ensure alignment with the observed interactional behavior.

4. Results

4.1 Quantitative Results

Both dyads achieved significant gains from pre- to post-test (see Table 2). For the collaborative dyad, total scores rose from M = 7.72 (SD = 1.75) to M = 9.44 (SD = 1.38), $t(35) = 11.03$, $p < .001$, Cohen's d=1.83. For the expert-novice dyad, the increase was from M = 7.04 (SD = 1.69) to M = 9.51 (SD = 1.64), $t(35) = 11.73$, $p < .001$, Cohen's d=1.96. Effect sizes ranging from Cohen's d = 1.83 to 1.96 represent large within-group gains in speaking performance during participation in the instructional program.

Table 2

Descriptive statistics and paired-samples T-test results, pre-test and post-test

	Pre-test		Post-test			Significance	Effect size
	Mean	SD	Mean	SD	t(35)	P	(Cohen's d)
Collaborative dyad	*7.72*	*1.75*	*9.44*	*1.38*	*11.03*	*.000*	*1.83*
Expert-novice dyad	*7.04*	*1.69*	*9.51*	*1.64*	*11.73*	*.000*	*1.96*

The one-way ANOVA on pre-test scores (Table 3) revealed the two dyads were equivalent at baseline. Differences were small and non-significant for grammar and vocabulary ($F(1, 70) = 3.60$, $p = .06$, $\eta^2 = .05$), pronunciation ($F(1, 70) = 1.52$, $p = .22$, $\eta^2 = .02$), and

interactive communication ($F(1, 70) = 2.43$, $p = .12$, $\eta^2 = .03$). These results indicate that both dyads began the intervention with equivalent speaking proficiency.

Table 3

Descriptive Statistics and One-Way ANOVA Results (with Effect Sizes) for Pretest Speaking Performance by Dyad

Speaking Dimension	Dyad	N	M	SD	F(1, 70)	P	η^2
Grammar & Vocabulary (Pretest)	Collaborative	*36*	*2.54*	*0.61*	*3.60*	*.06*	*.05*
	Expert–Novice	*36*	*2.28*	*0.57*			
Pronunciation (Pretest)	Collaborative	*36*	*2.63*	*0.59*	*1.52*	*.22*	*.02*
	Expert–Novice	*36*	*2.44*	*0.65*			
Interactive Communication (Pretest)	Collaborative	*36*	*2.56*	*0.64*	*2.43*	*.12*	*.03*
	Expert–Novice	*36*	*2.32*	*0.65*			

A one-way analysis of variance (ANOVA) was conducted to examine differences between the collaborative dyad and the expert–novice dyad in gain scores for grammar and vocabulary, pronunciation, and interactive communication. Levene's tests indicated violations of the homogeneity of variance assumption for all three outcomes; therefore, Welch's ANOVA was used for dyad comparisons (see Table 4).

For grammar and vocabulary, descriptive statistics showed that the expert–novice dyad (M = 0.96, SD = 0.67, n = 36) achieved higher gains than the collaborative dyad (M = 0.58, SD = 0.39, n = 36). Welch's ANOVA revealed a statistically significant difference between dyads, $F(1, 56.07) = 8.46$, $p = .005$, with a medium-to-large effect size ($\eta^2 = .13$), indicating that expert–novice dyads demonstrated significantly greater gains in grammar and vocabulary.

For pronunciation, a similar trend was observed. The expert–novice dyad (M = 0.90, SD = 0.49) outperformed the collaborative dyad (M = 0.54, SD = 0.39) as the difference was statistically significant, Welch's $F(1, 66.28) = 12.08$, $p = .001$, with a large effect size ($\eta^2 = .15$). This result suggested that expert–novice interaction was associated with greater improvement in pronunciation accuracy.

However, gains in interactive communication between the two dyads were comparable. The collaborative dyad (M = 0.60, SD = 0.41) and the expert–novice dyad (M = 0.61, SD = 0.55) demonstrated similar improvement. Welch's ANOVA indicated that there was no statistically significant difference, $F(1, 64.84) = 0.02$, $p = .904$, with a very small effect size ($\eta^2 < .01$). This result indicated that dyad type accounted for minimal differences in interactive communication gains.

