Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2023). Effects of Automatic Speech Recognition Technology on EFL Learners' Willingness to Communicate and Interactional Features. *Educational Technology & Society*, *26*(3), 37-52. https://doi.org/10.30191/ETS.202307_26(3).0004

Effects of Automatic Speech Recognition Technology on EFL Learners' Willingness to Communicate and Interactional Features

Michael Yi-Chao Jiang[1], Morris Siu-Yung Jong[2,3], Wilfred Wing-Fat Lau[2], Ching-Sing Chai[2] and Na Wu[4*]

[1]School of Foreign Languages, Shenzhen Technology University, China // [2]Department of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong SAR // [3]Centre for Learning Sciences and Technologies, The Chinese University of Hong Kong, Hong Kong SAR // [4]College of International Business, Shenyang Normal University, China // mjiang@sztu.edu.cn // mjong@cuhk.edu.hk // wwflau@cuhk.edu.hk // cschai@cuhk.edu.hk // wunacib@synu.edu.cn

[*]Corresponding author

ABSTRACT: This study examined the effects of using automatic speech recognition (ASR) technology on Chinese students' willingness to communicate (WTC) in oral English and the development trajectories of their interactional features in a flipped EFL context. One hundred sixty undergraduates from a Chinese university participated in the 14-week quasi-experiment. Both groups were taught in a flipped fashion. The treatment group was required to use the ASR technology for oral practice in their pre-class self-learning, while the control group conducted their self-learning without the ASR technology. The results found that the ASR-based oral practice led to a significant between-group difference in students' WTC with teacher and class and WTC with non-Chinese, showing that the ASR technology may contribute to improving the Chinese students' WTC in oral English. Conversely, except for the between-group effect on negotiation for meaning, there was no significant difference between the two groups on the other measures of interactional features. Moreover, none of the interactional features of the students in the treatment group changed significantly over time, indicating a limited role of the ASR technology on Chinese students' interactional features. Discussions were conducted regarding the contradictory effects of the ASR technology on WTC and peer interaction.

Keywords: Automatic speech recognition, English as a foreign language, Interactional features, Willingness to communicate

1. Introduction

Automatic speech recognition (ASR) is known as a computer-based process of decoding and transcribing oral language usually into text form (Kim, 2006). It is a specialized application of artificial intelligence in natural language processing and has been broadly incorporated into various scenarios in modern daily life (Evers & Chen, 2020). In recent years, free speech-to-text ASR-based technology (i.e., dictation ASR) has gained growing attention in the domain of foreign language (FL) education. Owing to the real-time feedback on FL learners' oral performance without time and space restrictions (Wang & Young, 2014), mobile-based dictation ASR technology can be integrated into out-of-class self-learning to enhance students' preparedness for in-class activities (Jiang et al., 2022a). As such, flipped classrooms that realize the switch of in-class lectures and out-of-class assignments seem to be "a good match" for dictation ASR technology in FL learning (e.g., Jiang et al., 2021a). In view of Bloom's taxonomy (Anderson & Krathwohl, 2001), the in-class activities in a flipped classroom are mostly for developing higher-order thinking skills in terms of applying, analyzing, evaluating, and creating (Jong, 2019). Contrastingly, the out-of-class learning is remembering- and understanding-oriented (Jong, 2017; Jong et al., 2019). Bloom's taxonomy contextualized into flipped FL classrooms, the in-class activities are generally composed of output-oriented tasks that can elicit language use in learners collaboratively, while the out-of-class self-learning abounds with input-oriented content knowledge learning such as vocabulary and reading comprehension.

Typically, under the influence of Chinese culture which advocates learners' conformity (Lee & Yin, 2011), Chinese students are often portrayed as obedient learners who are reluctant to express opinions. Particularly, in an English as a foreign language (EFL) classroom, probably due to their limited English, Chinese students do not seem to have strong willingness to communicate (WTC) in participating in in-class activities (Lou & Noels, 2021). Their peer interactions are typically limited to short phrases and even single words (Clark & Gieve, 2006). Tang et al. (2020) argued that the Chinese learners' reticence was a learnt behavior driven by intentions consisting of behavioral, normative and control beliefs towards classroom participation. While not all the students possess homogeneous learning styles (Tran, 2013), some studies have claimed that Chinese students are not accustomed to participating actively in the classroom (e.g., Tang et al., 2020). They argued the traditional

Chinese teaching and learning atmosphere did not create situations where raising questions and challenging peers was regarded as necessary (Durkin, 2011).

Owing to its flexibility, dictation ASR technology can facilitate free and unguided speech practice and allows users to dictate anything in the target language that interests them (Evers & Chen, 2020). In other words, it affords learner autonomy which is generally conducive to learning (Ryan & Deci, 2020). In addition, dictation ASR allows flexibility in the models of speech and students may not get discouraged if they speak with a different English accent or variety (Hincks, 2015). Furthermore, dictation ASR can be better integrated into the design of pedagogical activities, as "teachers can assign practice with sound or topics that they have been working on in class instead of sending students off to follow a program that may not correspond clearly with the class." (McCrocklin, 2019, p. 103) Those benefits of dictation ASR are thought to better facilitate students' pre-class self-study in an EFL class, thus resulting in a higher level of learner preparedness (Jiang et al., 2022a). Consequently, it is hypothesized that the ASR-based oral practice is likely to boost students' willingness to interact with their peers in class. On the other hand, because of the ample opportunities of speaking practice assisted by dictation ASR, the practice effect may have a significant impact in that students' speaking anxiety may be reduced and self-efficacy in speaking may be increased. They may, therefore, be more willing to conduct in-class peer interaction in the target language.

Although researchers have revealed that ASR-based technology may enhance users' EFL learning in terms of affective and behavioral variables, (e.g., Ahn & Lee, 2016; Jiang et al., 2022b), few studies examined the effects of ASR-based technology on learners' EFL learning from the perspective of peer interaction, and little is known about whether the ASR-based technology may contribute to the development of students' WTC and peer interaction. In a flipped FL classroom, the in-class activities are designed to be highly interactive featuring higher order language skills (e.g., analyzing- and evaluating-oriented tasks). However, Chinese learners' reticent non-participating behaviors in class might lead to insufficient in-class interactions (Jiang et al., 2022a). Hence, there is a need to raise the effectiveness of the ASR-based technology for enhancing peer interaction in the Chinese EFL context. Therefore, the current study was conducted to examine the effects of the ASR-based technology on Chinese students' WTC and peer interaction in a flipped EFL classroom.

2. Related works

2.1. ASR-based technology in flipped language classrooms

Intelligent computer assisted language learning (iCALL) alongside online resources can promote learner autonomy "by enabling experimentation through self-access work outside of class" (McCrocklin, 2016, p. 27). Compared with practicing with native speakers or language teachers, iCALL practice is not subject to time and space constraints (Jiang et al., 2021a), thus providing ample opportunities for developing learner autonomy. Pedagogically, such advantages of ASR coincide with the tenets of flipped learning for FL classrooms. The pre-class self-study in flipped learning requires adequate learner autonomy, indicating that the student should play a role of autonomous learners taking full charge of their learning. In other words, the students need to take responsibility for all the decisions in relation to their pre-class self-learning. The ASR technology transcribes students' speech into text, providing immediate and more importantly, visualized feedback on their speech (Levis, 2007), and students can see immediately what they say in the target language. In flipped language classrooms, therefore, the ASR-based technology is assumed to help students monitor their flow of speech and support to self-assess their output.

Learner preparedness is identified as a crucial factor for the success of the flipped classroom approach (Sun & Xie, 2020), and with the ASR-based technology, students in a flipped FL classroom are expected to be prepared with a better mastery of the contents for in-class higher order activities. Consequently, those with higher level of learner preparedness tend to be more willing to communicate with their peers. According to Tai and Chen (2020), interaction with the ASR-based application enhanced EFL learners' confidence in communication and reduced their speaking anxiety, resulting in a higher level of WTC. Furthermore, they also claimed such positive effects of ASR-based technology on WTC were attributable to increased interaction and engagement, and the less threatening environment the ASR-based application established. Overall, in flipped FL classrooms, the ASR-based technology provides an avenue for developing learner autonomy and enhances students' pre-class preparedness, which in turn may boost their WTC and ready them to engage in more peer interaction while performing higher-order tasks in class.

