



Castledown

 OPEN ACCESS

Technology in Language Teaching & Learning

ISSN 2652-1687

<https://www.castledown.com/journals/tltl/>

Technology in Language Teaching & Learning, 8, 104943 (2026)

<https://doi.org/10.29140/tltl.2026.104943>

Not Human, Still Social: Thai EFL Learners' Willingness to Communicate with ChatGPT



NARATIP JINDAPITAK^a 

MUNIR LAEHA^b  (Corresponding author)

^aPrince of Songkla University, Thailand
naratip.jin@psu.ac.th

^bPrince of Songkla University, Thailand
munir.l@psu.ac.th

Abstract

As generative artificial intelligence (AI) becomes increasingly used for second language speaking practice, willingness to communicate (WTC) (MacIntyre et al., 1998) needs to be understood not only in relation to task conditions, but also in relation to how learners experience AI as an interactional counterpart. This study examined how their WTC was sustained across AI-mediated speaking tasks and how their social positioning of AI shaped that willingness. Eighty English as a foreign language (EFL) learners completed five speaking tasks with ChatGPT, each followed by a post-task WTC self-rating and written reflection. Results showed no statistically significant differences in WTC ratings across the five tasks, suggesting overall stability at the aggregate level. However, the reflective data revealed variation in how learners experienced the tasks in relation to their WTC ratings. Four recurring domains were identified: language and cognitive manageability, affective state, task and topic fit, and AI interaction experience. Interview findings further showed that learners' WTC was linked to four positionings of ChatGPT: as socially present while recognizably artificial, as able to follow intended meaning, as nonjudgmental and at times emotionally sustaining, and as a perceived adaptive listener. Interpreted through a posthumanist lens (Pennycook, 2018), the findings suggest that WTC in AI-mediated speaking emerges relationally through the learner–AI encounter and that interactional affordances are essential to sustaining communicative willingness in AI-supported ELT contexts.

Keywords: Willingness to communicate, posthumanist, social presence, AI, ChatGPT.

Copyright: © 2026 Naratip Jindapitak & Munir Laeha. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Data Availability Statement: All relevant data are within this paper.

Introduction

Learning to speak in second language (L2) is rarely a smooth or linear process. Progress in speaking depends not only on what learners know, but also on whether they are prepared to use that knowledge in interaction. This readiness to enter communication at a particular moment is known as willingness to communicate, or WTC (MacIntyre et al., 1998). Because language development depends heavily on use, learners' willingness to speak becomes central to understanding how opportunities for interaction are taken up, avoided, or sustained. Learners who choose not to speak, even when they have adequate linguistic knowledge, lose valuable opportunities for meaningful interaction. By contrast, learners who are willing to speak, even with hesitation or imperfect language, are more likely to benefit from communicative practice. Yet sustaining such willingness is especially difficult in English as a foreign language (EFL) settings, where opportunities to use English beyond the classroom are often limited and speaking is frequently shaped by anxiety, low confidence, and uneven motivation (Chaisiri, 2023). As a result, many learners remain reluctant to speak even when they have the linguistic resources to do so (Sato, 2023).

The emergence of generative AI-powered conversational tools such as ChatGPT has opened a potentially significant space for EFL speaking practice (Xing & Saeed, 2025). Compared with human interlocutors, generative AI offers availability, patience, and a relatively nonjudgmental environment for practice (Jindapitak & Tantiwich, 2025), qualities associated with reduced speaking anxiety, improved confidence, and enhanced communicative willingness among EFL learners (Xin & Derakhshan, 2025). Yet, the promise of AI-mediated speaking cannot be understood through affordances alone. WTC is not simply readiness to speak in general, but readiness to communicate with a particular interlocutor, at a particular moment, in a particular context (MacIntyre et al., 1998). When the interlocutor is AI, learners' willingness may therefore depend not only on what the system can do linguistically, but also on how learners interpret its role in the exchange. This raises the question of how AI is positioned during spoken interaction and how such positioning relates to learners' willingness to continue speaking (Jeon & Lee, 2023; Lee & Jeon, 2024).

The present study shifts attention from WTC as a learner-internal variable and AI as a mere instructional tool to the interactional and relational conditions of AI-mediated speaking. Through a mixed-methods investigation of Thai EFL university learners using ChatGPT for speaking tasks, it examines patterns of WTC across a sequence of tasks and how learners describe their task experiences in relation to their WTC ratings. It further examines how learners position ChatGPT as an interactional counterpart, and how these positionings shape their willingness to continue speaking. In doing so, the study contributes to emerging work on WTC in AI-mediated contexts by illuminating the experiential and relational conditions through which WTC takes shape (Kirkpatrick et al., 2024; Peng & Liang, 2026). It also informs ongoing discussion of how AI-mediated speaking tasks are understood and designed in EFL pedagogy (Tai & Chen, 2023).

Literature Review

WTC in the Digital Era

AI chatbots have become increasingly relevant to WTC research as L2 communication moves beyond face-to-face classrooms into online and AI-mediated environments. Speaking is particularly constrained in EFL settings by limited authentic practice opportunities, insufficient exposure, and the evaluative pressures of human-witnessed interaction (Xing & Saeed, 2025). AI chatbots address these constraints through features that are especially well matched to the demands of speaking practice. As conversational agents capable of sustained, contextually responsive dialogue, they can simulate real-world

speaking scenarios, provide immediate and personalized linguistic feedback, and support extended practice without the social stakes of communicating with a human interlocutor (Jeon & Lee, 2024; Xing & Saeed, 2025). More broadly, they expand opportunities for practice, reduce social exposure through anonymity, allow flexible participation across modalities, and give learners more time to process and respond (Kruk, 2022; Tai & Chen, 2023). These features allow them to function as conversational partners, role-play simulators, feedback providers, and sources of personalized support, roles that are difficult to replicate in resource-limited EFL classrooms (Huang et al., 2022). Empirical studies have reported improvements in aspects of learners' oral performance, including fluency, vocabulary use, and speaking proficiency, following AI-mediated speaking practice (Jeon & Lee, 2024; Tai, 2024). However, less is known about how learners' WTC is sustained, disrupted, or reshaped across a sequence of AI-mediated speaking tasks (Waluyo & Isti Pratiwi, 2025).

