*Quest Journals*
*Journal of Education, Arts, Law and Multidisplinary*
*Volume 16 ~ Issue 2 (Mar. – Apr. 2026) pp: 85-96*
*ISSN(Online): 2347-2895*
www.questjournals.org

**Research Paper**

# The Application of ChatGPT Role Play in Improving Students' Willingness to Communicate: An Investigation into First-Year English Students at Fofl – Hust

Hoang Thi Minh Hanh, Tran Anh Huy, Ta Minh Quan, Ngo Tran Khanh Linh, Vu Phuong Anh
*Hanoi University of Science and Technology*

***Abstract***

*This study was conducted by second-year English students at Hanoi University of Science and Technology to investigate the application of ChatGPT role play in improving first-year students of faculty of foreign languages' willingness to communicate. Data were collected through online ChatGPT-based roleplay involving 15 participants with fundamental English proficiency. The research subjects will receive guidance on applying given prompts based on previous research on willingness to communicate to examine the change in their willingness to exchange information. The research includes several phrases, starting with choosing and gathering the right items, helping generate prompts and guiding participants on how to apply them correctly. A four-week experiment will be designed to track and follow participants' progress and data will be collected via pre and post test outcomes. Finally, a small survey will be conducted after the weekly experiment to gather student's feedback.*

***Keywords***
*ChatGPT, roleplay, prompts, four-week experiment*

## I. Introduction

The rapid rise of Artificial Intelligence (AI) promises to transform many aspects of life, from education and work to personal connections (Colleen & Brian, 2025). Over the next decade, artificial intelligence (AI) is expected to make substantial contributions across numerous sectors, yet its potential in education remains less frequently discussed compared to other fields (Holmes et al., 2019, p. 12). He also noted that AI had the potential to personalize learning experiences and provide data-driven feedback that could support both students and teachers. Recent findings about these fields have only focused on written languages but not on verbal communications. Writing apps with AI capabilities, including Grammarly, tools assist in improving writing style and tone, making the content more logical and suited ProWritingAid, and Paperrater, have transformed traditional teaching methods by providing students with individualized learning experiences and real-time feedback (Jati, 2022). These are for academic settings. Because AI can give immediate feedback, it is worth considering the potential advantage it may offer to language speakers. Therefore, the purpose of this research is to experiment the efficiency of ChatGPT role play in improving the willingness to communicate of first-year students within English major.

The purpose of this study is to analyze the influence of ChatGPT-facilitated role-play activities on the willingness of first-year students to communicate in English. Particular attention is given to variations in their confidence levels, anxiety, and interest in language use. Additionally, the research evaluates the efficacy of these activities in alleviating students' speaking-related anxiety.

## II. Literature review

### 2.1. ChatGPT

#### 2.1.1. Background information

ChatGPT is an artificial intelligence-based conversational AI created by OpenAI and introduced to the public in November 2022. It is based on the Generative Pre-trained Transformer (GPT), which includes GPT-3.5 and GPT-

4 models, all pre-trained on large amounts of text to produce human-like responses to given inputs. It applies transformer-based deep learning methods, such as the self-attention mechanism, to process contextual information and produce relevant and contextually appropriate text.

In recent years, the development of AI, specifically the voice chat and role-play features of the ChatGPT platform, has opened up a new approach to enhancing English as a Foreign Language (EFL) communication skills. The role-play feature allows learners to engage in diverse, real-life simulated scenarios with real-time, personalized feedback, thereby creating a context-rich practice environment. Notably, recent studies indicate that interacting with a "non-judgmental" AI system significantly reduces language anxiety—one of the most substantial psychological barriers for foreign language learners. For instance, according to Salsabil et al. (2025), AI Voice Chat cultivates a low-pressure learning space that directly enhance Willingness to Communicate WTC, transforming passive learners into more confident speakers. Although the pedagogical potential of this tool has been acknowledged, an in-depth empirical investigation into how ChatGPT Role-play specifically impacts the WTC levels of first-year English majors—who are in a transitional phase of building their foundational academic expertise—remains an area requiring further detailed study.

**2.1.2. The use of ChatGPT role play in English language learning**

There was a significant growth in the use of artificial intelligence (AI) chatbots in education (Hwang & Chang, 2023). A notable example in the field is ChatGPT, a pre-trained language model developed by OpenAI (Flanagin et al., 2023; Roumeliotis & Tselikas, 2023). "OpenAI proposed the GPT (Generative Pretrained Transformer) model based on the stacked Decoder structure in the Transformer architecture" (Liu et al., 2024, p.2). It is able to generate multimodal content from textual prompts, via an intermediate code generation step (Yejin, 2023). Role play is a form of experiential learning (Russell & Shepherd, 2010). A systematic review of empirical studies - finds that some of the uses of ChatGPT in English language teaching include using it as an interactive dialogue simulator which is essentially a role-play context; for example, allowing students to simulate conversations. (Philip et al., 2024).

One of the most commonly cited applications of ChatGPT's ability is that it provides real time feedback to students, a feature that enhances engagement and tailors learning experiences (Hong, 2023). Koraishi (2023), for example, underscored the efficiency of ChatGPT in automating assessment processes, providing immediate feedback and adapting to individual learner needs, thereby enhancing its overall effectiveness in education.