The benefits of ASR-based technology consist in the significant amount of practice, consistency, the unbiased nature of feedback and diverse forms of visual representations (Levis, 2007). Especially, the real-time assessment of learners' utterances is beneficial for acquiring listening and speaking skills (Wang & Young, 2014). Recently, a growing number of empirical studies have obtained positive evidence of ASR-enhanced FL learning (e.g., Ahn & Lee, 2016; Jiang et al., 2021a). It was concluded that ASR-based technology could provide language learners with (1) more opportunities to produce and extensive interaction in the target language, (2) immediate feedback in various direct and indirect forms, and (3) more control over their learning with increased confidence. On the learners' part, ASR-based technology can create a less threatening self-paced environment for them to learn to speak the target language (McCrocklin, 2016).

However, a closer look at those studies reveals that most of them relied on participants' affective and behavioral data to evaluate the effectiveness of ASR-based technology, and few studies investigated the potential changes from the interactional features of peer interaction. For example, Ahn and Lee (2016) investigated junior students' attitudes towards the use of an ASR-based application using close- and open-ended questionnaires. Evers and Chen (2020) surveyed adult EFL learners' learning styles through questionnaires and used human rating method to evaluate the effects of an ASR system on students' pronunciation performance. Although they are widely used in FL educational research, affective and behavioral data are mostly self-reported and subjective, which tend to be biased to some extent (Wilson & Zietz, 2004). In that case, the effectiveness of ASR-based technology on students' language performance may not be sufficiently evaluated. Moreover, the social dimension is also crucial in understanding the impact of ASR-based technology on students' FL learning. Indicators such as interactional features of the participants' peer interaction can provide a more comprehensive perspective of explaining students' language learning and may contribute to diversifying FL pedagogy.

2.2. WTC in EFL classrooms

In the domain of language learning, some researchers regard WTC to be a personality trait, while by many others WTC is taken as a context-dependent domain-specific variable (Tavakoli & Davoudi, 2017). Despite the complex nature of WTC which manifests itself in its diverse conceptualizations (Pawlak & Mystkowska-Wiertelak, 2015), it is one of the crucial determinants of language classroom communication (Tavakoli & Davoudi, 2017). Generally, WTC is conceptualized as a readiness to speak in an FL or second language (L2) in a particular situation with a specific person (MacIntyre, 2007) and thus perceived as part of the broader concept of FL learning motivation. As the "final psychological step to the initiation of L2 communication" (MacIntyre & Doucette, 2010, p. 161), WTC can indicate how well the students are engaged in the collaborative tasks in EFL classrooms on the social dimension.

A plethora of primary studies have obtained empirical evidence that WTC is highly associated with FL learners' communication in the target language (e.g., Zhang et al., 2020). Factors such as individual's language and communicative competence, affective and environmental factors were reported to have significant impact on language learners' WTC in various pedagogical contexts. For example, Sato (2020) revealed that individuals' target language proficiency might moderate the impact of various factors on their WTC. Low-intermediate speakers' WTC was more likely to be affected by their interest in the topic and self-confidence, whereas advanced speakers' WTC was influenced by the opportunity to talk about themselves and their opinions. Wang et al. (2019) found that learners' perception of group interaction and interaction with the language teacher had a significant effect on students' WTC and their classroom communication in the target language. Zare et al. (2020) investigated the interplay of oral corrective feedback (explicit correction, recasts, and prompts) and L2 WTC across different L2 proficiency levels. They found that learners preferred prompts most and elicitation-oriented feedback were the most contributory to L2 WTC.

While numerous empirical studies have been conducted to investigate the interplay of WTC and other relevant factors in traditional EFL classrooms, little is known in literature about how ASR-based technology could influence learners' WTC in EFL classrooms. Given the Chinese students' typical reticence in class, it is worth investigating whether the ASR-based technology could enhance their WTC for in-class peer interaction.

2.3. Peer interaction in flipped EFL learning

In a student-centered language classroom, peer interaction or learner-learner interaction is the most significant interaction that leads to language learning and development. Therefore, how students interact to co-construct meaning and knowledge of language form is a crucial issue in FL learning research (Loewen & Sato, 2018). Ever since the formulation of the interactionist hypothesis (Long, 1981), which claims that interaction, particularly

when it involves negotiation for meaning and feedback, facilitates L2 acquisition, interactionist approach has long been considered a theoretical underpinning for task-based language instruction. Numerous empirical studies based on the interactionist approach have provided "the main psycholinguistic underpinnings for task-based language teaching" (Long, 2015, p. 61). To date, studies on task-based peer interaction in language learning have realized that interactional features provide learning opportunities that are theoretically posited to be beneficial to EFL learning. As a review study found (Mackey & Goo, 2007), the occurrence of a variety of interactional features in task-based peer interaction can facilitate second language acquisition. In the past decade, a growing number of empirical EFL studies have investigated a range of interactional features (e.g., clarification requests, recasts and uptakes) and their effects on students' language learning.

Following Philp et al. (2014), peer interaction provides language learners "a context for experimenting with the language" (p. 17) and tends to be less threatening than teacher-led interaction, because students do not have to worry about making errors while talking to their peers. As with the other types of interaction in a language classroom, peer interaction is comprised of input, negotiation, output and noticing (Loewen & Sato, 2018). Correspondingly, a range of interactional features are established by interactionist researchers to operationalize those four constructs. Negotiation for meaning (NfM) is at the core and acts as a response to a communication breakdown, including clarification requests, confirmation checks and comprehension checks (Pica, 1994). When they cannot get themselves across, language learners may signal that a communication breakdown has occurred. However, NfM does not always occur frequently in classroom contexts. L2 or FL learners also have negotiation of form as a result of a desire for linguistic accuracy rather than as a result of communication breakdown (Gass et al., 2011). They may talk about the language they are producing or question their language use, which was defined by Swain and Lapkin (1998) as language-related episodes (LREs). LREs are identified in a myriad of primary studies as learning opportunities during interaction. One particular type of input-providing LRE is recast, which is a "correct restatement of a learner's incorrectly formed utterance" (Nicholas et al., 2001, p. 720). A recast is didactic in nature and provides the correct linguistic form for the learner. It is often provided by the interlocutor who is more proficient in the target language as a type of corrective feedback. Extant studies have demonstrated that learners can both notice and benefit from recasts despite individual and contextual variability (e.g., Rassaei, 2022), indicating recast plays an indispensable role in the interactionist approach.

In EFL classrooms, collaborative tasks are typically designed to elicit peer interaction in the target language, wherein learners may notice the necessary language forms and focus on meaning. Compared with the activities in a lecture-based EFL classroom, the higher-order tasks in flipped EFL classrooms are usually more complex and students need to allot more attentional resources to perform those tasks. For example, those higher-order tasks may involve more elements and verb tenses or more perspective taking. Therefore, according to the Cognition Hypothesis (Robinson & Gilabert, 2007), those tasks would promote more interactional features, which in turn lead to language development. Consequently, more peer interaction could be witnessed when students are performing more complex and higher-order tasks.

However, in the eyes of most Chinese learners, only their teachers comment on their academic performance and provide them with corrective feedback. Such conventionally deep-rooted perceptions are closely associated with the key merit of rituals/etiquette (*Li*) in Chinese culture, which refers to the ethic of propriety and the prescription of social relationship structures (Yum, 2007). If they replaced their teacher's role and provided feedback directly to their peers, the students would be considered as a "show-off" by their peers. Besides, acting as a teacher may result in an unnecessary deviation in power relations with their peers. Therefore, Chinese students psychologically resist playing a teacher's role in front of their peers (Xu & Kou, 2018) and may only participate in a discussion when they have something "safe" to say (Wu, 2015). On the other hand, the long-held test-orientated learning may "push" the Chinese EFL learners to make every effort only to score (Gao, 2008), losing interest in mastering English as a means of communication. Most of them were so accustomed to the exam-oriented and didactic English teaching that they might find themselves speechless when they were assigned free-talking tasks in EFL classrooms.