Evidence on affective responses to AI-mediated speaking is generally positive, but not uniform. Zhang et al. (2024) found that an AI speaking assistant enhanced Chinese EFL learners' WTC and enjoyment while reducing anxiety, while Kim and Su (2024) reported WTC gains, reduced anxiety, and improved confidence among Korean university learners using an AI chatbot. Related evidence from synchronous online speaking also suggests that the design of interaction matters. In a one-to-one online conversation setting, Tabuchi et al. (2024) found that repeated speaking practice with individualized corrective feedback improved speaking performance and reduced anxiety in a low-proficiency learner. However, Xin and Derakhshan (2025) showed that AI-mediated classrooms can generate both motivating emotions (e.g., excitement and confidence), and inhibiting ones, including frustration and anxiety linked to unfamiliarity with AI and limited human-like encouragement.

Beyond examining whether AI-mediated speaking generally enhances WTC, recent research has begun to trace how learners' WTC varies across time and in response to changing interactional conditions. Peng and Liang (2026) found that learners' WTC and perceptions of ChatGPT shifted across a ten-week period, shaped by the chatbot's changing roles as advisor, language helper, or conversational counterpart. Kruk (2022) similarly found that WTC fluctuated with boredom, topic familiarity, past experiences, interlocutor responsiveness, and the shielding effect of virtual avatars. Lee and Lee (2026) further showed that learners with higher WTC tended to experience more task enjoyment and treat AI interaction as practice, whereas those with lower WTC approached tasks with greater reluctance. Other studies point to the importance of AI and interlocutor characteristics. Wang et al. (2024) found that a human-like avatar supported stronger WTC gains, while Tai and Chen (2023) reported that an Intelligent Personal Assistant increased WTC, confidence, and reduced anxiety by creating a less evaluative environment than traditional classrooms. These studies show that WTC in AI-mediated speaking is shaped not only by learners' readiness, but also by the affective, social, and interactional qualities of the AI-mediated encounter.

A Posthumanist Perspective on WTC and Social Presence in AI-Mediated Speaking

Conventional accounts of WTC locate communicative willingness primarily within the individual learner, as a trait, disposition, or readiness shaped by confidence, anxiety, and motivation (MacIntyre et al., 1998). These accounts have been productive, but they also tend to assume that communicative agency originates in the human subject, with interlocutors and contexts facilitating or constraining a willingness that ultimately resides in the learner. A posthumanist perspective challenges this assumption by treating agency, cognition, and meaning as distributed across human and non-human relations (Liang & Wu, 2024; Pennycook, 2018). From this perspective, WTC is not simply brought into an interaction by the learner but takes shape through the relational and material conditions of the encounter. This shifts attention from learner-internal readiness to how communicative willingness is produced in interaction with particular interlocutors, technologies, and environments (Matthews, 2024; Peng & Liang, 2026).

Two concepts help specify how this relational view applies to AI-mediated speaking. The first is distributed cognition, which views thinking as extending beyond the individual mind into social and material environments where external artefacts participate in cognitive activity (Hutchins, 1995; Ou et al., 2024). ChatGPT can therefore be understood not merely as an instructional tool or neutral medium, but as part of the cognitive and communicative system through which learners organize speech, interpret responses, and sustain participation (Pennycook, 2018). The second is assemblage, which foregrounds the temporary configuration of human and non-human actors and the capacities that arise through their relations (Liang & Wu, 2024; Ou et al., 2024). Matthews (2024) helps explain these relations by distinguishing between active affects, an actor's capacities to initiate, adapt, or resist, and passive affects, through which self-determination is shaped by external forces. In AI-mediated speaking, learners exercise active affects when they initiate, adapt, or sustain interaction, while they may simultaneously be subject to passive affects when AI responses shape their subsequent willingness, confidence, or desire to continue. These processes occur together: learners shape the exchange through their own actions, while the unfolding interaction can in turn shape their willingness to remain in it. WTC therefore develops through this ongoing interplay between learner action and the possibilities or constraints created in the AI-mediated encounter. This is consistent with Peng and Liang's (2026) argument that WTC with ChatGPT is contingent on evolving interactional experience and that AI participates actively in shaping communicative engagement.

Social presence provides a way to examine how this distributed account of WTC is experienced in interaction. It refers to the extent to which an interlocutor is perceived as socially available, with intimacy and immediacy shaping users' sense of closeness and distance (Short et al., 1976). In human-AI interaction, this can be understood through the Computers Are Social Actors (CASA) paradigm, which argues that people apply social rules to technologies displaying social cues (Nass & Moon, 2000), and through anthropomorphism, or the attribution of human-like qualities to non-human entities (Epley et al., 2007). These frameworks suggest that social presence is not simply a property of AI, but an interactional effect produced when conversational cues invite users to treat the system as socially responsive.

Empirical work supports this view. Learners' perceptions of chatbot human-likeness and social presence have been shown to predict motivation and engagement, with chatbots often described as supportive, polite, emotionally responsive, credible, or rapport-building (Lee & Jeon, 2024). Studies with younger users similarly show that conversational behavior can lead learners to attribute intelligence, emotional capacity, and even friendship potential to chatbots (Brandtzaeg et al., 2022; Xu & Warschauer, 2020). At the feature level, socially appropriate chatbot responses have been found to increase perceived social presence, anthropomorphic attribution, and engagement, while perceived warmth, ability, and social presence predict users' trust and willingness to interact (Deng & Yan, 2025; Schuetzler et al., 2020). These findings indicate that the social qualities learners attribute to AI are not incidental to WTC but central to it, consistent with the posthumanist position that WTC is shaped through the relational conditions of the encounter rather than residing solely in the learner.

The studies reviewed above leave two issues insufficiently addressed. First, although recent WTC research has examined change over time (Kruk, 2022; Peng & Liang, 2026), less is known about the extent to which learners' WTC ratings remain stable or vary across a structured sequence of AI-mediated speaking tasks, and what kinds of task experiences accompany those ratings. Second, research on social presence in AI-mediated interaction has shown that conversational agents can shape motivation, engagement, and affective experience. However, less attention has been given to how learners position AI during spoken interaction and how these positionings shape their WTC. Underlying both gaps is a broader tendency to treat WTC as a learner-internal variable and AI as an external tool, rather than attending to how willingness and social perception of AI are co-constituted within interaction.

A posthumanist perspective, as developed by Pennycook (2018) and applied to AI-mediated contexts by Peng and Liang (2026), offers a framework for addressing this tendency by attending to how WTC and social meaning emerge through the learner-AI encounter. To address these gaps, the study was guided by two research questions:

1. To what extent did learners' post-task WTC ratings vary across the ChatGPT-mediated speaking tasks, and how did learners describe their task experiences in relation to these ratings?
2. How did learners position ChatGPT as an interactional counterpart, and how did these positionings shape their WTC?