Van Horn (2024) discussed the impact of ChatGPT, highlighting advantages such as personalized feedback and increased learner engagement. Notwithstanding, these earlier investigations predominantly employ document research methods, encompassing existing literature reviews, while lacking experimental or qualitative data from the perspectives of teachers and learners. In addition, ChatGPT's ability to provide translated explanations of grammar mistakes may facilitate learners' understanding of grammatical structure, thereby supporting English language acquisition. In fact, over 85% of the first-year students in Schmidt-Fajlik (2023)'s research advocated its ease of use and usefulness in English writing learning. In line with this finding, Bonsu and Baffour-Koduah(2023) discovered that students expressed support for the integration of ChatGPT at universities, accompanied by an effective management system, as students found ChatGPT valuable to alleviate the stress of generating ideas while doing scientific research.

However, these previous studies have yet to adequately consider students' attitudes towards English language learning. During speaking tasks, L2 learners frequently experience foreign language anxiety (FLA), an affective barrier characterized by situational tension and fear that significantly reduces their oral fluency and communicative engagement (Tania et al., 2025). There is a correlation between high levels of FLA and learners' WTC, with anxiety being associated with decreased initiating or continuing discourse in the target language. Empirical studies indicate that communication anxiety, a key component of FLA, decreases participation in classroom speaking activities and lowers overall WTC. Additionally, studies based on WTC modeling of EFL learners demonstrate that face-to-face and digital communication anxiety are both associated with reduced comprehensibility in communicating (Zarrinabadi, 2014; Khajavy, MacIntyre & Barabati: 2018). The study indicates that reducing speaking anxiety could be crucial in increasing learners' active involvement and confidence in L2 communication (Dewaele & MacIntyre, 2016).

**2.2. Willingness to Communicate (WTC)**

**2.2.1. Definition of Willingness to Communicate (WTC)**

WTC is the chance that a person chooses to speak when they have the opportunity (MacIntyre et al., 1988). In language learning, it is not fixed but can change depending on both personal traits and the situation (MacIntyre et al., 1998). The Pyramid Model explains that speaking is affected by many layers of factors, from stable ones like personality and learning background to more immediate ones such as confidence, mood, and who the person is talking to. Because of this, even learners with good language knowledge may still stay silent. In the end, they will only speak when they feel ready and willing at that moment.

**2.2.2 WTC Framework**

The concept of WTC is used as the theoretical foundation of this study. It was first introduced by James C. McCroskey and Virginia P. Richmond in 1985. Both are well-known researchers in the field of communication studies, focusing on how individuals behave in different communication situations, especially in relation to anxiety and communication confidence. Later, the concept was developed further by Peter D. MacIntyre and his colleagues in 1998. They applied WTC to second language learning and defined it as a learner's readiness to communicate in a second language in specific situations. This development made WTC an important concept in language learning research.

The WTC framework explains how a learner decides to speak in a second language at a given moment. According to Peter D. MacIntyre et al. (1998), WTC is influenced by multiple factors rather than a single cause. The framework is often described as a pyramid, with actual communication at the top, followed by willingness to communicate. This decision is directly affected by immediate factors such as confidence and anxiety. It is also influenced by situational factors, such as who the learner is speaking with and the type of activity. At the base are more stable factors like personality, motivation, and language ability. Overall, the framework shows that a learner's willingness to speak depends on both short-term and long-term influences.

This framework is appropriate for the study for several reasons. First, it directly measures the main focus of the research, which is students' willingness to communicate. Second, it allows the researcher to examine communication in different scenario-based situations, which matches the design of role-play activities. Finally, ChatGPT-based role-play can help students practice in a less stressful and more supportive environment, which may increase their confidence and reduce anxiety. As a result, WTC is a suitable framework to evaluate the effectiveness of this approach in improving students' communication at FOFL - HUST.

In addition, WTC focuses on factors like how confident students feel, how nervous they are when speaking, and how well they think they can communicate. These are the same aspects that GPT-based role-play activities may improve, since students can practice in a more relaxed and supportive environment.

Finally, using a pre-test and post-test helps show whether students become more willing to speak after the activities. This makes it possible to clearly evaluate how effective the ChatGPT-based role-play is in improving their participation in communication.

**2.3 The application of ChatGPT role-play in improving WTC.**

With the rapid development of Al technology, ChatGPT is becoming a useful tool for people in many fields in which it equips people with medical information, provides useful scientific data and become a steadfast companion for language enthusiasts. Previous findings on chatbots as conversational partners have contributed to our understanding of how learners engage in computer-mediated oral practice (Ayedoun et al., 2015). AI role-play chatbot helps reduce anxiety in speaking and create a more comfortable space for learners to practice communication (Ayedoun et al.,2015; Carayannopoulos, 2018; Kim & Su, 2024).

Earlier research has been conducted in diverse educational settings, including a university EFL classroom, high school, and online instructional environments. Participants engaged in ChatGPT-mediated role-play tasks over several weeks. According to Du and Daniel (2024), by giving feedback and acting as partners, chatbots create a continuous language learning experience.

Overall, findings consistently show that ChatGPT role-play improves confidence, reduces anxiety. Learners also report positive attitudes toward AI. Moreover, an accessible approach allows learners to practice speaking anytime and anywhere, encouraging independent learning and engagement beyond the traditional classroom (Dewi, Nur'Aini, & Suwarti, 2022). However, learners sometimes rely too much on chatbot prompts instead of constructing their own expressions ( Ayedoun et al., 2015). Despite the drawback, the article suggests that ChatGPT role-play is a promising and pedagogically valuable tool for fostering WTC in language learning environments.