This study aimed to examine the effects of the ASR-based technology on Chinese students' WTC and peer interaction in a flipped EFL classroom. Specifically, it was conducted to (1) investigate whether the integration of the ASR-based practice into pre-class self-learning could contribute to the development of Chinese EFL students' WTC and their interactional features, and (2) reveal the development trajectories of Chinese students' interactional features over time in a flipped EFL classroom. Accordingly, three research questions were formulated:

RQ 1. Does the ASR-based technology lead to any difference in Chinese EFL learners' WTC in a flipped classroom?

RQ 2. Does the ASR-based technology lead to any difference in Chinese EFL learners' interactional features in a flipped classroom?

RQ 3. How does the ASR-based technology influence the development trajectories of Chinese EFL learners' interactional features in a flipped classroom?

3. Methods

3.1. Participants and research background

Four classes of 160 freshmen from a public university in China consented to participate in this study. They majored in a range of disciplines, including mathematics, physics, education, psychology, Chinese language and arts, chemistry, and computer science. According to a pre-course survey, the participants had learnt English for approximately 10.9 years on average; 69.4% reported little or no training in oral English, and 92.5% had little or no experience of flipped learning. They all registered for the same EFL course, but according to the pre-course placement test, most were disfluent in oral English. The four classes were randomly assigned into a control group (CG) of two classes and a treatment group (TG) of the other two classes.

3.2. Instructional design

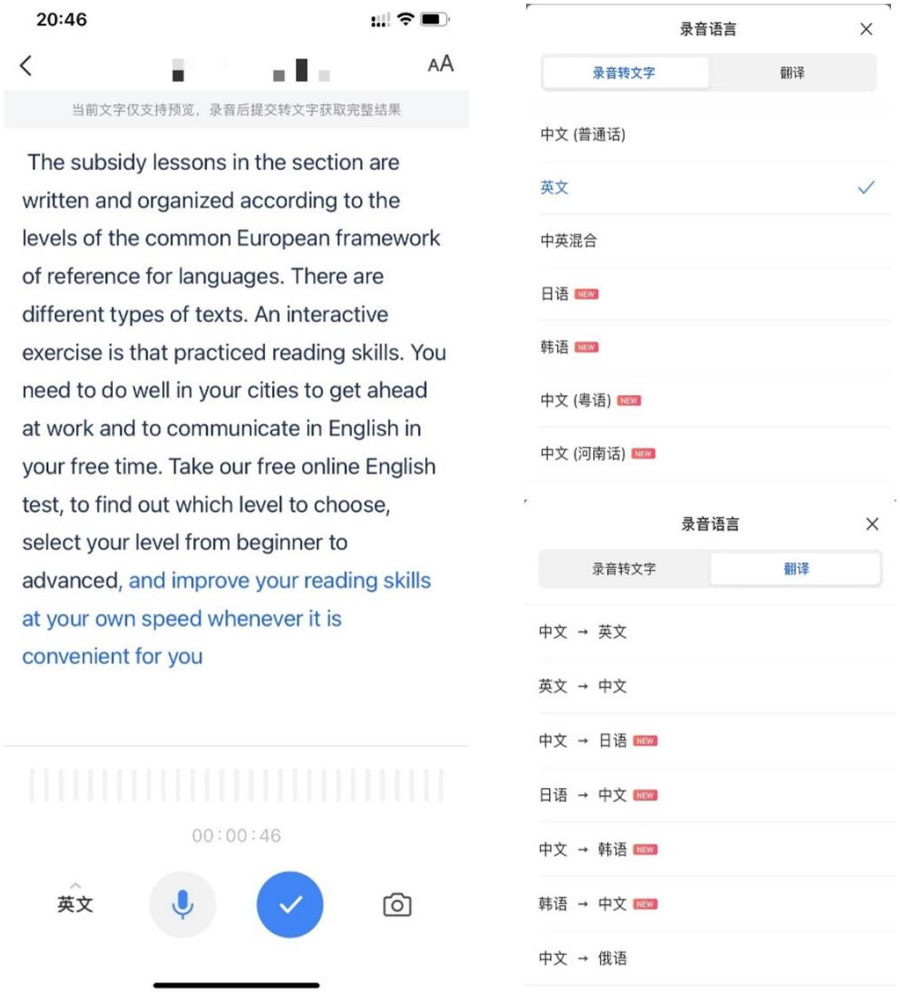
The course lasted for 14 weeks in the 2019 fall semester. It was a compulsory EFL course for Year 1 and Year 2 students which aimed to develop learners' English proficiency, especially to foster the students' oral language skills. Before the course started, a placement test was administered to measure students' overall English proficiency. The test was made up of a computer-based section (composed of reading comprehension, listening comprehension and essay writing) and an in-person individual interview. The computer-based section lasted for 100 minutes, and the oral interview was approximately five minutes per candidate.

Week 1 was the course orientation, and Weeks 2–13 covered the instruction of eight learning units. Each week, the students in both groups had one 90-minute face-to-face session. In Week 14, both groups took an end-of-semester test as a summative assessment of the course. Within each class, the students were randomly assigned into subgroups of three or four for in-class activities. Both groups were taught by the same teacher with a flipped classroom approach. An e-learning platform called Unipus (http://learn.unipus.cn/, Unipus https://u.unipus.cn/index.html/, is an online learning platform developed and owned by Foreign Language Teaching and Research Press in China, which provides hybrid teaching solutions for foreign language teaching in colleges and universities. This platform provides a one-for-all package of learning, practicing, testing, and evaluating that feature interactive experience.) was utilized to implement the flipped instruction. One week before each face-to-face session, the flipped materials were delivered to the students via Unipus for pre-class self-learning. In terms of Bloom's taxonomy (Anderson & Krathwohl, 2001), the flipped contents were remembering- and understanding-oriented and served as lead-ins for in-class higher order activities. Those materials were developed and organized in the form of a massive open online course by the textbook publisher and could be accessed conveniently by the students. The students were required to master the flipped content using Unipus and raise questions through WeChat (a social media/instant messaging platform) when needed. The textbook featured guided dialogues to enable participatory communication and interaction. The course teacher determined the learning materials from the textbook for both pre-class and in-class tasks. Generally, all the reading-focused sections alongside the vocabulary- and grammar-associated sections were "flipped" out of class for students' self-learning because they could serve as input to prepare the students for the in-class higher-order activities that concern the unit themes.

Particularly for the pre-class self-study, both groups were required to conduct a free-response oral practice to check the students' mastery level of the flipped content. A unit theme-focused free-response question was given to the students, and they needed to base their responses on their pre-class self-study. The students in both groups were encouraged to practice performing that oral task repeatedly. Then they needed to record and upload their responses to Unipus and their speech had to last for at least two minutes. The CG students performed this task with no additional requirements, while the TG students were required to utilize an ASR-based application called iFlyRec (https://www.iflyrec.com) to perform this task (Figure 1). This application is developed by iFlyTek, a renown artificial intelligence company, and it is free to download and can run on iOS and Android.

Contrary to computer assisted pronunciation training (CAPT) systems that are specifically designed for FL learners, iFlyRec is a mobile-based dictation ASR application originally developed for native speakers. Since they are not designed for pedagogical purposes, mobile-based dictation ASR applications do not provide analysis of speech or artificial intelligence-based interaction (e.g., Google Assistant) with users. However, because of its flexibility, dictation ASR applications would be more effective in FL teaching combined with scaffolded activities (Evers & Chen, 2020). More specifically, dictation ASR applications allow users to try practicing any word, phrase, or topic that interests them (Hincks, 2015), while CAPT programs are not able to facilitate free, unguided speech practice (McCrocklin, 2019). Besides, the flexibility of dictation ASR applications is also reflected in the models of speech (e.g., Received Pronunciation or American Standard), meaning ASR-based oral practice allows FL learners to speak with different accents and varieties. Apart from the above-mentioned advantages that all dictation ASR-based applications have, iFlyRec has some exceptional features that distinguish it from other dictation ASR-based applications. It realizes real-time conversion from speech to text across multiple languages and even some Chinese dialects. Additionally, it can also be used for interlingual translation based on ASR technology in several languages such as Chinese, English, and Russian.