Table 4

Descriptive Statistics and Welch's One-Way ANOVA Results (with Effect Sizes) for Speaking Gain Scores by Dyad Type

Speaking Dimension	Dyad	*N*	*M*	*SD*	Welch's *F*	*P*	η^2
Grammar & Vocabulary Gain	Collaborative	36	0.58	0.39	$F(1, 56.07) = 8.46$	.005	.13
	Expert–Novice	36	0.96	0.67			
Pronunciation Gain	Collaborative	36	0.54	0.39	$F(1, 66.28) = 12.08$	.001	.15
	Expert–Novice	36	0.90	0.49			
Interactive Communication Gain	Collaborative	36	0.60	0.41	$F(1, 64.84) = 0.02$	.904	< .01
	Expert–Novice	36	0.61	0.55			

Even though both types of dyad showed clear, measurable progress in speaking, how they improved was different for each speaking aspect. The expert-novice groups made larger gains in grammar, vocabulary, and pronunciation. This may be related to what we saw during their conversations, as shown in later interview data and speaking recordings. In these conversations, near-peer partners more frequently provided quick corrections, feedback, and pronunciation modeling, which could have made learners pay more attention and adjust how they spoke.

In contrast, gains in interactive communication were the same for both collaborative and expert-novice dyads. This pattern is consistent with the notion that interactional development was more strongly facilitated by regular conversational practice, familiarity with interactional routines, and sustained speaking opportunities, all of which were shared across both dyads.

4.2 Qualitative results

Analysis of the interviews produced five benefit themes and three challenge themes, showing how students experienced the interaction of CBLT tasks, dyadic work, and Zalo practice. Rather than simply reflecting positive attitudes, the themes illuminate processes that help explain the quantitative outcomes.

4.2.1 Benefits of the Integrated Model *(S1-S36: secondary school students in the collaborative dyad; S37-S72: secondary school students in the expert-novice dyad)*

Improved speaking performance

The first benefit shared by grade eight participants in both groups was improved speaking performance. Students in both dyad types reported noticeable improvement although the source of this improvement differed between dyads. Specifically, expert–novice learners repeatedly explained the presence of explicit correction and modeling:

> "When I said, "*He go to school"*, my buddy repeated, "*He goes to school*." Then, I tried again." (S44)

These moments showed the process of active language use and awareness of language forms, which enabled learners to practice them immediately. This helps explain why accuracy gains were stronger in this dyad compared to the collaborative dyad's.

On the other hand, collaborative students, lacking a stronger model, relied on discussion instead:

> "Sometimes we didn't know, so we guessed and said it again in another way." (S16)

Here, fluency developed through repeated practice and rephrasing, while grammatical accuracy improved at a slower speed.

Increased speaking opportunities

Another benefit reported by participants was increased speaking opportunities thanks to Zalo. Zalo created a semi-private space where learners could try, fail, and try again without embarrassment. For example, one participant mentioned:

> "In class, I speak only one or two sentences. On Zalo, I can talk many times until I feel okay." (S28)

> Or another student shared "When talking one-on-one, I wasn't scared of making mistakes as much as when I was speaking in front of the whole class." (S34)

> Students in both dyads also described self-regulated rehearsal:
> "I recorded my voice again and again until it sounded right" (S33)

These findings showed that improvement did not result simply from "liking" the app; instead, it came from extended practice cycles that classroom time alone could not provide.

Greater Autonomy and Self-Correction

The next benefit, reported by grade eight students, was greater autonomy and self-correction. Students began to monitor their own language with independence, as a participant reported:

> "Before, I waited for teacher correction. Now I listen again, write down my mistakes and fix them myself." (S64)

It was also reported that the feedback was internalized, especially by the expert-novice learners:

> "When I speak now, I remember what my buddy corrected last week." (S52)

This change from being controlled by others to being self-controlled is a key mechanism suggested by sociocultural theory.