Methodology

Participants

Eighty first-year English major students from a large public university in southern Thailand participated in the study. Convenience sampling was employed, given that the study targeted the entire accessible cohort. This specific group was selected because their communicative needs and expectations for interactional complexity are higher than those of general students, making their WTC in a digital environment a critical area of study. Participation was voluntary, and they could withdraw at any time without consequence. Participants were 19 to 20 years old and had intermediate to upper-intermediate English proficiency (CEFR B1–B2), as determined by the university's in-house English placement examination.

For the follow-up semi-structured interviews, 17 participants were selected from the larger sample based on their availability and willingness to participate. The interview sample was considered adequate for thematic analysis given the relatively homogeneous participant group and the focused scope of the research questions (Creswell, 2003).

Instruments

Three instruments were used: five communicative speaking tasks, a post-task WTC rating scale with self-reflection notes, and a semi-structured interview protocol. Participants were not graded for completing the tasks. All instruments were piloted with 30 students of similar proficiency level who did not participate in the current study.

Speaking Tasks

The speaking tasks consisted of five ungraded communicative activities (see Table 1), with one speaking task assigned per week throughout the five-week period. Each session lasted approximately 8 to 10 minutes. Tasks specified a communicative goal without prescribing how participants should manage the exchange, allowing interaction to unfold naturally. Built around familiar university-life topics, the five tasks covered distinct communicative functions: describing personal experience, seeking advice, negotiating, problem-solving, and giving recommendations. This variation kept content accessible to first-year students while allowing the study to examine WTC across different types of AI-mediated speaking activity. ChatGPT was selected as the interlocutor for its accessibility and conversational suitability (Jindapitak & Tantiwich, 2025).

A Post-Task WTC Self-Rating and Written Reflection

A self-rating scale ranging from 0 to 100 was used to capture participants' perceived WTC following each speaking task. State WTC, which is situational and transient by nature, is frequently investigated

Table 1 *Speaking Tasks and Their Communicative Function*

Task	Topic	Communicative Function
1	My First Week	Describing personal experiences about the first week on campus
2	Survival Guide	Seeking advice on managing university life
3	Activity Planning	Planning and negotiating a campus activity with friends or classmates
4	Student Problems	Explaining and discussing a problem and solutions
5	Campus Recommendations	Giving recommendations to foreign students about facilities and activities on campus

through self-ratings (Peng & Liang, 2026), and this approach was appropriate here given the absence of validated instruments specifically designed for WTC assessment in ChatGPT-mediated speaking contexts. The self-rating was accompanied by a written reflection in which participants explained and justified their score by describing the experiences they associated with their WTC during the task. The reflection therefore provided contextual information about the linguistic, affective, task-related, and interactional experiences associated with each rating. The use of Thai was permitted to avoid language barriers. The two data sources were complementary: the numerical rating provided a summary measure of perceived willingness, while the written reflection offered participants' own accounts of the conditions shaping that willingness.

Semi-Structured Interview

The semi-structured interview was used to examine how participants positioned ChatGPT as an interactional counterpart and how these positionings related to their willingness to continue speaking. The protocol consisted of five open-ended questions organized from broad positioning to interaction-specific accounts. Participants were first asked to describe what kind of partner ChatGPT felt like during speaking, with prompts such as tool, helper, or conversation partner used only as examples rather than fixed response categories. They were then asked to recall a specific moment that illustrated this positioning, allowing the interview to move beyond general evaluation and toward situated accounts of interaction.

Subsequent questions explored how participants' views of ChatGPT were sustained or adjusted across the five tasks and whether they experienced moments when ChatGPT felt merely machine-like or more socially responsive. Participants were also asked to connect these experiences to their willingness to continue speaking. This allowed the interviews to capture both stability and movement in learners' positioning of ChatGPT, while also eliciting responses of how such positioning shaped WTC. The semi-structured format ensured consistency across interviews while allowing follow-up probes when participants referred to specific responses, task moments, emotional reactions, or changes in their engagement.

Data Collection Procedure

Ethical approval was obtained before data collection, and all participants provided written informed consent. Data collection began with a brief training session to familiarize participants with the task format and ChatGPT's voice interface, including initiating voice input and managing turn-taking. For the main data collection, participants completed each speaking task at their own convenience, undertaking one task per week over the five-week period. They video-recorded their interactions and submitted the recordings together with their WTC rating scales and self-reflection notes directly to the researchers

by email to ensure confidentiality. As a guideline, each task was expected to last between 8 and 10 minutes. After that, 17 participants were approached to be involved individually in a semi-structured interview, which lasted approximately 35 minutes. The interview was audio-recorded and transcribed verbatim for analysis.

Data Analysis

Quantitative data were first summarized using descriptive statistics to examine learners' post-task WTC ratings across the five speaking tasks. Differences across tasks were examined using the Friedman test (Friedman, 1937), a non-parametric repeated-measures test appropriate for comparing related observations across more than two time points.

The written reflection notes were analyzed using qualitative content analysis. The analysis was conducted in relation to the task for which each reflection was written so that learners' explanations could be linked to the corresponding post-task WTC rating. The notes were read iteratively, and meaning units were identified and coded close to participants' own wording. Through progressive comparison across participants and tasks, the codes were organized into four domains: language and cognitive manageability, affective state, task and topic fit, and ChatGPT interaction experience. Each coded meaning unit was further classified as reflecting a positive or negative experience within the relevant domain. For instance, perceived improvement in fluency was coded as a positive occurrence under language and cognitive manageability, whereas reported anxiety was coded as a negative occurrence under affective state. A single reflection could contain multiple meaning units; therefore, the frequencies reported represent coded occurrences rather than numbers of participants. These occurrences were subsequently compared across domains and tasks.

Interview data were analyzed using thematic analysis (Braun & Clarke, 2022). Transcripts were read repeatedly for familiarization, after which initial meaning units were identified close to participants' own wording. Through iterative comparison across the 17 participants, recurring patterns were progressively grouped into broader categories and then into themes.

Two researchers were involved in checking both the written reflection notes and interview transcripts. For the reflection notes, the coding framework and a sample of coded comments were reviewed, including theme assignment and positive/negative classification. Discrepancies were discussed and used to refine the coding before the full dataset was finalized. For the interview data, both researchers independently coded the transcripts, compared emerging categories, and resolved inconsistencies through discussion. As the reflections and interviews were produced in Thai, selected excerpts used in the findings were translated into English and returned to participants for member checking to confirm that the translations reflected their intended meanings (Merriam & Tisdell, 2016).

Findings

Learners' Willingness to Communicate Across Tasks and the Factors Shaping It

As shown in Table 2, learners' post-task WTC ratings were relatively similar across the five AI-mediated speaking tasks. The mean score increased slightly from 69.40 ($SD = 18.17$) in Task 1 to 73.10 ($SD = 15.87$) in Task 5, while the means for Tasks 2 to 4 remained closely clustered between 70.45 and 71.23.