Figure 1. Screen capture of iFlyRec

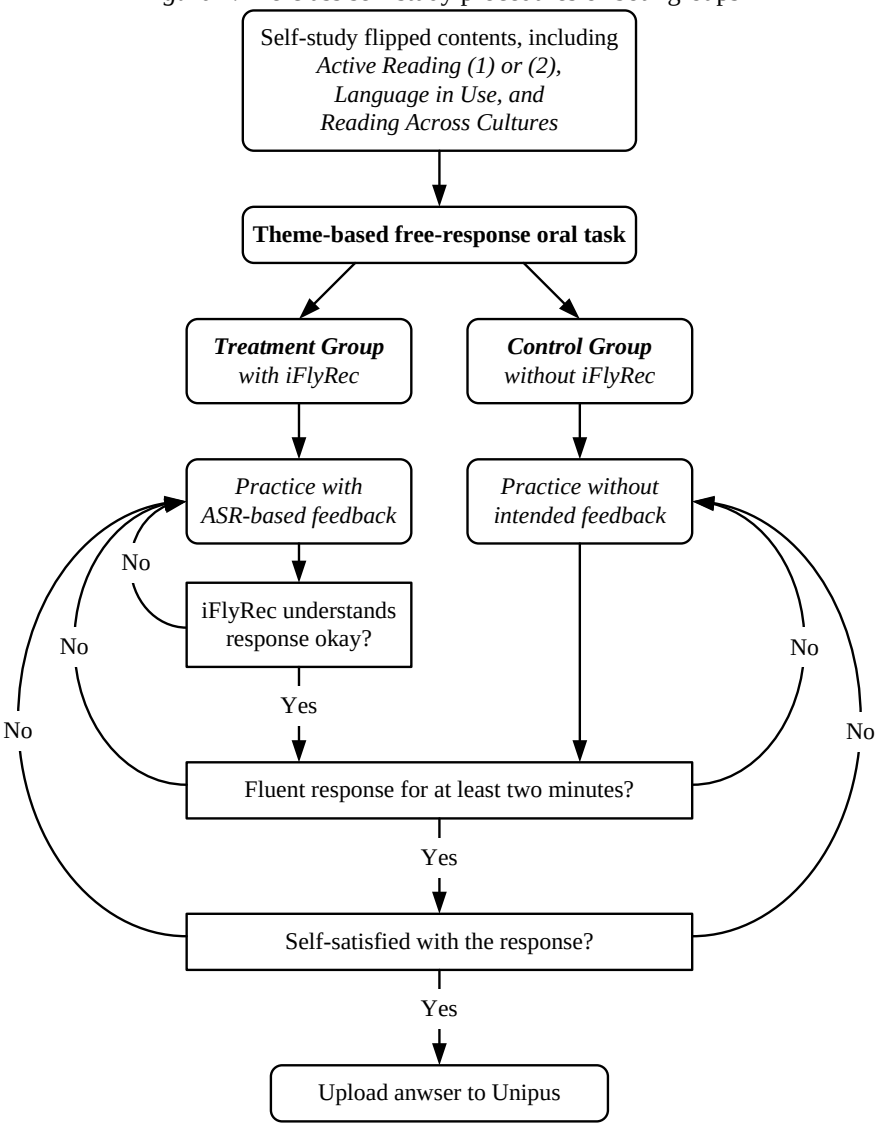


When performing the ASR-assisted oral tasks, the TG students could refer to the transcribed texts as immediate feedback to easily identify morphosyntactic or pronunciation-related errors in their speech. Thus, the TG students could improve their task performance by correcting themselves until they could make themselves understood by the application fluently. Conversely, the CG students did not have any particular visualized feedback to support their task-performing. Since they were given one week's time to perform the oral task, both groups were recommended to keep on practicing until they were satisfied with their oral answers. (see Figure 2 for the procedures of pre-class self-study of both groups).

For each unit of in-class learning, the students in both groups undertook one shared Unit Task (UT), which was designed to elicit their language use in the classroom setting (see Figure 3 for a sample UT). They were allowed five minutes to communicate with their subgroup members and perform the task collaboratively. The students

were required to record their performance using their mobile devices and upload their recordings to Unipus. Among the eight units, the recordings of Units 2, 4, 6 and 8 were chosen for data analysis, but the course teacher and the students were informed that recordings of all the eight units would be analyzed for the sake of their performance consistency.

Figure 2. Pre-class self-study procedures of both groups



3.3. Instrument and measures

3.3.1. Measuring WTC

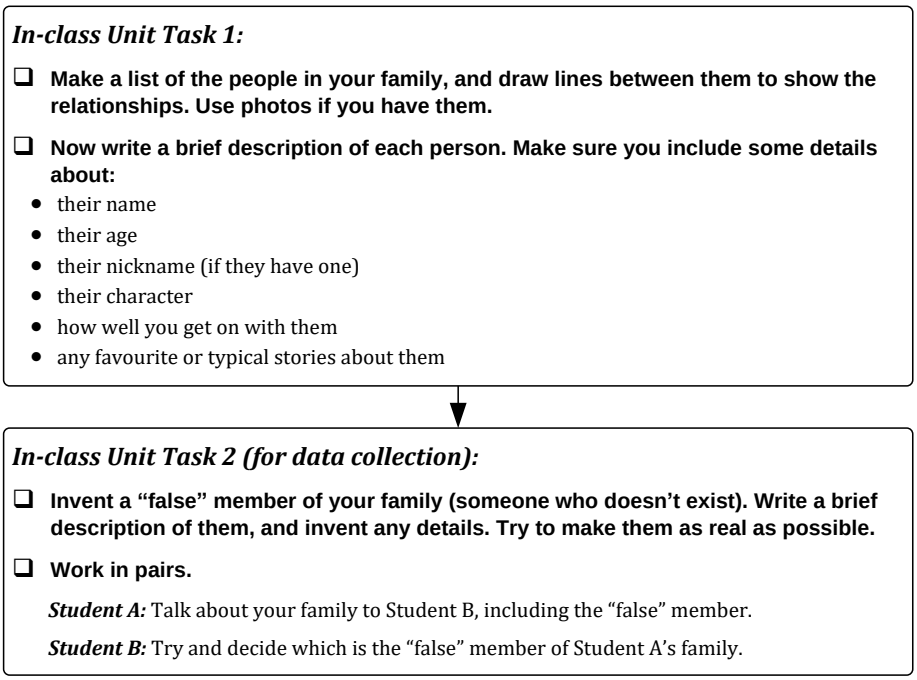
To survey participants' WTC in oral English, a WTC questionnaire developed by Tavakoli and Davoudi (2017) was adopted in this study because it was specifically created for the Asian EFL learning context (Appendix A). To ensure respondents' full understanding of the items, the original questionnaire was translated from English into Chinese. Then, backward translation was conducted to guarantee the Chinese version was equivalent in meaning to the original questionnaire. A panel of three professors in the domain of FL education were consulted, and a pilot study (n_p = 16) was conducted among another homogeneous group of the participants. Modifications were made following the feedback from the panel and the pilot students before the questionnaire was finalized.

The questionnaire consists of 27 items, assessing WTC in oral English in three different contexts: WTC with teacher and class (13 items), WTC with friends (7 items) and WTC with non-Chinese (7 items). Reliability analysis of the data showed that the Cronbach α of the pre-intervention survey was 0.845, 0.789 and 0.744 on the

three sub-questionnaires, and that of the post-intervention survey was 0.874, 0.795 and 0.816, respectively, indicating good internal reliability of both measurements.

Figure 3. A sample UT for data collection

Unit 4 *Family Affairs*



3.3.2. Measuring interactional features

A between- and within-subjects design was adopted to analyze the effects of the ASR-based technology on interactional features and the development trajectories. The independent variables were group (i.e., between-subjects factor) and time (i.e., within-subjects factor). The dependent variables were the participants' interactional features coded from their in-class task-performance at four timepoints (Units 2, 4, 6, and 8). Repeated measures ANCOVA was conducted to analyze the data. Following Gass et al. (2011), the current study operationalized interactional features as a set of measures, a practice that has been widely applied in interactionist studies. The measures include (a) NfM, which is further operationalized as clarification requests, confirmation checks and comprehension checks; (b) LREs, and (c) recasts (Appendix B).

