Journal of Technology and Chinese Language Teaching Volume 15 Number 1, June 2024
http://www.tclt.us/journal/2024v15n1/lizhanglauliang.pdf pp. 25-48

Predicting Chinese Language Learners' ChatGPT Acceptance in Oral Language Practices: The Role of Learning Motivation and Willingness to Communicate
(预测中文学习者在口语练习中对 ChatGPT 的接受度：学习动机和交流意愿的作用)

Li, Nuoen
(李诺恩)
The Chinese University of Hong Kong
(香港中文大学)
nuoenli@link.cuhk.edu.hk

Zhang, Lan
(张岚)
The Chinese University of Hong Kong
(香港中文大学)
1155128312@link.cuhk.edu.hk

Lau, Kit Ling
(刘洁玲)
The Chinese University of Hong Kong
(香港中文大学)
dinkylau@cuhk.edu.hk

Liang, Yu
(梁宇)
Beijing Language and Culture University
(北京语言大学)
liangyu@blcu.edu.cn

Abstract: Despite an increased interest in the potential benefits of ChatGPT for foreign language education, learners' intentions to use ChatGPT as a learning tool have so far received little research attention. This study aims at exploring Chinese language learners' acceptance of ChatGPT in oral language practices and its influencing factors based on the technology acceptance model (TAM). Data were collected from 375 Mongolian learners who learned Chinese as a foreign language (CFL) and analyzed by means of partial least squares structural equation modeling (PLS-SEM). The results indicated that learning motivation and willingness to communicate are critical antecedents of ChatGPT acceptance, and willingness to communicate has a critical mediating role on the link between the three motivational determinants (self-efficacy, utility value, and attainment value) and TAM variables. Prediction-oriented segmentation (POS) was further carried out and found unobserved heterogeneity among CFL learners' formation of ChatGPT acceptance rooted in the years of Chinese learning. The findings suggest the theoretical strengths of TAM in explaining CFL learners' adoption of AI-assisted language practices. Meanwhile, it underlies the importance to understand learners' psychological attributes before introducing technology-assisted speaking practices. Pedagogical insights into how to enhance ChatGPT acceptance among different learner populations were also offered.

摘要：尽管 ChatGPT 在外语教育中的潜在优势逐渐成为热点话题，但学习者将 ChatGPT 作为学习工具的意愿至今未受到充分的研究关注。

本研究旨在基于技术接受模型（ TAM ），探究中文学习者在口语练习中对 ChatGPT 的接受度及其影响因素。研究数据收集自 375 名蒙古中文学习者，并采用偏最小二乘结构方程模型（ PLS-SEM ）进行分析。结果表明，学习动机和交流意愿是影响 ChatGPT 接受度的关键前因，而交流意愿在三个动机因素（自我效能、实用价值与成就价值）与 TAM 变量之间的关联路径中发挥着重要的中介作用。预测导向分割（ POS ）分析进一步揭示了不同学习年限的中文学习者在 ChatGPT 接受度形成机制中存在异质性。上述研究结果证实了 TAM 在解释中文学习者采纳 AI 辅助语言实践方面的理论优势。同时，这也强调在引入技术辅助中文口语练习活动前了解学习者心理的重要性。此外，本文针对如何增强 ChatGPT 在不同学习者群体中的接受度提出了教学建议。

Keywords: ChatGPT, technology acceptance, oral language practices, learning motivation, willingness to communicate

关键词：ChatGPT、技术接受、口语练习、学习动机、交流意愿

1. Introduction

Artificial intelligence (AI) creates a novel paradigm for promoting the efficiency, effectiveness, and outcomes of teaching and learning across a wide range of educational settings. In foreign language (FL) education, chatbots can serve as a tireless language partner for learners to practice speaking with, or even an effective tutor or instructor to deliver extra knowledge that a language partner may not be able to provide due to their limited language proficiency level (Huang et al., 2022). Many studies have revealed the potential contributions that AI-powered chatbots could bring to oral language practices, such as improving language accuracy and fluency (Ruan et al., 2021), mitigating learners' anxiety (Hsu et al., 2023; Jeon, 2022), and enhancing engagement in speaking activities (Jeon, 2022; Ruan et al., 2021). Though these advantages have been generally acknowledged by researchers in the field, the integration of chatbots into FL learning practices greatly depends on learners' awareness of chatbots' practical value and their willingness to adopt them as regular learning tools. In this regard, the perceptions and acceptance of chatbots among FL learners warrant further research attention.

Previous research has sought to understand learners' acceptance of chatbots and its influencing factors. Several concerns, however, have appeared in the earlier investigations. First, there is a lack of empirical support for theoretical assumptions of information systems (IS) acceptance. The majority of relevant research was based on the technology acceptance model (TAM) (Davis, 1989), arguably the most popular yet parsimonious model in the IS field (Srite, 2006). However, certain theoretically conceptualized relationships between TAM variables have not been empirically validated yet (e.g., Liu & Ma, 2024), implying that further research is still necessary to determine the applicability of TAM in exploring chatbot acceptance. Second, the potential unobserved heterogeneity of learners' chatbot acceptance has received insufficient attention. According to Becker et al. (2013), unobserved heterogeneity captures situations where there is no clear theoretical account for heterogeneity in a certain population, in contrast to observed heterogeneity where prior knowledge about the group differences has been acquired. Existing literature has shown inconsistent findings regarding the relationships between TAM variables in the context of AI-assisted FL learning, which might be attributed to variations in participants' backgrounds across those studies. For instance, the hypothesized positive impact of perceived ease of use on attitudes failed to reach a statistically significant level in Liu and Ma (2024) with English language learners from various backgrounds in China, whereas it was supported by Belda-Medina and Calvo-Ferrer (2022) through a survey among college-level English language learners in Spain and Poland. Hence, the heterogeneity regarding the developing pattern of chatbot acceptance among learner populations with different backgrounds warrants further investigation. Third, little research effort was devoted to FL education, and there has been insufficient emphasis on learners' acceptance of chatbots for speaking practices. According to Petrović and Jovanović (2021), the most natural and effective application of chatbots is related to their fundamental nature—language practice. The capacity of chatbots could provide valuable learning opportunities, particularly for FL learners to practices their language either in text-based or oral-based manner, which requires special attention in the FL field. The recently released ChatGPT considerably expands the technological affordances of GenAI-powered chatbots in enabling better oral-based communication by providing customized feedback, answering follow-up questions, and generating more authentic and natural conversations (Kohnke et al., 2023; Tlili et al., 2023). Therefore, it is worthwhile to investigate FL learners' acceptance of ChatGPT, especially in oral language practices.

Motivation has been found to be an important source of users' acceptance and usage behavior of information technologies (Venkatesh, 2000; Venkatesh et al., 2003). To date, a range of motivational determinants has been identified as essential external variables for chatbot adoption and usage behaviors among FL learners, such as hedonic motivation (Strzelecki, 2023), perceived enjoyment (Chen et al., 2020), and perceived autonomy, relatedness, and competence (Xia et al., 2023). It has also been found that in FL learning, learners who have stronger motivation towards the language they learn are more inclined to actively seek out advantageous technology to optimize their learning experience (Hsu, 2017). Thus, understanding technology acceptance from an academic-learning motivational perspective would offer valuable insights into the matter. Furthermore, learners' willingness to communicate (WTC), which refers to their readiness to enter into discourse using a second or foreign language, is particularly crucial for their decisions to initiate communication as a volitional process (MacIntyre et al., 1998). When it comes to FL oral language practices, whether learners voluntarily commit to the advantages of technology in offering communication opportunities might also be greatly determined by their willingness to enter into discourse using the target language. Therefore, WTC should be taken into consideration as a critical individual difference factor influencing FL learners' adoption and usage, especially for oral-based interaction-enabling technologies.