Enhanced emotional engagement

Enhanced emotional engagement was the fourth benefit shared by grade eight students. Their confidence increased not because they were told they were good, but because they could hear improvement by themselves, as one student mentioned:

"Old recordings sound terrible. New ones are better. That makes me want to continue." (S40)

This showed that students' progress became visible (or audible), which sustained their efforts even when the tasks were difficult.

Improved motivation

The last benefit reported by grade eight participants was improved motivation. They described motivation in terms of real speaking, not technology excitement, as one participant shared:

"I wasn't just practicing English. I was talking to someone who answered back. Speaking with my buddy felt like a real friendship" (S30)

Some expert–novice students also added responsibility:

"I didn't want my buddy to think I was lazy, so I practiced more." (S47)

Motivation, therefore, was connected to social commitment, purpose, and perceived competence.

Supplementary Evidence from Speaking Recordings

The findings from the interviews were confirmed when the researchers observed the weekly speaking recordings on Zalo. Learners usually performed self-correction, particularly with verb forms (e.g., "*My sister like… sorry, likes drawing*") and final consonants (e.g., "*I have two dog… dogs*"). These suggested that they were noticing their mistakes and trying to correct them. In expert–novice dyad, English-major buddies offered clear scaffolding, such as, "*Try saying: He usually has lunch at school*", which were immediately adopted by novices. Later recordings showed longer, and more fluent speaking in both dyads, particularly during shopping role-plays (e.g., "*Can I help you? … Yes, I'm looking for a blue jacket*"). These speaking recording observations reinforced themes of improved accuracy, fluency, and confidence.

4.2.2 Challenges of the new instructional model of English-speaking interaction

Language and communication barriers

The first challenge reported by grade eight participants was language and communication barriers, especially in collaborative dyads. Collaborative pairs often relied on Vietnamese, as one participant mentioned:

"Sometimes we finished the task in Vietnamese and only translated." (S12)

Or stated by another participant:

"Sometimes I didn't know whether I was practicing English or Vietnamese" (S6)

This suggested that without a more knowledgeable partner, the task difficulties exceeded language resources, and led to weaker gains in accuracy.

In addition, low-proficiency pairs in the collaborative dyad also expressed uncertainty about providing or accepting peer feedback, with S8 sharing, *"I wasn't confident giving feedback."* These difficulties suggested that when both partners lacked language knowledge, sustaining English-only communication was harder, and opportunities for negotiation of meaning were reduced.

Imbalanced participation

Imbalanced participation was the second challenge noted by grade eight students, especially in expert-novice dyads. Some novices felt overshadowed by their more proficient partners, as one student (S55) noted, *"My partner spoke most of the time, so I just listened."* This pattern showed the uneven participation in mixed-proficiency pairs, where the expert's language confidence might unintentionally limit the novice's participation. Even though these problems didn't stop learners from making progress in general, they showed that mobile-based interactions might need more structured turn-taking guidance.

Technology challenges

The last challenge shared by participants in both groups was related to technology. Students noted that small smartphone screens made it hard to do multiple tasks at once, as a participant shared: "It was difficult to see the video call and the messages at the same time" (S14). Unstable internet connectivity occasionally interrupted speaking practice, asking some learners to find different locations with stronger Wi-Fi (S58). Zalo's limited recording feature was another technological problem, with S48 explaining, "I needed another phone to record my speaking." Some shy students also wanted privacy-enhancing options, such as face filters or blurred backgrounds, to feel more comfortable during video calls; however, these options are not available in the Zalo app yet.

5. Discussion

This study investigated how dyadic interaction, Competence-Based Language Teaching (CBLT), and Zalo-mediated communication worked together to support the speaking development of Vietnamese EFL grade eight students. Instead of treating dyad type as a separate causal factor, the study positioned it within a broader instructional context. Here, dyadic interaction supported language development, CBLT structured practice, and Zalo extended opportunities for low-pressure speaking. Consequently, differences between collaborative and expert–novice dyads are better understood in terms of differences in scaffolding, feedback timing, and interactional roles, rather than proficiency pairing alone. This view aligns with sociocultural perspectives, which see learning as developing through mediated activity rather than a single instructional factor.