Despite these minor numerical fluctuations, the Friedman Test revealed no statistically significant differences in WTC levels across the five tasks ($\chi^2(4) = 6.46, p = .168$, Kendall's $W = .020$), indicating a

Table 2 Learners' WTC Self-Rating across Five Speaking Tasks ($N = 80$)

Task	Topic	Mean	SD	Mean Rank (MR)	Friedman Test	Kendall's W
1	My First Week	69.40	18.17	2.84	.168	.020
2	Survival Guide	70.46	18.27	2.89		
3	Activity Planning	70.45	15.89	2.92		
4	Student Problems	71.23	16.28	3.03		
5	Campus Recommendations	73.10	15.87	3.32		

Note. * $p < .05$.

negligible effect size. This indicates that the learners' WTC with ChatGPT remained relatively stable throughout the study. While a slight descriptive increase in mean scores can be observed, the statistical evidence suggests that the varying communicative demands and task purposes did not significantly alter the participants' overall WTC levels.

Findings further showed that learners' reflections on their post-task WTC ratings clustered around four interrelated domains: language and cognitive manageability, affective state, task and topic fit, and AI interaction experience (see Table 3). When aggregated across tasks, positive occurrences exceeded negative occurrences in each domain. Positive occurrences were most frequent for affective state (612 occurrences), followed by language and cognitive manageability (553 occurrences), AI interaction experience (406 occurrences), and task and topic fit (388 occurrences).

Table 3 Positive and Negative Reflections on Task Experiences in Relation to WTC Ratings

Task	Language & Cognitive Manageability		Affective State		Task & Topic Fit		AI Interaction Experience	
	P	N	P	N	P	N	P	N
1	63	68	100	17	57	32	55	53
2	98	94	115	50	62	34	64	51
3	119	86	137	50	76	22	74	51
4	121	80	114	63	88	55	70	89
5	152	78	146	56	105	42	143	69
Total	553	406	612	236	388	185	406	313

Note. P = Positive occurrences, N = Negative occurrences.

Language and Cognitive Manageability

Positive occurrences related to language and cognitive manageability increased across the five tasks. In T1 and T2 (63 and 98 occurrences, respectively), participants could partially follow ChatGPT's output and reported that practicing pronunciation and vocabulary strengthened their engagement with AI. By T3 and T4 (119 and 121 occurrences, respectively), gains in processing and structural control emerged, including participants self-corrected after feedback, spoke more fluently, and produced better-organized ideas. By T5, the occurrences rose to 152, with more advanced development evident, including greater discourse control, longer utterances, and more contextually appropriate responses.

Negative occurrences increased from T1 (68 occurrences) to T2 (94 occurrences) before declining gradually across T3, T4, and T5 (86, 80, and 78 occurrences, respectively). In the earlier tasks, participants commonly described difficulties generating ideas, processing vocabulary, and organizing information. Although similar difficulties persisted in the later tasks, participants increasingly reported problems with real-time lexical retrieval, while comprehension breakdowns involving more complex structures remained evident in T5.

Affective State

Participants' affective states shifted considerably across the five tasks, moving from initial nervousness toward a growing sense of rapport with AI. In T1 (100 occurrences) and T2 (115 occurrences), positive experiences centered on excitement and accomplishment from engaging with ChatGPT during the early tasks, with participants describing the experience as enjoyable and fun. In T3 (137 occurrences) and T4 (114 occurrences), these occurrences shifted toward increased confidence and reduced anxiety, with participants expressing greater WTC. By T5 (146 occurrences), stronger feelings of pride and satisfaction emerged due to task completion, and participants became more proactive in initiating interaction with AI.

Negative occurrences fluctuated throughout the tasks. In T1 (17 occurrences), they were minimal, mainly involving fear due to unfamiliarity with voice-based AI interaction. Negative occurrences remained at 50 in T2 and T3 before increasing to 63 in T4, driven by emotional regulation difficulties, including panic, frustration, and mental blocks. These were often triggered by complex or unexpected AI outputs. By T5, negative occurrences declined slightly to 56 occurrences, although concerns shifted toward the perceived unnaturalness of AI interaction, particularly the lack of genuine emotional engagement.

Task and Topic Fit

Positive occurrences rose steadily from T1 to T5, reflecting growing alignment between tasks and participants' needs. In T1 and T2 (57 and 62 occurrences, respectively), participants noted that topics matched their interests and that role-playing with AI was appropriately challenging. Task relatability increased notably after T2, as university life-themed content allowed participants to engage more naturally. From T3 to T5, positive occurrences increased steadily from 76 to 105, with topics such as activity planning, student problems, and campus recommendations seen as highly relevant to daily life, deepening engagement and transforming interactions into something resembling social support.

Negative occurrences fluctuated across T1 to T3, increasing slightly from 32 in T1 to 34 in T2 before declining to 22 in T3. These occurrences primarily concerned content depth and knowledge gaps, with participants noting that unfamiliar scenarios caused cognitive blocks. An additional concern identified in T4 (55 occurrences) was topic sensitivity. Tasks involving the disclosure of personal experiences, such as loneliness or group conflict, posed challenges for participants in expressing their thoughts clearly. Although negative comments decreased in T5 (42 occurrences), themes of personal mismatch, limited topic relatability, and topic unfamiliarity persisted.

AI Interaction Experience

Participants initially struggled with technical constraints but progressively developed strategies to manage AI interactions. Positive occurrences increased from T1 to T3, rising from 55 to 74, and centered on exploring the AI's capabilities, including its accessibility and support for different accents. Although occurrences decreased slightly in T4 (70 occurrences), participants appreciated the

increasingly human-like voice responses, on-screen text support, and the AI's ability to sustain conversation, which encouraged greater participation. By T5, positive experiences reached 143, marking a clear shift toward greater user agency, with participants demonstrating the ability to direct the length and style of the AI's responses, resulting in smoother and more effective interactions.

Negative occurrences remained relatively stable between T1 and T3, decreasing slightly from 53 to 51 occurrences. Common issues included the AI interrupting or producing overly long responses. Negative experiences were noticeable in T4 (89 occurrences), where additional problems such as delayed responses, rapid speech, insufficient pausing, and system freezing disrupted conversational flow. By T5 (69 occurrences), negative occurrences declined, reflecting participants' growing resilience and improved ability to navigate technical limitations.

The Positioning of ChatGPT as a Conversational Counterpart

Social Enough to Speak to, Artificial Enough to Remain a System

Findings reveal how learners positioned ChatGPT as a nonhuman counterpart that could still feel social in interaction. Participants did not need ChatGPT to be a real person for the exchange to count as speaking. What mattered was that it could respond in ways that made the interaction feel conversational, while still remaining recognizably artificial.