In light of the above discussions, this study aims to explore CFL learners' acceptance of ChatGPT in oral language practices, and the role of learning motivation and WTC in affecting their acceptance by means of partial least squares structural equation modeling (PLS-SEM). Prediction-oriented segmentation (POS), a distance-based segment detection method in PLS path models, will be employed to investigate if there is any unobserved heterogeneity among learners' ChatGPT acceptance. The specific research questions that guide this study are:

1. Are the theoretical assumptions between TAM variables supported in the context of using ChatGPT in CFL oral language practices?
2. Do learning motivation and willingness to communicate have significant effects on CFL learners' ChatGPT acceptance in oral language practices?
3. Is there any unobserved heterogeneity among CFL learners regarding their ChatGPT acceptance in oral language practices?

2. Theoretical foundation and model development

2.1 Technology acceptance model

TAM is a well-established model that aims to explain and predict how users accept and use information technologies. According to TAM (Figure 1), the most proximal antecedent of technology use is behavioral intention, and whether an individual intend to use or reject the technology is determined by his/her attitude toward using the given technology. The attitude of the individual was considered to be affected by two key factors: (1) perceived ease of use (PEOU), which refers to ‘the degree to which a person believes that using a particular system would be free of effort’; and (2) perceived usefulness (PU), which refers to ‘the degree to which a person believes that using a particular system would enhance his/her job performance’ (Davis, 1989, p. 320). It also posits that PEOU has a significant positive effect on PU. In addition, Davis et al. (1989) suggested that an individual might form a strong behavioral intention towards a certain behavior they believe will increase their job performance, thus deriving the hypothesis regarding the direct positive effect of PU on behavioral intention. Hence, the following hypotheses were proposed on the basis of TAM's theoretical underpinnings:

H1: Perceived ease of use (PEOU) has a positive effect on perceived usefulness (PU).
H2: Perceived ease of use (PEOU) has a positive effect on attitude toward using (ATU).
H3: Perceived usefulness (PU) has a positive effect on attitude toward using (ATU).
H4: Perceived usefulness (PU) has a positive effect on behavioral intention to use (BIU).
H5: Attitude toward using (ATU) has a positive effect on behavioral intention to use (BIU).

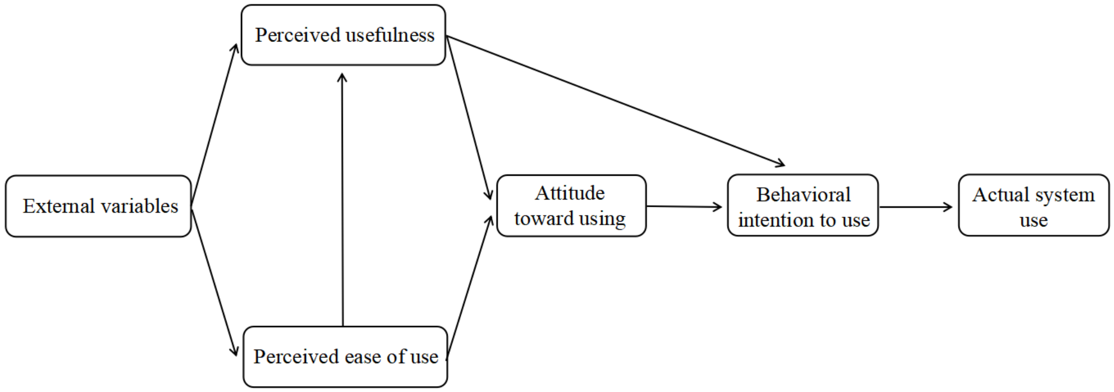

Figure 1 Technology acceptance model (Davis et al., 1989)

2.2 Relations between learning motivation and willingness to communicate

Learning motivation is an important individual difference variable that have been extensively investigated in FL research, which refers to an amalgamation of desires, attitudes, and efforts that encourage learners to learn the target language (Gardner, 1985). Among current motivation theories, expectancy-value theory (EVT) (Eccles-Parsons et al., 1983) demonstrated valuable promise in analyzing academic learning motivation with its overarching theoretical construct (Loh, 2019; Wang & Xue, 2022). Specifically, the conceptualized motivational determinants in EVT not only concerned learners' ability beliefs with *expectancy*, but also various motivational valences of the subjective task with *task values* in terms of shaping self-schema, achieving instrumentality, and offering enjoyment or pleasure.

In a meta-analysis of WTC with 64 studies, Elahi Shirvan et al. (2019) identified motivation as a key variable that influences foreign/second language learners' WTC. The predictive role of learning motivation in WTC has also been found with motivational determinants conceptualized in the EVT framework. In EVT, *expectancy* refers to individuals' beliefs about how they would do on upcoming tasks (Eccels-Parsons et al., 1983). It is highly related to *self-efficacy* that proposed in Bandura (1997), which comprises learners' beliefs on their competence to accomplish a certain task (Wigfield & Eccles, 2000). Therefore, self-efficacy has common be used as one important variable to measure expectancy component in the EVT in empirical research (Bai et al., 2020). The potential positive impact of self-efficacy on WTC has been theoretically reflected in MacIntyre et al.'s (1998) pyramid model of L2 WTC, which conceptualized learners' L2 self-confidence as a critical antecedent of WTC. Empirically, the positive influence of self-efficacy on WTC has also been supported in both traditional language classrooms (e.g., Yang & Lian, 2023) and digital language learning contexts (e.g., Soyoof, 2023; Zadorozhnyy & Lee, 2023).

Another essential aspect of EVT motivation, *task value*, refers to the incentives and reasons for choosing to do a certain work or activity (Eccles-Parsons et al., 1983). There are four components consist of task values: *utility value*, *attainment value*, *intrinsic value*, and *cost*. Utility value, or usefulness, has been defined as how well a particular task fits into individuals' present or future plans; attainment value is the importance of doing well on a given task; and intrinsic value refers to the enjoyment that one gains from doing a task (Eccles-Parsons et al., 1983; Wigfield & Eccles, 2000). Finally, cost is conceptualized as any negative aspect of engaging in a task (e.g., losing alternative opportunities, spending extra efforts, and causing negative emotions). Since cost is a multifaceted mechanism that greatly varies across individuals and still lacks detailed measures exhaustively listing its sources (Rosenzweig et al., 2019), this study solely focused on the former three types of task values.

Prior studies have linked the former three aspects of task values to WTC and revealed a significant relationship between the two constructs. Integrating utility value, attainment value, and intrinsic value as a composite variable, MacIntyre and Blackie (2012) found a significant relationship between task values and WTC among high school L2 French learners. Nagle (2021) also identified attainment value and intrinsic value as significant predictors of WTC with college-level L2 Spanish learners. Based on the comprehensive descriptive insights that EVT could offer into learners' motivational systems and the aforementioned empirical evidence about the influences of EVT motivational determinants on WTC, we formulated the following hypotheses:

H6: Self-efficacy (SE) has a positive effect on willingness to communicate (WTC).
H7: Utility value (UV) has a positive effect on willingness to communicate (WTC).
H8: Attainment value (AV) has a positive effect on willingness to communicate (WTC).
H9: Intrinsic value (IV) has a positive effect on willingness to communicate (WTC).