The current study quantified each measure of interactional features in terms of their relative frequency, which was computed by dividing the frequency of a given measure over a production unit (Norris & Ortega, 2009). In this study, the analysis of speech unit (AS-unit) was chosen as the production unit of the participants' utterances. An AS-unit is "a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s)" (Foster et al., 2000, p. 365). It was revised based on the extant production units available to particularly deal with the fragmentary nature of oral data and provide a solution to the fuzziness and complexity of the spoken language (Jiang et al., 2021a).

3.4. Data collection and analysis

The WTC questionnaire was administered right before and after the intervention. A total of 155 responses (out of 160) were collected because five participants were absent for either the pre- or the post-intervention survey. To minimize possible data contamination caused by "careless" respondents, following Jiang's et al. (2021b) method, we added three "filtering items" to the questionnaire which served as an indicator of the participants' consistency of their responses. Consequently, 24 participants were removed from the sample, leaving 131 participants for the analysis of the WTC data (68 CG students and 63 TG students). One-way ANCOVA was conducted to examine the effect of using ASR-based technology on participants' WTC in oral English. The participants' pre-intervention English proficiency and WTC in oral English were controlled for in the data analysis as covariates.

Because of classroom noise, dropout and absence issues, a sum of 32 participants were removed from the recordings sample, leaving 128 participants for the analysis of interactional features (60 CG students and 68 TG students). The recordings were transcribed verbatim and then coded with ELAN (https://archive.mpi.nl/tla/elan), a professional software for annotating audio and video data in linguistic studies. Two authors were involved in the coding process, and the inter-coder reliability estimated by Krippendorff's α (Hayes & Krippendorff, 2007) was calculated to be 0.818 (> 0.8), indicating that the co-coding was deemed consistent between the two coders. Any disagreement was resolved through discussion until an agreement was reached. Annotated examples of NfM, LRE and recast are provided in Appendix B.

4. Results

4.1. Change in WTC

Descriptive statistics showed that compared with their CG counterparts, the TG students had higher post-intervention WTC scores in all the three contexts (Table 1). One-way ANCOVA was then conducted with the participants' pre-intervention overall English proficiency and WTC scores as covariates. With regard to RQ 1, the results revealed that the TG students had statistically higher scores on the dimensions of WTC with teacher and class ($F_{(1, 125)} = 12.743$, $p < .001$) and WTC with non-Chinese ($F_{(1, 125)} = 34.709$, $p < .001$). However, there was no significant between-subjects difference on WTC with friends ($F_{(1, 125)} = 0.929$, $p = .337$).

Table 1. Descriptive statistics of WTC mean scores

WTC	Pre-intervention		Post-intervention	
	TG	CG	TG	CG
With teacher and class	36.49	37.90	43.87	40.10
With friends	16.33	16.21	19.32	18.53
With non-Chinese	18.16	17.65	21.78	17.79

4.2. Change in interactional features

As for RQ 2, descriptive statistics showed that the TG students outperformed their CG counterparts on all the three measures of interactional features (Table 2). Moreover, the mean values of NfM for both groups demonstrated a perceptible downward trend over time, and the means of LRE for both groups showed an oscillating pattern. As for "recast," the TG students demonstrated no specific pattern, while their CG counterparts showed an evident decreasing trend.

Table 2. Between-subjects effects on NfM, LRE and recast

Measures	Unit	TG	CG	*p*	η^2_p
NfM	U2	0.011	0.006	.018	.044
	U4	0.014	0.002		
	U6	0.005	0.002		
	U8	0.003	0.001		
LRE	U2	0.110	0.081	.115	.020
	U4	0.178	0.136		
	U6	0.153	0.112		
	U8	0.164	0.163		
recast	U2	0.005	0.005	.176	.015
	U4	0.003	0.002		
	U6	0.004	0.001		
	U8	0.007	0.000		

Note. $\alpha < .05$.

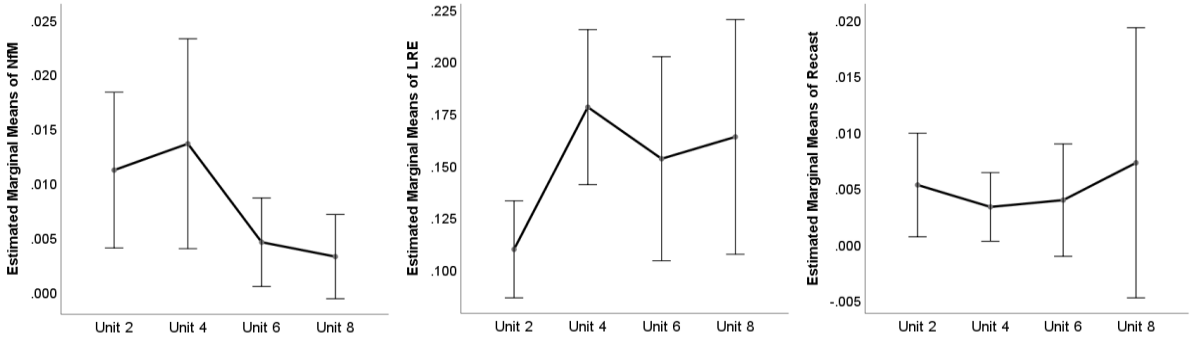
Levene's tests showed that the equality of error variances was not assumed on all the measures across the four units. However, a violation of this assumption is not considered an issue with roughly equivalent sample sizes on condition that the ratio of the largest sample size to smallest sample size is less than the threshold of 1.5 (Pituch & Stevens, 2016). The main effect of the between-subjects factor (i.e., group) on the mean values of NfM over time was statistically significant ($F_{(1, 125)} = 5.754$, $p = .018 < .05$) with a small to medium effect size ($\eta^2_p = 0.044 > 0.01$). In other words, the TG students generated more NfM per AS-unit than their CG counterparts. However, the TG students did not outperform their counterparts in the CG on LRE ($p = .115 > .05$) or recast ($p = .176 >$

.05) significantly, even though descriptive statistics showed that most of the TG mean scores appeared to be greater than the ones of the CG students.

4.3. Development trajectories of interactional features

In response to RQ 3, this study investigated the development trajectories of the TG students' interactional features. The results of Mauchly's test of sphericity showed that sphericity was not assumed ($p < .001$) for the three measures. The Greenhouse-Geisser epsilon for LRE was 0.783, greater than 0.75. Conversely, the Greenhouse-Geisser epsilon for NfM and recast was 0.724 and 0.552, respectively, both less than 0.75. Therefore, the Huynh-Feldt adjustment with the univariate tests was used for LRE, while the Greenhouse-Geisser adjustment was used for NfM and recasts. The tests of within-subjects effects revealed that the main effect of the within-subjects factor (i.e., time) was not statistically significant for NfM ($F_{(2.171,\ 143.302)} = 1.075$, $p = .361 > .05$), LRE ($F_{(2.478,\ 163.564)} = 0.110$, $p = .930 > .05$) or recast ($F_{(1.655,\ 109.213)} = 0.299$, $p = .700 > .05$), sphericity not assumed. In conclusion, the time factor did not lead to any statistically significant effect on any of the interactional features in the TG. Figure 4 illustrates the development trajectories of the three measures of the TG students' interactional features. Graphically, no particular pattern could be identified with regard to the effect of using ASR-based technology on the development of students' interactional features.

Figure 4. Development of trajectories of TG students' NfM, LRE and recast



Covariates appearing in the model are evaluated at the following values: placement = 82.25
Error bars: 95% CI

5. Discussion

5.1. Effect of ASR-based technology on Chinese EFL learners' WTC

The results of this study showed that owing to the use of the ASR-based technology in pre-class self-learning, the TG students' WTC with teacher and class and with non-Chinese was significantly improved. We argue that the ASR-enhanced practice changes the general notion pre-class self-study in a flipped classroom which is mainly passive absorption of knowledge (e.g., video-based self-learning). It allows the flipped EFL pedagogy to include an active component of interacting with ASR technology. Pedagogically, our approach enriches the current design for flipped classroom, especially for language-based flipped classroom approach.

