Journal of Digital Education and Learning Engineering ปีที่ 2 ฉบับที่ 1 (เดือน มกราคม – มีนาคม 2569) หน้า 54-73
Journal of Digital Education and Learning Engineering, Volume 2 No. 1 (January – March 2026) Page 54-73
บทความวิจัย (Research Article)

Journal of Digital Education and Learning Engineering
ดำเนินการวารสารโดย สมาคมการศึกษาดิจิทัลและวิศวกรรมการเรียนรู้

The Effect of Virtual Reality–Based Learning on Thai Students' Chinese Listening and Speaking Skills: A Pretest–Posttest Study in Ayutthaya

Changyan Wu[1] & Jirapipat Thanyaphongphat[2*]

Faculty of Multidisciplinary and Interdisciplinary School, Chiang Mai University , Thailand[1], College of Arts,Media and Technology, Chiang Mai University, Thailand[2]

Received: February 28, 2026 Revised: March 17, 2026 Accepted: March 24, 2026

Abstract

This study examines the impact of Virtual Reality (VR)–based instruction on Thai secondary school students' Chinese listening and speaking skills. A one-group pretest–posttest research design was employed. The participants were 45 Thai students aged 13–18 who had studied Chinese for at least five years. Their listening and speaking performances were measured before and after the implementation of VR-supported learning activities. The results indicate a significant improvement in listening performance, with mean scores increasing from 8.36 in the pretest to 12.62 in the posttest. The reduction in standard deviation suggests not only overall progress but also a narrowing of performance gaps among students. Speaking scores improved from 2.02 to 2.76, showing a positive trend, although the magnitude of improvement was smaller compared to listening. Overall scores rose from 10.38 to 15.38, accompanied by a decrease in standard deviation, indicating enhanced overall language proficiency and more consistent performance. The findings suggest that VR-based immersive learning can effectively enhance students' Chinese listening and speaking abilities, particularly in listening comprehension. By providing authentic contextual environments and interactive experiences, VR technology increases learner engagement and supports meaningful language input and output. Although improvements in speaking were comparatively moderate, the overall results demonstrate the potential of VR-assisted instruction in improving language learning outcomes. This study offers empirical evidence supporting the integration of VR technology into Teaching Chinese as a Foreign Language and highlights its value in immersive, context-based language education.

Keywords: Virtual Reality (VR), Technology-Enhanced Language Learning (TELL), Chinese as a Foreign Language (CFL), Immersive Learning Environment, Pretest–Posttest Design

*Corresponding author
Email address: jirapipat.than@cmu.ac.th

Introduction

With the rapid advancement of digital technologies, Virtual Reality (VR) has emerged as a promising tool for enhancing second language acquisition by creating immersive and interactive learning environments. A growing body of research indicates that VR-assisted language learning can significantly improve learners' linguistic performance and affective engagement compared to traditional instructional approaches. For example, a meta-analysis by Chen, Wang, and Wang (2022) reported medium effect sizes of VR interventions on language learning outcomes, particularly in vocabulary acquisition and oral skills. From a theoretical perspective, immersive VR environments align with constructivist learning theory and cognitive load theory by situating learners in meaningful contexts that promote active knowledge construction while managing cognitive processing demands (Song, Shin, & Shin, 2023). Empirical studies have further demonstrated that VR-based instruction can enhance oral proficiency, increase learner motivation, and reduce foreign language anxiety by providing authentic yet low-risk practice environments (Li et al., 2024).

In the context of Teaching Chinese as a Foreign Language (TCFL), the integration of cultural elements is considered essential for developing communicative competence and intercultural awareness. Authentic cultural settings can facilitate meaningful language input and output, thereby strengthening learners' contextual understanding. The ancient city of Ayutthaya, once a major political and commercial center in Southeast Asia, played a crucial role in historical exchanges between Siam and China. As a UNESCO World Heritage Site, Ayutthaya Historical Park preserves architectural remains that reflect centuries of Sino-Thai cultural interaction. The long-standing diplomatic and economic relationship between Thailand and China has further increased the importance of Chinese language education in Thailand, particularly in tourism-related and culturally enriched learning contexts.

From an intercultural perspective, the use of Ayutthaya as a contextual backdrop also contributes to the development of intercultural communicative competence (ICC). Byram (1997) conceptualizes ICC as the ability to interact effectively with people from different cultural backgrounds through the integration of linguistic competence, cultural knowledge, and critical cultural awareness. By situating Chinese language learning within a historically significant site that embodies centuries of Sino–Thai exchange, the instructional design extends beyond linguistic form to intercultural meaning-making. As Kramsch (1993) argues, language learning is inseparable from the cultural contexts in which meanings are constructed and negotiated. The historical and cultural symbolism embedded in Ayutthaya provides learners with opportunities to interpret language within a cross-cultural framework, thereby fostering not only communicative competence but also intercultural awareness.

Despite the growing interest in VR-assisted language learning, limited empirical research has examined its effectiveness in improving Chinese listening and speaking skills among Thai secondary school students within culturally immersive environments. To address this gap, the present study employs a one-group pretest–posttest design to investigate the impact of VR-based instruction on Thai students' Chinese listening and speaking performance in a culturally contextualized setting inspired by Ayutthaya. By

combining immersive technology with heritage-based learning, this study seeks to provide empirical evidence for the pedagogical value of VR in Chinese language education.

■ Research Questions

1) Does VR-based instruction significantly improve Thai secondary students' Chinese listening skills?
2) Does VR-based instruction significantly improve Thai secondary students' Chinese speaking skills?
3) Is there a significant difference between students' pretest and posttest scores after the implementation of VR-based Chinese instruction?

■ Research Objectives

1) To evaluate the effectiveness of VR-based instruction in enhancing Thai secondary students' Chinese listening skills.
2) To examine the impact of VR-based instruction on students' Chinese speaking performance.
3) To determine whether there is a statistically significant difference between pretest and posttest scores following the VR intervention.