For some participants, ChatGPT remained clearly nonhuman. P10 described it as “a kind of search engine, but a more advanced one,” explaining that it could “speak and respond like a human” but did “not have feelings.” P15 similarly emphasized, “I always knew I was talking to a program. I never forgot that.” These learners did not position ChatGPT as a real conversation partner. Still, they did not treat it as a simple machine either. Its ability to speak, respond, and assist gave it a role in the interaction, even when its artificial status remained clear.

Other participants described a more ambiguous status. P11 called ChatGPT “something half-machine and half-human,” noting that it felt “friendly and warm during the tasks, but not like a real friend.” P17 expressed a similar stance, describing it as “just a machine or a program” that could still be “a good conversation partner because it feels social.” For these learners, ChatGPT's social value did not require belief in human feelings. It was enough that the interaction carried signs of social presence.

This status shaped WTC by affecting whether ChatGPT felt worth continuing with as an interactional counterpart. P1 explained that when ChatGPT felt “somewhere in between,” they “wanted to keep going because it was interesting,” but when it felt “clearly like a machine,” they “just completed the task” and did not feel like continuing. In this sense, continued speaking depended partly on whether ChatGPT could maintain a middle status: social enough to sustain exchange, but artificial enough to remain a system rather than a human interlocutor.

When Uptake Holds: Feeling Followed and Wanting to Continue

WTC often turned on whether learners felt that ChatGPT had grasped what they were trying to say. Participants described moments where ChatGPT seemed to pick up their intention, emotional stance, or dissatisfaction even when these were only indirectly expressed. In those moments, the exchange felt less like answer retrieval and more like being followed in conversation.

P14 stated this link most directly: “If I said something and the response matched what I meant, it felt like a real conversation.” The effect on speaking was immediate: “When I felt understood, I wanted

to keep talking. I had more things to say.” Other participants described similar moments of uptake. P3 said ChatGPT was “tracking my situation, not just giving general advice,” while P5 recalled that when they said, “What? I don’t get it,” ChatGPT seemed to understand that they were unhappy with the explanation. P8 gave an even more minimal example: after the participant simply said “Nah,” ChatGPT treated it as a cue to look for the information again.

When uptake failed, the interaction quickly lost its conversational pull. P9 noted that ChatGPT sometimes seemed to be “just picking up on the keywords,” although their “real intention was not always tied to those keywords.” P14 similarly reported that when the response “did not fit” what they had said, they “lost track” and “just wanted to finish quickly.” P15 described another version of this breakdown, which is an empathetic response could feel “rehearsed,” as if ChatGPT had “learned how to sound empathetic.”

From No Judgment to Real Warmth: The Floor and Ceiling of Emotional Safety

Emotional safety worked in two different ways. At the lower level, ChatGPT removed the threat of being judged. At the higher level, it created a warmer form of engagement that made learners want to continue speaking. This distinction is important because the two levels shaped WTC differently: the first made speaking possible, while the second made speaking desirable.

At the floor level, ChatGPT was valued because it did not mark learners’ English as deficient. Participants repeatedly contrasted this with human interaction, where judgment could be communicated without words. P4 explained that when speaking English with people, judgment could appear through “facial expressions,” even if no one directly criticized the language. With ChatGPT, this threat was absent: “It does not judge my poor English.” P6 described a similar moment after producing an ungrammatical sentence. ChatGPT “incorporated the correct form into its reply without pointing out my mistake directly.” The correction was therefore present but not socially exposed. This low-risk condition changed how learners approached speaking. P9 stated that they felt “less pressure than when I talk to real people,” which made them “brave enough to experiment with new words and phrases.” P17 made the same link more directly: “I can speak longer with AI because I do not worry about being judged.” At this level, ChatGPT made it easier for learners to continue speaking because mistakes were not singled out or treated as embarrassing.

Other responses moved beyond this baseline. For some participants, ChatGPT was not only nonjudgmental but affectively sustaining. P6 drew this distinction most clearly: “The safe feeling made me willing to start speaking and to keep going even when I made mistakes. But the warmer moments made me want to continue, not just feel okay continuing...” This contrast shows that safety and warmth did different kinds of work. Safety lowered inhibition, but warmth generated motivation. The ceiling appeared most clearly when learners shared personal or emotionally loaded concerns. P7 recalled saying, “I want to cry,” and receiving a response such as, “It’s okay to want to cry. People cry sometimes.” P12 similarly described ChatGPT’s support as something that made them “want to continue, not just practice.” In these moments, WTC was no longer only about tolerating the act of speaking. The interaction itself became rewarding enough to sustain further talk.

The Learning Listener: Perceived Adaptation and Sustained Speaking

A further pattern was that some learners positioned ChatGPT as a counterpart that could learn from the interaction and adjust itself to them. This was different from simply receiving a good response in one moment. Participants described ChatGPT as remembering preferences, avoiding disliked styles, apologizing after misalignment, and becoming more fitted to their way of speaking. These responses made the interaction feel less fixed and more capable of continuing after trouble.

This perception appeared clearly when ChatGPT changed after learners expressed dissatisfaction. P4 recalled that when they said they were not comfortable with a certain suggestion, ChatGPT “apologized and did not suggest that kind of thing again.” P14 described a similar moment: after they told ChatGPT to change something and expressed dislike, it began with “sorry,” which felt “like it was adjusting to me personally.” In these cases, apology was not reported only as politeness. It also marked a shift in how learners perceived ChatGPT, as the system seemed able to register a problem and modify its next response.

Other participants described this adjustment as a more cumulative quality. P11 explained that when they said they disliked “wordy or flowery language, ChatGPT tried to avoid that style later.” P16 similarly said that ChatGPT seemed to be “learning my behavior, including what I liked and did not like.” P8 also described the longer interaction as becoming better fitted to the learner’s expression: “The longer I talked to it, the more its responses felt suited to what I was trying to express.” This made learners more willing to continue because ChatGPT was not experienced as a fixed machine response, but as a counterpart that could become more responsive to them over time.

Discussion

Post-task WTC ratings did not differ significantly across the five tasks, suggesting broadly stable willingness at the aggregate level despite variation in task purpose and communicative demand. This stability, however, should not be read as uniform experience. Cao and Philp’s (2006) distinction between general self-reported communicative tendencies and situationally contingent WTC helps explain why stable ratings can coexist with more varied task-level experiences. Aggregate ratings may remain stable even when learners’ immediate experiences of speaking vary across tasks, and the reflective data in the present study point to this distinction. Across all four domains, learners reported varied task-level experiences in relation to their WTC rating, with positive occurrences generally increasing over tasks while negative ones fluctuated.