Extant studies revealed that the use of the ASR-based technology could credibly simulate understanding by a native speaker, allowing learners to bridge the intelligibility gap and develop a sense of what successful interpersonal communication entails (Mroz, 2018). Moreover, as the ASR technology can immediately demonstrate the consequence of users' speech input, learners reported positive attitudes towards the use of the ASR-based technology for learning to speak (Ahn & Lee, 2016). Therefore, the ASR-based oral training is likely to establish a more friendly environment for learners to practice at their own pace (Wang & Young, 2014). As such, integrating ASR technology into FL oral training may contribute to lowering FL learners' anxiety and increasing their perceived competence, which helps foster learners' WTC in the target language (MacIntyre, 2007; MacIntyre, 2020).

Conversely, no significant difference in the participants' WTC with friends was observed. Attending an EFL class or talking with non-Chinese are specific occasions where the Chinese students must use English for communication. In contrast, talking with their friends in everyday life is a different occasion where the students do not necessarily have to speak in English. Since Asian societies are mostly monolingual and homogeneous, a

great majority of Asian students do not speak English in daily life. In other words, on most occasions, they will only initiate daily conversations in their first language. Therefore, we speculate that the effects of the ASR-based technology on learners' WTC in oral English may only confine itself to occasions where the students must speak English and could not override the influence of the Chinese culture on interpersonal communication. But on a general basis, it is concluded that the use of ASR-based technology has positive impact on Chinese learners' WTC in oral English.

Recent research in other cultural contexts also obtained positive effects of technology-enhanced EFL learning on students' WTC (e.g., Lee & Sylvén, 2021), indicating that EFL teachers in the Chinese context should encourage students to use ASR technology for oral practice as frequently as possible in out-of-class learning. Accordingly, EFL teachers in monolingual societies such as the Chinese context should make better use of the ASR-based applications to support students' self-study of English. For example, EFL course developers and teachers should design more ASR-based oral practice for students' self-learning, which in turn may enhance their WTC and likely improve English oracy. On the other hand, policymakers and researchers in the Chinese EFL context should acknowledge that the integration of the ASR-based technology may not improve students' WTC in English with their friends. The reasons for that may be deeply rooted in the cultural impact and educational context, which may need to be further explored in future studies.

5.2. Effect of ASR-based technology on Chinese EFL learners' interactional features

Although the participants' WTC was enhanced through the adoption of the ASR-based technology, no significant improvement was observed in their in-class peer interaction, and the development trajectories of the TG students' interactional features showed no specific pattern. Such findings may indicate that the use of the ASR-based technology may only exert a limited effect on the participants' in-class peer interaction.

Peer interaction is a group-based and social behavior that involves at least two interlocutors in a communication. The dynamics among those interlocutors may be influenced by a range of factors from language competence, individual learning style to interpersonal relationship and local culture (MacIntyre, 2020). Inconsistent with the findings in previous studies where WTC was labelled as the decisive step to initiate a communication in the target language (MacIntyre & Doucette, 2010), the current study revealed that the improvement in WTC might not necessarily lead to more interactional features in peer interaction in the Chinese EFL classrooms. In other words, WTC may not function as a direct indicator of how well the Chinese students are engaged in in-class peer interaction in an EFL classroom. Future studies need to investigate the relationship between Chinese EFL learners' WTC and their actual behaviors of peer interaction in class. Particularly, since the Chinese educational practice and local culture seem to have a more direct effect on the students' peer interaction than WTC, follow-up research is expected to determine what contextual factors may mediate the effect of WTC on peer interaction in the Chinese EFL context.

Given the students' inexperience in learning spoken English, the integration of the ASR-based technology into their pre-class self-learning may not facilitate the improvement of their English oracy within such a short period of 14 weeks. Besides, since the input of the ASR-based tasks was mainly self-produced by the students, most of whom were not as competent in oral English, both the quality and the quantity of such input might be inadequate and limited. Therefore, the TG students using ASR-based technology might need to take longer to attain improvements similar to those typically observed in naturalistic contexts (Hanzawa, 2021). In EFL pedagogy, the ASR-based pre-class tasks may need re-designing to better prepare the students for in-class peer interaction. Although it is a dictation ASR-based application which does not provide intended feedback as Google Assistant, iFlyRec may be used together with Google Assistant in future research to see whether the two kinds of ASR technologies could jointly improve Chinese EFL learners' peer interaction and further enhance iCALL in the Chinese EFL context. Moreover, to understand the holistic process of the technology use, future studies may study how the students use the ASR-based applications for oral practice and whether there are unforeseen usage behaviors on the learners' side.

Social context may also act as a critical factor in classroom-based interaction among Chinese EFL learners. In the Chinese context, communicating in English and providing corrective feedback to peers can easily be interpreted as an act of showing off, which may make their peers lose face. This was corroborated by Tomita and Spada's (2013) study among Japanese EFL learners, whose culture is deeply influenced by the Chinese culture. Conversely, the non-significant effect may also result from the test-oriented educational practice in the Chinese context. The Chinese students tend to focus on "knowledge on test papers," meaning what is tested in examinations or even quizzes will be considered as "useful knowledge" by them simply because such knowledge can score and make them rank higher. There is no doubt that the prevalent examination-oriented learning culture

substantially impacts teachers' teaching and students' learning behavior (Hu & West, 2014). Most Chinese students in high school learned English as an academic subject, of which the examination score served as a "ladder" for them to enter university. Since speaking is not tested in most middle schools in China, the students spent almost all their time on reading and writing practice (Liu & Chen, 2018). This kind of learning experience bears a considerable gap in tertiary EFL education. The gap between students' perceived mastery of exam-oriented English and the competence of authentic language use may constitute a major obstacle for most of the students to interact with each other orally in class.

In view of the above, it is worth mentioning that the use of educational technologies in EFL classrooms should be given full consideration of the Chinese educational context in practice. While the advantages of the ASR-based technology seem evidently conducive to developing students' English oracy in extant studies, as found in this study, its effect on facilitating task-based peer interaction in class may not be as expected. The discrepancy may be probably attributable to students' unforeseen and even distorted use of the ASR technology, which is beyond the scope of the current study. Accordingly, it is recommended that EFL teachers try to design and implement learning tasks that fit into the Chinese educational context. As Chinese university students are more likely to remain reticent in class (Tang et al., 2020), the tasks in Chinese EFL classrooms should be deliberately designed to have a meaningful connection with the students' lives. Besides, authenticity is regarded as a critical feature in task design. When the tasks cannot resemble real-life language use, the students may be less inclined to get involved. Meanwhile, to reduce Chinese students' speaking anxiety, teachers need to elaborate scaffolded tasks (Evers & Chen, 2020) by providing relevant organizational structures and flexible templates that best align to students' learner preparedness in a flipped EFL classroom. The scaffolded tasks should be contingent on students' task readiness, which requires skillful assessment for instructor to peg the task at the right level of zone of proximal development to optimize sustained engagement for the flipped approach. More importantly, helping Chinese EFL students to develop an English user identity may lead them to engage in positive peer interactions.

6. Conclusion, implications and limitations

The present study found that integrating the ASR-based technology into pre-class self-learning in a flipped EFL classroom could improve Chinese students' WTC with teacher and class as well as with non-Chinese. Contrarily, no significant difference was noticed in students' WTC with friends. With respect to the task-based peer interaction, despite the significant between-subjects difference in NfM, this study revealed that the Chinese students' interactional features in the EFL classroom did not improve significantly over time, indicating that the effects of the ASR-based technology on Chinese EFL learners' interactional features seemed to be quite limited.