■ Literature Review

Virtual Reality (VR) in Language Education

Virtual Reality (VR) technology has gained increasing attention in education as an innovative medium that can create immersive, interactive environments and simulate real-world scenarios that are otherwise difficult to reproduce in traditional classroom settings. This capacity to simulate authentic contexts has been shown to benefit language learners by providing contextualized exposure to language use (Chen, Wang, & Wang, 2022). In their meta-analysis, Chen et al. consolidated findings from multiple VR-assisted language learning studies and found moderate to strong positive effects of VR interventions on vocabulary, listening, and speaking outcomes.

The authors argue that VR's immersive nature supports embodied cognition, thereby enabling learners to process linguistic input more meaningfully. In a study examining learner affective factors, Kaplan-Rakowski and Gruber (2023) reported that high-immersion VR environments significantly reduce foreign language anxiety, which is often a barrier to productive use of the target language. Empirical evidence further demonstrates that VR environments can promote repeated practice, immediate feedback, and interactive dialogues—all crucial elements in second language acquisition (Chen et al., 2022; Kaplan-Rakowski & Gruber, 2023). These findings establish VR as a promising technological tool in language education, particularly for developing communicative competence.

Technology-Enhanced Language Learning (TELL)

Technology-Enhanced Language Learning (TELL) refers to the application of digital technologies to facilitate and enrich second language acquisition processes. TELL evolved from Computer-Assisted Language Learning (CALL) and now encompasses a broader array of interactive, networked, and immersive tools including VR, mobile learning, and corpus-assisted systems. Golonka et al. (2014) conducted a comprehensive review of technology applications in language learning and highlighted that technology can improve learner motivation, create authentic learning contexts, and extend communicative opportunities beyond the classroom. Their synthesis underscores that successful TELL implementations typically integrate pedagogical design with technology affordances, rather than merely using technology as an add-on. Chamot and O'Malley's (1994) cognitive academic language learning approach also supports the integration of multimodal digital tools in scaffolding complex language tasks.

In the context of immersive technologies, VR is an extreme form of TELL that situates learners within simulated environments where they can engage in meaningful linguistic interactions with dynamic contexts. Research in TELL has shown that such immersive tools encourage learner autonomy, provide multimodal input, and support collaborative learning tasks (Golonka et al., 2014). Thus, TELL provides the broader theoretical foundation in which VR applications are positioned, emphasizing not only technological novelty but pedagogically sound practice.

Chinese as a Foreign Language (CFL) and Technology Integration

Research on technology integration in Teaching Chinese as a Foreign Language (CFL) has expanded with the development of digital learning environments. Technology-enhanced language learning emphasizes the role of digital tools in supporting authentic communication and intercultural competence (Chun, Kern, & Smith, 2016). In CFL contexts, language learning is closely connected to cultural understanding, as Chinese linguistic forms are deeply embedded in sociocultural practices (Everson & Shen, 2010). Therefore, technology integration is increasingly viewed as a means to provide contextualized and culturally meaningful learning experiences.

Among emerging technologies, virtual reality (VR) has shown particular promise. Liu (2023), in a systematic review of VR applications in CFL, found that VR environments can construct immersive cultural scenarios and simulated communicative situations that promote listening and speaking development. VR enables learners to engage in contextualized interaction and experience cultural settings that are otherwise inaccessible in traditional classrooms.

In addition to linguistic competence, intercultural competence has become a central objective in foreign language education. Liddicoat and Scarino (2013) argue that intercultural language teaching requires learners to engage with cultural meanings, perspectives, and practices rather than merely acquire factual cultural knowledge. Within the context of Chinese as a Foreign Language (CFL), this perspective is particularly relevant, as Chinese language use is deeply embedded in sociocultural norms and historical

narratives (Everson & Shen, 2010). By integrating VR-based simulations of culturally significant environments such as Ayutthaya, instruction can facilitate experiential engagement with intercultural content. Such contextualization enables learners to interpret linguistic input through the lens of shared historical interactions between Thailand and China, thereby strengthening intercultural communicative competence alongside listening and speaking skills.

Overall, existing literature suggests that immersive VR aligns with pedagogical principles emphasizing communicative competence and cultural immersion, making it a promising approach for enhancing CFL instruction.

Immersive Learning Environment and Listening–Speaking Development

Immersive learning environments are grounded in constructivist and situated learning theories, which emphasize the role of contextualized interaction and active engagement in knowledge construction. Virtual Reality (VR), as a technologically mediated immersive medium, operationalizes these principles by placing learners in realistic, interactive environments where they can process audiovisual input and produce language in meaningful contexts. From a second language acquisition perspective, Krashen (1985) argues that comprehensible input is essential for language development. VR environments are capable of delivering rich, contextualized input that supports comprehension while maintaining learner engagement. Complementing this view, Swain and Lapkin (1995) propose that language output plays a critical role in facilitating internalization and cognitive processing, as learners become aware of gaps in their linguistic knowledge through production.

Empirical evidence supports the pedagogical value of VR in language learning. A meta-analysis by Chen, Wang, and Wang (2022) found that VR-assisted language instruction produced moderate to strong positive effects on listening and speaking outcomes compared to traditional methods. The authors suggest that immersive interaction and embodied experience enhance linguistic processing and retention. Furthermore, VR environments have been shown to reduce foreign language anxiety and increase learners' willingness to communicate (Kaplan-Rakowski & Gruber, 2023), thereby lowering affective barriers to oral production. A recent systematic review on presence in VR learning contexts also highlights that immersive environments enhance engagement and task involvement, which are key factors in effective language practice (Wei et al., 2025).

Collectively, these theoretical and empirical findings indicate that immersive VR environments provide favorable conditions for developing listening and speaking competence through contextualized input, meaningful output, and reduced affective constraints.