2.3 Relations between willingness to communicate and technology acceptance

Individual difference is an important sort of antecedent that determines learners' adoption and use of information technologies, which specifically influences PEOU and PU (Venkatesh & Bala, 2008). Previous studies have identified WTC as an important individual characteristic that affects learners' communication behaviors, either in in-class, out-of-class, or digital settings (e.g., Balouchi & Samad, 2021; Lee & Hsieh, 2019; Lee & Lee, 2020). Specifically, Lee and colleagues (Lee & Hsieh, 2019; Lee & Lee, 2020; Lee & Drajati, 2019) found that learners with higher levels of WTC were more likely to attach greater value to language communication, have positive perceptions about information technologies, and seek more opportunities to practice their language communicative skills in assistance with educational technologies. Such influencing links may arise from the nature of WTC as a final psychological step before actual language communication (Lee, 2020). The great value that communication-oriented learners attach to computer-assisted language interaction might also lead to their active cognitive involvement in the meaningful construction of the provided language input, the adaptation of communication strategies they used, and the close attention to functional features of assisted technologies in the speaking tasks (e.g., Mystkowska-Wiertelak, 2021), and as a result, foster positive perceptions about the usefulness and usability of communication-enabling technologies. Especially when it comes to oral language practices, a challenging task for many language learners that requires high cognitive engagement to understand spoken language and initiate communication immediately and effectively (Hsu et al., 2023), WTC might serve an even more crucial role in learners' perceptions towards chatbots and thereafter adoption decisions. Therefore, we proposed that:

> **H10:** Willingness to communicate (WTC) has a positive effect on perceived ease of use (PEOU).
> **H11:** Willingness to communicate (WTC) has a positive effect on perceived usefulness (PU).

2.4 The hypothesized model

Based on the above hypotheses, we developed a research model to predict FL learners' ChatGPT acceptance in oral language practices (Figure 2). Individual differences were explicitly targeted as the external variables, which include learning motivation and willingness to communicate. Learning motivation was further conceptually specified based on EVT, which consists of self-efficacy, utility value, attainment value, and intrinsic value. The technology acceptance construct was developed based on TAM, which consists of perceived ease of use, perceived usefulness, attitude toward using, and behavioral intention to use. In addition, Lee and Lu (2023) showed that learning motivation significantly predicted WTC in both classroom and digital learning environments, and learners with high WTC tend to enthusiastically seek opportunities for text-based or oral-based interactions with information and communication technologies. Thus, we also hypothesized that WTC has a mediating effect on the link between learning motivation and technology acceptance.

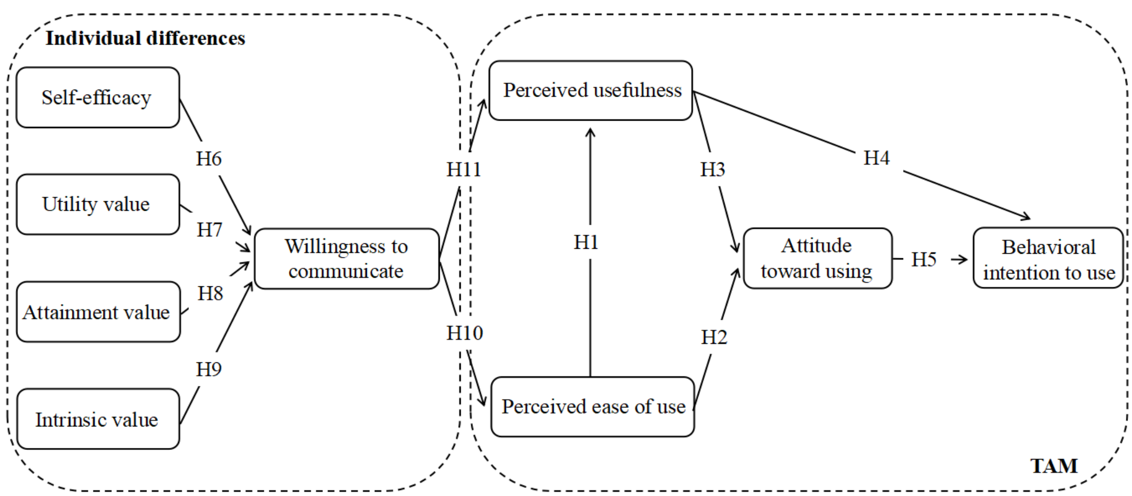

Figure 2 Proposed research model

3. Research method

3.1 Research context and participants

This study was conducted at two comprehensive universities in Mongolia with the target to college-level learners who learned Chinese-as-a-foreign-language (CFL) as a compulsory course. To ensure all the potential participants formed clear perceptions towards ChatGPT, this study conducted group-based oral Chinese practice activities with ChatGPT-3.5 using the Chrome extension ‘Voice for Control ChatGPT’ in a total of 12 CFL classes prior to data collection, with the following steps: (1) learners were randomly divided into groups of three to five by their CFL instructor at first; (2) five minutes were then given to learners to discuss the topic they intended to speak about with ChatGPT, which was either based on personal interests or referenced the topics provided by their instructor. The oral practice topics that the instructor provided were designed according to learners' CFL textbooks and differentiated by learners' average level of language proficiency across classes, as shown in Table 1; (3) each group took turns participating in the discussion activity with ChatGPT (five to ten minutes) during class, and every learner in the group was required to take at least two conversational turns with ChatGPT in this process. A discussion example was illustrated in Figure 3, where the group of learners were curious about the best place to visit in China. Through the discussion with ChatGPT, learners in the group finally gained more knowledge about the Hutongs and reached an agreement to travel to Beijing; (4) the activities were repeated twice in one week with the above-mentioned group format and activity procedures. Learners were also encouraged to explore the use of ChatGPT as a chatbot in their extracurricular time to get a clearer understanding of the functions and features of ChatGPT.

Table 1 The samples of oral practice topic

Target learner	Topic	Sample initiating question
Beginner level	See doctors	What should I do if I am sick?
Intermediate level	Traveling	What is the best place to visit in China?
Advanced level	Chinese New Year	How do Chinese celebrate Chinese New Year?

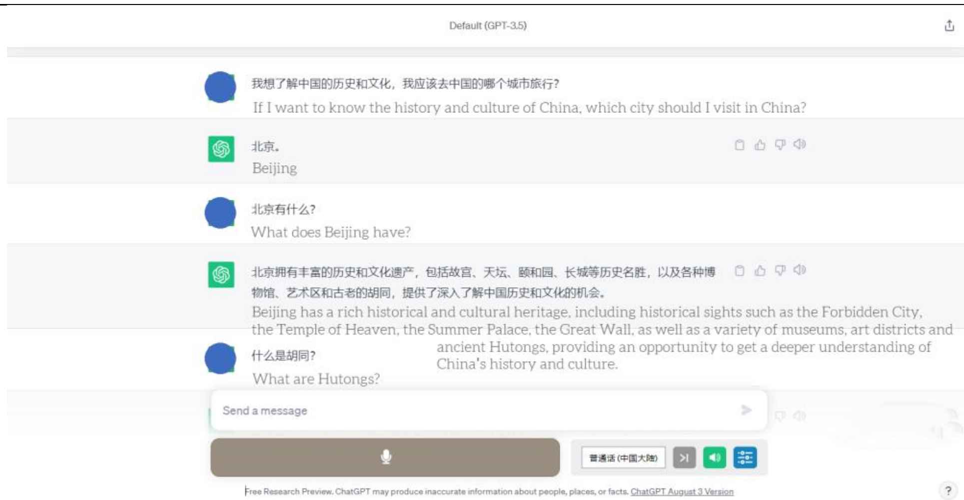

Figure 3 An example of group work in the discussion activities with ChatGPT

Following the completion of the activities, all CFL learners were invited to complete an unidentifiable questionnaire through the online questionnaire tool Wen Juan Xing. Only those learners who completed both the group-based oral practice activities and the questionnaire survey were regarded as final research participants in this study. The demographic information of the final 375 participants is presented in Table 2.