The findings from the current study may prompt ASR proponents in EFL learning to revisit its effectiveness in the Chinese culture and the overriding impact of local culture on peer interaction in English class. Although task-based peer interactions are regarded as learning opportunities in EFL classrooms, the interactional features of the Chinese students' in-class peer interaction may not occur as expected. Therefore, course developers and practitioners need to take the cultural influence into serious consideration when designing EFL learning activities for Chinese learners. Chinese EFL teachers may need to come up with learning and teaching strategies to mitigate the effects of the local culture. Orienting the students to the benefits of task-based peer interactions in EFL classrooms is also necessary.

One major limitation of this study is the sample size. The participants all came from a single university in China, which might lead to a representativeness issue for the generalization of the findings. Another limitation of the present study is that, owing to the university rules, we were not allowed to collect data in person in the classroom. Therefore, the participants recorded their own task performance in class, which may have resulted in the loss of some raw data. Moreover, longer term research is needed to examine whether the novelty effect may contribute to the positive effect of the ASR-based technology on students' WTC revealed in the present study.

Acknowledgement

Our special thanks go to the Chinese University of Hong Kong Stanley Ho Big Data Decision Analytics Research Centre and the Chinese University of Hong Kong Teaching Development and Learning Enhancement Grant awarded to Professor Helen Meng for supporting this interdisciplinary research.

References

Ahn, T. Y., & Lee, S. M. M. (2016). User experience of a mobile speaking application with automatic speech recognition for EFL learning. *British Journal of Educational Technology, 47*(4), 778-786. https://doi.org/10.1111/bjet.12354

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for learning, teaching, and assessing: A Revision of Bloom's taxonomy of educational objectives.* Longman.

Clark, R., & Gieve, S. N. (2006). On the discursive construction of "the Chinese learner." *Language, Culture and Curriculum, 19*(1), 54-73. https://doi.org/10.1080/07908310608668754

Durkin, K. (2011). Adapting to Western norms of critical argumentation and debate. In L. Jin, & M. Cortazzi (Eds.), *Researching Chinese Learners: Skills, Perceptions, and Intercultural Adaptations* (pp. 274-291). Palgrave Macmillan.

Evers, K., & Chen, S. (2020). Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2020.1839504

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A Unit for all reasons. *Applied Linguistics, 21*(3), 354-375. https://doi.org/10.1093/applin/21.3.354

Gao, X. S. (2008). You had to work hard 'cause you didn't know whether you were going to wear shoes or straw sandals! *Journal of Language Identity and Education, 7*(3-4), 169-187. https://doi.org/10.1080/15348450802237798

Gass, S., Mackey, A., & Ross-Feldman, L. (2011). Task-based interactions in classroom and laboratory settings. *Language Learning, 61*(S1), 189-220. https://doi.org/10.1111/j.1467-9922.2011.00646.x

Hanzawa, K. (2021). Development of second language speech fluency in foreign language classrooms: A Longitudinal study. *Language Teaching Research.* https://doi.org/10.1177/13621688211008693

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77-89. https://doi.org/10.1080/19312450709336664

Hincks, R. (2015). Technology and leaning pronunciation. In M. Reed, & J. Levis (Eds.), *The Handbook of English Pronunciation* (pp. 505-519). John Wiley & Sons, Inc.

Hu, B., & West, A. (2014). Exam-oriented education and implementation of education policy for migrant children in urban China. *Educational Studies, 41*(3), 249-267. https://doi.org/10.1080/03055698.2014.977780

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2021a). Using automatic speech recognition technology to enhance EFL learners' oral language complexity in a flipped classroom. *Australasian Journal of Educational Technology, 37*(2), 110-131. https://doi.org/10.14742/ajet.6798

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Meng, Y. L., Chai, C. S., & Chen, M. Y. (2021b). Validating the general extended technology acceptance model for e-learning: Evidence from an online EFL course amid COVID-19. *Frontiers in Psychology, 12*, 671615. https://doi.org/10.3389/fpsyg.2021.671615

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., Liu, K. S. X., & Park, M. (2022a). A scoping review on flipped classroom approach in language education: Challenges, implications and an interaction model. *Computer Assisted Language Learning, 35*(5-6), 1218-1249. https://doi.org/10.1080/09588221.2020.1789171

Jiang, M. Y. C., Jong M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2022b). Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom. *Journal of Computers Assisted Learning*. https://doi.org/10.1111/jcal.12732

Jong, M. S. Y. (2017). Empowering students in the process of social inquiry learning through flipping the classroom. *Educational Technology & Society, 20*(1), 306-322.

Jong, M. S. Y. (2019). To flip or not to flip: Social science faculty members' concerns about flipping the classroom. *Journal of Computing in Higher Education, 31*(2), 391-407. https://doi.org/10.1007/s12528-019-09217-y

Jong, M. S. Y., Chen, G. W., Tam, V., & Chai, C. S. (2019). Adoption of flipped learning in social humanities education: The FIBER experience in secondary schools. *Interactive Learning Environments, 27*(8), 1222-1238. https://doi.org/10.1080/10494820.2018.1561473

Kim, I. S. (2006). Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. *Educational Technology & Society, 9*(1), 322-334.

Lee, J. C. K., & Yin. H. B. (2011). Teachers' emotions and professional identity in curriculum reform: A Chinese perspective. *Journal of Educational Change, 12*(1), 25-46. https://doi.org/10.1007/s10833-010-9149-3

Lee, J. S., & Sylvén, L. K. (2021). The Role of informal digital learning of English in Korean and Swedish EFL learners' communication behaviour. *British Journal of Educational Technology, 52*, 1279-1296. https://doi.org/10.1111/bjet.13082

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27*, 184-202. https://doi.org/10.1017/S0267190508070098

Liu, S., & Chen, Y. J. (2018). A Practical exploration on NMET(Shanghai)-based English listening and speaking teaching. *Technology Enhanced Foreign Language Education, 5*, 30-34. https://doi.org/CNKI:SUN:WYDH.0.2018-05-006.

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 285–329. https://doi.org/10.1017/S0261444818000125

Long, M. H. (1981). Input, interaction, and second language acquisition. *Annals of the New York Academy of Sciences, 379*(1), 259-278. https://doi.org/10.1111/j.1749-6632.1981.tb42014.x

Long, M. H. (2015). *Second language acquisition and task-based language teaching.* Wiley.

Lou, N. M., & Noels, K. A. (2021). Western and heritage cultural internalizations predict EFL students' language motivation and confidence. *International Journal of Bilingual Education and Bilingualism, 24*(5), 636-650. https://doi.org/10.1080/13670050.2018.1508277

MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *Modern Language Journal, 91*, 564-576. https://doi.org/10.1111/j.1540-4781.2007.00623.x

MacIntyre, P. D. (2020). Expanding the theoretical base for the dynamics of willingness to communicate. *Studies in Second Language Learning and Teaching, 10*(1), 111-131. https://doi.org/10.14746/ssllt.2020.10.1.6

MacIntyre, P. D., & Doucette, J. (2010). Willingness to communicate and action control. *System, 38*, 161-171. https://doi.org/10.1016/j.system.2009.12.013

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A Meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational Interaction in Second Language Acquisition* (pp. 407-452). Oxford University Press.

McCrocklin, S. (2016). Pronunciation learner autonomy: The Potential of automatic speech recognition. *System, 57*, 25-42. https://doi.org/10.1016/j.system.2015.12.013

McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation, 5*(1), 98-118. https://doi.org/10.1075/jslp.16034.mcc

Mroz, A. (2018). Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. *Foreign Language Annals, 51*(3), 617-637. https://doi.org/10.1111/flan.12348

Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*, 719-758. https://doi.org/10.1111/0023-8333.00172

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The Case of complexity. *Applied Linguistics, 30*(4), 555-578. https://doi.org/10.1093/applin/amp044

Pawlak, M., & Mystkowska-Wiertelak, A. (2015). Investigating the dynamic nature of L2 willingness to communicate. *System, 50*, 1-9. https://doi.org/10.1016/j.system.2015.02.001

Philp, J., Adams, R., & Iwashita, N. (2014). *Peer interaction and second language learning.* Routledge.

Pica, T. (1994). Questions from the language classroom: Research perspectives. *TESOL Quarterly, 28*, 49-79. https://doi.org/10.2307/3587198

Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences* (6th ed.). Routledge.