In this study, dyadic interaction was a key factor in learners' speaking development. Both dyads showed improvement, but the expert–novice pairs made greater gains in grammar & vocabulary, and pronunciation. These findings reflect sociocultural theory, which supports

the idea that learning is driven by scaffolded interaction within the Zone of Proximal Development (Lantolf, 2000; Vygotsky, 1978). In these pairs, near-peer mentors frequently offered modeled target forms, provided prompts, and encouraged self-correction, which allowed less proficient learners to go beyond what they could previously do.

These findings also support earlier research, suggesting that mixed-proficiency pairs promote greater language awareness and active use of English, while collaborative pairs tend to rely more on L1 to manage tasks (Kim & McDonough, 2008; Philp et al., 2010; Tian & Jiang, 2021). This helped explain why both dyads improved similarly in interactive communication, which seems to depend more on repeated conversational routines and turn-taking practice than on accuracy. Overall, the findings suggest that no single pairing model is always better; instead, different dyad types support different learning processes.

Beyond peer interactions, the 10-step CBLT cycle helped students practice with clear goals, receive support, and use language relevant to real-life situations. With specific objectives, model dialogues, vocabulary preparation, pronunciation practice, and structured self-assessment, CBLT guided dyadic interaction so that it was purposeful rather than unfocused. This step-by-step procedure also helped students to go beyond the limited practice and speak more freely (Dooly & Sadler, 2016; López-Medina & Casado, 2024; Pambudi et al., 2024; Richards & Rodgers, 2014; DeKeyser, 2007).

Qualitative evidence further showed that CBLT fostered autonomy and self-monitoring: students repeatedly described checking their recordings, tracking their progress, and using vocabulary lists or learning objectives to prepare for interaction. These behaviors did not arise from dyadic interaction alone; they were facilitated by the CBLT design. Thus, the pedagogical architecture of CBLT shaped learners' linguistic engagement and sustained improvement over the 12-week intervention.

Zalo played a mediating role and did not merely provide some technological convenience to the learners. Its synchronous video call features and asynchronous voice messages and recording features extended speaking practice far beyond the educational institution classroom limitations. These conditions appeared to reduce anxiety and enhance participation in mobile-assisted environments, aligning with findings from Han and Keskin (2016), Shih et al. (2015), and Akkara et al. (2020).

Zalo provided the students with additional exposure time, which was closely associated with their extended speaking practice. Classroom interaction is often constrained by having large classes and very limited speaking turns, which makes learning difficult, but mobile exchanges permit extended and valuable one-to-one conversation among the students (Mao, 2014; Sun & Yang, 2015).

In other words, beyond being a simple communication tool, Zalo provided a low-pressure environment for speaking practice. Through Zalo, learners were able to re-record responses, practice individually, or interact with a single partner, thereby reducing their anxiety. These features are particularly valuable in exam-oriented EFL contexts, where opportunities for in-class speaking are limited.

Overall, these findings indicated that speaking development in technology-mediated, competence-based environments was shaped by the interplay of pedagogical structure, interactional scaffolding, and mediational tools. Specifically, dyadic scaffolding gave students chances to build understanding through interaction, the CBLT cycle helped their learning develop step by step with clear goals, and Zalo supported continued practice across time and

different settings. Removing any single layer would likely weaken the overall system: interaction without structure may become unfocused, structured tasks without sustained dialogue limit authentic practice, and both are constrained without a digital environment that extends participation. This interplay reflects the mutual relationship, which was described in sociocultural perspectives on mediated learning (Ohta, 2001; Swain, 2000).

Without a traditional control group, the study cannot clearly distinguish the effects of CBLT, dyadic pairing, and Zalo. It is also difficult to determine how much near-peer mentors influenced the results in the expert–novice dyad. The conclusions therefore focus on how these elements worked together in an integrated instructional model, rather than treating them as separate factors.

Several design principles were identified based on the model's effective use in real classrooms. Peer work was effective because students' interaction followed a clear sequence of CBLT tasks with shared goals, helping them understand what they were doing and why. Additionally, using Zalo gave students a private space to practice, make mistakes, and improve before speaking in class. Reviewing their recordings and reflecting on errors also encouraged them to take more responsibility for their learning instead of relying on teacher feedback. Finally, expert-novice dyads worked best when teachers monitored turn-taking to make sure that more proficient partners supported novices, rather than overshadowed them.