Table 2 Demographic information of research participants (*N* = 375)

	Category	Frequency	%
Age (*M*±*SD*)		20.53±2.15	—
Gender	Male	99	26.40%
	Female	276	73.60%
Years of Chinese learning	≤ 1 year	150	40.00%
	1-3 years	156	41.60%
	≥ 3 years	69	18.40%
Chinese language proficiency	Beginner level (level 1-2)	4	1.07%
	Intermediate level (level 3-4)	145	38.67%
	Advanced level (level 5-6)	119	31.73%
	Never participated	107	28.53%

Chinese language proficiency was referenced with the participants' passing level in Hanyu Shuiping Kaoshi (HSK, see Peng et al., 2020 for a detailed description of HSK).

3.2 Instruments

Two questionnaires were developed based on existing valid instruments: (1) the first questionnaire comprised 16 items that measured five variables of individual differences to reflect learners' EVT-based academic learning motivation and WTC. Items for self-efficacy (SE) were adapted from Shaaban & Ghaith (2000) to measure learners' general self-efficacy for Chinese speaking and self-efficacy for Chinese academic learning, while items for attainment value, utility value, and intrinsic value were adapted from Gaspard et al. (2017) to measure learners' overall task values in Chinese learning. Items concerning WTC were adapted from Lee and Lee (2020), which specifically focused on inside classroom situations; and (2) the second questionnaire consisted of 14 items that measured four variables in TAM to reflect learners' ChatGPT acceptance in oral language practices. Items for perceived ease of use and perceived usefulness were adapted from Davis (1989), while items for attitude toward using and behavioral intention to use were adapted from Venkatesh et al. (2003). The adaptation on the two instruments was mainly about phrasing the items within the CFL learning and ChatGPT-assisted oral language practice contexts. All the items were measured on a 5-point Likert scale (1: Strongly disagree to 5: Strongly agree).

3.3 Data analysis

PLS-SEM was implemented to assess the measurement model and the hypothesized structural model in SmartPLS 4.0. The reasons to use PLS-SEM rather than covariance-based structural equation modeling (CB-SEM) were: (1) little research have used CB-SEM to investigate FL learners' EVT-based motivation, willingness to communicate, and ChatGPT acceptance in oral language practices, therefore, the high level of statistical power of PLS-SEM would benefit the theory developing from a predictive standpoint (Hair et al., 2019); (2) PLS-SEM is not subject to data distribution restrictions, while CB-SEM can produce abnormal results with non-normal data (Hair et al., 2019); and (3) PLS-SEM offers better solutions with a small sample size (Hair et al., 2019; 2022). The inverse square root method proposed by Kock and Hadaya (2018) was used to determine whether our sample size was sufficient for PLS-SEM analysis. Given the anticipated effect size of 0.20 and the desired probability of 0.05, 155 samples would be required to detect the effect. Thus, 375 samples in this study met the minimum sample size requirements for PLS-SEM. Prediction-oriented segmentation (POS), a method for detecting unobserved heterogeneity that was specifically developed to fit PLS path modeling, was further employed to test whether the examined research model significantly differed among research participants. Compared to the traditional approach to segmentation in SEM by assigning samples to predefined segments on the basis of demographic variables, POS is especially beneficial in identifying potential heterogeneity in a case where there is a lack of ground rationale for distinguishing subgroups within a population, allowing for more efficient capturing of heterogeneity while avoiding under- or over-segmenting (Hair et al., 2016; Rigdon et al., 2010). Given that there has been little previous research on the disparities in the influencing relationships between learning motivation, WTC, and technology acceptance across different FL learner populations, we were thus conducting POS in an attempt to detect any unobserved heterogeneity from a predictive perspective. The demographic backgrounds of learners in different groups were also compared based on the POS results.

4. Results

4.1 The measurement model

The reliability and convergent validity were assessed through factor loadings, Cronbach' s alpha, composite reliability (CR), and average extracted variance (AVE). The factor loadings of each indicator ranged from 0.78 to 0.94 (Table 3). Both the Cronbach's alpha and CR (rho_c) for each latent variable were higher than the recommended value of 0.70, and the AVE for each latent variable exceeded the minimum requirement of 0.50, which corroborates the reliability and convergent validity of the measurement model.

Table 3 Reliability and convergent validity of the measurement model

Variable	Item	Factor loading	Mean	SD	α	CR	AVE
Self-efficacy (SE)	SE1	0.789	3.25	1.16	0.826	0.882	0.652
	SE2	0.784	3.29	1.12			
	SE3	0.848	3.54	1.04			
	SE4	0.807	3.84	1.07			
Utility value (UV)	UV1	0.901	4.24	0.98	0.859	0.914	0.780
	UV2	0.898	4.16	1.03			
	UV3	0.849	4.30	0.97			
Attainment value	AV1	0.934	4.30	0.97			

(AV)	AV2	0.931	4.21	1.00	0.920	0.949	0.862
	AV3	0.920	4.25	1.05			
Intrinsic value (IV)	IV1	0.941	3.90	1.11	0.924	0.951	0.867
	IV2	0.930	3.93	1.10			
	IV3	0.922	3.91	1.02			
Willingness to communicate (WTC)	WTC1	0.885	3.96	1.07	0.850	0.909	0.769
	WTC2	0.911	3.94	1.06			
	WTC3	0.834	3.78	1.09			
Perceived ease of use (PEOU)	PEOU1	0.904	3.84	0.96	0.917	0.941	0.800
	PEOU2	0.907	3.85	0.95			
	PEOU3	0.889	3.75	0.97			
	PEOU4	0.878	3.76	0.97			
Perceived usefulness (PU)	PU1	0.882	3.79	0.89	0.901	0.931	0.771
	PU2	0.889	3.85	0.83			
	PU3	0.872	3.81	0.90			
	PU4	0.870	3.79	0.97			
Attitude toward using (ATU)	ATU1	0.894	3.94	0.96	0.850	0.909	0.770
	ATU2	0.863	3.96	0.95			
	ATU3	0.874	3.94	0.95			
Behavioral intention to use (BIU)	BIU1	0.899	3.63	1.10	0.891	0.933	0.822
	BIU2	0.931	3.55	1.18			
	BIU3	0.889	3.65	1.10			

Discriminant validity was analyzed by using the Fornell-Larcker criterion (Fornell & Larcker, 1981) and the Heterotrait-Monotrait method (Henseler et al., 2015). The square root of AVE for each variable exceeded the correlations among latent variables (Table 4), which fulfilled the Fornell-Larcker criterion. The HTMT ratios of all latent variables were less than the criteria of 0.85 in Hair et al. (2019) (Table 5), demonstrating adequate discriminant validity of the measurement model. Furthermore, the variance inflation factor (VIF) values of all indicators ranged from 1.68 to 4.24, which is less than the suggested cut-off value of 5.0 in Hair et al. (2022), indicating high collinearity was not an issue in this study.