Among the task-level experiences reported in the reflections, affective state and language and cognitive manageability were especially prominent in learners’ explanations of their WTC ratings. Affective state emerged as the more pervasive influence, with positive occurrences consistently outnumbering negative ones, moving from initial excitement and growing mid-task confidence to occasional disappointment toward the end, consistent with conceptualizations of affective state as dynamic and situationally unstable rather than accumulating toward a single dominant orientation (Kruk, 2022; Liu et al., 2026). This reinforces MacIntyre et al.’s (1998) argument that WTC is subject to transient affective fluctuation even when broader dispositional tendencies remain stable. Alongside affective state, participants reported increasing capacity to manage linguistic and cognitive demands as tasks progressed, often attributing this to ChatGPT’s corrective feedback and interactional support. This perceived scaffolding function is consistent with findings from AI-mediated L2 learning contexts (Jindapitak & Tantiwich, 2025). Greater perceived linguistic manageability was also accompanied by more positive WTC-related accounts in the later tasks. A similar process was observed by Tabuchi et al. (2024), where short, individualized feedback and clarification support helped the learner draw more effectively on existing linguistic knowledge and become more willing to communicate in the target language. The present findings also illustrate what Ou et al. (2024) describe as assemblage, a mutually constitutive relationship between learner and system in which communicative capacity is co-produced rather than residing solely in the learner. Persistent difficulties, primarily around vocabulary retrieval when topics shifted, indicate that this co-production remained partial and task-dependent.

Two further domains reflected the interactional and contextual conditions of the AI-mediated format itself. With respect to AI interaction experience, early positive comments centered on the novelty

and accessibility of ChatGPT's affordances, with participants' openness to technological exploration appearing to buffer initial uncertainty (Lee & Liu, 2024). As tasks progressed, positive experiences shifted in character: they reflected growing user agency and strategic familiarity with how to navigate the system, consistent with Lee and Dressman's (2018) account of learner adaptation to digital interactional environments. A modest dip in later tasks corresponded to technical disruptions, suggesting that positive AI interaction experience is sensitive to system responsiveness rather than purely to learner orientation. Task and topic fit operated in a similar direction: participants frequently associated greater willingness with tasks they perceived as relatable or connected to real-life situations, resonating with evidence that topic relevance facilitates L2 speaking engagement (Sato, 2023). Conversely, emotionally sensitive content was sometimes described alongside more guarded participation, suggesting that topic design may shape learners' readiness to self-disclose in AI-mediated speaking.

The interview findings show that learners' WTC was shaped by how they positioned the AI as an interactional counterpart. The first positioning was marked by a productive tension. Participants knew that ChatGPT was artificial, yet some still treated it as socially present enough to speak with. This supports Xu et al.'s (2022) account of social responses to technology that awareness of artificiality does not necessarily prevent social responses to technology. It also resonates with Jeon and Lee (2023) and Lee and Jeon (2024), who show that learners may place conversational agents in a relational space between tool and person. In the present study, learners did not need to believe in ChatGPT's humanness for speaking to continue. Continued engagement depended instead on whether ChatGPT could sustain a conversational presence. When it became too visibly mechanical, WTC weakened and learners shifted toward task completion.

This conversational presence was sustained most clearly when ChatGPT appeared to follow learners' intended meanings. Learners were more willing to continue when ChatGPT appeared to grasp what they meant, including implied emotion, dissatisfaction, or an incomplete utterance. This finding extends situational accounts of WTC (Cao & Philp, 2006; Kang, 2005) by showing that willingness was not only shaped by broad task conditions, but also by small moments within the exchange. A response that followed the learner's intended meaning could make the interaction feel conversational; a response that relied on keywords or sounded formulaic could quickly return it to a machine-like mode. This is consistent with Schuetzler et al.'s (2020) finding that conversational skill affects chatbot engagement and with Wang et al.'s (2024) finding that social presence in AI interaction predicts WTC and learning outcomes. The present data specify the mechanism more closely, as learners continued when they felt that ChatGPT was following their point, not merely processing their words.

The findings also refine how emotional safety is understood in AI-mediated speaking. WTC was shaped not only by reducing evaluative pressure, but also by creating moments of warmth and reassurance. At the floor level, its nonjudgmental response style made speaking easier to begin, aligning with communication apprehension research (MacIntyre, 1994) and recent evidence that AI-mediated speaking environments can reduce anxiety in EFL contexts (Liu et al., 2026; Zhang et al., 2024). At the ceiling level, some learners described warmth, comfort, or support that made them want to continue beyond task completion. This second level goes beyond anxiety reduction. These warmer responses gave some learners a more positive reason to remain in the exchange. Positive affect can strengthen communicative engagement (Yashima et al., 2004). In human-AI interaction, emotionally supportive chatbots may also be perceived as more human-minded (Lee & Hahn, 2024), while companion-like responses can make users experience AI as a socially meaningful presence (Brandtzaeg et al., 2022). In the present study, the floor made speaking easier to begin; the ceiling made continued speaking more desirable.

A further contribution concerns perceived adaptation. Learners did not only describe ChatGPT as responsive in isolated moments. Some also described it as able to register dissatisfaction, apologize,

avoid disliked styles, and become more fitted to their preferences over time. This perception gave the interaction a sense of continuity after trouble. A poor or unsuitable response did not always end the exchange because learners could experience ChatGPT as capable of adjusting in the next turn or later interaction. This finding should not be overstated as evidence that learners believed the system genuinely learned in a human sense. Nonetheless, it shows that perceived adaptability became part of how they made sense of ChatGPT as a speaking counterpart.

The four positionings locate WTC in the learner-AI configuration rather than in the individual learner alone. From a posthumanist perspective (Pennycook, 2018), WTC becomes possible through an assemblage of human and non-human elements, including the learner, task, platform, voice interface, and AI response. Whether ChatGPT was experienced as social enough to continue with, attuned to the learner's intent, emotionally safe, or able to adjust after misalignment, these were relational qualities of the encounter rather than properties of the learner alone. Such qualities are difficult to capture through learner-internal accounts of WTC. In the learner-AI interaction, learners may actively initiate, adapt, or sustain the interaction, while AI responses can simultaneously alter the conditions under which that willingness continues, reflecting the interplay of active and passive affects within human-AI assemblages (Matthews, 2024). At the same time, the learner and AI should not be treated as equivalent participants, nor should the AI be assumed to share the learner's communicative goals. Technical actors, such as AI systems, may shape the course of action without possessing the intentionality or responsibility attributed to human actors (Matthews, 2024). Stockwell and Wang (2025) similarly caution that human-AI interaction may involve functional interdependence and distribution of cognitive work while still lacking shared purpose, reciprocal understanding, and distributed accountability. In the present study, ChatGPT's contribution to WTC is therefore better understood in terms of what its responses made possible within the interaction than as evidence that the system held or pursued the learner's communicative intentions. Consistent with posthumanist accounts of human-AI assemblages in writing and academic performance contexts (Liang & Wu, 2024; Ou et al., 2024), the present findings extend this framework into spoken interaction, showing that WTC in AI-mediated speaking is co-produced through the encounter rather than simply brought to it.