Pretest–Posttest Design in Educational Research

The pretest–posttest design is one of the most widely used quantitative research methods in educational intervention studies because it allows researchers to assess changes in learner performance

over time. In this design, participants are measured on key variables before and after the intervention, enabling researchers to attribute differences in scores to the treatment effect when other variables are controlled. Creswell (2014) notes that the pretest–posttest design is particularly useful in studies without control groups, as long as threats to internal validity are considered and addressed.

Fraenkel, Wallen, and Hyun (2012) describe how paired-samples t-tests are commonly used in pretest–posttest studies to statistically determine whether a significant change has occurred. Educational technology research frequently adopts this design to measure outcomes related to learner achievement, motivation, or attitudes following the implementation of a digital tool or pedagogical innovation. For example, multiple VR language studies employ pretest–posttest assessments to evaluate the effectiveness of VR intervention (e.g., Tai, 2022; Fan, 2025). Additionally, this design aligns well with classroom practicalities in language settings where random assignment to control and experimental groups is often not feasible.

Overall, the pretest–posttest design provides a valid and practical framework for assessing the impact of technology-enhanced instruction such as VR on learner outcomes.

VR Instructional Design and Theoretical Alignment

To establish a clear connection between theoretical foundations and instructional practice, the VR-based learning activities in this study were designed to align with key principles of second language acquisition through specific task sequences.

First, during the vocabulary preview stage, learners were introduced to directional terms (e.g., 东, 南, 西, 北) through multimodal input, including written text, audio pronunciation, and English translations. This stage reflects Krashen's (1985) input hypothesis, as learners were provided with comprehensible input supported by visual and auditory cues. The ability to replay audio and access translations further ensured that the input remained within learners' comprehension level (i+1), facilitating effective intake of new linguistic forms.

Second, in the guided VR exploration stage, learners navigated panoramic scenes and listened to contextualized explanations provided by a virtual guide. This stage embodies principles of situated learning (Lave & Wenger, 1991), as language input was embedded within an authentic cultural environment modeled on Ayutthaya. The integration of spatial context and cultural elements enabled learners to construct meaning through interaction with the environment, consistent with constructivist learning theory (Vygotsky, 1978). Additionally, the combination of visual, spatial, and auditory information supports deeper cognitive processing, aligning with research on multimodal learning in VR environments (Chen, Wang, & Wang, 2022).

Third, during the interactive navigation task, learners were required to select the correct direction to proceed within the virtual environment. This task operationalizes Swain and Lapkin's (1995) output hypothesis, as learners actively processed linguistic input and produced responses by applying directional vocabulary in context. The requirement to make decisions based on comprehension encourages learners to notice gaps in their understanding and engage in deeper cognitive processing.

Fourth, the immediate feedback mechanism, where learners were prompted to retry incorrect answers after reviewing the guide's explanation, reflects principles of formative assessment and cognitive reinforcement. This iterative process supports learning by allowing learners to refine their understanding through repeated exposure and correction, which is consistent with theories of feedback in language acquisition.

Finally, the overall VR environment contributed to the reduction of affective barriers, as learners engaged with the tasks in a low-pressure, interactive setting. Unlike traditional classroom speaking activities, the VR environment allowed learners to explore and respond without direct social evaluation, which may reduce anxiety and increase willingness to engage with the target language. This aligns with findings by Kaplan-Rakowski and Gruber (2023), who highlight the role of VR in lowering foreign language anxiety and enhancing learner engagement.

In summary, the VR-based instructional design in this study operationalizes theoretical constructs through a structured sequence of tasks, where each stage (i.e., input, contextualized comprehension, output, and feedback) systematically supports language acquisition within an immersive and culturally meaningful environment.

Research Method

This study adopted a quantitative research methodology and employed a pre-experimental design in the form of a one-group pretest–posttest design. In this design, the same group of participants was assessed before and after the VR-based instructional intervention, allowing the researcher to examine changes in listening and speaking performance attributable to the treatment. The one-group pretest–posttest design is commonly used in educational research when random assignment and control groups are not feasible (Creswell, 2014).

The one-group pretest – posttest design was selected due to practical constraints in the school setting, where random assignment and the establishment of a control group were not feasible. This design is commonly used in educational research to examine within-group changes over time (Creswell, 2014).

To partially address potential threats to internal validity, several measures were taken. First, the intervention was conducted within a relatively short time frame, which helps reduce the likelihood of maturation effects. Second, identical test formats were used in the pretest and posttest to ensure consistency in measurement. Third, all participants received the same instructional treatment under controlled classroom conditions.

In addition, to strengthen the validity of the findings, qualitative data were collected through student reflections following the VR learning session. These data provide supplementary evidence to support the interpretation of quantitative results.

Participants

The participants in this study were 45 Thai secondary school students aged between 13 and 18 years old who come from Niva American International School in Bangkok. The mean age of the participants was 14.33 years (SD = 1.638), with a median age of 13, indicating that the sample tended toward the lower end of the secondary school age range and primarily consisted of lower secondary learners.

In terms of Chinese language learning experience, participants had studied Chinese as a foreign language for between 5 and 11 years. The average length of study was 6.11 years (SD = 1.584), with a median of 5 years. These figures suggest that most students possessed a foundational level of Chinese proficiency, while variation in years of study remained moderate across the sample.

The sample was selected using purposive sampling, as the researcher intentionally recruited students who met the study's criteria and were accessible for participation in the VR-supported instructional activities.

■ *Research Design*

Conceptual framework

As shown in Figure 1, this study is guided by a conceptual framework in which VR-based instruction serves as the independent variable influencing students' Chinese language learning outcomes. The framework is grounded in constructivist learning theory, as well as input and output hypotheses, which emphasize the importance of meaningful interaction, comprehensible input, and language production in second language acquisition.