Table 4 Discriminant validity based on Fornell-Larcker criterion

	SE	UV	AV	IV	WTC	PEOU	PU	ATU	BIU
SE	0.807								
UV	0.163	0.883							
AV	0.069	0.480	0.928						
IV	0.003	0.396	0.564	0.931					
WTC	0.291	0.389	0.385	0.318	0.877				
PEOU	0.175	0.155	0.341	0.214	0.340	0.895			
PU	0.255	0.350	0.457	0.335	0.374	0.529	0.878		
ATU	0.205	0.143	0.331	0.263	0.330	0.506	0.515	0.877	
BIU	0.095	0.269	0.440	0.421	0.299	0.385	0.450	0.448	0.906

Table 5 Discriminant validity based on Heterotrait-Monotrait method

	SE	UV	AV	IV	WTC	PEOU	PU	ATU
UV	0.185							
AV	0.090	0.546						
IV	0.062	0.452	0.616					
WTC	0.333	0.451	0.430	0.351				
PEOU	0.204	0.173	0.369	0.231	0.382			
PU	0.292	0.401	0.501	0.364	0.428	0.577		
ATU	0.244	0.171	0.376	0.294	0.384	0.573	0.585	
BIU	0.112	0.311	0.486	0.461	0.340	0.426	0.500	0.515

4.2 The structural model

As the reliability and validity of the measurement model have been established, we examined the structural model to evaluate model quality and test the proposed hypotheses. The R^2 values of endogenous variables and the Stone-Geisser test (Q^2) were applied to ensure the predictive relevance of the model. According to Hair and Alamer (2022), R^2 values between 0 to 0.10, 0.11 to 0.30, 0.30 to 50, and > 0.50 indicate weak, modest, moderate, and strong explanatory power in L2 research. The R^2 values of endogenous variables in the structural model ranged between 0.12 to 0.34, indicating modest to moderate explanatory power. Q^2 values should be greater than zero for a particular endogenous variable to indicate predictive accuracy (Hair et al., 2022). The Q^2 values of all endogenous variables in the structural model were above zero, ranging from 0.09 to 0.26, thus establishing the predictive accuracy of the model.

The structural relationships between the latent variables are presented in Table 6. Eight hypotheses (H1-6, H10, and H11) were supported at a significant level of $p < .001$, and two hypotheses (H7 and H8) were supported at a significant level of $p < .01$. H9 was supported at a significant level of $p < .05$ but had a path coefficient below 0.02, and thus should be eliminated from the nested model. Indirect effects of the four motivational determinants on TAM variables were also examined under maximum likelihood estimation with 5,000 bootstrap samples (Table 7). Except for IV, all indirect paths from SE, UV, and AV to the four TAM variables reached significance, revealing the critical mediating role of WTC in the structural model.

Table 6. Path coefficients of the proposed research model

	Path	β	T statistics	f^2	p	Results
H1	PEOU -> PU	0.455	8.527	0.270	$< .001^{***}$	Supported
H2	PEOU -> ATU	0.325	5.196	0.115	$< .001^{***}$	Supported
H3	PU -> ATU	0.343	5.591	0.128	$< .001^{***}$	Supported
H4	PU -> BIU	0.298	5.172	0.089	$< .001^{***}$	Supported
H5	ATU -> BIU	0.295	5.623	0.087	$< .001^{***}$	Supported
H6	SE -> WTC	0.243	5.474	0.078	$< .001^{***}$	Supported
H7	UV -> WTC	0.204	3.370	0.040	$.001^{**}$	Supported
H8	AV -> WTC	0.200	2.830	0.033	$.004^{**}$	Supported
H9	IV -> WTC	0.123	2.024	0.014	$.043^{*}$	Supported
H10	WTC -> PEOU	0.340	6.211	0.131	$< .001^{***}$	Supported
H11	WTC -> PU	0.220	4.422	0.063	$< .001^{***}$	Supported

$^{*} p < .05$, $^{**} p < .01$, $^{***} p < .001$.

Table 7 Indirect effects of learning motivation

	β	T statistics	Bias-corrected 95% CI		p
			Lower	Upper	
SE -> PEOU	0.083	3.833	0.046	0.128	$< .001^{***}$
SE -> PU	0.091	4.058	0.050	0.138	$< .001^{***}$
SE -> ATU	0.058	3.759	0.031	0.089	$< .001^{***}$
SE -> BIU	0.044	3.513	0.023	0.071	$< .001^{***}$
UV -> PEOU	0.069	3.174	0.032	0.118	$.002^{**}$
UV -> PU	0.076	3.238	0.033	0.125	$.001^{**}$
UV -> ATU	0.049	3.096	0.021	0.083	$.002^{**}$
UV -> BIU	0.037	2.992	0.015	0.064	$.003^{**}$
AV -> PEOU	0.068	2.377	0.019	0.132	$.018^{*}$
AV -> PU	0.075	2.452	0.021	0.142	$.015^{*}$
AV -> ATU	0.048	2.344	0.013	0.093	$.020^{*}$
AV -> BIU	0.036	2.249	0.010	0.075	$.025^{*}$
IV -> PEOU	0.042	1.895	0.002	0.089	.058
IV -> PU	0.046	1.933	0.002	0.096	.053
IV -> ATU	0.029	1.897	0.001	0.062	.058
IV -> BIU	0.022	1.861	0.001	0.048	.063

$^{*} p < .05$, $^{**} p < .01$, $^{***} p < .001$.

4.3 Prediction-oriented segmentation (POS)

PLS-POS was performed in an attempt to find any unobserved heterogeneity among the samples. The sum of all constructs weighted R^2 was chosen as the optimization criterion. Considering the above-mentioned minimum sample requirement, we opted for a 2-segment solution with 1000 iterations and a search depth of 375 to perform PLS-POS. The demographic information of learners in the two segments is displayed in Table 8, and the SEM results in segment 1 (N = 200) and segment 2 (N = 175) are presented in Figure 4 and Figure 5, respectively.

Table 8 Sample demographics in the two segments

	Category	Frequency			
		Segment 1 (*N* = 200)	% within group	Segment 2 (*N* = 175)	% within group
Age (*M*±*SD*)		20.50±2.07		20.56±2.25	
Gender	Male	57	28.5%	42	24.0%
	Female	143	71.5%	133	76.0%
Years of Chinese learning	≤ 1 year	86	43.0%	64	36.6%
	1-3 years	88	44.0%	68	38.9%
	≥ 3 years	26	13.0%	43	24.6%
Chinese language proficiency	Beginner level (level 1-2)	3	1.5%	1	0.6%
	Intermediate level (level 3-4)	77	38.5%	68	38.9%
	Advanced level (level 5-6)	65	32.5%	54	30.9%
	Never participated	55	27.5%	52	29.7%