Rassaei, E. (2022). Recasts during mobile-mediated audio and video interactions: Learners' responses, their interpretations, and the development of English articles. *Computers Assisted Language Learning*. https://doi.org/10.1080/09588221.2019.1671461

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis, and second language learning and performance. *International Review of Applied Linguistics, 45*(3), 161-176. https://doi.org/10.1515/IRAL.2007.007

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sato, R. (2020). Examining fluctuations in the WTC of Japanese EFL speakers: Language proficiency, affective and conditional factors. *Language Teaching Research.* https://doi.org/10.1177/1362168820977825

Sun, Z. R., & Xie, K. (2020). How do students prepare in the pre-class setting of a flipped undergraduate math course? A Latent profile analysis of learning behavior and the impact of achievement goals. *The Internet and Higher Education, 46*, 100731. https://doi.org/10.1016/j.iheduc.2020.100731

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal, 82*, 320-337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

Tai, T. Y., & Chen, H. H. J. (2020). The Impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments.* https://doi.org/10.1080/10494820.2020.1841801

Tang, X. W., Wang, Y., & Wong, D. (2020). Learning to be silent: Examining Chinese elementary students' stories about why they do not speak in class. *Language, Culture and Curriculum, 33*(4), 384-401. https://doi.org/10.1080/07908318.2020.1715999

Tavakoli, E., & Davoudi, M. (2017). Willingness to communicate orally: The Case of Iranian EFL learners. *Journal of Psycholinguistic Research, 46*, 1509-1527. https://doi.org/10.1007/s10936-017-9504-0

Tomita, Y., & Spada, N. (2013). Form-focused instruction and learner investment in L2 communication. *The Modern Language Journal, 97*(3), 591-610. https://doi.org/10.1111/j.1540-4781.2013.12031.x

Tran, T. T. (2013). Is the learning approach of students from the Confucian heritage culture problematic? *Educational Research for Policy and Practice, 12*(1), 57-65. https://doi.org/10.1007/s10671-012-9131-3

Wang, C. C., Tseng, W. T., Chen, Y. L., & Cheng, H. F. (2019). Classroom interactions in the target language: Learners' perceptions, willingness to communicate, and communication behavior. *Asia-Pacific Education Researcher, 29*(5), 393-404. https://10.1007/s40299-019-00492-y

Wang, Y. H., & Young, S. S. C. (2014). A Study of the design and implementation of the ASR-based iCASL system with corrective feedback to facilitate English learning. *Educational Technology & Society, 17*(2), 219-233.

Wilson, M. L., & Zietz, J. (2004). Systematic bias in student self-reported data. *Journal for Economic Educators, 4*(4), 13-19.

Wu, Q. (2015). Re-examining the 'Chinese learner': A Case study of mainland Chinese students' learning experiences at British universities. *Higher Education, 70*(4), 753-766. https://doi.org/10.1007/s10734-015-9865-y

Xu, J. F., & Kou, J. N. (2018). Group interaction strategies and students' oral performance in Chinese EFL classrooms. *TESOL Quarterly, 52*(1), 198-209. https://doi.org/10.1002/tesq.398

Yum, J. O. (2007). Confucianism and communication: Jen, Li and Ubuntu. *China Media Research, 3*(4), 15-22. http://chinamediaresearch.net/index.php/back-issues?id=34

Zare, M., Shooshtari, Z. G., & Jalilifar, A. (2020). The Interplay of oral corrective feedback and L2 willingness to communicate across proficiency levels. *Language Teaching Research, 26*(6). https://doi.org/10.1177/1362168820928967

Zhang, J. Y., Beckmann, N., & Beckmann, J. F. (2020). More than meets the ear: Individual differences in trait and state willingness to communicate as predictors of language learning performance in a Chinese EFL context. *Language Teaching Research.* https://doi.org/10.1177/1362168820951931

Appendix A. Items in the WTC questionnaire

1. WTC with teacher and class (13 items)

WTCTC1. When the teacher's instruction for a task in classroom is not clear, I feel relaxed to ask in English from the teacher.
WTCTC2. For speaking in English in the class, I like to wait for my own turn and wait for the teacher to ask me to speak.
WTCTC3. When the teacher asks a question from all the students (for instance when asking the class opinion), I am willing to immediately answer in English.
WTCTC4. I am willing to participate in classroom discussions in English voluntarily.
WTCTC5. In classroom discussions, I am more willing to wait for the time the teacher asks me in person.
WTCTC6. I am not among the students who voluntarily start speaking in English in the classroom.
WTCTC7. I am willing to take part in classroom group works and speak in English.
WTCTC8. After classroom group discussions/speaking tasks, I am willing to be a volunteer to report the results.
WTCTC9. When I start speaking in English in the class in front of all my friends, I lose my confidence and concentration.
WTCTC10. I am willing to express my thoughts, opinions and even emotions in English in the class to all.
WTCTC11. I feel relaxed to share my opinions and even emotions with my teammates in English.
WTCTC12. I volunteer to orally present the lessons or talk about a topic in English to all the students.
WTCTC13. I prefer to keep silent in the class because speaking in English makes me agitated.

2. WTC with friends (7 items)

WTCF1. I am willing to speak in English with my classmates before the class begins.
WTCF2. If I have questions about the assigned homework for the next class, I prefer to ask in English from

the students next to me.

WTCF3. When the teacher's instruction for a task in classroom is not clear, I feel relaxed to ask in English from my friends sitting next to me.

WTCF4. I have the desire to communicate in English with my classmates on the first day of the class.

WTCF5. I like to use every opportunity in the class (like the break time or the spare time of the group works) to speak in English with my friends.

WTCF6. I have the desire to speak in English with my former classmates or teachers outside of the class at the language school.

WTCF7. For group speaking tasks, I am more willing to have students in my group who let me speak more.

3. WTC with non-Chinese (7 items)

WTCNS1. When I see a tourist in street, hotel, restaurant, or park, I try to find an excuse so that I can approach them and speak with them in English.

WTCNS2. If I travel to an English-speaking country, I feel relaxed to approach the people in the street, parks etc. and start communicating in English.

WTCNS3. I am willing to communicate in English with people who speak English as their first language.

WTCNS4. If I encounter people who can speak English as fluently as a native speaker, I will easily start speaking with them despite not knowing them in advance.

WTCNS5. I like to learn speaking English so that I can communicate orally with the English speakers.

WTCNS6. I am willing to be a tour guide for some days (even free of charge) so that I can communicate in English with the native speakers who have come to visit my city (such as Australians, Canadians).

WTCNS7. I am willing to be a tour guide for some days (even free of charge) so that I can communicate in English with the non- native speakers who have come to visit my city (such as Pakistanis, Japanese).

Appendix B. Definitions of the measures of interactional features (Gass et al., 2011)

Measures	Definitions	Working examples*
NfM	An attempt to overcome comprehension problems, such as confirmation checks, clarification requests and comprehension checks.	**Student 1:** … and I don't, I don't like to eat some French, eh, French things, like French fries in the morning because it isn't very healthy. **Student 2:** French fries? ***(Confirmation check)*** **Student 1:** Yeah. But I think the healthy diet is to, eh, keep your meals in order, … *[Class 1, Group 5, Unit Task 2: Food and diet]*
LRE	Any part of a dialogue in which students talk about the language they are producing, question their language use or other- or self-correct, such as corrective feedback and self-correction.	**Student 1:** Many people like west, eh, western ***(Self-correction)*** food, but I think Chinese food is very delicious, like, eh, noo…noodles. **Student 2:** Rice. *[Class 3, Group 2, Unit Task 2: Food and diet]*
recast	Correct restatement of a learner's incorrectly formed utterance.	**Student 1:** Oh, my parents like some, some place where it is quite, quite. **Student 2:** Quite? What the meaning? **Student 1:** Quite, such as quite, peace. **Student 2:** Quiet! ***(Recast)*** **Student 1:** Quiet. Okay, fine, quiet. *[Class 4, Group 1, Unit Task 6: Dream trip]*

Note. *The examples were directly obtained from the data collected in this study.