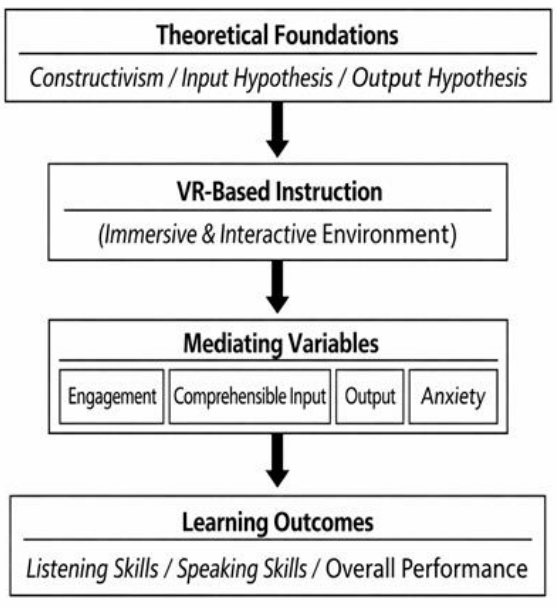

Figure *1*. Conceptual framework

Within this framework, the effects of VR instruction are mediated by several key learning mechanisms, including learner engagement, exposure to comprehensible input, opportunities for

meaningful output, and reduced language anxiety. These mediating variables explain how immersive and interactive VR environments facilitate language learning.

As a result, these processes contribute to improvements in students' listening skills, speaking skills, and overall language performance, which are measured through pretest and posttest assessments.

Instructional Procedure

Figure 2 illustrates the instructional procedure employed in this study, organized as a structured sequence within a single experimental cycle. The process consisted of six stages: pretest, lesson introduction, VR-based learning tasks, posttest, interactive discussion, and statistical analysis.

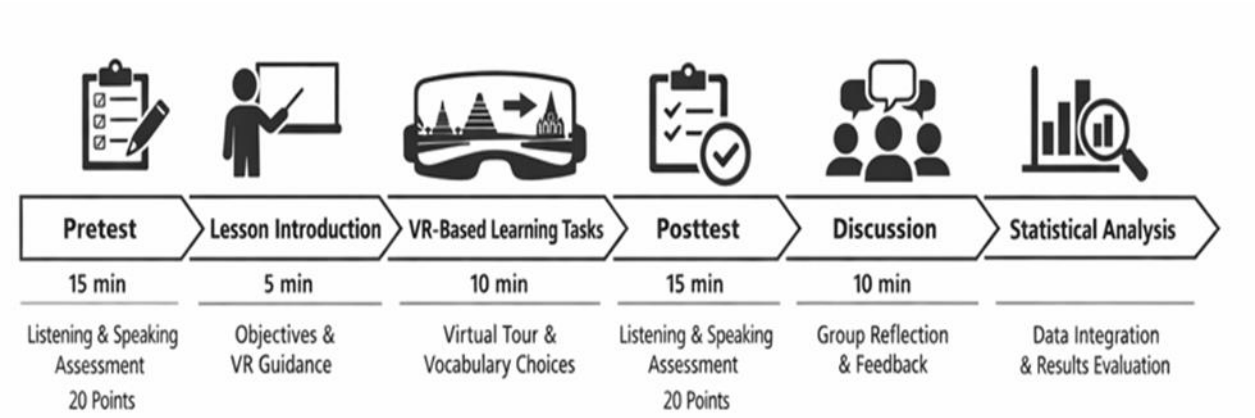

Figure *2*. Instructional Procedure

First, participants completed a 15-minute pretest assessing their Chinese listening and speaking abilities. This was followed by a 5-minute lesson introduction, during which the teacher explained the learning objectives and provided guidance on the use of the VR system.

Next, students engaged in a 10-minute VR-based learning session, where they completed guided exploratory tasks in a screen-based virtual environment. During this stage, learners navigated through panoramic scenes, interpreted directional instructions, and selected appropriate spatial directions, with immediate feedback provided to support learning.

Following the VR session, students completed a 15-minute posttest using the same format as the pretest to allow for direct comparison of performance changes. After the assessment, a 10-minute interactive discussion was conducted to facilitate reflection and reinforce learning through teacher feedback and peer interaction.

Finally, all collected data were compiled and subjected to statistical analysis to examine differences between pretest and posttest performance.

Each stage of the VR-based instructional procedure was designed to reflect key principles of second language acquisition, including comprehensible input, contextualized learning, meaningful output, and iterative feedback.

Specifically, the vocabulary preview and guided VR narration provided multimodal comprehensible input, while the directional decision-making tasks required learners to actively process input and produce responses within a meaningful context.This structured alignment between instructional design and theoretical principles enhances the pedagogical validity of the VR-based learning activities.

VR System Interface and Learning Tasks

As shown in Figure 3, the VR-based instructional session was designed to facilitate the acquisition and application of Chinese locational vocabulary, specifically the directional terms *east (东), south (南), west (西), north (北), southeast (东南), southwest (西南), northeast (东北),* and *northwest (西北).* The instructional design integrated vocabulary learning, contextualized comprehension, guided navigation tasks, and interactive elements within a culturally situated virtual environment modeled on Ayutthaya.