### 2.4. Research Gap

In a different glance, despite the above-mentioned benefits of ChatGPT, several limitations of current studies remain. Yıldız (2024) demonstrated that ChatGPT-assisted communicative activities significantly improved EFL learners' speaking self-efficacy; however, the study did not examine whether this increased confidence translated into greater WTC in authentic interaction contexts. In a similar vein, Núñez et al. (2025) observed that while ChatGPT Voice significantly enhanced learners' pronunciation, fluency, vocabulary, and grammar, their research lacked the incorporation of structured role-play tasks—a method well-known for fostering authentic communicative engagement and reducing anxiety. Across these studies, scholars have mainly focused either on psychological effects (e.g., self-efficacy) or on voice-based affordances for speaking practice, leaving the combined effect of ChatGPT-mediated role-play activities on learners' communicative willingness largely unexplored. Notably, within the specific educational context of the Faculty of Foreign Languages (FOFL) at Hanoi University of Science and Technology (HUST), this topic remains empirically under-researched. Consequently, there is a critical need to examine how role-play with ChatGPT might influence first-year English majors' WTC, a gap this study seeks to address to fulfill both theoretical and contextual voids.

## III. Research Methodology

### 3.1. Research method

#### 3.1.1. Mix method approach

This study employs a mixed-methods approach to comprehensively evaluate the impact of ChatGPT role-play activities on students' WTC (Creswell & Plano Clark, 2018). To measure specific variations in students' interaction levels and willingness to communicate across specific scenarios—based on MacIntyre's (2001) modified questionnaire framework—quantitative data is collected through a pre-test and post-test system utilizing a Likert-scale questionnaire. To decode the psychological dynamics underlying the statistical figures, qualitative data is simultaneously gathered through two channels: chat log analysis to determine interaction frequency, utterance length, and conversation maintenance strategies, alongside semi-structured interviews to explore students' perceptions regarding the alleviation of their fear of making mistakes and their enhanced self-confidence (Patton, 2015). This triangulation between statistical data and personal experiences ensures validity and provides a multidimensional perspective on the transformation of learners' communicative behavior.

### 3.2. Data Collection Method

#### 3.2.1. Experiment Design

A four-week experiment will be conducted in which 15 students gather three times a week for a brief speaking session. Each session is planned to last between 45 minutes to 1 hour maximum. Throughout the four-week experiment, participants are provided with 8 practical scenario-based items that are covered in" Willingness to Communicate, Social Support, and Language-learning Orientations of Immersion Students" (MacIntyre, 2001). In the beginning of every session, the leader of the team will clearly and briefly explain basic information about ChatGPT roleplay and introduce the prompt that participants will use that day. The instruction in each session has 3 main components. The first one is greeting, the second one is the demonstration of how to use prompts before the participants begin the activities and the last one is giving participants suggestions, open-ended questions to overcome communication breakdowns with ChatGPT.

Moreover, Students will be encouraged to express freely and each session will be recorded for subsequent analysis. The process is carried out under our supervision and there will be immediate support from the conductor if any troubles arise during the session.

### 3.2.2. Semi-structured Interview

In the final stage of data collection, semi-structured interviews were conducted after the completion of the entire intervention with five participants. Individual interviews were employed to collect in-depth data on each participant's experiences and perceptions.

This method was selected because it provides a balance between consistency and flexibility, allowing the researchers to ask a set of predetermined questions while also probing deeper into individual responses when necessary (Creswell, 2012).

The interviews were carried out after the post-test phase. Vietnamese was used as the primary language during the interviews to ensure that participants could fully express their thoughts without being limited by their English proficiency. This choice helped enhance the depth and authenticity of the collected data.

The interview protocol was adapted from Xu et al. (2024) consisted of a series of open-ended questions developed based on the research objectives. The questions focused on several key aspects, including: (1) psychological pressure and anxiety in traditional speaking classes, (2) differences between interacting with ChatGPT and real-life communication, (3) the influence of a non-judgmental environment on confidence and fear of making mistakes, (4) changes in willingness to initiate communication, (5) perceptions of feedback and learning support, (6) technological and interactional limitations, and (7) future intentions to continue using ChatGPT.

During the interviews, follow-up questions were used when necessary to clarify participants' responses and encourage further elaboration. All interviews were audio-recorded and later transcribed to ensure accuracy and reliability in data handling.

## 3.3. Sampling

### 3.3.1. Participants

The participants in this program are 15 first-year students from the Faculty of Foreign Languages (FOFL) who are learning basic English speaking skills. To find these students, the study uses the Snowball sampling method. According to Nikolopoulou (2022), snowball sampling is a way of choosing people where the current participants find and invite new people to join the group. This means that the first few students in the program will ask their classmates or friends to become part of the sample as well. The group gets bigger as more students recruit others, just like a snowball growing as it rolls. This is a simple way to gather a specific group of students who are all focused on the same goal of improving their speaking.

### 3.3.2. Snowball Sampling

A snowball sampling technique, a form of non-probability sampling (Patton, 2002), was strategically employed for participant recruitment. Initial participants were identified within the HUST freshman community; these individuals then referred peers who shared similar linguistic barriers and a desire to improve their oral proficiency. This method was chosen to ensure a high degree of participant commitment and homogeneity, which is vital for maintaining the integrity of a focused, intensive intervention (Cohen, Manion, & Morrison, 2017). By utilizing peer networks, the researchers were able to reach students who might otherwise be hesitant to participate in a study focusing on their perceived communicative weaknesses (Flick, 2018).