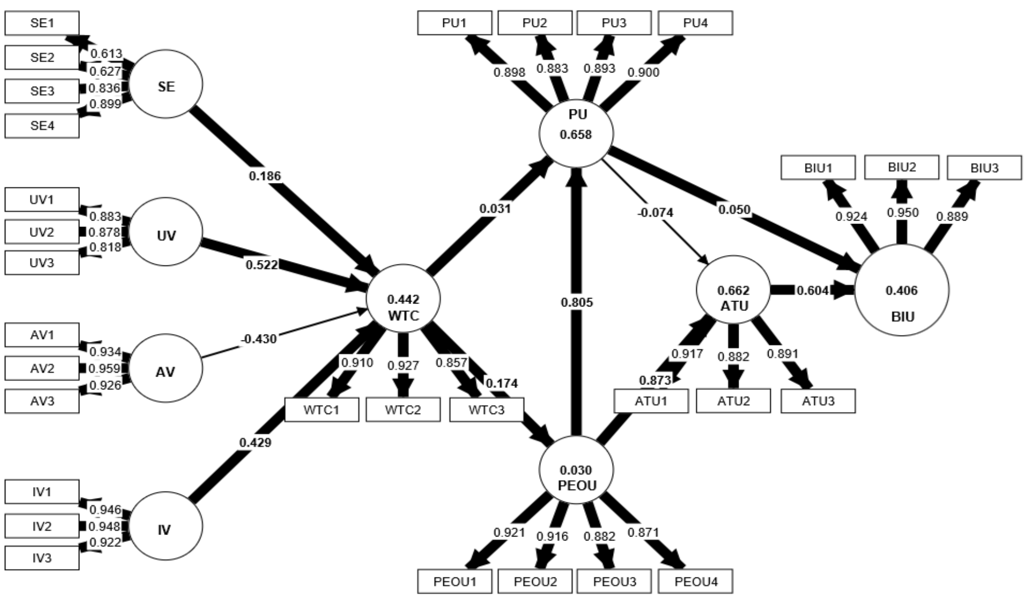

Figure 4 The structural model in segment 1

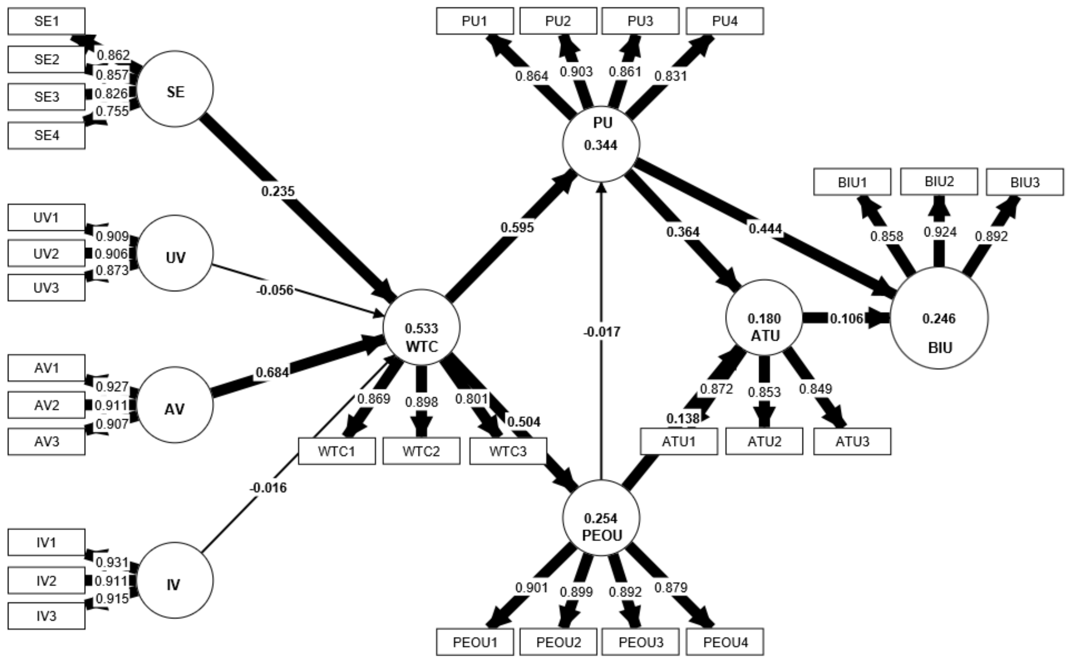

Figure 5 The structural model in segment 2

A multi-group analysis (MGA) was performed to test if there are any significant differences in the path coefficients across the two segmental models. The results indicated that all other paths in the model showed significant differences in their path coefficients, with the exception of the influence of SE on WTC (Table 9). Mann-Whitney U tests were further carried out to evaluate whether learners' age, gender, years of Chinese learning, and language proficiency differed across the two segments. The results indicated that the years of Chinese learning among learners in segment 2 was significantly longer than those among segment 1 ($Z = 2.27$, $p = .023$), but no statistically significant difference was found in other demographic background variables between the two segments.

Table 9 Differences in the path coefficients across the two segments

Path	Segment 1	Segment 2	Path coefficient difference	*p*
PEOU -> PU	0.805	-0.017	0.822	$< .001^{***}$
PEOU -> ATU	0.873	0.138	0.735	$< .001^{***}$
PU -> ATU	-0.074	0.364	-0.438	$< .001^{***}$
PU -> BIU	0.050	0.444	-0.393	$< .001^{***}$
ATU -> BIU	0.604	0.106	0.498	$< .001^{***}$
SE -> WTC	0.186	0.235	-0.049	.598
UV -> WTC	0.522	-0.056	0.578	$< .001^{***}$
AV -> WTC	-0.430	0.684	-1.114	$< .001^{***}$
IV -> WTC	0.429	-0.016	0.445	$.001^{**}$
WTC -> PEOU	0.174	0.504	-0.330	$.003^{**}$
WTC -> PU	0.031	0.595	-0.564	$< .001^{***}$

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

5. Discussion

5.1 The effectiveness of TAM in predicting ChatGPT acceptance

To explore the effectiveness of TAM in the context of accepting ChatGPT as a learning tool in FL oral language practices, we developed a research model in which involved four TAM variables. The five hypotheses (H1-H5) between TAM variables were supported by the PLS-SEM results in this study, which were consistent with the theoretical assumptions in Davis et al. (1989). Specifically, FL learners' PEOU has a significant positive influence on PU, and both of the two positively affect BIU through ATU. The previous investigation on FL learners' ChatGPT acceptance based on TAM in Liu and Ma (2024) dissolved the positive influence of PEOU on ATU. However, in this study, both PEOU and PU were found to be significant predictors of ATU, forming a more holistic picture of TAM's theoretical strengths in predicting FL learners' ChatGPT acceptance. In addition to ATU, PU could also directly affect BIU, even having a stronger influence than ATU, which empirically supports the assumption in Davis et al. (1989) that ‘people form intentions toward using computer systems based largely on a cognitive appraisal of how it will improve their performance’ (p. 986).

5.2 The role of learning motivation and willingness to communicate

To explore the role of learning motivation and WTC in FL learners' ChatGPT acceptance, six hypotheses (H6-H11) were proposed in the research model. SE was found to be a significant predictor of WTC, which supports the theoretical assumption in MacIntyre et al.'s (1998) pyramid model of L2 WTC that learners' positive belief about their language ability is a critical antecedent of their willingness to communicate. Concurred with previous results in MacIntyre and Blackie (2012), AV and UV were also found as significant predictors of WTC. However, the positive effect of IV on WTC, though supported in PLS-SEM, failed to reach a sufficient effect size and thus cannot be accepted in this study. The eliminated influence of IV on WTC might result from the co-existence of other affective or emotional factors (e.g., L2 anxiety, shyness) as restraining forces for language communication (Pavelescu, 2023), which has especially been commonly reported among East Asian language learners under the influence of their cultural system and educational practices (for a detailed review, see Shao & Gao, 2016).

Furthermore, WTC significantly mediated the influences of SE, UV, and AV on the four TAM variables. In other words, FL learners with higher learning motivation are more willing to communicate in the target language and thus inclined to accept ChatGPT as a learning tool in oral language practices. This echoed Eccles-Parsons et al.'s (1983) argument with regard to learning motivation as a critical psychological antecedent of learners' subsequent academic task choices and achievement-related decision making. The findings further concretize the above argument in the context of ChatGPT-assisted oral language practices and highlight the significance of willingness to communicate in such academic decision-making process.

5.3 The unobserved heterogeneity among CFL learners

Due to the complexity of social and behavioral phenomena, heterogeneity in the samples is likely to exist (Becker et al., 2013). PLS-POS results in this study demonstrated that the formation pattern of ChatGPT acceptance is characterized by heterogeneity among CFL learners rooted in their years of Chinese learning. Two main findings can be drawn from the examined heterogeneity of ChatGPT acceptance between learners with different CFL learning experiences:

First, WTC has a greater impact on ChatGPT acceptance among learners with longer Chinese learning experiences compared to their counterparts, since the effects of the two paths from WTC to PEOU and to PU were significantly stronger in segment 2 than in segment 1. The reason for this might be that learners who had been learning Chinese for a longer time had more opportunities to speak the language while attributing their prior academic success to language communication (Wen & Piao, 2020), therefore attaching a greater value on language communication in FL learning and being more open to interacting with ChatGPT. Second, for learners with longer Chinese learning experiences, their BIU benefited more from PU, as the effect of PU on BIU was significantly higher in segment 2 than in segment 1. In contrast, for those learners with shorter Chinese learning experiences, their BIU was more influenced by PEOU through ATU, as the effects of the relevant two paths were significantly higher in segment 1 than in segment 2. Compared to CFL beginners, learners with longer learning experience may have already experimented with different educational technologies, acquired more effective technology-assisted learning techniques, and thus been able to interact with technologies more efficiently (Durndell & Haag, 2002; Luo, 2020). As a result, those long-term CFL learners may place more emphasis on ChatGPT's effectiveness for oral language practice than its efficiency. This finding also reveals that ChatGPT might play different roles among CFL learners. Given that the PU of ChatGPT is more important in forming acceptance for learners with longer Chinese learning experience, they may regard ChatGPT as a tutor or instructor with whom they expect to learn extra language knowledge; on the contrary, those beginners might consider ChatGPT simply as a convenient language partner to interact with because the PEOU of ChatGPT is more crucial for developing their behavioral intentions.