6. Conclusion

This study explored how a competence-based, dyadic, mobile-supported model could help Vietnamese EFL lower-secondary learners develop speaking skills over the semester of 16 weeks. Quantitative results showed that the model was linked to students' enhancements in vocabulary, grammar, pronunciation, and interactive communication for both dyads; however, expert-novice dyads showed better gains in grammar, vocabulary, and pronunciation. The qualitative results helped explain these gains: learners were not just speaking more, but were also noticing errors, rehearsing, receiving feedback, and gradually taking more responsibility for their own progress.

Overall, these findings suggest that the integrated instructional model was linked with substantial speaking development, because all three of its components reinforced one another. Dyadic interaction played a central role in the learning process, as it created opportunities for learners to discuss and refine language through their dialogue, while CBLT provided a clear structure for interaction through defined goals and tasks, along with opportunities to track learners' progress. At the same time, Zalo allowed learning to continue beyond the classroom by giving students space to practice privately and repeatedly without the pressure of speaking in front of others. Thus, the model can be interpreted as an integrated instructional model in which mechanism, pedagogy, and environment worked together.

However, these findings should be interpreted with caution because the study's design did not include a non-intervention control group and the observed gains represent the combined effect of the intervention rather than the isolated impact of any single component. Future research could look more closely at specific mechanisms and explore how the model works across different age groups, pairing types, and technological platforms.

7. About the Authors

Tran Anh Vu is a full-time English lecturer at Faculty of Foreign Languages, Thu Dau Mot University, Vietnam. He is currently pursuing a Ph.D. in English Language Studies of English Department, Faculty of Liberal Arts, Thammasat University. His research focuses on TESOL, mobile assisted language learning, English language assessment, and intercultural communication. ORCID: https://orcid.org/0000-0001-7439-3255

Kwanjira Chatpunnarangsee, Ph.D., is an Assistant Professor at Faculty of Liberal Arts, Thammasat University, Thailand. Her research areas include bilingualism and bilingual education, corpus linguistics, and English for specific purposes.

8. Acknowledgement

I am deeply grateful to the students and teachers at An Dien Secondary School, who welcomed this project into their busy classrooms and gave their time with enthusiasm. I also thank to the English-major students from Thu Dau Mot University, because their patience, kindness, and willingness to support younger learners were central to this study. I am especially grateful to my supervisor, Prof. Kwanjira Chatpunnarangsee, for her steady guidance and for helpful feedback to strengthen this research. Finally, I would like to acknowledge the families who supported the students' participation, as their trust made this research possible.

9. Declaration of AI Use

The authors declare that ChatGPT 5.0 was used only for proofreading purposes (i.e., checking spelling and grammatical errors). All ideas, analyses, and interpretations presented in this study are the authors' own.

10. References

Adeyeye, B., Ojih, S. E., Bello, D., Adesina, E., Yartey, D., Ben-Enukora, C., & Adeyeye, Q. (2022). Online learning platforms and Covenant University students' academic performance in practical related courses during COVID-19 pandemic. *Sustainability, 14*(2), 878. https://doi.org/10.3390/su14020878

Akkara, S., Anumula, V., & Mallampalli, M. (2020). Impact of WhatsApp interaction on improving L2 speaking skills. *International Journal of Emerging Technologies in Learning, 15*(3), 250–259. https://doi.org/10.3991/ijet.v15i03.11534

Avci, H., & Adiguzel, T. (2017). A case study on mobile-blended collaborative learning in an English as a foreign language context. *International Review of Research in Open and Distributed Learning, 18*(7). https://doi.org/10.19173/irrodl.v18i7.3261

Boahin, P., & Hofman, W. H. A. (2014). Perceived effects of competency-based training on the acquisition of professional skills. *International Journal of Educational Development, 36*, 81–89. https://doi.org/10.1016/j.ijedudev.2013.11.003