## 3.4. Data analysis

### 3.4.1. Quantitative data analysis

A four-week experiment was conducted with first-year language students to measure their WTC. Two instructional sessions were implemented per week, involving 15 participants, in order to examine patterns of change in their self-reported willingness to participate in communication during class time. Data were collected across eight sessions to capture potential differences over time within a short-term instructional context. A control group was not included, as the experiment was designed to observe within-group fluctuations rather than to establish cause–effect relationships. The overall design reflects an exploratory approach to understanding the dynamic and situational characteristics of willingness to communicate in language learners.

To analyze the collected data, SPSS (Statistical Package for the Social Sciences) was employed due to its effectiveness in handling quantitative data and conducting statistical tests. The researchers applied a paired-samples t-test to evaluate changes in participants' willingness to communicate before and after their participation in the instructional sessions. This procedure was selected to compare data obtained from the same group of students at two different time points, thereby accounting for within-subject variation. By examining the differences between pre- and post-intervention measures, the analysis aimed to determine whether any observed change in WTC was statistically significant rather than attributable to chance.

In addition, SPSS was used to compute descriptive statistics such as mean scores and standard deviations, providing an overview of participants' performance across sessions. The use of SPSS ensured accurate data processing, minimized calculation errors, and enhanced the reliability and transparency of the findings. The paired-samples design, combined with SPSS analysis, was therefore appropriate for examining changes within the same 15 participants over the course of the study.

**3.4.2 Qualitative data analysis**

To comprehensively evaluate the impact of ChatGPT role-play on students' situational WTC, the qualitative data analysis is conducted in two distinct but complementary phases, enabling robust methodological triangulation.

First, the study analyzes the two-way chat history using a modern Human-AI Conversation Analysis (CA) framework (Choi & Oh, 2026). This approach adapts traditional CA (González-Lloret, 2011) to the inherently asymmetrical nature of human-AI interaction. Drawing upon Mehan's (1979) IRE framework and Schegloff's (2007) principles of sequence organization, student turns are systematically coded into three distinct communicative acts: Initiation (I), Response (R), and Evaluation (E). In accordance with the theoretical stances of Markee (2000) and Choi & Oh (2026), a developmental shift from merely producing passive responses (R) to proactively generating initiating (I) and evaluating (E) turns serves as tangible behavioral evidence of enhanced interactional competence and situational WTC. Specifically, the student's use of Evaluation (E)—traditionally a teacher-led act in classroom discourse—demonstrates a proactive agency in assessing and negotiating meaning with the AI system.

Second, to decode the underlying cognitive drivers behind these behavioral fluctuations, the semi-structured interview dataset is evaluated utilizing Thematic Analysis (Braun & Clarke, 2006). This systematic framework identifies and interprets patterns of meaning regarding the students' psychological experiences. By distilling themes related to interactional confidence, psychological barriers, and perceived safety, this phase captures the subjective factors influencing the learners' willingness to engage.

Ultimately, this dual-method approach allows for methodological triangulation, cross-validating the objective behavioral shifts observed in the chat logs with the students' subjective, self-reported psychological experiences. This provides a comprehensive and empirically grounded understanding of the students' WTC development in an AI-mediated environment.

**3.4.2.1 Chat history**

To provide a deeper understanding of the psychological shifts underlying the survey scores, this study conducted a qualitative analysis of the chat histories generated between the students and ChatGPT. These logs provide direct evidence of participants' interactional competence (Markee, 2000) and allow for the observation of longitudinal changes in communicative behavior, such as utterance length and lexical precision (Skehan, 1998). Following Purposive Sampling (Patton, 2015), three representative participants—ID 4, ID 10, and ID 5—were selected to illustrate three distinct WTC transformation trajectories: remarkable progress, moderate fluctuation, and minimal change.

The study employs a specialized Human-AI Conversation Analysis (CA) framework (Choi & Oh, 2026) to examine Communicative Agency—the learner's capacity to exercise agency in directing the interaction, rather than remaining passive (Ahearn, 2001). The coding procedure follows an adaptation of Mehan's (1979) IRE framework, integrated with Schegloff's (2007) principles of sequence organization. Every student turn is meticulously categorized into three functional codes:

- Response (R): Passive turns where the student merely provides conditional replies to AI prompts to fulfill the assigned tasks. A high frequency of "R" codes often signifies a low threshold of WTC.
- Initiation (I): Proactive acts where the student introduces new topics or seeks clarification. This indicates reduced linguistic inhibition and a transition toward taking control of the interaction.
- Evaluation (E): The most distinct indicator of interactional competence. When a student evaluates or reacts to the AI's input (e.g., *"I see," "That is a great point"*), they are engaging in the negotiation of meaning. This role reversal suggests that the student feels psychologically safe enough to interact with the AI as a peer-like partner.

Finally, the data is cross-analyzed across Social and Academic contexts to validate the theoretical argument that WTC is a situational construct, shifting dynamically according to the perceived complexity and safety of the interactional environment (MacIntyre et al., 2001). This triangulation process confirms whether the students' self-reported confidence translates into functional communicative performance within the digital domain.