6. Implications, and limitations, conclusions

This study sought to predict CFL learners' acceptance of ChatGPT in oral language practices with learning motivation and willingness to communicate, as well as explore any potential heterogeneity of ChatGPT acceptance among CFL learners. The results of this study provide evidence on the effectiveness of TAM in investigating ChatGPT acceptance in the context of CFL oral language practices. TAM has recently been employed and validated in AI-assisted language learning, with a focus on automated writing evaluation (e.g., Li et al., 2019), intelligent tutoring systems (e.g., Ni & Cheung, 2023), and AI-powered chatbots (e.g., Belda-Medina & Calvo-Ferrer, 2022; Chen et al., 2020; Liu & Ma, 2024). This study further contributes to the TAM literature by concentrating on ChatGPT-assisted oral language practices. The technological tool examined in this study, ChatGPT, could be utilized for different learning purposes among foreign language learners, such as providing feedback for essays, generating assessment tasks, performing language translation, and recommending specific learning materials (Lo, 2023). The supported effectiveness of TAM in this study implies future IS research to contextualize measurements within specific learning purposes in order to accurately evaluate learners' technology acceptance, particularly when the targeted technology offers a variety of technical affordances.

Results also highlight the antecedental role of learning motivation and willingness to communicate in ChatGPT acceptance, which offer valuable practical insights from a pedagogical perspective. The findings demonstrated that situational analysis regarding learners' psychological attributes is necessary before delivering formal instructions with the aid of educational technologies in FL classrooms. With a thorough understanding of CFL learners' motivation towards academic learning and how potential sociocultural factors exert impacts in the learning contexts, teachers could utilize effective motivational strategies and learning activities as incentives to promote learners' acceptance of educational technologies (i.e., designing topics that learners are familiar with, incorporating cultural elements, and providing clear language structure for the scaffolding purpose), further leading to active engagement in technology-assisted language learning and producing meaningful educational outcomes. Moreover, while willingness to communicate has been extensively explored in traditional language learning contexts, it has received insufficient attention in technology-assisted language learning contexts. In our study, willingness to communicate was found to be a significant mediator between learning motivation and ChatGPT acceptance, which suggests future research focus more on learners' willingness to communicate and its impacts on technology adoption and usage, especially when the learning contexts require oral-based interaction in the target language. From a pedagogical standpoint, this finding also highlights the critical role of foreign language teachers in encouraging East Asian learners' willingness to communicate with effective pedagogical strategies and sufficient talking opportunities before implementing technology-assisted language practices.

The formation of ChatGPT acceptance appeared heterogeneity among CFL learners. This result offers possible explanations for why certain theoretically supported relationships between TAM variables had been dissolved in prior relevant investigations. To enhance ChatGPT acceptance among CFL learners, suitable instructional strategies should be carefully chosen when designing ChatGPT-assisted language learning activities, while different features of ChatGPT should be purposefully promoted throughout the process with consideration of learners' past learning experiences. When facing long-term CFL learners, more emphasis should be placed on linking ChatGPT-assisted oral language practices to their previous knowledge constructions, demonstrating the great potential of ChatGPT in enhancing their speaking performance. Whereas for those CFL beginners, teachers may start by providing more guidance on learner-technology interaction techniques that could be applied in AI-assisted language learning, supporting learners in generating operable and favorable interacting experiences, and thus developing positive attitudes towards technologies in oral language practices.

There are certain limitations in this study. First, due to the survey nature of our work and time constraints, the interaction time allotted to each learner in the speaking activities prior to the survey was relatively limited. To reach a more comprehensive understanding of learners' adoption of and interaction with chatbots, interventional or observational studies are thus advised for future research to reveal the interaction patterns and strategies utilized by learners with various levels of learning motivation and WTC. Second, our findings were solely based on self-reported survey data. Future research is expected to incorporate data from classroom observations or interviews to provide additional triangulation reference. Additionally, it is valuable to identify other individual-level, task-level, teacher-level, and organization-level influencing factors that may impact learners' acceptance of GenAI-powered chatbots with qualitative data. Last, the sample size in this study was somewhat small and limited to Mongolian CFL learners. Survey studies with larger sample sizes or include other CFL learner populations are thus recommended to enhance the generalizability of our findings and to improve the statistical power of the analysis on the intricate relationships between learning motivation, WTC, and TAM variables among different learner populations.

Acknowledgment: This work was supported by the Key Research Institute of Humanities and Social Sciences at Universities of the Ministry of Education, China (Grant No. 22JJD740016). The corresponding author for this paper is Liang Yu.

References

Bai, B., Nie, Y., & Lee, A. N. (2020). Academic self-efficacy, task importance and interest: Relations with English language learning in an Asian context. *Journal of Multilingual and Multicultural Development*, *43*(5), 438–451.

Balouchi, S., & Samad, A. A. (2021). The effect of perceived competence on second language communication frequency: The mediating roles of motivation, willingness to communicate, and international posture. *Education and Information Technologies*, *26*(5), 5917–5937.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. Freeman.

Becker, J. M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly*, *37*(3), 665–694.

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, *12*(17), 8427.

Chen, H. L., Vicki Widarso, G., & Sutrisno, H. (2020). A chatbot for learning Chinese: Learning achievement and technology acceptance. *Journal of Educational Computing Research*, *58*(6), 1161–1189.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003.

Durndell, A., & Haag, Z. (2002). Computer self efficacy, computer anxiety, attitudes towards the internet and reported experience with the internet, by gender, in an East European sample. *Computers in Human Behavior*, *18*(5), 521–535.

Eccles-Parsons, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., & Midgley, C. (1983). Expectancies, values and academic behaviors. In Spence, J. T. (Ed.), *Achievement and Achievement Motives: Psychological and Sociological Approaches* (pp. 75–146). Freeman.

Elahi Shirvan, M., Khajavy, G. H., MacIntyre, P. D., & Taherian, T. (2019). A meta-analysis of L2 willingness to communicate and its three high-evidence correlates. *Journal of Psycholinguistic Research*, *48*(6), 1241–1267.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*. Edward Arnold Publishers.

Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology*, *48*, 67–84.

Hair, J. F., & Alamer, A. (2022). Partial least squares structural equation modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics*, *1*(3), 100027.

Hair, J. F., Sarstedt, M., Matthews, L. M., & Ringle, C. M. (2016). Identifying and treating unobserved heterogeneity with FIMIX-PLS: part I – method. *European Business Review*, *28*(1), 63–76.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3rd ed.). SAGE Publications.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135.

Hsu, L. (2017). EFL learners' acceptance of technology in a computer-assisted language learning (CALL) context: The role of intrinsic-extrinsic motivation in English learning. *International Journal of Information and Education Technology*, *7*(9), 679–685.

Hsu, M. H., Chen, P. S., & Yu, C. S. (2023). Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments*, *31*(7), 4297–4308.

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.