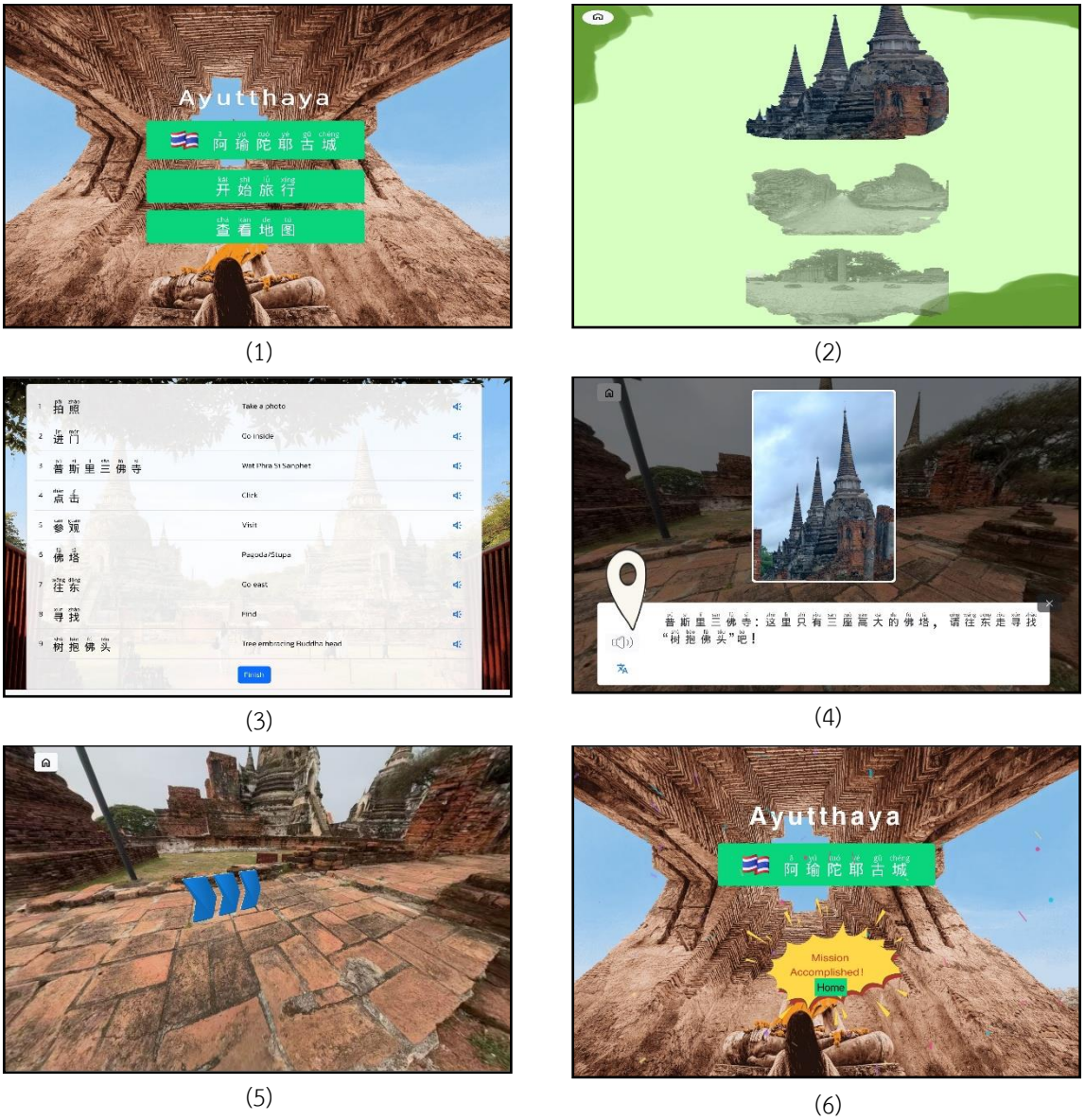

Figure *3*. User Interface of the VR-Based Chinese Learning Environment of Ayutthaya Historical Park

Upon entering the application interface, learners were presented with two primary options: “开始学习Start Learning” and “查看地图View Map.” Selecting “开始学习Start Learning” directed learners to the first panoramic VR map. Each map represented a different area within the virtual Ayutthaya environment, and a total of five interconnected maps formed the complete learning route. This dual-option interface supported learner autonomy by allowing flexible navigation prior to formal task engagement.

At the beginning of each map, learners were introduced to the target directional vocabulary relevant to spatial navigation within the environment. Learners could repeatedly listen to the pronunciation of each word and access its English translation. This multimodal presentation (visual text, audio input, and bilingual support) was designed to reinforce phonological recognition and semantic understanding.

After completing the vocabulary preview, learners proceeded to the interactive map environment. An AI-supported virtual guide provided contextualized explanations of the architectural site where the learner was positioned. The guide’s narration could be replayed by clicking an audio icon, and learners were able to view the English translation of the explanatory script. This stage emphasized listening comprehension within a culturally meaningful context.

Following the guided explanation, learners were required to apply their understanding of directional vocabulary by selecting the correct spatial direction to proceed to the next location. If the selected direction was correct, the learner advanced to the subsequent map. If incorrect, the learner was prompted to review the guide’s explanation and attempt the directional selection again. This immediate feedback mechanism functioned as formative assessment and reinforced vocabulary application in context.

To enhance learner engagement and motivation, additional interactive elements were embedded within the environment, including mini-tasks such as locating a hidden mobile phone and clicking on architectural structures to view authentic images of the real buildings. These gamified components were intended to sustain attention and encourage exploratory interaction.

After successfully navigating through all five maps of the virtual ancient city, learners completed the learning sequence. Throughout the session, learners retained the flexibility to revisit maps, review vocabulary, or return to the main interface as needed. This design supported self-paced learning and repeated exposure to linguistic input.

Overall, the instructional procedure followed a structured progression from vocabulary input to contextualized listening, guided comprehension, applied spatial decision-making, and interactive reinforcement within a culturally immersive VR environment.

Instruments

Motivation Questionnaire

The motivation questionnaire was designed to assess students' learning motivation in the VR-supported Chinese learning environment. Content validity was established through expert evaluation, with I-CVI values of 1.00 for all items and an S-CVI of 1.00, indicating excellent content validity.

The internal consistency reliability of the questionnaire was examined using Cronbach's alpha, yielding a value of 0.81, which indicates good reliability.

Satisfaction Questionnaire

The satisfaction questionnaire was used to evaluate students' perceptions of the VR-based learning experience. Similar to the motivation questionnaire, content validity was established through expert evaluation. The I-CVI values ranged from 0.67 to 1.00, and the S-CVI was 0.95, indicating an acceptable to high level of content validity. Items with lower I-CVI values were reviewed and revised prior to final implementation.

In addition, the internal consistency reliability of the questionnaire was examined using Cronbach's alpha, yielding a value of 0.97, which indicates excellent reliability.

