Journal of Information Technology Education: Research

An Official Publication of the Informing Science Institute
InformingScience.org

JITEResearch.org

Volume 24, 2025

ENHANCING EFL ORAL PRODUCTION THROUGH MOBILE-ASSISTED TASK-BASED LANGUAGE TEACHING: A STUDY IN EFFECTIVENESS

Shijiao Jia	Hebei University, Baoding, China	2455431447@qq.com
Zhaoxia Lu*	Cangzhou Jiaotong College, Cangzhou, China	Luzx86@163.com

* Corresponding author

ABSTRACT

Aim/Purpose	This study examines the effects of the mobile-assisted task-based language teaching (M-TBLT) approach on EFL learners' oral production. It evaluates three key second language acquisition measures: complexity (syntactic and lexical), accuracy (error-free clauses and correct verb forms), and fluency (unpruned and pruned speech rates). Additionally, it explores learners' perceptions of the approach to gain deeper insights into its effectiveness.
Background	Task-Based Language Teaching (TBLT) provides a communicative framework but often faces challenges such as insufficient oral practice, delayed feedback, and limited authenticity. Mobile-assisted language learning (MALL), with its flexibility and real-time interaction, offers potential solutions. However, empirical research integrating MALL with TBLT in oral English instruction remains scarce. This study introduces a structured six-step speaking instruction model to assess the impact of M-TBLT on learners' oral performance.
Methodology	A quasi-experimental pretest-posttest control group design was employed. Participants were university students from two intact speaking classes in Hebei, China. The intervention incorporated two mobile applications—WeChat and Liulishuo—within a structured six-step speaking instruction model. Data collection combined quantitative measures (pretest-posttest assessments) with qualitative insights from interviews and a teacher journal.
Contribution	This study fills a research gap by providing empirical evidence on the integration of mobile applications with TBLT in oral English instruction. Unlike existing studies that focus on spontaneous and unstructured speaking activities, this research integrates a well-sequenced instructional model. Findings contribute to

Accepting Editor Martin D Beer | Received: December 16, 2024 | Revised: February 26, March 19, 2025 | Accepted: March 27, 2025.
Cite as: Jia, S., & Lu, Z. (2025). Enhancing EFL oral production through mobile-assisted task-based language teaching: A study in effectiveness. *Journal of Information Technology Education: Research, 24,* Article 12.
https://doi.org/10.28945/5482

	both theoretical discussions and practical applications in mobile-assisted language learning.
Findings	Results indicate that M-TBLT significantly improved students' speaking complexity (dependent clauses per T-unit), accuracy (error-free clauses and correct verb forms), and fluency (unpruned and pruned speech rates). However, no significant improvement was observed in lexical complexity (mean segmental type-token ratio). Qualitative findings reveal that participants had positive perceptions of M-TBLT, appreciating mobile affordances (ease of use, synchronous/asynchronous communication, repetition, instant feedback, and authenticity) and educational benefits (increased speaking practice and learning autonomy). Students particularly valued WeChat's voice recording and group functions, as well as Liulishuo's automatic speech recognition and shadowing features.
Recommendations for Practitioners	The structured speaking-instruction model offers educators a practical framework for implementing M-TBLT, enhancing both instructional design and students' speaking performance.
Recommendations for Researchers	This study provides empirical support for M-TBLT in EFL settings. Researchers may explore its application in diverse linguistic and cultural contexts and examine additional instructional variables.
Impact on Society	These findings benefit educators, researchers, and mobile technology developers, encouraging collaboration in advancing mobile-assisted task-based learning.
Future Research	Further research could investigate M-TBLT in varied learning environments, incorporate additional performance measures, and assess new mobile-assisted speaking tools to enhance task-based learning.

Keywords	mobile-assisted task-based language teaching, speaking, complexity, accuracy, fluency, perception

INTRODUCTION

As an international language, English plays a critical role in facilitating communication worldwide. Therefore, acquiring proficient English-speaking skills has become an issue of utmost importance for non-English-speaking countries (Chen et al., 2019). As per Hafour (2022), the acquisition of speaking proficiency demands substantial time and commitment, involving efforts in designing, practicing, evaluating, and providing feedback. In many English-speaking courses in China, the conventional teaching approach was adopted, employing exam-focused and grammar-translation methods and viewing students as passive recipients of knowledge (Sun et al., 2017). Consequently, a large number of university graduates exhibit limited proficiency in spoken English, rendering them incapable of attaining a satisfactory level in oral communication.

As a pedagogical framework fostering meaningful, real-world interaction, Task-Based Language Teaching (TBLT) has gained considerable attention over the past three decades, which has been widely accepted by educators in countries of the Asia-Pacific region (Fang et al., 2021). Willis (2021) defines a classroom task as a purposeful communicative endeavor, emphasizing the exchange of meaning to achieve specific outcomes. Notwithstanding its pedagogical significance, TBLT encounters various challenges, including insufficient oral practices, lack of instant feedback, and deficiency in authenticity (see more details in the next section).

To overcome these challenges, researchers have been drawn to incorporating mobile technology into their oral instructions. Fang et al. (2021) and Hwang et al. (2022), among others, assert that due to the salient features of flexibility, portability, connectivity, and context sensitivity, mobile-assisted language learning (MALL) aligns with TBLT principles, thus can effectively tackle the challenges in

terms of compensating for limited opportunities for speaking practice (Hafour, 2022), engaging students with diverse authentic audio-video resources (Asratie et al., 2023), and facilitating instant feedback (Dai & Wu, 2022; Xu & Peng, 2017). However, by reviewing the previous studies, it was found that scarce studies provided empirical evidence of using mobile applications in oral English teaching in China, and research on the combination of mobile applications with TBLT is limited (Fang et al., 2021; Xue, 2020). In addition, teacher instruction in speaking class is mostly spontaneous, asystematic, and not well-sequenced. Limited empirical studies were found to report on speaking instructions with a well-developed model. In this study, the mobile-assisted TBLT (M-TBLT) approach was adopted, and a self-designed six-step speaking instruction model was followed to implement the approach. The overarching objective of the study is to examine the effects of the M-TBLT approach on EFL learners' oral production in terms of three principal measures in second language acquisition (SLA), i.e., complexity, accuracy, and fluency (CAF). A more comprehensive elucidation of CAF and the variables is provided in Section 2. Meanwhile, participants' perceptions towards the M-TBLT approach were also examined to have a more lucid understanding of the effects of the approach.

In this study, the effects of the M-TBLT approach on EFL learners' oral production in terms of CAF were examined, and the following research questions have been formulated to guide the study:

1) What are the effects of the M-TBLT approach on the complexity of EFL students' oral production in terms of: a) syntactical complexity DC/T? b) lexical complexity MSTTR?
2) What are the effects of the M-TBLT approach on the accuracy of EFL students' oral production in terms of: a) error-free clauses? b) correct verb forms?
3) What are the effects of M-TBLT approach on the fluency of EFL students' oral production in terms of: a) unpruned speech Rate A? b) pruned speech Rate B?
4) What are participants' perceptions towards the M-TBLT approach?

LITERATURE REVIEW

Proficiency in spoken English is an essential skill for language learners, which facilitates effective communication among students and their interlocutors (Chen et al., 2019). The conventional teaching approach constitutes a significant method in FL speaking instruction, and a common method employed in conventional teaching approach is the PPP model, i.e., presentation, practice, and production, which is a teacher-centered approach that requires students to remember the grammar rules and then apply them in translating from the source language to the target language (Ahmad, 2016). Guided by the conventional PPP model, teachers put a strong focus on the accurate use of language, resulting in students adopting a passive learning role with insufficient fluency and complexity practice (Sun et al., 2017). TBLT approach provides an alternative to the conventional teaching approach. It is regarded as a subset of communicative language teaching, which accentuates the principle that language learning occurs through extensive interactions and maximal engagement with the target language (Yuan & Ellis, 2003). As a communication-oriented and student-centered teaching approach, since its introduction, it has been widely used in developing students' oral production (Ellis et al., 2020; Xue, 2020).

Despite the pedagogical significance of TBLT, numerous challenges still need to be addressed, which will otherwise cause a diminishing effect on language learning. One major challenge is a lack of oral practice. EFL students encounter limited occasions to employ the target language beyond the classroom walls, and within the classroom, opportunities for practicing speaking are notably scarce, leading to constrained expressions in their oral communication (Akkara et al., 2020; Hwang et al., 2022). Insufficient instant feedback and limited exposure to authentic materials are the two other significant challenges faced by TBLT. Hwang et al. (2016) highlighted that a lack of instant feedback could have the potential to frequently result in communication breakdowns and further diminish the impact of task performance. Ko (2017) and Xue (2020), among others, stressed that a lack of authenticity was identified as a major challenge faced by conventional TBLT classrooms, potentially leading to an increased use of stilted English.