Braun, V., & Clarke, V. (2021). Can I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research, 21*(1), 37–47. https://doi.org/10.1002/capr.12360

Bui, T. T. N., & Nguyen, H. T. M. (2016). Standardizing English for educational and socio-economic betterment: A critical analysis of English language policy reforms in Vietnam. In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 363–388). Springer. https://doi.org/10.1007/978-3-319-22464-0_17

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research, 42*(3), 294–320. https://doi.org/10.1177/0049124113500475

Chan, S. C., Wan, C. J., & Ko, S. (2019). Interactivity, active collaborative learning, and learning performance: The moderating role of perceived fun by using personal response systems. *The International Journal of Management Education, 17*(1), 94–102. https://doi.org/10.1016/j.ijme.2018.12.004

DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Dooly, M., & Sadler, R. (2016). Becoming little scientists: Technologically enhanced project-based language learning. *Language Learning & Technology, 20*(1), 54–78. https://doi.org/10.64152/10125/44446

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Ablex.

Education First. (2024). *EF English proficiency index*. https://www.ef.com/wwen/epi/

Eltanskaya, E., Arzhanovskaya, A., & Linkova, Y. (2019). Teaching English by means of competence-based approach to the students of secondary vocational education. In *Proceedings of the 1st International Scientific Practical Conference: The Individual and Society in the Modern Geopolitical Environment (ISMGE 2019)* (pp. 177–181). Atlantis Press. https://doi.org/10.2991/ismge-19.2019.35

Fouz-González, J. (2017). Pronunciation instruction through Twitter: The case of commonly mispronounced words. *Computer Assisted Language Learning, 30*(7), 631–663. https://doi.org/10.1080/09588221.2017.1340309

Ha, V. X., Tran, N. G., & Tran, N. H. (2021). Teachers' beliefs and practices regarding assessment in EFL classrooms in Vietnam. *The Qualitative Report, 26*(11), 3457–3475. https://doi.org/10.46743/2160-3715/2021.5063

Han, T., & Keskin, F. (2016). Using a mobile application (WhatsApp) to reduce EFL speaking anxiety. *GIST Education and Learning Research Journal, 12*, 29–50. https://doi.org/10.26817/16925777.243

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue among Korean as a second language learners. *Language Teaching Research, 12*(2), 211–234. https://doi.org/10.1177/1362168807086288

Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 1–26). Oxford University Press.

Le, V. C. (2011). *Form-focused instruction: A case study of Vietnamese teachers' beliefs and practices* (Doctoral dissertation, University of Waikato).

Lin, C.-H. (2012). *Language learning through social networks: Perceptions and reality* (Doctoral dissertation, University of California, Irvine).

López-Medina, E. F., & Casado, J. M. (2024). CLIL as a competence-based strategy against foreign language anxiety. *Discover Psychology, 4*(1), 68. https://doi.org/10.1007/s44202-024-00187-6

Mao, J. (2014). Social media for learning: A mixed methods study on high school students' technology affordances and perspectives. *Computers in Human Behavior, 33*, 213-223. https://doi.org/10.1016/j.chb.2014.01.002

McDonough, K. (2004). Learner–learner interaction during pair and small-group activities in a Thai EFL context. *System, 32*(2), 207–224. https://doi.org/10.1016/j.system.2004.01.003

Nguyen, G. V. (2013). *Orienting to pedagogical innovation: A case study of Vietnamese teachers' beliefs and practices* (Doctoral dissertation, University of Waikato).

Nguyen, V. (2022). *Vietnamese university EFL teachers' beliefs and practices: A cultural historical activity theory analysis* (Doctoral dissertation, Victoria University of Wellington). https://doi.org/10.26686/wgtn.19492178

Ohta, A. S. (2001). Peer interactive tasks and assisted performance in classroom language learning. In A. S. Ohta (Ed.), *Second language acquisition processes in the classroom: Learning Japanese* (pp. 73–128). Lawrence Erlbaum. https://doi.org/10.4324/9781410604712-9