Listening and Speaking Tests

The listening and speaking tests were developed based on the instructional content of the VR learning activities. The listening test assessed students' ability to comprehend directional vocabulary and contextualized instructions, while the speaking test evaluated their oral production in terms of pronunciation, accuracy, and fluency.

Content validity of the test items was established using the Index of Item-Objective Congruence (IOC). Three experts evaluated the alignment between test items and learning objectives. The IOC values ranged from 0.67 to 1.00, indicating acceptable to high consistency with the instructional objectives. Items with lower IOC values were reviewed and revised prior to final implementation.

Items with lower IOC values were reviewed and revised prior to final implementation. To ensure consistency in evaluation, a scoring rubric (Figure 4) was used for the speaking assessment, focusing on pronunciation, accuracy, and fluency.All responses were evaluated consistently based on the predefined rubric.

Score	Pronunciation	Accuracy	Fluency
5 (Excellent)	Clear and accurate pronunciation with minimal errors; easily understood.	Uses correct vocabulary and sentence structures appropriate to the task; responses are relevant and complete.	Speaks smoothly with little hesitation; ideas are well connected.
4 (Good)	Generally clear pronunciation with minor errors that do not affect understanding.	Mostly correct vocabulary and structures; minor grammatical errors.	Some hesitation, but speech is generally smooth and understandable.
3 (Satisfactory)	Noticeable pronunciation errors but meaning is generally understandable.	Limited vocabulary; some grammatical errors that occasionally affect clarity.	Frequent pauses; speech is somewhat fragmented.
2 (Limited)	Frequent pronunciation errors that make understanding difficult.	Frequent errors in vocabulary and grammar; responses are partially relevant.	Hesitant and slow; difficulty forming complete sentences.
1 (Poor)	Pronunciation errors severely hinder understanding.	Inappropriate or incorrect use of vocabulary and structures; responses are unclear or irrelevant.	Speech is very limited or incomplete.

Figure *4*. Scoring rubric of speaking part

Qualitative Data

To complement the quantitative findings, qualitative data were collected through brief student reflections following the VR-based learning session. Students were asked to describe their learning experiences, perceived benefits, and challenges when using the VR environment.

The qualitative responses were analyzed thematically to identify common patterns related to engagement, comprehension, and speaking practice. These qualitative insights were used to triangulate the quantitative results and provide a more comprehensive understanding of the effectiveness of the VR-based instruction.

For example, one student reported that “The immersive real-life scenarios made learning Chinese more practical and engaging, motivating me to practice more.” highlighting the role of multimodal input in supporting comprehension. Another student noted that “the digital tools,like VR,allow more creativity” suggesting a reduction in language anxiety.

■ *Data Analysis*

Justification of Method

The one-group pretest–posttest design was chosen because it allows researchers to measure learning gains within the same group of participants, which is particularly useful when random assignment and control groups are not feasible in classroom settings. Although this design has limitations in controlling all threats to internal validity, it remains a widely accepted method in educational technology research where experimental control is constrained.

Theoretical Support for VR Intervention

The use of Virtual Reality (VR) as an instructional tool in language education is supported by existing research indicating that virtual learning environments can enhance learner engagement and communicative competence. Although fully immersive VR systems using head-mounted displays (HMDs) have been widely discussed, research also suggests that non-immersive, screen-based virtual environments can provide meaningful contextualized interaction and experiential learning opportunities.

Prior studies have shown that virtual environments, whether immersive or desktop-based—can increase interaction, promote contextualized language practice, and support oral language development in second language learning settings. By situating learners in simulated communicative contexts, VR-based activities facilitate authentic input and output processes that are essential for language acquisition.

■ Results

Descriptive Statistics

Descriptive statistics were calculated to examine students' listening and speaking performance before and after the VR-based instructional intervention.

It can be seen from Table 1 that: For listening, the pretest mean score was 8.36 (SD = 2.773), with scores ranging from 3 to 15. Following the intervention, the posttest mean increased to 12.62 (SD = 2.146), with scores ranging from 8 to 15. The increase in mean score suggests notable improvement in listening comprehension. Moreover, the reduction in standard deviation indicates decreased variability in students' performance after the intervention.

Table 1.

Statistical Analysis of Variables

	N	Minimum	Maximum	Mean	Std. Deviation	Median
Age	45	13.000	18.000	14.333	1.638	13
Length of Chinese learning experience	45	5.000	11.000	6.111	1.584	5
Listening (Pre)	45	3.000	15.000	8.356	2.773	8
Speaking (Pre)	45	1.000	5.000	2.022	1.011	2
Total (Pre)	45	4.000	20.000	10.378	3.645	10
Listening (Post)	45	8.000	15.000	12.622	2.146	13
Speaking (Post)	45	1.000	5.000	2.756	1.048	3
Total (Post)	45	9.000	20.000	15.378	3.070	16

For speaking, the pretest mean score was 2.02 (SD = 1.011), with a range of 1 to 5. In the posttest, the mean score increased to 2.76 (SD = 1.048), with the same score range. Although the magnitude of improvement was smaller than that observed in listening, descriptive results indicate positive development in students' oral production ability.

Paired-Samples t-Test Results

To assess whether students' language proficiency significantly changed following the intervention, paired-samples t-tests were performed to compare pretest and posttest scores in listening, speaking, and overall performance. The paired-samples design was selected because the same group of participants was measured at two time points, allowing for direct within-subject comparison. Prior to conducting the analyses, assumptions of normality were examined to ensure the appropriateness of parametric testing.

Table 2.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Listening (Pre)	8.356	45	2.773	0.413
	Listening (Post)	12.622	45	2.146	0.320
Pair 2	Speaking (Pre)	2.022	45	1.011	0.151
	Speaking (Post)	2.756	45	1.048	0.156
Pair 3	Total (Pre)	10.378	45	3.645	0.543
	Total (Post)	15.378	45	3.070	0.458

Listening scores showed a statistically significant increase from pretest (M = 8.36, SD = 2.773) to posttest (M = 12.62, SD = 2.146), $t(44) = -15.460$, $p < .001$. Similarly, speaking scores increased significantly from pretest (M = 2.02, SD = 1.011) to posttest (M = 2.76, SD = 1.048), $t(44) = -11.000$, $p < .001$.