Whereas, scholars (e.g., Fang et al., 2021; Hsu & Liu, 2021; Hwang et al., 2022) mentioned that many of the challenges are caused due to the constraints of physical classroom settings with inadequate context-based practice, which could be addressed with the aid of technologies, especially mobile technology. MALL is a subtype of technology-assisted language learning (TALL). With the major features of portability, flexibility, ubiquity, etc., MALL has gained substantial attention in language learning (Karakaya & Bozkurt, 2022). Multiple researchers (e.g., Asratie et al., 2023; Hafour, 2022) found that integrating technology in teaching methods boosts students' speaking skills. Compared with computer-assisted language learning (CALL), MALL can be more deeply integrated into FL learning, as mobile devices possess more conspicuous advantages such as self-paced portability, more immediate Internet access, compatibility with various mobile applications, and flexibility with extended battery life, etc. (Karakaya & Bozkurt, 2022; Ko, 2017). The benefits of integrating MALL into TBLT-based speaking courses can be illustrated in three ways. First, it facilitates speaking practice. According to Hwang et al. (2016), the incorporation of mobile applications into speaking tasks serves as a valuable educational aid, facilitating the extension of students' formal learning to informal settings and allowing students to engage in more oral practice with fragmented time in their preferred places. Second, MALL assures instant feedback from peers and teachers when implementing speaking tasks. As Xu and Peng (2017) stressed, via community groups on mobile applications like WeChat, immediate feedback can be achieved both inside and outside the classroom, addressing the recognized issue of inadequate feedback in conventional classroom learning. Lim and Churchill (2016) also held that MALL expands instructors' capacity to track learning progress by providing real time feedback. Third, MALL could provide an authentic learning environment. According to Hsu and Liu (2021), the foremost pedagogical benefit of mobile devices is their ability to support learning in real-world settings, enhancing the authenticity and situated meaning-making during the learning process. In spite of the usability and rising popularity, there has been limited attention paid to the role of M-TBLT in the development of speaking skills (Kartal, 2022).

Complexity, Accuracy, and Fluency

Numerous researchers (e.g., Ellis, 2003; Skehan, 2003) have pointed out the multi-componential nature of FL speaking performance, which could be analyzed by a framework of three principal research variables, that is, complexity, accuracy, and fluency (CAF). A common definition of CAF that was agreed upon by most researchers is that complexity refers to the ability to produce a variety of diversified and refined items (Ellis, 2003). Accuracy means the ability to produce error-free speech (Foster et al., 2000), and fluency is the ability to approach the target language with English L1-like rapidity (Lennon, 1990) with few pauses, repairs or hesitations.

Conventionally, complexity has been measured in two major dimensions, i.e., syntactic complexity and lexical complexity. In this study, dependent clauses per T-unit (DC/T) was adopted for syntactic complexity and mean segmental type-token ratio (MSTTR) for lexical complexity. T-unit remains to be the most popular unit in measuring syntactic complexity, which is commonly used for non-interactive tasks (Ortega, 2003). Dependent clauses were defined as subordinate clauses, which contained a subject and a verb but cannot stand alone as a sentence. They commonly commence with subordinate conjunctions, such as if, although, as, who, whenever, etc. Lexical complexity is typically measured using the type-token ratio (TTR). Type refers to the basic forms of words, such as nouns, verbs, adjectives, etc., while token refers to the total number of words. However, TTR is extremely sensitive to the length of output, i.e., a higher token count would result in increased word types (Yuan & Ellis, 2003). MSTTR is a modification of TTR, which could offset the effects of differences in output length by calculating the mean ratios of TTR in all segments.

When measuring accuracy, both general and specific aspects were considered. For general aspects, the most widely adopted measures are error-free clauses (EFC), i.e., the ratio of error-free clauses to the total number of clauses (Hashemifardnia et al., 2021; Yuan & Ellis, 2003). The higher the ratio is, the more accurate the oral production is supposed to be. A clause was defined as a group of words that

is composed of at least one verb element (finite or non-finite) and a clause element (Foster et al., 2000). For specific aspects, the measure of correct verb forms (CVF) is commonly used (Hashemi-fardnia et al., 2021; Yuan & Ellis, 2003). Verbs were chosen due to their critical role in narrating stories. It was claimed to be a challenge for many Chinese FL learners to master the target-like use of verbs, especially in terms of verb tenses or other respects (Sun et al., 2017).

In terms of fluency, unpruned speech Rate A and pruned speech Rate B were adopted. Rate A mainly deals with the total number of syllables in the output, which considers both the amount of speech and pauses. Rate B refers to all syllables with repetition, reformulation, or replacement excluded, which is a measure of meaningful speech per time unit (Foster et al., 2000; Yuan & Ellis, 2003). The prominent merit of Rate A is that it contains both the length of speech and the length of pauses, encompassing the total number of syllables and overall time. It was found that part of the reason for students' pauses, hesitations, reformulations, and self-repairs was, in part, attributed to stress. The adoption of Rate B helped circumvent the confounding effects of learners' affective factors in task performance (Yuan & Ellis, 2003). CAF adopted in this study and the corresponding calculation methods are presented in Table 1.

Table 1. CAF measures and calculations

Measures	Aspects	Calculations
Complexity	DC/T	The number of dependent clauses/ number of T-units
	MSTTR	TTR of each segment/ number of segments
Accuracy	EFC	The number of error-free clauses/ total number of clauses
	CVF	The number of accurately used verb forms/ total number of verbs
Fluency	Rate A	(The total number of syllables/ the total amount of seconds) ×60
	Rate B	(The total number of meaningful syllables/ the total amount of seconds) ×60

Limitations of CAF Measures

While widely used, the CAF measures exhibit several limitations that warrant attention. First, complexity measures such as DC/T and MSTTR may not fully capture the depth of linguistic proficiency, as they often fail to account for variations in task demands or individual learner strategies. For example, while MSTTR is less sensitive to text length than TTR, it may still be influenced by the learner's overall fluency in a given context. Furthermore, DC/T overlooks the complexity of syntactic structures beyond dependent clauses, such as coordination or subordination.

Regarding accuracy, while EFC is a robust measure of overall grammatical correctness, it fails to address specific types of errors, such as those in word order or article usage. CVF is similarly limited, as it does not account for the learner's mastery of irregular verb forms, which are a known challenge in many FL contexts. The fluency measures, particularly Rate A and Rate B, may be influenced by affective factors such as anxiety, which could lead to inconsistencies across different learner profiles. Thus,

more comprehensive measures of fluency that address pause length, intonation, and cognitive processing would offer a more nuanced evaluation.

Theoretical Frameworks Underpinning CAF

The theoretical underpinnings of the CAF framework are grounded in cognitive and interactionist theories of second language acquisition. Cognitive theories, such as Levelt's (1993) model of speech production, suggest that fluency and accuracy emerge from the interaction of mental processing systems during speaking tasks. These cognitive models emphasize the mental effort involved in constructing grammatically accurate and lexically complex speech. Interactionist theories, on the other hand, highlight the role of interaction in language acquisition, where learners engage in conversation and use feedback from their interlocutors to improve accuracy and fluency (Long, 1983). The CAF framework thus integrates these perspectives by quantifying speaking performance across cognitive and interactive dimensions.

METHOD

PARTICIPANTS

This study adhered to pretest-posttest control group quasi-experimental design. Participants were selected from two intact classes undertaking the same speaking course among the ten parallel classes of the second academic year at the Faculty of Foreign Language (FFL) in a university situated in Hebei province, China. Students' mean English scores in the National College Entrance Examination (NCEE) were attained to ascertain the selection of two homogeneous groups. Subsequently, two classes, exhibiting the smallest difference value with the overall mean score of the ten parallel classes (0.1 and 0.06 respectively), were selected, and no significant difference (p=0.991, >0.05) existed in the two classes. One class was randomly assigned as the experimental group (EG, n=30) and the other as the control group (CG, n=30). The participants all aged between 18 and 21 years old with 9-13 years of experience in learning English, and no students in the EG or CG reported experience living or traveling in other countries. Prior to the study, all participants willingly provided their informed consent by signing the consent form.