Conclusion and Implications

This study examined Thai EFL learners' WTC with ChatGPT across five speaking tasks. The quantitative results showed no significant change in post-task WTC ratings, indicating broad stability at the aggregate level. The reflective and interview data, however, showed that this stability did not mean that learners experienced the tasks in the same way. Their accounts pointed to changing task-level experiences involving language manageability, affective response, topic fit, and interaction with ChatGPT. The interviews further showed that learners engaged with ChatGPT not as a human substitute, but as an interactional counterpart whose social presence, uptake, emotional support, and perceived adaptability helped sustain communicative willingness. The findings support a relational view of WTC in AI-mediated speaking. Learners' WTC was not located solely in their prior confidence, anxiety, or linguistic readiness; it emerged through an assemblage of learner, task, platform, voice interface, and AI response. From a posthumanist perspective (Pennycook, 2018), ChatGPT was not a neutral channel through which learners simply practiced English, but part of the communicative arrangement through which speaking was initiated, sustained, repaired, or abandoned. This suggests that WTC with AI should be understood as a distributed and interactional phenomenon, formed through the learner-AI encounter.

Several limitations should be noted. The study examined one AI platform and five prescribed speaking tasks, so the findings may not transfer directly to other tools, interaction modes, or open-ended uses of AI. The WTC ratings captured learners' overall post-task impressions rather than moment-to-moment

changes during interaction. Future research could use idiodynamic methods or experience sampling to trace how WTC rises and falls within AI-mediated talk (Khajavy et al., 2021; MacIntyre & Legatto, 2011). Further studies could also compare different AI voices, response styles, or levels of personalization, and include human interlocutor conditions to clarify what is distinctive about AI-supported speaking.

For ELT educators, the findings suggest that AI speaking activities should be designed as interactional practice, not only as opportunities for language production. ChatGPT can provide a space where learners rehearse, try out imperfect language, and continue speaking with less concern about immediate evaluation, a possibility consistent with recent studies on AI-supported speaking and reduced language anxiety (Liu et al., 2026; Zhang et al., 2024). This also means that pedagogical attention should extend beyond the speaking task itself to the conditions under which the learner–AI interaction unfolds. Because learner agency is shaped by both the possibilities and constraints of the surrounding human–technical arrangement (Matthews, 2024), teachers may need to consider features such as response length, turn-taking, repair opportunities, voice settings, and the degree of learner control over the exchange. Alongside this, teachers should help learners develop a critical stance toward AI feedback, which can be overgeneralized, inconsistent across platforms, and insufficiently attuned to individual proficiency levels and backgrounds (Chan, 2025). The capacity to evaluate and selectively use AI output, rather than accept it uncritically, is part of the communicative competence that AI-mediated practice alone cannot develop. This positions the teacher as an active pedagogical agent rather than merely a facilitator, responsible for maintaining instructional direction, evaluating AI-supported learning, and providing sustained interpersonal engagement that AI-mediated interaction cannot replicate. These responsibilities remain fundamentally human regardless of how capable AI systems become (Stockwell & Wang, 2025). AI speaking tasks are consequently most effective when embedded within a broader instructional design that includes clear learning goals and human-mediated interaction, rather than used as isolated practice disconnected from the curriculum (Chan, 2025). Integrated in this way, AI-mediated speaking may help learners build confidence, communicative initiative, and greater awareness of how different interlocutors shape their WTC.

Use of Generative AI

This manuscript made use of ChatGPT and Claude AI to assist in language editing, clarity, and organization. All AI-assisted output was reviewed and revised by the authors, who take full responsibility for the final content.

Funding Source

This research was funded by the Faculty of Liberal Arts, Prince of Songkla University “2025 Research Grant” [LIA6804097S].

References

- Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI friend: How users of a social chatbot understand their human–AI friendship. *Human Communication Research*, 48(3), 404–429. <https://doi.org/10.1093/hcr/hqac008>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. Sage Publications.
- Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behaviour in whole class, group and dyadic interaction. *System*, 34(4), 480–493. <https://doi.org/10.1016/j.system.2006.05.002>

- Chaisiri, P. (2023). Promoting Thai EFL learners' willingness to communicate in the virtual classroom through technology-mediated oral tasks. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 97–120. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/263435>
- Chan, K. L. R. (2025). Exploring generative artificial intelligence-aided language learning tools in language classrooms: Insights from language teachers of English in Hong Kong. *Technology in Language Teaching & Learning*, 7(2), 103217. <https://doi.org/10.29140/tltl.v7n2.103217>
- Creswell, J. (2003). *Research design: Qualitative, quantitative and mixed methods approaches*. Sage Publications.
- Deng, Z., & Yan, J. (2025). The effect of perceived warmth, competence, and social presence of AI-Driven chatbots on consumers' engagement and satisfaction. *Sage Open*, 15(3). <https://doi.org/10.1177/21582440251365438>
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. <https://doi.org/10.1037/0033-295X.114.4.864>
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675–701. <https://doi.org/10.2307/2279372>
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28(12), 15873–15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Jeon, J., & Lee, S. (2024). Can learners benefit from chatbots instead of humans? A systematic review of human-chatbot comparison research in language education. *Education and Information Technologies*, 29(17), 23329–23360. <https://doi.org/10.1007/s10639-024-12725-9>
- Jindapitak, N., & Tantiwich, K. (2025). I'm from Singapore lah: AI-simulated interlocutors' responses during in-the-cloud interviews with Thai EFL learners. *Ampersand*, 15, 100241. <https://doi.org/10.1016/j.amper.2025.100241>
- Kang, S. J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System*, 33(2), 277–292. <https://doi.org/10.1016/j.system.2004.10.004>
- Khajavy, G. H., MacIntyre, P. D., Taherian, T., & Ross, J. (2021). Examining the dynamic relationships between willingness to communicate, anxiety and enjoyment using the experience sampling method. In N. Zarrinabadi & M. Pawlak (Eds.), *New perspectives on willingness to communicate in a second language* (pp. 169–197). Springer.
- Kim, A., & Su, Y. (2024). How implementing an AI chatbot impacts Korean as a foreign language learners' willingness to communicate in Korean. *System*, 122, 103256. <https://doi.org/10.1016/j.system.2024.103256>
- Kirkpatrick, R., Vafadar, H., & Mohebbi, H. (2024). A scoping review of willingness to communicate in language education: Premises to doubt, lessons to learn, and future research questions to examine. *Language Testing in Asia*, 14(1), 9. <https://doi.org/10.1186/s40468-024-00284-2>
- Kruk, M. (2022). Dynamicity of perceived willingness to communicate, motivation, boredom and anxiety in second life: The case of two advanced learners of English. *Computer Assisted Language Learning*, 35(1–2), 190–216. <https://doi.org/10.1080/09588221.2019.1677722>
- Lee, H., & Lee, J. H. (2026). How do willingness to communicate and its high evidence factors shape learners' task emotions in chatbot-assisted language learning? *Language Awareness*, 35(2), 692–711. <https://doi.org/10.1080/09658416.2025.2574010>
- Lee, I., & Hahn, S. (2024). On the relationship between mind perception and social support of chatbots. *Frontiers in Psychology*, 15, 1282036. <https://doi.org/10.3389/fpsyg.2024.1282036>