Jeon, J. (2022). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, *31*(7), 4297–4308.

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*, *28*(1), 227–261.
Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550.
Lee, J. S. (2020). The role of grit and classroom enjoyment in EFL learners' willingness to communicate. *Journal of Multilingual and Multicultural Development*, *43*(5), 452–468.
Lee, J. S., & Drajati, N. A. (2019). Affective variables and informal digital learning of English: Keys to willingness to communicate in a second language. *Australasian Journal of Educational Technology*, *35*(5), 168–182.
Lee, J. S., & Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. *System*, *82*, 63–73.
Lee, J. S., & Lee, K. (2020). Affective factors, virtual intercultural experiences, and L2 willingness to communicate in in-class, out-of-class, and digital settings. *Language Teaching Research*, *24*(6), 813–833.
Lee, J. S., & Lu, Y. (2023). L2 motivational self system and willingness to communicate in the classroom and extramural digital contexts. *Computer Assisted Language Learning*, *36*(1–2), 126–148.
Li, R., Meng, Z., Tian, M., Zhang, Z., Ni, C., & Xiao, W. (2019). Examining EFL learners' individual antecedents on the adoption of automated writing evaluation in China. *Computer Assisted Language Learning, 32*(7), 784–804.
Liu, G., & Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, *18*(2), 125–138.
Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), 410.
Loh, E. K. Y. (2019). What we know about expectancy-value theory, and how it helps to design a sustained motivating learning environment. *System*, *86*, 102119.
Luo, B. (2020). The influence of teaching learning techniques on students' long-term learning behavior. *Computer Assisted Language Learning*, *33*(4), 388–412.
MacIntyre, P. D., & Blackie, R. A. (2012). Action control, motivated strategies, and integrative motivation as predictors of language learning affect and the intention to continue learning French. *System*, *40*(4), 533–543.
MacIntyre, P. D., Dörnyei, Z., Clément, R., & 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.
Mystkowska-Wiertelak, A. (2021). The link between different facets of willingness to communicate, engagement and communicative behaviour in task performance. In Budzińska, K., & Majchrzak, O. (Eds.), *Positive Psychology in Second and Foreign Language Education* (pp. 95–113). Springer.
Nagle, C. (2021). Using expectancy value theory to understand motivation, persistence, and achievement in university-level foreign language learning. *Foreign Language Annals*, *54*(4), 1238–1256.
Ni, A., & Cheung, A. (2023). Understanding secondary students' continuance intention to adopt AI-powered intelligent tutoring system for English learning. *Education and*

Information Technologies, *28*(3), 3191–3216.

Pavelescu, L. M. (2023). Emotion, motivation and willingness to communicate in the language learning experience: A comparative case study of two adult ESOL learners. *Language Teaching Research*. Advance online publication. http://doi.org/10.1177/13621688221146884

Peng, Y., Yan, W., & Cheng, L. (2020). Hanyu Shuiping Kaoshi (HSK): A multi-level, multi-purpose proficiency test. *Language Testing*, *38*(2), 326–337.

Petrović, J., Jovanović, M. (2021). The role of chatbots in foreign language learning: The present situation and the future outlook. In Pap, E. (Eds.), *Artificial Intelligence: Theory and Applications* (pp. 313–330). Springer.

Rigdon, E. E., Ringle, C. M., & Sarstedt, M. (2010). Structural modeling of heterogeneous data with partial least squares. In N. K. Malhotra (Ed.), *Review of Marketing Research* (volume 7, pp. 255–296). Sharpe.

Rosenzweig, E. Q., Wigfield, A., Eccles, J. S., Renninger, K. A., & Hidi, S. E. (2019). Expectancy-value theory and its relevance for student motivation and learning. In Renninger, K. A., & Hidi, S. E. (Eds.), *The Cambridge Handbook of Motivation and Learning* (pp. 617–644). Cambridge University Press.

Ruan, S., Jiang, L., Xu, Q., Liu, Z., Davis, G. M., Brunskill, E., & Landay, J. A. (2021). EnglishBot: An AI-powered conversational system for second language learning. *Proceedings of 26th International Conference on Intelligent User Interfaces*, 434–444. Association for Computing Machinery.

Shaaban, K. A., & Ghaith, G. (2000). Student motivation to learn English as a foreign language. *Foreign Language Annals*, *33*(6), 632–644.

Shao, Q., & Gao, X. (2016). Reticence and willingness to communicate (WTC) of East Asian language learners. *System*, *63*, 115–120.

Soyoof, A. (2023). Iranian EFL students' perception of willingness to communicate in an extramural digital context. *Interactive Learning Environments*, *31*(9), 5922–5939.

Srite, M. (2006). Culture as an explanation of technology acceptance differences: An empirical investigation of Chinese and US users. *Australasian Journal of Information Systems*, *14*(1), 5–25.

Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*. Advance online publication. https://doi.org/10.1080/10494820.2023.2209881

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15.

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research*, *11*(4), 342–365.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273–315.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478.

Wang, Q., & Xue, M. (2022). The implications of expectancy-value theory of motivation in language education. *Frontiers in Psychology*, *13*, 992372.

Wen, X., & Piao, M. (2020). Motivational profiles and learning experience across Chinese language proficiency levels. *System*, *90*, 102216.
Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.
Xia, Q., Chiu, T. K., Chai, C. S., & Xie, K. (2023). The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot. *British Journal of Educational Technology*, *54*(4), 967–986.
Yang, H., & Lian, Z. (2023). Ideal L2 self, self-efficacy, and pragmatic production: The mediating role of willingness to communicate in learning English as a foreign language. *Behavioral Sciences*, *13*(7), 597.
Zadorozhnyy, A., & Lee, J. S. (2023). Informal digital learning of English and willingness to communicate in a second language: Self-efficacy beliefs as a mediator. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2023.2215279

Appendix

Self-efficacy (SE)
I can speak Chinese fairly fluently.
I can communicate with Chinese speakers in Chinese.
I can receive a good grade from my Chinese course.
I can master the knowledge in my Chinese course.

Utility value (UV)
Being good at Chinese will bring me many benefits in my future daily life.
The things I learn in the Chinese course will be applicable in my future life.
In general, learning Chinese is practical for my future plans.

Attainment value (AV)
It is important to me to be good at Chinese.
Being good at Chinese means a lot to me personally.
In general, learning Chinese well is important to me.

Intrinsic value (IV)
I like learning Chinese.
I am fascinated by Chinese.
In general, learning Chinese is interesting to me.

Willingness to communicate (WTC)
I am willing to communicate in Chinese when I have a chance to talk freely in Chinese classes.
I am willing to communicate in Chinese when I have a chance to talk in front of other students in Chinese classes.
I am willing to communicate in Chinese when I have a group discussion in Chinese

classes.
Perceived ease of use (PEOU)

It is easy to learn how to use ChatGPT to practice oral Chinese.
It is easy to become proficient in practicing oral Chinese using ChatGPT.
It is easy to orally interact with ChatGPT.
The interaction with ChatGPT is clear and understandable.

Perceived usefulness (PU)
Using ChatGPT could improve my oral Chinese learning performance.
Using ChatGPT could enhance my oral Chinese learning effectiveness.
Using ChatGPT could increase my Chinese language output in oral practices.
Using ChatGPT could help me complete oral Chinese practice tasks more quickly.
Attitude toward using (ATU)

I believe that using ChatGPT is a good idea.
I believe that using ChatGPT is advisable.
I agree with the practice of using ChatGPT for oral Chinese practices.

Behavioral intention to use (BIU)
I intend to use ChatGPT in oral Chinese practices in the future.
I intend to use ChatGPT regularly to practice my oral Chinese in the future.
I intend to use ChatGPT to practice on more topics in oral Chinese practices in the future.