INSTRUMENTS

This study utilized both quantitative and qualitative instruments to collect comprehensive data on the effects of the M-TBLT approach on students' oral production. The quantitative aspect involved a pretest-posttest design, while the qualitative aspect incorporated semi-structured interviews and a teacher journal (TJ).

Pretest and posttest

Students' FL oral production was measured respectively before and after the experiment by conducting a picture narrative test. Narrative tasks refer to a sequenced set of pictures provided for participants to elicit language performance. According to Yuan and Ellis (2003), narratives are monologues rather than dialogues, the influence of interactions while performing the speaking task could be avoided. Therefore, it has been employed by numerous researchers. A total of four sets of pictures were prepared and piloted with two experienced EFL teachers and two students who did not participate in the actual study. Ultimately, the picture set from E. Q. Plauen's (1903-1944) *Father and Son* was selected for the pre and posttest, as it incorporated more figures with a clearer storyline, which needed more verified vocabulary when narrating the story.

The pretest and posttest were conducted in the language laboratory of FFL. First, each student in EG and CG was given one sheet of empty paper to take notes, and then the picture narration task was presented to them. The paper was taken away before the task performance to avert the mixing

of speaking and writing performance. Students were allocated 10 minutes for preparation. Recordings began after preparation.

After the pre and posttest, the speech discourses were first transcribed verbatim and digitized using Xunjie (https://app.xunjiepdf.com/voice2text/), an intelligent voice transcription tool. Next, the transcriptions were proofread manually by two raters (see Table 2) to ensure their accuracy and consistency, which were then saved into the file "Unpruned Speech". A pruned form of the copy was saved after removing all repetitions, reformulation, replacement, and filled pauses. Examples of unpruned and pruned speech in each pruning aspect is presented in Appendix A. Next, speaking complexity, accuracy, and fluency were measured respectively to examine the effects of the teaching approach.

Table 2. Demographic information of the two raters

Pseudonym	Age	Gender	Qualification	Position	Background in linguistic teaching and learning
Rong	55	Female	Doctorate	Professor	> 20 years
Lu	39	Female	Doctorate	Lecturer	12 years

Rater training and calibration

Prior to scoring, both raters underwent thorough training and calibration to ensure consistent application of the scoring criteria for EFC, CVF, and total verb forms. This training involved reviewing example scripts, discussing the nuances of each criterion, and practicing the rating process on a sample of pretest recordings. Discrepancies between raters were resolved through discussion to achieve alignment on how specific errors or variations in speech should be scored.

Measuring tools and scoring criteria

As mentioned earlier, speaking complexity was measured by DC/T and MSTTR. The calculation of DC/T was operated on Second Language Syntactic Complexity Analyzer (L2SCA), and MSTTR was calculated on Lexical Complexity Analyzer (LCA), which were both developed by Xiaofei Lu (2017) at The Pennsylvania State University. When measuring fluency, an online syllable counter (https://www.syllablecount.com/) was used to calculate Rate A and Rate B. Both unpruned and pruned discourses were separately counted for the syllables, and then speech Rate A and speech Rate B were manually calculated. In terms of accuracy, when calculating EFC, the number of clauses were first determined on L2SCA and then manually segregated. CVF was also manually calculated for both correct verb forms and total verb forms. Discrepancies were resolved through discussions. The scoring criteria for EFC, CVF, and total verb forms is presented in Table 3. The raters first independently rated the same 25% randomly selected speech discourses to test inter-rater reliability. Pearson correlation coefficient was identified, which was above 0.8 for each measure (see Table 4), indicating a high consistency.

Table 3. Scoring criteria for EFC, CVF, and total verb forms

EFC (No repeated errors are counted in one clause)		
Scoring Aspects		Examples
Morphology:	Word form	These birds leaving the father and the son. (left)
	Word class	The boy played happy with the birds. (happily)
	S-V agreement	The story tell us we should be… (tells)
Syntax:	Tense	When they pass by a tree, the little boy was… (passed)
	Singular/plural	One day, the father and the son met a birds. (a bird)
	Collocation & expression	The little boy looked the birds in the tree. (looked at)
		He said cry. (cried)
Lexical Choices:		The birds chipped. (chirpped)
Unidentifiable places:		The children used their ** snacks to feed the birds.

CVF	
Scoring Aspects	Examples
S-V Agreement:	The birds was innocent. (were)
Tense:	The little boy took out his biscuits and feed the birds. (fed)
Modality:	To live in harmony, everyone will set an example. (should)
Morphology (word form):	The little boy felt very sad and crying. (cried)
Lexical Choices:	…, they heard several birds chopping. (chirpping)
Collocations:	…and his grandpa continued to fed the birds. (feed)
Unidentifiable places:	and then they dr** the birds away and left happily.

Total Verb Forms	
Scoring Aspects	Examples
Auxiliary word (counted together with the verb that followed):	was walking, were attracted, didn't leave, had gone, … (Each counted as ONE verb)
Modality (counted together with the verb that followed):	should protect, could divide, would be blessed, … (Each counted as ONE verb)
Collocations:	heard some birds chirping, need us to feed them, kept shouting, … (Each counted as TWO verbs)

Note. All examples in the table were from students' oral discourses in the oral test.

Table 4. Pearson correlation coefficient for inter-rater reliability

	Correlation between Rater 1 and 2	Sig. (2-tailed)
EFC	.992**	.000
CVF	.989**	.000
Total Verb Forms	.989**	.000

Note. **Correlation is significant at the 0.01 level (2-tailed).

Data Analysis

Quantitative data from the pretest and posttest were analyzed using SPSS (Version 25). The analysis followed a series of steps to ensure accurate and reliable results.

First, data cleaning and normality checks were performed. Initial checks revealed no missing values, outliers, or exceptional data points, and the distribution of data was found to be normal. This allowed for the use of parametric tests, which provide robust statistical power when assumptions of normality are met. Next, an independent samples t-test was conducted to compare the baseline CAF scores between the EG and the CG before the intervention. This test was chosen because it assesses whether the two groups had statistically significant differences in their initial oral production abilities, ensuring comparability between groups at the outset. This test helped assess whether there were any significant differences in CAF scores at the beginning of the study between the two groups. Following that, a paired samples t-test was used to analyze the changes in CAF scores within each group before and after the intervention. This test enabled the evaluation of how each group's performance had changed over time due to the intervention. Lastly, an ANCOVA (Analysis of Covariance) was employed to compare the changes in CAF scores between the two groups while controlling for potential covariates. ANCOVA was chosen because it enhances the accuracy of comparisons by accounting for initial differences between groups and controlling extraneous variables that could affect the results. This approach provides a more precise estimate of the intervention's effect on oral production.

Qualitative data were analyzed using an inductive content analysis approach to gain deeper insights into students' experiences and perceptions of the M-TBLT approach. Since qualitative data, such as transcribed interviews and teacher journals (TJs), capture participants' subjective views and experiences, inductive content analysis was selected to allow themes to emerge naturally from the data rather than being pre-determined. This method ensures that the findings reflect the authentic perspectives of participants. After the data were transcribed and proofread, the transcripts and TJ entries were subjected to manual coding to identify recurring themes and categories. The analysis involved reviewing and refining these themes to ensure they accurately captured participants' perceptions of the M-TBLT approach. To enhance validity, member checking was conducted, where participants verified the interpretations of their responses. Additionally, a peer reviewer was consulted to identify any themes that might have been overlooked, ensuring the reliability and credibility of the findings.

By integrating both quantitative and qualitative approaches, this study aimed to provide a comprehensive understanding of the effects of M-TBLT on EFL oral production. The quantitative analysis offered objective, statistical evidence of improvements in oral production, while the qualitative analysis provided rich, contextualized insights into learners' experiences, allowing for a more holistic interpretation of the findings.

Research Procedure

The experiment lasted for 12 weeks. In the first week, a picture narrating speaking pretest was conducted in both EG and CG. From Week 2 to Week 11, the M-TBLT approach was implemented in

EG, and CG conducted using a conventional teaching approach, without using mobile technologies. Immediately after the experiment, a speaking posttest was arranged for EG and CG, and a semi-structed interview was conducted in EG in Week 12.