**3.4.2.2 Thematic analysis**

The researchers decided to employ thematic analysis to scrutinize the qualitative data for several compelling reasons. According to Braun and Clarke (2006), thematic analysis is a highly flexible approach that allows researchers to identify, analyze, and report patterns or "themes" within data without being tied to a specific, rigid theoretical framework. This flexibility was particularly beneficial for this study as it allowed for a comprehensive exploration of the diverse psychological factors influencing students' WTC when interacting with ChatGPT. This method is particularly appropriate for the current study as the data were collected through semi-structured interviews, which generate rich and complex qualitative responses. Since the research aims to explore students' perceptions, experiences, and psychological changes in WTC rather than test predefined hypotheses, a flexible and interpretive approach like thematic analysis is highly suitable. Furthermore, thematic analysis is widely recognized for its accessibility, making it an ideal choice for researchers to systematically organize and summarize large volumes of interview data while highlighting key insights (Braun & Clarke, 2006). By using this method, the study could effectively capture the nuances of students' perceptions, from their reduced anxiety to the technical challenges they faced during the AI role-play sessions.

To ensure a rigorous and transparent analysis, the study followed Braun and Clarke's six-phase framework. The procedure was conducted manually to maintain a close connection with the participants' original voices. First, the researchers familiarized themselves with the data by transcribing the audio recordings and reading the transcripts multiple times while taking initial notes. In the second phase, initial codes were generated by identifying and labeling meaningful segments of the text with descriptive tags related to WTC and AI interaction. These codes were then sorted and combined into potential themes that captured broader patterns across the five participants (P1-P5). During the fourth and fifth phases, these themes were reviewed, refined, and clearly defined to ensure they accurately represented the data set. Finally, the analysis concluded with the production of a narrative report, where the identified themes were illustrated with direct quotes from the students to provide a cohesive and persuasive account of the findings.

## IV. Research finding

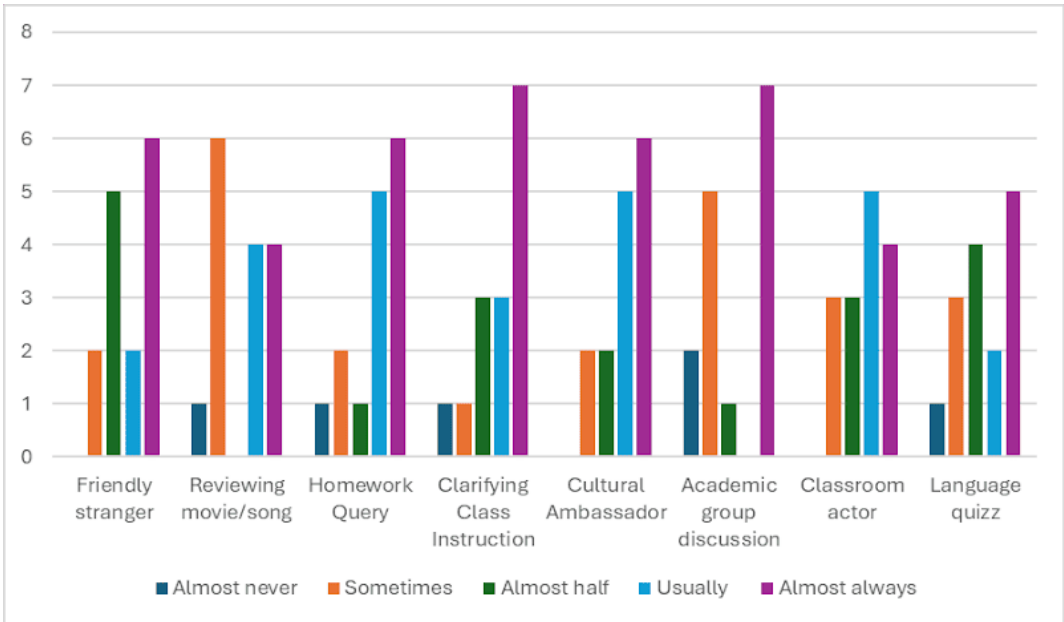


**Figure 1:** ***Students' WTC Levels before the ChatGPT Role-play Intervention (Pre-test).***

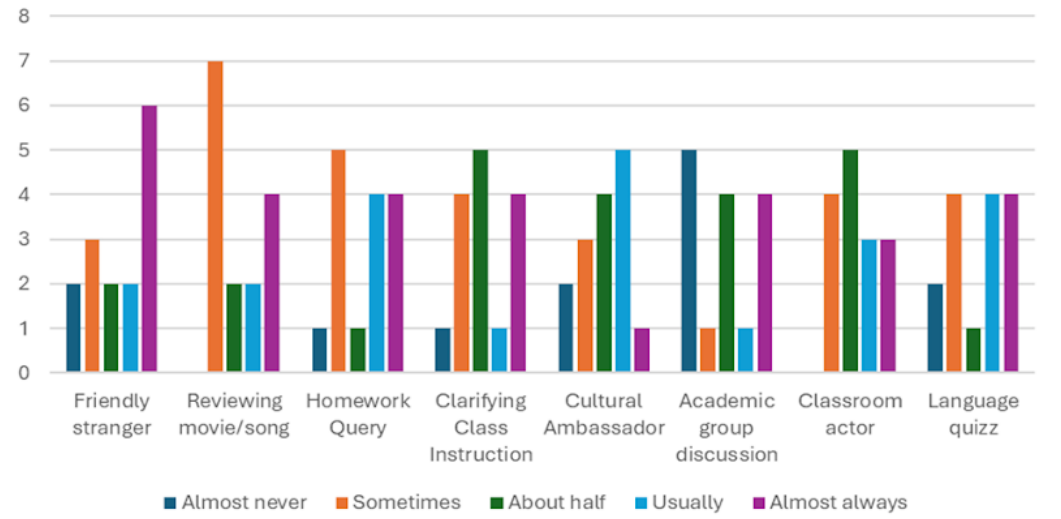


**Figure 2:** ***Students' WTC Levels after the ChatGPT Role-play Intervention (Post-test).***

| Assessment | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|
| Pre-test | 3.21 | 0.93 | 0.24 |
| Post-test | 3.67 | 0.75 | 0.19 |