- Lee, J. S., & Dressman, M. (2018). When IDLE hands make an English workshop: Informal digital learning of English and language proficiency. *TESOL Quarterly*, 52(2), 435–445. <http://www.jstor.org/stable/44986999>
- Lee, J. S., & Liu, L. (2024). Dynamicity of EFL learners' willingness to communicate in an online class. *Journal of Multilingual and Multicultural Development*, 45(6), 2078–2096. <https://doi.org/10.1080/01434632.2022.2039677>
- Lee, S., & Jeon, J. (2024). Visualizing a disembodied agent: young EFL learners' perceptions of voice-controlled conversational agents as language partners. *Computer Assisted Language Learning*, 37(5–6), 1048–1073. <https://doi.org/10.1080/09588221.2022.2067182>
- Liang, W., & Wu, Y. (2024). Exploring the use of ChatGPT to foster EFL learners' critical thinking skills from a post-humanist perspective. *Thinking Skills and Creativity*, 54, 101645. <https://doi.org/10.1016/j.tsc.2024.101645>
- Liu, Z., Zhang, W., & Yang, P. (2026). Can AI chatbots effectively improve EFL learners' learning effects?—A meta-analysis of empirical research from 2022–2024. *Computer Assisted Language Learning*, 39(4), 1176–1202. <https://doi.org/10.1080/09588221.2025.2456512>
- MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports*, 11(2), 135–142. <https://doi.org/10.1080/08824099409359951>
- 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. <https://doi.org/10.1111/j.1540-4781.1998.tb05543.x>
- MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, 32(2), 149–171. <https://doi.org/10.1093/applin/amq037>
- Matthews, B. (2024). Finding the human in an era of machine intelligence: A flat ontological analysis of generative AI and language learning. *Technology in Language Teaching & Learning*, 6(3), 1612. <https://doi.org/10.29140/tltl.v6n3.1612>
- Merriam S. B., & Tisdell E. J. (2016). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Ou, A. W., Stöhr, C., & Malmström, H. (2024). Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System*, 121, 103225. <https://doi.org/10.1016/j.system.2024.103225>
- Peng, J. E., & Liang, W. (2026). Willingness to communicate with artificial intelligence (AI)? insights from tracking EFL learners' perceived acceptance and chat output. *Computer Assisted Language Learning*, 39(5–6), 1684–1712. <https://doi.org/10.1080/09588221.2025.2486147>
- Pennycook, A. (2018). *Posthumanist applied linguistics*. Routledge.
- Sato, R. (2023). Examining fluctuations in the WTC of Japanese EFL speakers: Language proficiency, affective and conditional factors. *Language Teaching Research*, 27(4), 974–994. <https://doi.org/10.1177/1362168820977825>
- Schuetzler, R. M., Grimes, G. M., & Scott Giboney, J. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal of Management Information Systems*, 37(3), 875–900. <https://doi.org/10.1080/07421222.2020.1790204>
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.
- Stockwell, G., & Wang, Y. (2025). Framing human-AI collaboration in language education. *Technology in Language Teaching & Learning*, 7(4), 103824. <https://doi.org/10.29140/tltl.v7n4.103824>
- Tabuchi, K., Kobayashi, S., Fukuda, S. T., & Nakagawa, Y. (2024). A case study on reducing language anxiety and enhancing speaking skills through online conversation lessons. *Technology in Language Teaching & Learning*, 6(3), 1497. <https://doi.org/10.29140/tltl.v6n3.1497>

- Tai, T. Y. (2024). Comparing the effects of intelligent personal assistant-human and human-human interactions on EFL learners' willingness to communicate beyond the classroom. *Computers & Education*, 210, 104965. <https://doi.org/10.1016/j.compedu.2023.104965>
- Tai, T. Y., & Chen, H. H. J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 31(3), 1485–1502. <https://doi.org/10.1080/10494820.2020.1841801>
- Waluyo, B., & Isti Pratiwi, D. (2025). AI chatbot-assisted English learning and willingness to communicate: A narrative meta-synthesis of evidence from Asian English as a foreign language contexts. *The JALT CALL Journal*, 21(3), 102884. <https://doi.org/10.29140/jaltcall.v21n3.102884>
- Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System*, 127, 103533. <https://doi.org/10.1016/j.system.2024.103533>
- Xin, Z., & Derakhshan, A. (2025). From excitement to anxiety: Exploring English as a foreign language learners' emotional experiences in the artificial intelligence-powered classrooms. *European Journal of Education*, 60(1), e12845. <https://doi.org/10.1111/ejed.12845>
- Xing, C., & Saeed, M. A. (2025). A systematic review on artificial intelligence (AI) technologies in ESL/EFL speaking skills. *International Journal of TESOL Studies*, 8(2), 240–270. <https://doi.org/10.58304/ijts.250908>
- Xu, K., Chen, X., & Huang, L. (2022). Deep mind in social responses to technologies: A new approach to explaining the computers are social actors phenomena. *Computers in Human Behavior*, 134, 107321. <https://doi.org/10.1016/j.chb.2022.107321>
- Xu, Y., & Warschauer, M. (2020, April). What are you talking to?: Understanding children's perceptions of conversational agents. In *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).
- Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, 54(1), 119–152. <https://doi.org/10.1111/j.1467-9922.2004.00250.x>
- Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, 121, 103259. <https://doi.org/10.1016/j.system.2024.103259>