To ensure comparability, the treatment received in EG and CG was ensured fair in aspects of coursebook utilization, teaching process, mediation and assessment, as well as supplementary materials. First, the same coursebook and speaking topics were employed during the same sessions, ensuring uniformity in content coverage. The teacher assigned the topics on a biweekly basis with two consecutive sessions of 100 minutes each week. The speaking tasks, which were meticulously designed to reflect real-world communicative situations, were mainly carried out through topic discussion or picture narrations. They gradually increase in complexity to accommodate learners' speaking skill development (see Appendix B for samples of the speaking tasks).

In the teaching process, attempts were made to teach based on a prespecified six-step instructional model (see Figure 1). Both EG and CG followed identical instructions and guidelines.

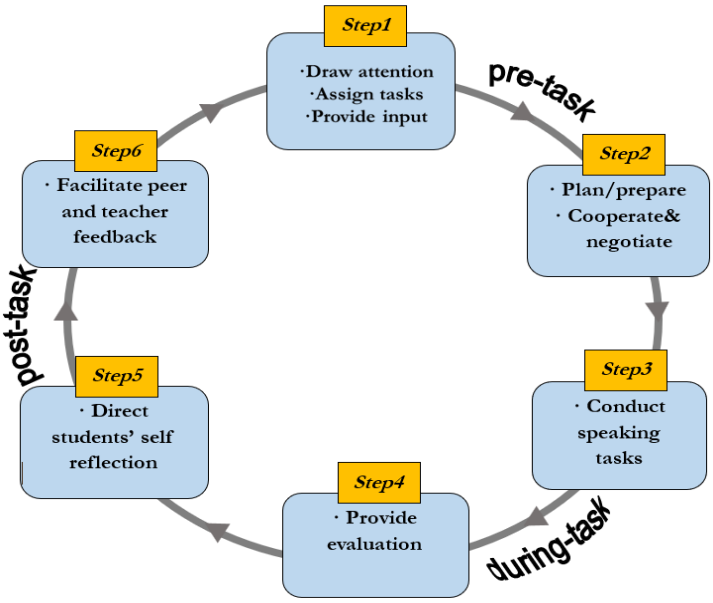

Figure 1. Speaking instructional model

As seen in the figure, the model includes three stages. In the pre-task stage, the teacher first draws students' attention, assigns the speaking task, elaborates on task requirements, assessment criteria, etc., and provides input of related linguistic and background knowledge. Next, students were allocated some time to plan for the task. Students were given sufficient time for pre-task planning, which, according to Yuan and Ellis (2003), could reduce their cognitive burden on message processing and lead to an increased fluency. Both EG and CG were divided into groups of five to discuss the speaking topics and get involved in discussion activities. They noted down the brainstormed ideas for subsequent task performance. Teachers in both EG and CG acted as facilitators, monitoring students' discussions, providing support, and offering suggestions when necessary to engage students in speaking activities. In the during-task stage, first, students independently performed the speaking task, during which the teacher refrained from interrupting to promote and ensure their speaking fluency. After students' task performance, the teacher would instantly evaluate all students in terms of pronunciation, speaking CAF, discourse strategies, etc. The focus shifted from fluency to accuracy by prompting students to notice and analyze their errors. In the post-task stage, students would first conduct self-reflection and peer feedback by completing a reflection form (refer to Appendix C).

This process elicits thoughts on their learning experience and offers peer scaffolding to help classmates improve. In the final step, the teacher would provide feedback concerning students' overall performance in previous steps. Out of class, students could do more oral practice.

Although EG and CG followed the same instructional process, learners in EG utilized mobile apps of WeChat and Liulishuo to assist their language learning experience, whilst CG learners refrained from using any mobile apps for English-speaking learning. Major functions of WeChat and Liulishuo are respectively presented in Figures 2 and 3. As a mobile social media app, WeChat is equipped with prominent functions such as voice messaging, voice recording, and group building. With the world's most extensive speech corpus for Chinese people speaking English, the AI-powered system of Liulishuo includes functions of automatic speech recognition (ASR), shadowing reading, and providing authentic learning materials. Students in EG mainly used the mobile apps for out-of-class learning.

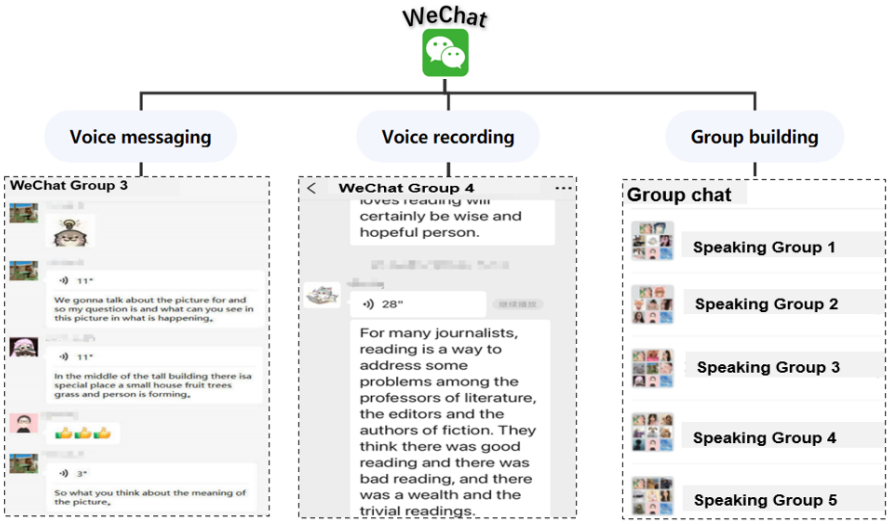

Figure 2. Major functions of WeChat

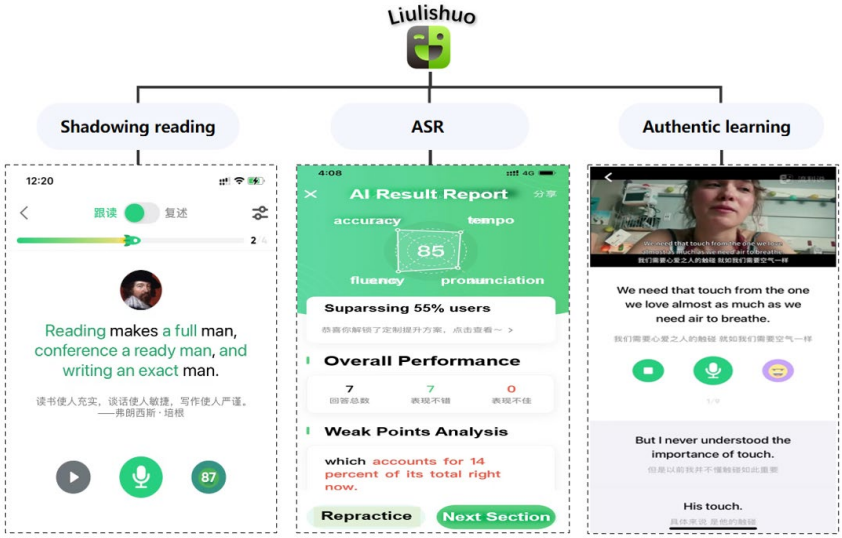

Figure 3. Major functions of Liulishuo

On the one hand, student-student and teacher-student discussion process was facilitated by WeChat, given its capability for both synchronous and asynchronous communication. In EG, students could engage in virtual meetings using video conferencing tools or communicating through voice or text messages. When engage in WeChat interactive communications, the learners were required to discuss

in English. The teacher was also actively involved in these sessions, offering guidance and support to learners when necessary. Students could also record their speaking task by using voice recording function in WeChat, they can repeatedly listen to their own and their peers' voices, enhancing awareness of their speaking deficiencies. Meanwhile, teacher evaluation process was facilitated by sharing digital resources, such as videos featuring competent speakers, in students' WeChat groups. Additionally, community groups on WeChat made both instant peer and teacher feedback possible. The CG learners were also encouraged to engage in out-of-class speaking activities by meeting up their peers in person. They could also receive out-of-class teacher mediation and assessment face-to-face.

On the other hand, students could further do independent oral practice on Liulishuo by using its AI-based functions at their convenience. Before class, the teacher would send some authentic video/audio materials pertinent to the speaking topic on Liulishuo for students to familiarize with the speaking topic. Liulishuo is specifically designed for speaking learning, which could provide real-time scores and automatically generate a comprehensive result report, pinpointing areas for improvement, including pronunciation, tone, pausing, etc. Students could also do shadowing reading by repeating after English L1 speakers on the platform, enabling them to identify the discrepancy in their own speech. In addition, authentic video and audio materials featuring various topics and themes, which happen in authentic social context, are available on Liulishuo. Learners in CG were also provided with a set of out-of-class supplementary materials that were similar in scope and type to those available to the MALL group.