**Figure 3** ***Descriptive Statistics for Pre-test and Post-test WTC Scores (N = 15)***

| Pair | Mean Difference | t | df | Sig.(2-tailed) |
|---|---|---|---|---|
| wtc_pre - wtc_post | -0.46 | -2.615 | 14 | 0.020 |

**Figure 4** ***Paired-Samples T-test Results for WTC Scores***

**4.1. Research Question 1: Specific Aspects of Willingness to Communicate Affected by ChatGPT Role-play**

Synthesized data from both quantitative and qualitative sources show that speaking anxiety and communication confidence are the two aspects of WTC most strongly affected by the ChatGPT role-play practice.

The main evidence for this change comes from the student interviews. Participants consistently said the AI was a safe environment that removed the fear of being judged by others. For example, P1 and P2 clearly stated that they no longer worried about making grammar mistakes or forgetting their ideas when talking to a computer. This helped them feel much more confident.

These interview findings match the statistical results from Lesson 6 (Academic Group Discussion). In this lesson, the number of students who were "Almost never" willing to communicate dropped from six to just two. This shows a huge improvement in situations that used to cause the highest levels of anxiety.

The increase in confidence is also clear when looking at how students actively participated in the chat logs. Case 1 (Student ID 4) is a great example of this confidence in action. Instead of just giving short answers (Code R), this student actively asked questions (Code I) and shared personal opinions (Code E), even during difficult academic topics. This active participation proves that the student's fear of making language mistakes had been successfully reduced.

In addition, the immediate language help from the AI—like suggesting better words—directly improved the students' vocabulary. P3 said learning new words was the main reason she felt more ready to start a conversation, showing how technical help leads to better communication.

However, the results regarding personal interest (intrinsic motivation) were mixed. Some students were motivated by the AI's game-like feedback and high scores. But the chat log of Student ID 5 showed someone just trying to finish the assignment. Even though this student had good grammar skills, they only answered the AI to complete the task. They refused to expand the conversation or show any real interest.

This finding is supported by comments from P4 and P5, who felt the AI was robotic and unnatural. In the end, this suggests that while ChatGPT is great at reducing fear and building confidence, its ability to create genuine interest is limited because it lacks real human empathy and body language.

**4.2. Research Question 2: The extent of Speaking Anxiety Reduction through ChatGPT Role-play**

The study confirms that the use of ChatGPT role-play helps reduce students' speaking anxiety to a substantial and statistically significant extent. This change is objectively proven by the paired-samples t-test results, which show a significant rise in WTC mean scores from 3.21 to 3.67 ($p = 0.020$). Most importantly, the calculated Cohen's d value of 0.675 signifies a medium-to-large effect size. This provides empirical confirmation that the intervention had a considerable practical impact on the students' affective states. Furthermore, the reduction in standard deviation from 0.93 to 0.75 suggests that the students' communication confidence became more stable and cohesive as a group following the intervention.

Visual data from the eight-lesson distribution charts provides a granular look at the effectiveness of this support across varying levels of social and academic pressure. In high-anxiety scenarios, such as Lesson 4 (Clarifying Class Instruction), the number of students choosing the "Almost always" willing category rose from only two in the pre-test to six in the post-test. This quantitative shift demonstrates that the intervention successfully moved the majority of the cohort away from moderate hesitation toward active initiative. The reduction in anxiety is also highlighted by Lesson 2 (Reviewing a movie/song), where an initial mental block felt by eight students was removed, leading to a substantial migration toward the "Usually" and "Almost always" willing categories.

Qualitative evidence reinforces that the decrease in anxiety was sufficient to facilitate real-world transferability. The confidence built in the simulated environment was not confined solely to the AI interaction. For example, P2 reported a clear behavioral shift, stating he is now ready to take the initiative and talk to foreigners. Similarly, P3 mentioned being less hesitant to answer her teacher during actual class discussions.

While technical limitations such as voice recognition errors (noted by P1) and repetitive prompts (noted by P5) presented some hurdles, they did not negate the overall positive trend. Therefore, it can be concluded that

ChatGPT role-play serves as a functional preparatory tool that effectively alleviates communicative apprehension and empowers first-year English majors to engage in L2 interactions more frequently and with less fear.

## V. Discussion

The study was conducted to explore the impact of ChatGPT role-play activities on the Willingness to Communicate (WTC) of English majors at Hanoi University of Science and Technology (HUST). Both qualitative and quantitative results indicate that these activities have a significantly positive effect on learners' confidence and their ability to lead conversations. Specifically, ChatGPT facilitates a low-pressure language practice environment, enabling students to transition from passive responses to proactive interaction initiation.