Pambudi, N., Syam, H., Rustam, U., Dewi, R. A. P. K., & Amir, H. S. (2024). The effect of a competency-based learning approach on improving English speaking skills. *International Journal of Language and Ubiquitous Learning, 2*(1), 46–56. https://doi.org/10.70177/ijlul.v2i1.753

Peterson, M. (2012). Learner interaction in an MMORPG: A sociocultural discourse analysis. *ReCALL, 24*(3), 361–380. https://doi.org/10.1017/S0958344012000195

Phan, D. Q. T. (2021). *Vietnamese EFL teachers' experiences of the new English curriculum* (Doctoral dissertation, UNSW Sydney). https://doi.org/10.26190/unsworks/2262

Philp, J., Walter, S., & Basturkmen, H. (2010). Peer interaction in the foreign language classroom: What factors foster a focus on form? *Language Awareness, 19*(4), 261–279. https://doi.org/10.1080/09658416.2010.516831

Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, S. A., & Yousufi, S. Q. (2023). Factors affecting students' learning performance through collaborative learning and engagement. *Interactive Learning Environments, 31*(4), 2371–2391. https://doi.org/10.1080/10494820.2021.1884886

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781009024532

Saint, M. (2021). *Development of a curriculum based on content and language integrated learning and competency-based education for enhancing business English writing ability of undergraduate students* (Doctoral dissertation, Chulalongkorn University). https://doi.org/10.58837/chula.the.2021.123

Shih, R.-C., Lee, C., & Cheng, T.-F. (2015). Effects of English spelling learning experience through a mobile LINE app for college students. *Procedia – Social and Behavioral Sciences, 174*, 2634–2638. https://doi.org/10.1016/j.sbspro.2015.01.945

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52*(1), 119–158. https://doi.org/10.1111/1467-9922.00179

Storch, N. (2004). Using activity theory to explain differences in patterns of dyadic interactions in an ESL class. *Canadian Modern Language Review, 60*(4), 457–480. https://doi.org/10.3138/cmlr.60.4.457

Sun, Y.-C., & Yang, F.-Y. (2015). I help, therefore I learn: Service learning on Web 2.0 in an EFL speaking class. *Computer Assisted Language Learning, 28*(3), 202–219. https://doi.org/10.1080/09588221.2013.818555

Sundkvist, P., & Nguyen, X. N. C. M. (2020). English in Vietnam. In K. Bolton, A. Kirkpatrick, & J. Lee (Eds.), *The handbook of Asian Englishes* (pp. 683–703). Wiley. https://doi.org/10.1002/9781118791882.ch30

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford University Press.

Thao, L. T., & Mai, L. X. (2022). English language teaching reforms in Vietnam. *Innovation in Language Learning and Teaching, 16*(1), 29–40. https://doi.org/10.1080/17501229.2020.1846041

Tian, L., & Jiang, Y. (2021). L2 proficiency pairing, task type and L1 use: A mixed-methods study on optimal pairing in dyadic task-based peer interaction. *Frontiers in Psychology, 12*, 699774. https://doi.org/10.3389/fpsyg.2021.699774

Van, V. H. (2010). The current situation and issues in the teaching of English in Vietnam. *Ritsumeikan Language and Culture Research, 22*(1), 7–18.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Walls, L. C. (2018). The effect of dyad type on collaboration: Interactions among heritage and second language learners. *Foreign Language Annals, 51*(3), 638–657. https://doi.org/10.1111/flan.12356

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research, 11*(2), 121–142. https://doi.org/10.1177/1362168806070745

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Xodabande, I. (2017). The effectiveness of social media network Telegram in teaching English language pronunciation to Iranian EFL learners. *Cogent Education, 4*(1), 1347081. https://doi.org/10.1080/2331186X.2017.1347081

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal, 86*(1), 54–66. https://doi.org/10.1111/1540-4781.00136

Yen, Y.-C., Hou, H.-T., & Chang, K. E. (2015). Applying role-playing strategy to enhance learners' writing and speaking skills in EFL courses using Facebook and Skype as learning tools: A case study in Taiwan. *Computer Assisted Language Learning, 28*(5), 383–406. https://doi.org/10.1080/09588221.2013.839568