RESULTS

EFFECTS OF THE M-TBLT APPROACH ON COMPLEXITY, ACCURACY, AND FLUENCY OF EFL STUDENTS' ORAL PRODUCTION

To examine effects of the M-TBLT approach on students' oral production in terms of complexity, accuracy, and fluency, data collected from the speaking pretest and posttest were analyzed in SPSS (Version 25). An independent samples *t*-test was conducted to examine whether significant difference existed between CG and EG in the pretest in CAF. As observed in Table 5, mean values of speaking complexity in terms of DC/T and MSTTR, accuracy measured by EFC and CVF, and fluency accessed through Rate A and Rate B exhibit close proximity, and the p-values are all above 0.05, indicating the absence of a significant difference between CG and EG in CAF before the experiment.

To examine the effects of the M-TBLT approach on students' oral production in terms of complexity, accuracy, and fluency, data from the pretest and posttest were analyzed using SPSS (Version 25). An independent samples t-test revealed no significant differences between EG and CG in CAF before the experiment, as indicated by the similar mean values for complexity (DC/T, MSTTR), accuracy (EFC, CVF), and fluency (Rate A, Rate B), with p-values all above 0.05 (see Table 5).

Table 5. Independent samples t-test results of CAF in the pretest

		CG		EG		
		M	*SD*	*M*	*SD*	*Sig. (2-tailed)*
Complexity	DC/T	0.28	0.19	0.29	0.13	0.886
	MSTTR	0.67	0.04	0.66	0.04	0.582
Accuracy	EFC	0.61	0.18	0.63	0.14	0.564
	CVF	0.78	0.14	0.79	0.11	0.590
Fluency	Rate A	118.26	21.08	117.78	22.3	0.932
	Rate B	112.65	23.03	110.42	24.36	0.718

At the end of the experiment, both paired samples *t*-test (Table 6) and ANCOVA (Table 7) were conducted to compare the pretest and posttest results within each group and between groups. As shown in Table 6, within EG, the posttest mean values in EG are significantly higher ($p < 0.05$) than those in the pretest. In contrast, the posttest mean values in CG closely resembled those of the pretest, and no significant difference ($p > 0.05$) was observed across all CAF measures. Consequently, it can be inferred that the M-TBLT approach significantly improved oral performance while the conventional teaching approach did not yield significant improvement in EFL learners' oral production. In addition, the ANCOVA results indicated a statistically significant difference ($p < 0.05$) between EG and CG in DC/T, EFC, CVF, as well as Rate A and Rate B in favor of EG after controlling for the covariate in the pretest. No significant difference was observed in the measure of MSTTR ($p > 0.05$). It indicated that compared with conventional teaching approach, the M-TBLT approach could significantly enhance EFL students' oral production regarding the syntactical complexity of DC/T, speaking accuracy of EFC and CVF, and speaking fluency of Rate A and Rate B, despite that no significant effects can be produced in the lexical complexity of MSTTR.

Table 6. Paired samples t-test result of speaking CAF in CG and EG

		Pretest		Posttest			
		M	*SD*	*M*	*SD*	*t*	*Sig. (2-tailed)*
CG	DC/T	0.28	0.19	0.23	0.13	-4.754	.291
	MSTTR	0.67	0.04	0.68	0.06	-3.449	.226
	EFC	0.61	0.18	0.56	0.19	1.115	.274
	CVF	0.78	0.14	0.71	0.17	1.878	.070
	Rate A	118.26	21.08	114.29	19.55	0.766	.450
	Rate B	112.65	23.03	110.25	19.60	0.441	.662
EG	DC/T	0.29	0.13	0.52	0.22	1.075	.000
	MSTTR	0.66	0.04	0.68	0.03	-1.237	.002
	EFC	0.63	0.14	0.91	0.06	-11.538	.000
	CVF	0.79	0.11	0.96	0.03	-8.878	.000
	Rate A	117.78	22.30	131.69	21.62	-2.289	.030
	Rate B	110.42	24.36	125.64	26.07	-2.212	.035

Meanwhile, for a deeper insight into students' speaking fluency, the mean length of speaking time and mean values of Rate A and Rate B in the pretest and posttest were calculated (see Table 8). A graphical representation of the trendline for Rate A and Rate B over time is also depicted and shown in Figure 4. The red dotted lines with arrows are the exponential trendline drawn based on Rate B, i.e., the meaningful syllables.

Table 7. ANCOVA results comparing CG and EG after controlling covariates

	Type III Sum of Squares	df	Mean Square	F	*Sig.*	Partial Eta Squared
DC/T	1.225	1	1.225	37.766	0.000	0.399
MSTTR	0.001	1	0.001	0.403	0.528	0.007
EFC	1.809	1	1.809	95.314	0.000	0.626
CVF	0.876	1	0.876	61.522	0.000	0.519
Rate A	4527.633	1	4527.633	10.525	0.002	0.156
Rate B	3493.459	1	3493.459	6.475	0.014	0.102

Table 8. Mean time length and mean values of Rate A and Rate B

	Pretest		Posttest	
	CG	EG	CG	EG
Time	125.5	143.2	127.4	165.7
Rate A	118.26	117.78	118.29	131.69
Rate B	112.48	110.42	113.25	125.64

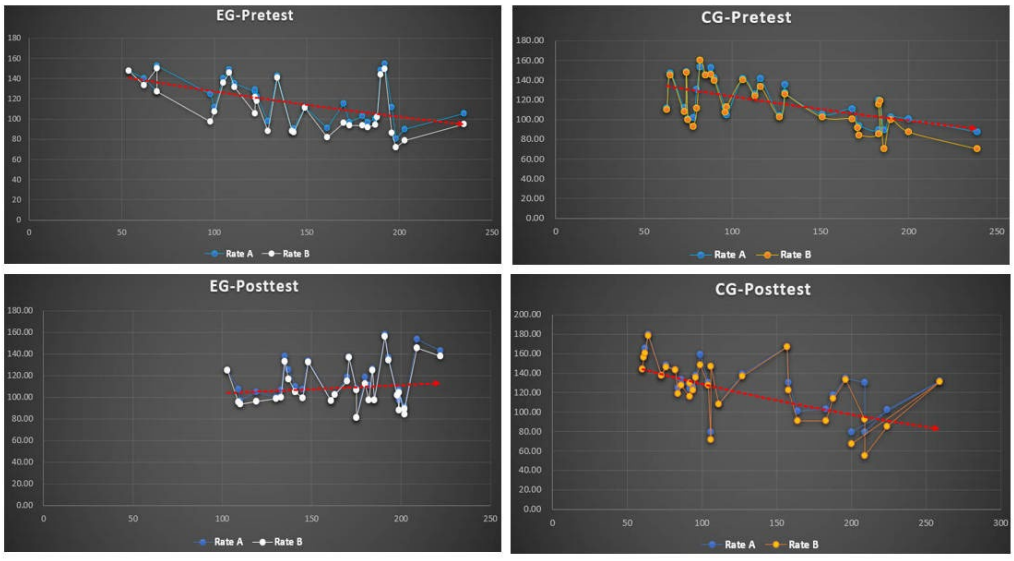

Figure 4. Trendline of rate A and rate B for both EG and CG

As seen in Figure 4, in the pretest, the exponential lines for both EG and CG, as well as in the post-test for CG, exhibit a downward trend. Conversely, the exponential trendline in EG (post) shows an upward trajectory, suggesting that the implementation of the M-TBLT approach in EG led to a growth in meaningful syllables over time. Meanwhile, the statistics in the table show that the mean length of time in EG increased from 143.2 to 165.7, and the mean values of Rate A and Rate B increased from 117.78 to 131.69 and 110.42 to 125.64, respectively. In contrast, in CG, the mean length of time and mean values of Rate A and Rate B had a minimal change.

Participants' Perceptions towards the M-TBLT Approach

As mentioned, participants' perceptions of the M-TBLT approach were explored through the semi-structured interview and TJ. Both interview transcriptions and the TJ were subjected to manual inductive content analysis for recurring themes. To enhance validity, the data were subjected to a peer reviewer who examined them looking for themes that may have been missed. The results showed that participants held positive attitudes towards the M-TBLT approach. The findings are given under the main categories (mobile affordances and educational affordances) and themes, which are shown in Table 9.