### 5.1. Argument 1: The Impact of ChatGPT on Specific Aspects of WTC

The research findings indicate that practicing role-play with ChatGPT has a positive impact on students' communication confidence and their ability to lead conversations. The shift from passive responses (Code R) to actively initiating questions and sharing opinions (Codes I and E) demonstrates that learners reduced their dependence on the AI's suggested prompts.

This increase can be explained through the Foreign Language Anxiety Theory (Horwitz et al., 1986). Data from P1 and P2 show that the non-personal nature of AI eliminates the pressure of being evaluated by teachers and peers. This result is consistent with the study by MacIntyre et al. (2001), which asserts that WTC is a situational construct and increases when learners feel psychologically safe. Additionally, ChatGPT's provision of immediate vocabulary helped participant P3 to resolve linguistic difficulties instantly, thereby facilitating the maintenance of the conversation. However, data from student ID 5 suggest that the impact of AI on intrinsic motivation remains limited, as this student interacted only to complete the assigned task rather than for genuine communicative purposes. This indicates that while AI is an effective support tool, it cannot yet replace the emotional connections found in human-to-human interaction.

### 5.2.Argument 2: The Extent of Speaking Anxiety Reduction through Role-play Activities

Statistical data, featuring a mean score increase from 3.21 to 3.67 and Cohen's d = 0.675, confirm that these activities significantly reduce speaking anxiety. The results show that the group of students with the highest initial anxiety levels made the most progress. This proves that ChatGPT is an appropriate practice tool for low-confidence learners, helping to bridge the communication gap between students within the same class.

The confidence gained from practicing with AI appears to be transferable to real-world contexts. Data from P2 and P3 show that students tended to be more willing to speak with foreigners and participate in class discussions after the intervention. This finding reinforces the argument that virtual practice environments serve as a preparation stage for real-life communication. Despite some technical issues regarding voice recognition, the study concludes that ChatGPT is an effective support tool for first-year students to overcome psychological barriers when speaking English.

## VI. Implications

### 6.1. Limitations

This study has several limitations. First, the researchers could not fully control how the students spoke to ChatGPT. Some students might have used prepared scripts or translation tools instead of speaking naturally. Because the researchers were not always watching, it is possible that some interactions were not spontaneous.

Second, the group was small. The study only included 15 first-year English students at HUST. Therefore, these results might be different for students in other majors or other universities who face different problems when they speak English.

Third, technical problems affected the experience. Sometimes ChatGPT did not understand the students' voices correctly. Also, ChatGPT often provides responses that are too long and lack naturalness. These issues made it difficult for students to have a normal conversation.

Fourth, ChatGPT does not have human features. It cannot make eye contact or show real emotions. This made the practice feel less real for some students. For those who need to see a person's face and feelings to feel confident, practicing with a computer might not be enough to prepare them for real-life social situations.

Finally, this study was conducted using only the standard, free version of ChatGPT. As a result, the role-play interactions may have been restricted by the technical limitations of this free plan, such as daily usage limits, shorter context memory. These factors could have interrupted the natural flow of the conversation. Future research should consider using premium versions (such as ChatGPT Plus or Pro), which provide access to more advanced language models and improved voice features. This could offer learners a smoother, more human-like, and uninterrupted communication experience.

### 6.2. Pedagogical Implications

At FOFL–HUST, ChatGPT role-play should be integrated as a supportive pedagogical tool in speaking classes to help students build confidence before engaging in real-life communication. This tool is particularly effective in creating a low-pressure environment where students can experiment with language, make mistakes, and expand their vocabulary without fear of judgment.

To maximize its effectiveness, instructors should provide clear guidance on how to construct effective and interactive prompts, rather than allowing students to rely on pre-written scripts. Emphasizing spontaneous and unpredictable exchanges can enhance the authenticity of the interaction and better prepare learners for real communicative situations.

In addition, ChatGPT role-play can be implemented as an initial classroom activity before the main speaking lesson. This approach can help lower students' affective filter and reduce psychological pressure, enabling them to participate more actively in classroom discussions. Overall, such integration can contribute to improving students' WTC and encourage more proactive use of English in academic context

## VII. Conclusion

This study offers novel insights into the use of ChatGPT-based role-play as an effective method for enhancing first-year English students' willingness to communicate (WTC). Rather than concentrating solely on speaking performance, the research emphasizes the role of AI-supported interaction in addressing critical psychological factors, particularly by alleviating speaking anxiety and fostering communication confidence. The findings illustrate that ChatGPT provides a safe, low-pressure environment that motivates students to actively engage in communication, facilitating a transition from passive to active participation. From a practical perspective, the results advocate for integrating ChatGPT role-play as a preparatory stage before students engage in real classroom interactions. Nonetheless, the study recognizes that ChatGPT cannot fully substitute human interaction, given that some students exhibit limited emotional engagement with AI. As a result, its application should be considered supplementary rather than standalone. Future research should investigate the potential of combining ChatGPT with voice-based systems or AI-driven avatars, as well as assess its long-term implications using larger sample sizes for greater generalizability.

## REFERENCES


[1]. Ahearn, L. M. (2001). Agency and language. *Anthropological Quarterly, 74*(3), 110–137. https://doi.org/10.1353/anq.2001.0026

[2]. Ayedoun, E., Hayashi, Y., & Seta, K. (2015). A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Procedia Computer Science, 60*, 1433–1442. https://doi.org/10.1016/j.procs.2015.08.219

[3]. Baghaei, P., Dourakhshan, A., & Salavati, O. (2012). The relationship between willingness to communicate and success in learning English as a foreign language. *Modern Journal of Applied Linguistics, 4*(2), 53–67.