To examine whether the observed improvements were statistically significant, paired-samples t-tests were conducted. As shown in Table 3, listening scores increased significantly from pretest (M = 8.36, SD = 2.77) to posttest (M = 12.62, SD = 2.15), $t(44) = -15.46$, $p < .001$, $d = 2.30$, indicating a very large effect. Similarly, speaking scores showed a significant increase from pretest (M = 2.02, SD = 1.01) to posttest (M = 2.76, SD = 1.05), $t(44) = -11.00$, $p < .001$, $d = 1.64$, indicating a large effect. In addition, total scores increased significantly from pretest (M = 10.38, SD = 3.65) to posttest (M = 15.38, SD = 3.07), $t(44) = -16.96$, $p < .001$, d = 2.53, indicating a very large overall effect. These results demonstrate that the VR-based instructional intervention led to statistically significant improvements in students' listening, speaking, and overall Chinese language performance.

Table 3.

Paired Samples Test

		Mean Difference	SD	Std. Error Mean	*t*	N	*P*	Cohen's *d*
Pair 1	Listening (Pre) – Listening (Post)	-4.267	1.851	0.276	-15.460	44	0.000	2.30
Pair 2	Speaking (Pre) – Speaking (Post)	-0.733	0.447	0.067	-11.000	44	0.000	1.64
Pair 3	Total (Pre) - Total (Post)	-5.000	1.977	0.295	-16.964	44	0.000	2.53

■ Discussion and Contribution

The present study examined the impact of VR-based instruction on Thai secondary school students' Chinese listening and speaking performance within a culturally contextualized learning environment. The findings revealed statistically significant improvements in listening, speaking, and overall scores, with particularly strong gains in listening comprehension. These results are consistent with recent meta-analytic evidence indicating that virtual reality–assisted language learning produces positive effects on linguistic outcomes (Chen, Wang, & Wang, 2022). The substantial effect sizes observed in the present study further support the pedagogical potential of VR in secondary language classrooms.

The marked improvement in listening performance may be attributed to the multimodal and contextualized nature of the VR-supported tasks. Chen et al. (2022) note that VR environments enhance language acquisition by integrating visual, spatial, and auditory input, thereby supporting deeper processing of linguistic information. In the present study, learners were exposed to situationally embedded input rather than decontextualized audio materials. Such contextualization likely facilitated comprehension by providing semantic and visual scaffolding. Moreover, the reduction in standard deviation suggests that VR-based instruction may help narrow performance gaps among learners. This observation resonates with Golonka et al. (2014), who argue that technology-enhanced language learning environments can provide differentiated support and expand learning opportunities beyond traditional classroom constraints.

A notable finding of this study is that the improvement in listening was substantially greater than that in speaking. This difference can be explained by the nature of VR-based learning environments, which tend to provide input-rich, multimodal experiences that more directly support receptive skills. In the present study, learners engaged with panoramic scenes, guided explanations, and repeated access to audiovisual input, all of which align with Krashen's (1985) input hypothesis emphasizing the role of comprehensible input in language acquisition. The integration of visual, spatial, and auditory cues in the VR environment likely facilitated deeper processing of linguistic input, thereby enhancing listening comprehension.

In contrast, the development of speaking skills requires not only exposure to input but also sufficient opportunities for meaningful output and interaction. According to Swain and Lapkin (1995), language production plays a critical role in promoting cognitive processing and helping learners notice gaps in their interlanguage. However, in the present study, speaking opportunities were relatively limited, as learners primarily engaged in comprehension-based navigation and decision-making tasks rather than extended verbal interaction. This may explain why improvements in speaking, although statistically significant, were smaller in magnitude. These findings are consistent with previous research suggesting that VR environments tend to yield stronger effects on receptive skills than productive skills (Chen et al., 2022).

Another noteworthy aspect of this study is the use of a non-immersive, screen-based VR system rather than high-cost immersive hardware. The significant gains achieved indicate that pedagogical effectiveness depends not only on technological sophistication but also on instructional design. As Chun, Kern, and Smith (2016) emphasize, technology integration in language learning is most effective when grounded in meaningful communicative practice rather than technological novelty. The culturally contextualized design of the VR tasks, inspired by local historical heritage, further aligns with perspectives in Chinese as a Foreign Language (CFL) research that stress the inseparability of language and cultural understanding (Everson & Shen, 2010).

Nevertheless, it is important to acknowledge methodological limitations. The one-group pretest–posttest design, as discussed by Creswell and Creswell (2017), does not fully control for threats to internal validity. Potential influences such as testing effects or maturation cannot be entirely ruled out. However, the short duration of the intervention and the consistency of testing procedures help mitigate some of these concerns. In addition, qualitative data from student reflections provided supplementary evidence supporting the quantitative findings. For example, some students reported that the visual and auditory cues in the VR environment made it easier to understand directions, while fewer opportunities for speaking limited their oral practice.

Despite these limitations, the consistent statistical improvements across skill areas provide meaningful evidence for the instructional value of VR-supported learning in this context. Future research incorporating control groups, extended intervention periods, and more interactive speaking tasks would further strengthen the evidence base and enhance the effectiveness of VR-based language instruction.

Contribution

This study contributes to the growing body of research on technology-enhanced language learning by extending empirical evidence of VR-assisted instruction to the context of Chinese as a Foreign Language (CFL) in Thailand. While existing meta-analytic and review studies (e.g., Chen et al., 2022; Golonka et al., 2014) have documented the general effectiveness of technology-supported language learning, relatively few studies have focused on secondary-level CFL learners in Southeast Asia. By demonstrating significant improvements in listening and speaking within a Thai educational setting, this research broadens the contextual scope of VR language learning scholarship.