Table 9. The list of categories, themes, and codes

Categories	Themes	Codes	*N*
Mobile affordances	Ease of use	convenient (10); anywhere, anytime (10); portability (7)	5
	Synchronous/asynchronous communication	voice messaging/chatting (18); group building on WeChat (20)	7
	Repetition	voice recording on WeChat (14); shadowing reading on Liulishuo (15)	5
	Instant feedback	technological feedback ASR (19); instant peer & teacher feedback on WeChat (22)	7
	Authenticity	authentic video/audio materials (15); providing authentic learning context (13)	6
Educational affordances	More speaking practice	expanding speaking opportunities (18); pre-class planning & out-of-class practice (12)	6
	Learning autonomy	more self-regulation (8); autonomous/independent learning (9)	4

Note. The number in parenthesis refers to the frequency with which the expression was repeated. One person could emphasize the expression more than once. *N* shows the number of participants mentioning the theme.

Mobile affordances

The analysis identified five major themes within mobile affordances that contributed to participants' positive attitudes toward the M-TBLT approach: ease of use, synchronous and asynchronous communication, repeated oral practice, instant feedback, and authenticity (see Appendix D for representative participant quotes supporting these findings).

First, participants highlighted the ease of use of mobile applications, emphasizing their quick accessibility, compatibility across devices, and flexibility for oral practice beyond classroom constraints. They noted that WeChat and Liulishuo enabled seamless interaction due to their user-friendly interface and cross-platform functionality.

Second, synchronous and asynchronous communication emerged as a key theme, with participants highlighting WeChat's role in facilitating real-time and flexible interactions among students and between students and the teacher. They noted that the platform's voice and text messaging functions enabled instant communication, overcoming physical constraints and ensuring seamless peer-to-peer and student-teacher exchanges. Additionally, the group-building function on WeChat was widely praised for fostering structured and interactive discussions. Participants emphasized that working in smaller groups promoted focused engagement, allowing for more in-depth conversations and encouraging active participation. The combination of synchronous messaging and structured group discussions significantly enhanced students' communication experiences in the M-TBLT approach.

Third, participants valued the opportunity for repeated oral practice provided by the WeChat voice recording function and Liulishuo's shadowing reading feature, leading to enhanced pronunciation and speaking performance. The ability to record and review their own speech allowed students to identify and correct mispronunciations, making the learning process more self-directed and reflective. Additionally, shadowing reading enabled students to imitate native English speakers, improving their accuracy and fluency through repeated practice and comparison. These affordances contributed to greater self-awareness and iterative improvement in their speaking skills.

Fourth, instant feedback emerged as another major factor contributing to participants' positive attitudes toward the M-TBLT approach, with mobile-assisted feedback occurring through ASR technology and peer and teacher feedback in WeChat groups. The ASR function on Liulishuo provided immediate evaluation reports, allowing students to receive scores and pinpoint areas for improvement right after speaking practice, which they found highly engaging and beneficial. Meanwhile, WeChat facilitated real-time feedback from both peers and teachers, enabling students to identify weaknesses in their oral output and enhance their speaking accuracy through immediate corrections and discussions. These affordances reinforced self-monitoring learning, making the feedback process both efficient and interactive.

Finally, participants reported that authenticity played a crucial role in their learning experience, Liulishuo provided students with exposure to real-life language use in situated contexts. Through engagement with authentic corpus materials, students were able to accumulate natural oral expressions and gain a deeper understanding of language appropriateness. This practical application of learned expressions allowed them to identify when and how to use language effectively in real-world interactions, thereby reinforcing both comprehension and communicative competence.

Educational affordances

In addition to the mobile affordances, the analysis yielded two educational affordances that benefit from the M-TBLT approach, including increased speaking opportunities outside class and enhanced learner autonomy (see Appendix E for representative participant quotes supporting these findings).

The M-TBLT approach significantly expanded speaking opportunities beyond the classroom, addressing the limitations of conventional speaking classes, where time constraints often restrict individual practice. Through mobile applications, students engaged in pre-class preparation, including exposure to authentic learning materials and collaborative discussions in WeChat groups. The structured tasks, such as submitting progress reports, ensured that students completed their speaking practice and arrived in class better prepared for communicative activities. As a result, students experienced more frequent and meaningful speaking practice, both individually and in group settings, contributing to their overall speaking development.

Additionally, the M-TBLT approach fostered greater student autonomy by enabling learners to take more control over their oral practice through mobile technologies. Without direct teacher supervision, students engaged in self-regulated learning, making use of mobile applications to practice

speaking at their own pace and seeking additional resources. The flexibility of mobile-assisted learning allowed them to select materials based on their individual needs and proactively seek support in peer discussions, leading to a more self-directed and efficient learning experience.

Furthermore, the teacher also noted in the TJ that students' attitudes toward the mobile applications changed from being unfamiliar and being pushed to use it to enjoying and habitually using it to practice speaking. Therefore, they dedicated tremendously more time to oral practice than at the beginning of the experiment.

DISCUSSION

This study aimed to explore the effects of the M-TBLT approach on EFL students' oral production, focusing on three key aspects: syntactical complexity, accuracy, and fluency. The results indicate significant improvements in all three dimensions, and these findings can be understood through the lens of previous research as well as the specific mechanisms inherent in the M-TBLT approach.

EFFECTS ON SYNTACTICAL COMPLEXITY (DC/T)

Our findings show that students' syntactical complexity, particularly in their use of dependent clauses, was notably improved. This is in line with previous studies, such as those by Wang and Han (2021), which reported that mobile-based learning supports syntactical development by increasing the use of complex sentence structures. In our study, the M-TBLT approach facilitated this improvement by providing students with extensive opportunities to practice speaking through mobile applications like Liulishuo. The task-based learning environment, coupled with repeated practice and immediate feedback, may have encouraged students to experiment with more complex sentence structures. As students continued to receive feedback on their syntactical errors via ASR and peer feedback through WeChat, they were able to identify and correct mistakes, thus enhancing their syntactical complexity over time.

EFFECTS ON ACCURACY (EFC AND CVF)

Accuracy was another area of significant improvement, particularly in terms of error-free clauses, specially in the accurate use of morphology and syntax, and the correct use of verb forms. Students showed a marked reduction in errors related to subject-verb agreement, tenses, and modality. By the end of the experiment, students demonstrated improved accuracy in word forms and classes, reducing errors in tenses as well as singular or plural forms. Students reported that correctly using verb forms was a primary challenge in their speaking. However, with technological ASR feedback and instant peer/teacher feedback on WeChat, they became more attentive to the accurate use of third-person verb forms and avoiding verb omissions.

The finding aligns with results reported by Dai and Wu (2022) and Tsai (2019), which highlighted the positive impact of combining ASR feedback with peer and teacher feedback on language accuracy. In this study, ASR feedback on Liulishuo provided students with real-time evaluations of their speech, making them more aware of errors in verb forms and encouraging them to correct these mistakes during practice. Peer and teacher feedback via WeChat further reinforced these corrections. This process resonates with Kartal (2022), who noted that peer feedback helps learners develop self-evaluation skills, crucial for improving accuracy in oral production. The students' reports of becoming more attentive to their verb forms reflect the value of such feedback in enhancing grammatical precision.

EFFECTS ON FLUENCY (SPEECH RATE A AND B)

Fluency improvements were evident in the increased speech rate and the reduction of filled pauses, which is consistent with findings by Fang et al. (2021) that mobile-supported task-based learning

contributes to better fluency through increased speaking opportunities. The M-TBLT approach created a learning environment where students had more chances to practice speaking outside the classroom, leading to greater fluency. Activities like shadow reading, voice recording, and peer feedback discussions on WeChat provided students with repetitive practice opportunities, which contributed to smoother, more coherent speech production. Moreover, as students engaged in these activities, they became more conscious of and less reliant on fillers like "um" or "uh," contributing to a more fluent speaking style. The teacher observed that students who had more practice through these mobile-assisted tasks were more likely to speak confidently without resorting to reformulations or pauses, demonstrating enhanced fluency.

Perceptions of M-TBLT

Participants expressed a strong positive attitude toward the M-TBLT approach, largely due to its mobile and educational affordances, which contributed significantly to enhancing their learning experiences. These affordances provided opportunities for increased speaking practice, real-time feedback, greater autonomy, and more authentic language use, ultimately fostering a more engaging and effective learning environment. The finding is consistent with previous studies (e.g., Chapelle, 2006; Godwin-Jones, 2019) that have emphasized the importance of learner-centered, technology-enhanced learning environments in improving language proficiency. The students appreciated the flexibility that mobile learning offered, allowing them to practice at their own pace and at times convenient for them. They reported that this increased their confidence and reduced the anxiety typically associated with speaking in front of others. As Kormos and Préfontaine (2017) suggested, reducing anxiety in language learners can improve fluency and accuracy, and our findings support this claim, as students mentioned that practicing in a less stressful virtual environment helped them overcome their speaking challenges.