[4]. Baker, M. (2018). *In other words: A coursebook on translation* (3rd ed.). Routledge.

[5]. Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Tiezzi, T., Shin, J., Madotto, A., & Fung, P. (2023). A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity. *arXiv*. https://doi.org/10.48550/arXiv.2302.04023

[6]. Bonsu, G. A., & Baffour-Koduah, D. (2023). ChatGPT for over-the-counter (OTC) drug information: An assessment of its accuracy and reliability. *Cureus, 15*(10), e47496. https://doi.org/10.7759/cureus.47496

[7]. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

[8]. Carrera Nuñez, A. A., Carrera Nuñez, M. S., Zambrano Pachay, J. F., & Carrera Bosquez, A. M. (2025). Uso de ChatGPT Voice para mejorar las habilidades de habla en estudiantes del idioma inglés. *Ciencia Latina Revista Científica Multidisciplinar, 9*(1), 7143–7161.

[9]. Choi, S. Y., & Oh, S. (2026). *Human-AI conversation analysis: Frameworks for the digital age*. [Publication place/Publisher].

[10]. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

[11]. Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). Routledge.

[12]. Dewaele, J. M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching, 4*(2), 237-274. https://doi.org/10.14746/ssllt.2014.4.2.5

[13]. Dewi, D. F., Nur'Aini, S., & Suwarti, T. T. (2022). Students' perception on the use of chatbot from Memrise site and their willingness to communicate in English. *Linguistics and Education Journal, 2*(1), 11–21.

[14]. Du, H., & Daniel, B. K. (2024). The role of AI chatbots in supporting continuous language learning: Feedback and interaction in practice. *Computers & Education: Artificial Intelligence, 6*, 100205.

[15]. Flanagin, A., Bibbins-Domingo, K., Berkwits, M., & Christiansen, S. L. (2023). Nonhuman "authors" and implications for the integrity of scientific publication and medical knowledge. *JAMA, 329*(8), 637-639. https://doi.org/10.1001/jama.2023.1344

[16]. Flick, U. (2018). *An introduction to qualitative research* (6th ed.). SAGE Publications.

[17]. González-Lloret, M. (2011). Conversation analysis in computer-mediated communication. *Annual Review of Applied Linguistics, 31*, 308–325.

[18]. Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

[19]. Hong, W. C. H. (2023). The impact of ChatGPT on foreign language learning confidence and motivation: A self-determination theory perspective. *Computers and Education: Artificial Intelligence, 5*, 100150.

[20]. Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132.

[21]. Hwang, G. J., & Chang, C. Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments, 31*(7), 4099-4112.

[22]. Jati, A. G. (2022). Envisioning the partnership of human and AI in English language teaching and learning. *Journal of English Language Teaching and Linguistics, 7*(2), 297-310.

[23]. Kang, S. J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System, 33*(2), 277–292.

[24]. Khajavy, G. H., MacIntyre, P. D., & Barabadi, E. (2018). Role of the emotions and classroom environment in willingness to communicate: Applying multilevel structural equation modeling in foreign language learning. *Studies in Second Language Acquisition, 40*(3), 605-624.

[25]. Kim, N. Y., & Su, Y. (2024). The effects of AI chatbot role-play on EFL learners' willingness to communicate. *Computers & Education: Artificial Intelligence, 6*, 100197.

[26]. Liu, M., Xu, S., Wu, L., Wu, J., & Cheng, J. (2024). A survey of Generative Pre-trained Transformer (GPT) in healthcare: Opportunities, challenges, and future directions. *Journal of Medical Systems, 48*(1), 2.

[27]. MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *Modern Language Journal, 91*(4), 564-576.

[28]. MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition, 23*(3), 369–388.

[29]. MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545-562.

[30]. Markee, N. (2000). *Conversation analysis*. Lawrence Erlbaum Associates.

[31]. McClain, C., & Kennedy, B. (2025). *How the U.S. Public and AI Experts View Artificial Intelligence*. Pew Research Center.

[32]. McCroskey, J. C., & Richmond, V. P. (1990). Willingness to communicate: Differing cultural perspectives. *Southern Communication Journal, 56*(1), 72-77.

[33]. Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Harvard University Press.

[34]. Newmark, P. (1988). *A textbook of translation*. Prentice-Hall International.

[35]. Nikolopoulou, K. (2022). *Exploring the role of AI in education: Challenges and opportunities*. [Check for specific journal/publisher details].

[36]. Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.

[37]. Patton, M. Q. (2002). Qualitative Research & Evaluation Methods (3rd ed.). Sage Publications.

[38]. Philip, J. M., Kim, N. J., & Park, J. H. (2024). A systematic review of ChatGPT in English language teaching and learning. *Educational Technology Research and Development*.

[39]. Salsabil, S. N., Gunawan, M. I., & Nurochim. (2025). From silent learners to confident speakers: The effect of AI voice chat with ChatGPT on EFL speaking skills. *Journal of English Education, 3*(1), 38-50.

[40]. Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis* (Vol. 1). Cambridge University Press.

[41]. Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

[42]. Tania, S., Rahman, M., & Islam, S. (2025). Psychological barriers in L2 oral communication: The role of AI chatbots in mitigating foreign language anxiety. *Language Learning & Technology, 29*(1), 112-130.

[43]. Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *Modern Language Journal, 86*(1), 54–66.