Importantly, the findings suggest that effective VR integration does not necessarily require fully immersive or high-cost hardware. The successful implementation of a non-immersive, screen-based virtual environment demonstrates that meaningful language gains can be achieved using widely accessible digital devices, thereby enhancing scalability and practical applicability in resource-constrained school contexts. In this regard, the study provides empirical support for pedagogically grounded technology integration, echoing Chun et al.'s (2016) argument that technology should serve communicative and educational purposes rather than function as an end in itself.

Pedagogically, the instructional design moves beyond simple vocabulary presentation by integrating spatial navigation tasks with directional language use. By requiring learners to interpret contextualized explanations and make directional decisions within a virtual environment, the system promotes embodied and contextually grounded language learning. The embedded immediate feedback mechanism further supports formative learning cycles, allowing learners to revisit explanations and refine their understanding through repeated exposure. This structured progression from input to guided comprehension and applied task performance offers a replicable instructional model for task-based CFL learning.

Furthermore, by embedding language learning within the historically and culturally meaningful setting of Ayutthaya, the study supports the view that CFL instruction should integrate linguistic and cultural dimensions (Everson & Shen, 2010). In this sense, VR functioned as a mediating tool that connected communicative practice with learners' sociocultural context, situating vocabulary learning within a heritage-based intercultural framework. Beyond measurable achievement gains, the observed reduction in performance variability and improvement in speaking performance suggest potential benefits for learner confidence and inclusive participation, aligning with research emphasizing the role of supportive technological environments in reducing language anxiety (Kaplan-Rakowski & Gruber, 2023).

Taken together, this study offers a context-sensitive, scalable, and pedagogically structured model for integrating VR into secondary CFL classrooms in Thailand and similar Asian educational settings. It demonstrates that thoughtfully designed virtual environments can enhance language proficiency while remaining accessible, sustainable, and adaptable to real-world classroom conditions.

Another important contribution of this study lies in the integration of VR technology with a culturally significant heritage context, namely Ayutthaya. Unlike many VR-based language learning studies that employ generic or decontextualized virtual environments, the present study situates language learning within a historically meaningful and locally relevant setting. As a former capital of Siam and a site of long-standing Sino-Thai cultural exchange, Ayutthaya provides a rich intercultural context that connects language learning with learners' own cultural and historical background.

This contextualization enhances the authenticity of the learning experience by allowing students to engage with language in a setting that is both meaningful and familiar. Through the VR environment, learners not only processed linguistic input but also interacted with culturally embedded content, which aligns with the view that language and culture are inseparable in foreign language education (Kramsch, 1993;

Everson & Shen, 2010). In this sense, VR functioned as a mediating tool that bridged linguistic knowledge and cultural understanding.

For Thai learners of Chinese, this approach is particularly valuable, as it situates the target language within a local cultural framework rather than a distant or abstract Chinese context. Such localization may enhance learner engagement, relevance, and intercultural awareness by highlighting historical connections between Thailand and China. Therefore, the combination of VR technology and cultural heritage-based design represents an innovative and context-sensitive approach to Chinese as a Foreign Language (CFL) education in Thailand.

■ Limitations and Recommendations

Despite the positive findings, several limitations should be acknowledged. First, the study employed a one-group pretest-posttest design without a control group, which may limit the ability to attribute improvements solely to the VR-based intervention. Potential threats to internal validity, such as testing effects and maturation, cannot be fully ruled out.

However, efforts were made to mitigate these limitations. The short duration of the intervention reduces the likelihood of substantial maturation effects, and consistent testing procedures were applied across both pretest and posttest. In addition, the inclusion of qualitative data provides supplementary evidence that supports the interpretation of the quantitative results.

The qualitative findings revealed that students perceived the VR environment as engaging and helpful for understanding contextualized language input, which aligns with the observed improvements in listening performance. Furthermore, students reported increased confidence when interacting in the VR environment, which may help explain the improvement in speaking scores.

Nevertheless, future research is encouraged to adopt more rigorous experimental or quasi-experimental designs, such as the inclusion of control groups or longitudinal data collection, to strengthen causal interpretations and examine the long-term effects of VR-based instruction.

■ References

Byram, M. (1997). Teaching and assessing intercultural communicative competence. Multilingual Matters.

Chamot, A. U., & O'Malley, J. M. (1994). The CALLA handbook: Implementing the cognitive academic language learning approach. Addison-Wesley Publishing Company.

Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. Sustainability, 14(6), 3147.

Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. The Modern Language Journal, 100(S1), 64–80.

Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). Sage Publications.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage Publications.

Everson, M. E., & Shen, H. H. (Eds.). (2010). Research among learners of Chinese as a foreign language (Vol. 4). National Foreign Language Resource Center.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2014). How to design and evaluate research in education (8th ed.). McGraw-Hill Education.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. Computer Assisted Language Learning, 27(1), 70–105.

Kaplan-Rakowski, R., & Gruber, A. (2023). The impact of high-immersion virtual reality on foreign language anxiety. Smart Learning Environments, 10(1), 46.

Kramsch, C. (1993). Context and culture in language teaching. Oxford University Press.

Li, J., Qiu, T., Li, C., Xu, C., Cheng, P., Tang, Y., & Georgiou, G. K. (2024). The effects of immersion in virtual reality environment on oral English learning for Chinese university students. Education and Lifelong Development Research, 1(1), 3–14.

Liddicoat, A. J., & Scarino, A. (2013). Intercultural language teaching and learning. Wiley-Blackwell.

Liu, W. (2025). A systematic review of the use of virtual reality in teaching Chinese as a foreign language. Journal of China Computer-Assisted Language Learning, 5(1), 130–155.

Song, C., Shin, S. Y., & Shin, K. S. (2023). Optimizing foreign language learning in virtual reality: A comprehensive theoretical framework based on constructivism and cognitive load theory (VR-CCL). Applied Sciences, 13(23), 12557.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. Applied Linguistics, 16(3), 371–391.