Lexical Complexity (MSTTR)

Despite significant improvements in fluency, accuracy, and syntactical complexity, no significant difference in lexical complexity (measured by MSTTR) was found. This result contrasts with Wang and Han's (2021) findings, which suggested that mobile-assisted learning promotes lexical development. Several factors may explain this discrepancy. One potential explanation is that students focused more on syntactical complexity and fluency during their practice, possibly at the expense of lexical variation. Some students reported relying on Chinese phrase structures when speaking, which could have constrained their ability to experiment with diverse vocabulary. Furthermore, the focus of the M-TBLT approach was primarily on improving accuracy and fluency, which may have limited opportunities to consciously engage with and expand lexical choices. To enhance lexical richness, future studies could incorporate specific vocabulary-focused tasks within the mobile-assisted environment, helping students expand their lexical range by interacting with authentic materials that introduce varied vocabulary.

Conclusion

This study provides new insights into the integration of MALL with TBLT, focusing on its impact on EFL students' oral production. By incorporating mobile applications such as Liulishuo and WeChat, which offer real-time feedback and facilitate increased speaking opportunities outside the classroom, this research demonstrates the potential of mobile technology to enhance language learning, particularly in the areas of fluency, accuracy, and syntactical complexity.

The study revealed that the M-TBLT approach led to notable improvements in speaking fluency, with students showing greater confidence and smoothness in their spoken English. Additionally, there was significant growth in syntactical complexity and accuracy, reflecting the positive influence of mobile-assisted, task-based interactions. However, the study also found that lexical complexity did not improve significantly, suggesting a gap in vocabulary development that warrants further attention. This result

suggests that the current implementation of M-TBLT, while effective in improving fluency and syntactic complexity, may require adjustments to specifically target lexical enhancement. Future studies could focus on incorporating specific tasks aimed at vocabulary growth.

The contribution of this study lies in its demonstration of how M-TBLT can be an effective pedagogical approach for enhancing EFL students' oral production, offering empirical evidence that bridges the gap between MALL and TBLT. The findings advance our understanding of how mobile technology can facilitate task-based learning, especially in terms of creating flexible, stress-free environments for language practice. This study contributes to the ongoing conversation about the integration of mobile technology in language education and provides insights into the types of feedback and practice that best support language development.

In terms of practical implications, educators can utilize mobile applications to provide real-time feedback and increase speaking opportunities for students, especially those who experience anxiety in conventional classroom settings. This approach allows for continuous interaction with the language, making it easier for students to practice and improve outside of class. The findings offer a framework for integrating mobile technology into language teaching, emphasizing the importance of creating supportive, flexible learning environments that facilitate student engagement and progress.

Future research should explore the long-term effects of the M-TBLT approach, as well as examine additional factors such as phonological accuracy and vocabulary use, to further assess the approach's impact on language acquisition. The inclusion of a broader range of linguistic variables and a longer study duration could provide more comprehensive insights into the potential of M-TBLT across different EFL contexts.

FUNDING

This research was supported by the Hebei Social Science Fund (No. HB21YY023).

REFERENCES

Ahmad, J. (2016). Technology-assisted language learning is a silver bullet for enhancing language competence and performance: A case study. *International Journal of Applied Linguistics and English Literature*, *5*(7), 118-131. https://doi.org/10.7575/aiac.ijalel.v.5n.7p.118

Akkara, S., Anumula, V. S. S., & Mallampalli, M. S. (2020). Impact of WhatsApp interaction on improving L2 speaking skills. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(3), 250–259. https://doi.org/10.3991/ijet.v15i03.11534

Asratie, M. G., Wale, B. D., & Aylet, Y. T. (2023). Effects of using educational technology tools to enhance EFL students' speaking performance. *Education and Information Technologies*, 1-21. https://doi.org/10.1007/s10639-022-11562-y

Chapelle, C. A. (2006). *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology*. John Benjamins Publishing Company.

Chen, C. M., Liu, H., & Huang, H. B. (2019). Effects of a mobile game-based English vocabulary learning app on learners' perceptions and learning performance: A case study of Taiwanese EFL learners. *ReCALL*, *31*(2), 170-188. https://doi.org/10.1017/s0958344018000228

Dai, Y., & Wu, Z. (2022). Mobile-assisted peer feedback on EFL pronunciation: Outcome effects, interactional processes, and shaping factors. *System*, *111*, 102953. https://doi.org/10.1016/j.system.2022.102953

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford university press.

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2020). *Task-based language teaching: Theory and practice*. Cambridge University Press.

Fang, W. C., Yeh, H. C., Luo, B. R., & Chen, N. S. (2021). Effects of mobile-supported task-based language teaching on EFL students' linguistic achievement and conversational interaction. *ReCALL*, *33*(1), 71-87. https://doi.org/10.1017/s0958344020000208

Foster, P., Tonkin, 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

Godwin-Jones, R. (2019). Emerging technologies: Mobile-assisted language learning. *Language Learning & Technology*, *23*(1), 1-9.

Hafour, M. F. (2022). Interactive digital media assignments: effects on EFL learners' overall and micro-level oral language skills. *Computer Assisted Language Learning*, 1-33. https://doi.org/10.1080/09588221.2022.2067180

Hashemifardnia, A., Shafiee, S., Esfahani, F. R., & Sepehri, M. (2021). Effects of massive open online course (MOOC) on Iranian EFL Learners' speaking complexity, accuracy, and fluency. *Computer-Assisted Language Learning Electronic Journal (CALL-EJ)*, *22*(1), 56-79. https://doi.org/10.1080/2331186X.2021.1987375

Hsu, K. C., & Liu, G. Z. (2021). A systematic review of mobile-assisted oral communication development from selected papers published between 2010 and 2019. *Interactive Learning Environments*, 1-17. https://doi.org/10.1080/10494820.2021.1943690

Hwang, W. Y., Guo, B. C., Hoang, A., Chang, C. C., & Wu, N. T. (2022). Facilitating authentic contextual EFL speaking and conversation with smart mechanisms and investigating its influence on learning achievements. *Computer Assisted Language Learning*, 1-27. https://doi.org/10.1080/09588221.2022.2095406

Hwang, W. Y., Shih, T. K., Ma, Z. H., Shadiev, R., & Chen, S. Y. (2016). Evaluating listening and speaking skills in a mobile game-based learning environment with situational contexts. *Computer Assisted Language Learning*, *29*(4), 639–657. https://doi.org/10.1080/09588221.2015.1016438

Karakaya, K., & Bozkurt, A. (2022). Mobile-assisted language learning (MALL) research trends and patterns through bibliometric analysis: Empowering language learners through ubiquitous educational technologies. *System*, *110*, 102925. https://doi.org/10.1016/j.system.2022.102925

Kartal, G. (2022). Evaluating a mobile instant messaging tool for efficient large-class speaking instruction. *Computer Assisted Language Learning*, 1-29. https://doi.org/10.1080/09588221.2022.2074463

Ko, M. H. (2017). Learner perspectives regarding device type in technology-assisted language learning. *Computer Assisted Language Learning*, *30*(8), 844-863. https://doi.org/10.1080/09588221.2017.1367310

Kormos, J., & Préfontaine, Y. (2017). Affective factors influencing fluent performance: French learners' appraisals of second language speech tasks. *Language Teaching Research*, *21*(6), 699-716. https://doi.org/10.1177/1362168816683562

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*(3), 387-417. https://doi.org/10.1111/j.1467-1770.1990.tb00669.x

Levelt, W. J. (1993). *Speaking: From intention to articulation*. MIT press.

Lim, C. P., & Churchill, D. (2016). Mobile learning. *Interactive Learning Environments*, *24*(2), 273-276. https://doi.org/10.1080/10494820.2015.1113705

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input1. *Applied Linguistics*, *4*(2), 126-141. https://doi.org/10.1093/applin/4.2.126

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing*, *34*(4), 493-511. https://doi.org/10.1177/0265532217710675

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*(4), 492-518. https://doi.org/10.1093/applin/24.4.492

Skehan, P. (2003). Task-based instruction. *Language Teaching*, *36*(1), 1-14. https://doi.org/10.1017/s026144480200188x

Sun, Z., Lin, C. H., You, J., Shen, H. J., Qi, S., & Luo, L. (2017). Improving the English-speaking skills of young learners through mobile social networking. *Computer Assisted Language Learning*, *30*(3-4), 304-324. https://doi.org/10.1080/09588221.2017.1308384

Tsai, P. H. (2019). Beyond self-directed computer-assisted pronunciation learning: A qualitative investigation of a collaborative approach. *Computer Assisted Language Learning*, *32*(7), 713–744. https://doi.org/10.1080/09588221.2019.1614069

Wang, Z., & Han, F. (2021). Developing English language learners' oral production with a digital game-based mobile application. *Plos One*, *16*(1), 1-11. https://doi.org/10.1371/journal.pone.0232671

Willis, J. (2021). *A framework for task-based learning*. Intrinsic Books Ltd.

Xu, Q., & Peng, H. (2017). Investigating mobile-assisted oral feedback in teaching Chinese as a second language. *Computer Assisted Language Learning*, *30*(3), 173-182. https://doi.org/10.1080/09588221.2017.1297836

Xue, S. (2020). A conceptual model for integrating affordances of mobile technologies into task-based language teaching. *Interactive Learning Environments*, 1-14. https://doi.org/10.1080/10494820.2019.1711132

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and online planning on fluency, complexity, and accuracy in L2 monologic oral production. *Applied Linguistics*, *24*(1), 1-27. https://doi.org/10.1093/applin/24.1.1

Appendix A. Example of Unpruned and Pruned Speech in Each Pruning Aspect

Pruning Aspects	Unpruned Speech	Pruned Speech
Repetition	The boy approached approached approached them again and fed them.	The boy approached ~~approached approached~~ them again and fed them.
Reformulation	When took, after taking photos, they left.	~~When took,~~ after taking photos, they left.
Replacement	One of the gentlemen took the food and gave it to them to birds.	One of the gentlemen took the food and gave it ~~to them~~ to birds.
Filled Pauses	Umm, well, umm, then the angel kissed the boy.	~~Umm, well, umm,~~ then the angel kissed the boy.

Note. All examples in the table were from students' oral discourses in the oral test.

APPENDIX B. SAMPLES OF SPEAKING TASKS

Speaking Topic	Complexity Level	Task Type	Samples of Speaking Tasks
Love	Simple	Topic Discussion	1) What is the ideal romantic relationship between men and women and how to build the ideal man-woman relationship? 2) State by providing more details, like describing stories happened around you.
Reading Experience	+Complex	Topic Discussion	1) Has your life ever been significantly changed by any particular book? Please share one of your reading experiences. 2) How has technology influenced people's way of reading? Please elaborate with details.
Grow Up and Grow Old	+Complex	Topic Discussion	1) What's your understanding of the sentence "Growing old is inevitable. Growing up is optional"? 2) Do your parents often tell you stories that happened in their times? What is the story that touches your heart and is rooted deeply in your memory? Please share. 3) As you grow up, your parents are growing older. Are there any conflicts between you and your parents due to different generation imprints? Elaborate with details.
Life Attitude/Philosophy	++Complex	Picture Narration	1) What are your attitudes towards fame and wealth? What is ideal life for you? 2) Two sets of pictures concerning attitudes toward fame and wealth are presented to students. They need to describe the pictures first, and then elaborate on their attitude/philosophy toward life with reasoning.
Environment	++Complex	Picture Narration	1) Out of his ignorance and lack of concern for the integrity of nature, man is tampering with nature by abusing chemicals, causing irrecoverable harm to the environment and people. Do you agree? State your views with reasonings and details about current environmental problems. 2) Four pictures about environmental destruction are presented to students. They are supposed to describe the pictures first and then elaborate more on environmental problems with reasoning.

APPENDIX C. REFLECTION FORM

Self-reflection
1. What do you think of the preview and pre-class oral practice? Can the provision of linguistic and background knowledge help you improve your speaking and how?
2. What do you think of your performance in the speaking tasks?
3. What do you think of the teacher's role in your class? Did the teacher effectively help you in your oral discussions or presentations? How?
4. Do you have any problems with your speaking performance? (If resolved, how?) And are there any places you think you need to further improve your speaking?
5. Other reflections:

Peer Feedback	
Comment Guide: 1. Speaking complexity, fluency, and accuracy, etc. 2. Places need to be improved.	
Group Member	Comments/Suggestions
Catherine	
…	

Appendix D: Representative Participant Quotes on Mobile Affordances

Ease of Use

- *"WeChat and Liulishuo are very user-friendly. The tools can be opened quickly, and they are compatible with various operating systems."* (Catherine)
- *"Wireless mobile devices are portable and flexible. Students can do oral practice at their convenience without the limitation of time and space."* (Teacher)

Synchronous and Asynchronous Communication

- *"WeChat can serve as a networking tool for synchronous voice chatting among students, unrestricted by physical proximity."* (Teacher)
- *"The voice messages on WeChat facilitated instant communication...We can communicate with each other anytime by sending or receiving voice messages."* (Devin)
- *"Speaking in a WeChat group is more focused, as we only have six students in one group."* (Richard)
- *"We can have a very thorough discussion in WeChat groups, and my classmates are very active in speaking."* (Ives)

Repeated Oral Practice

- *"The voice recording function on WeChat is well-suited for pronunciation practice, allowing students to repeatedly record their voices, read the speech-to-text transcription, and identify any mispronounced words."* (Teacher)
- *"While I was listening to my own voice, I realized that my pronunciation was not correct. Then I repeatedly recorded it."* (Verna)
- *"Shadowing was helpful because by repeating the words of English L1 speakers, I can speak more accurately and fluently."* (Devin)

Instant Feedback

- *"The ASR function on Liulishuo is instant. The system offers me evaluation results immediately after I finish my speaking practice."* (Ives)
- *"It generates a score for each sentence immediately after my speaking and pinpoints the places that need improvement."* (Verna)
- *"We can get both peer and teacher feedback instantly using WeChat. It helped me realize the problems in my speaking."* (Catherine)

Authenticity

- *"Being able to apply what I learned was useful because I was able to find out when and how to use expressions appropriately in real life."* (Ives)
- *"Liulishuo provides abundant authentic corpus. Students accumulate a large number of authentic oral expressions through practice."* (Teacher)

APPENDIX E: REPRESENTATIVE PARTICIPANT QUOTES ON EDUCATIONAL AFFORDANCES

Increased Speaking Opportunities Outside Class

- *"In the past, the speaking class was mostly based on the textbook, and few opportunities were provided for us to speak due to time and space restrictions. Now we have more chances to practice speaking, especially out of class, thanks to mobile applications."* (Catherine)
- *"Despite out-of-class assignments, students still used to come to class unprepared due to a lack of supervision and deficient collaborative learning. The new approach ensured they engaged in speaking practice beforehand."* (Teacher)
- *"Now, out of class, I always actively practice oral English on Liulishuo. It has triggered my learning interest and prepared me for the speaking task."* (Afra)

Enhanced Learner Autonomy

- *"I can learn now with more self-regulation by actively doing oral practice online without the teacher's supervision, which is a big improvement for me."* (Richard)
- *"I can make use of my free time more efficiently now. I always autonomously search for more materials on Liulishuo and actively seek help in WeChat groups when I have problems."* (Devin)
- *"MALL provides opportunities for independent oral practice. Students can flexibly choose the content of digital materials on mobile apps and practice according to their own pace and level."* (Teacher)

AUTHORS

Shijiao Jia obtained her PhD from SEGi University, Malaysia, and her study program is Doctor of Philosophy (Education). As an associate professor now attached with School of Foreign Studies, Hebei University, she has an EFL teaching experience in universities in China for more than 8 years. Her research interests include technology-assisted learning, mobile learning, task-based language teaching, collaborative learning, EFL teaching and learning, and applying big data tools in educational analysis, etc. She has now presented and published a number of papers on mobile-assisted learning and task-based language learning.

Zhaoxia Lu is an Associate Professor of English at the School of Foreign Languages at Cangzhou Jiaotong College in China. Her research interests are focused on translation studies, blended teaching and learning and EFL Teaching and learning. she has participated in two projects funded by the Hebei Provincial Department of Education in China, both of which have been successfully completed and accepted. She has also presided over one educational reform project in Hebei Province in China. She has published three translated works, three textbooks, and several articles.