EUROCALL · CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Examining the effectiveness of integrating corpus-based and AI approaches for English speaking practice

Hsueh Chu Chen[iD]

The Education University of Hong Kong, Hong Kong (hsuehchu@eduhk.hk)

Xiaona Zhou[iD]

The Education University of Hong Kong, Hong Kong (s1126341@s.eduhk.hk)

Jing Xuan Tian[iD]

The Education University of Hong Kong, Hong Kong (s1126315@s.eduhk.hk)

Abstract

This study developed and evaluated an online English speaking training approach that integrates corpora and artificial intelligence (AI) tools. The training integrated a self-developed spoken corpus, generative AI tools, and text-to-speech AI tools. Pre- and post-test results identified improvements in participants' speaking performances. Participants attempted to use more positive linguistic features (e.g. producing complex sentences more frequently) and avoid using negative linguistic features (e.g. reducing the number of vowel errors) after receiving the training. Participants showed positive attitudes towards this corpus-based and AI-integrated English oral ability learning approach and affirmed the importance of integrating both tools. The corpus helped raise participants' awareness of features that influence speaking performance and offered prompt engineering and feedback-checking functions, while the generative AI tools provided useful feedback and tailor-made sample responses. Additionally, text-to-speech AI tools offered learners with tailor-made native speaker samples for imitation and helped learners learn pausing. Results also revealed that this approach helped create an interactive oral ability learning environment, and the combination of corpora and AI tools provided more accurate feedback for each subskill of speaking.

Keywords: corpus-based speaking learning; AI in speaking learning; English speaking performance

1. Introduction

Speaking represents an active ability to construct significance, encompassing the transmission, reception, and organization of information (e.g. Brown & Abeywickrama, 2019). It is one of the crucial skills in language learning and communication (e.g. Luthfillah & Fauzia, 2023). In many English as a foreign language (EFL) regions such as China (Sun *et al.*, 2017) and South Korea (Ahn & Lee, 2016), the expanding circle in Kachru's (1992) Three Circles model where English is not an official or commonly spoken language, speaking is often considered the most challenging language skill. Speech fluency (Yan, Kim & Kim, 2021), accuracy (Nunan, 2015), pronunciation (Chau, 2021), and vocabulary (Afna, 2018) serve as critical components in evaluating learners' speaking proficiency. Take the International English Language Testing System (IELTS) speaking test, a

globally recognized standardized English test, as an example. This test evaluates four core competencies: fluency and coherence (FC), lexical resource (LR), grammatical range and accuracy (GA), and pronunciation (PN).

Traditional approaches to teaching and learning English speaking include dubbing (e.g. Jao, Yeh, Huang & Chen, 2024), learning rules (e.g. Talley & Tu, 2014), and presentation practice production (e.g. Criado, 2013) primarily targeting face-to-face learning environments. However, the post-pandemic era has witnessed a significant shift towards technology-mediated speaking instruction, with digital platforms (e.g. videoconferencing tools; Tran, Hoang, Gillespie, Yen & Phung, 2024) becoming a mainstream pedagogical medium (e.g. Cahyono, Fauziah, Santoso & Wulandari, 2024). Given the advancements in information and communication technology, the extensive integration of digital resources into educational practices represents a transformative shift in education (e.g. Shin & Yunus, 2021). As suggested by previous studies (e.g. Arkorful & Abaidoo, 2015), the use of electronic technologies is considered as one of the most effective approaches for instructing and acquiring speaking abilities, particularly due to its adaptability in accommodating constraints related to time and location. This study aims to develop and evaluate the effectiveness of a self-directed online oral ability learning approach to integrate corpus technology with multiple artificial intelligence (AI) tools to create an interactive learning environment for EFL learners.

1.1. Technologies in speaking teaching and learning

Previous studies integrated technologies like podcasting (e.g. Yeh, Chang, Chen & Heng, 2021), blogging (e.g. Parveen, 2016), videoconferencing (e.g. Bahadorfar & Omidvar, 2014), videos (Kohen-Vacs, Milrad, Ronen & Jansen, 2016), and corpora (e.g. Chen & Tian, 2022) in speaking teaching and identified their effectiveness.

Gablasova and Bottini (2022) encouraged the use of learner corpora in teaching. Corpora findings can be used to create resources like dictionaries, grammar guides, and language curricula (e.g. Feak, Reinhart & Rohlck, 2018). Teachers and students can also get access to corpora data directly (e.g. Gablasova & Bottini, 2022); furthermore, learners' awareness could be raised by working with the corpora data (e.g. Chen & Tian, 2022). Gablasova and Bottini (2022) also suggested that tailoring corpora data to students' language backgrounds could boost learning progress and motivation. Some of the technologies used are one-way, such as videos or podcasts that provide learners with a channel to access instructions and productions by native speakers or high-level learners. Videoconferencing tools offer a platform for learners to communicate with teachers or other learners; however, learners have limited access to feedback or communication opportunities, as most of their interactions and feedback occur during class, either from instructors or peers, with little to no engagement beyond in-class discussions. Studies integrating AI in speaking teaching and learning could help resolve this issue (e.g. Huang & Zou, 2024).

1.2. AI for speaking teaching and learning

The use of AI tools (e.g. ChatGPT) has been encouraged in language teaching and learning (e.g. Kostka & Toncelli, 2023; Mompean, 2024). AI-powered tools allow students to engage in realistic conversations with virtual native speakers and receive feedback (e.g. Zou, Guan, Shao & Chen, 2023). Fathi, Rahimi and Derakhshan (2024) found AI speaking tasks outperformed traditional methods in boosting EFL learners' oral skills and communication confidence.

Zou *et al.* (2023) examined the impact of AI-powered apps on Chinese EFL learners' English speaking. Learners viewed interactive activities positively, reporting enhancements in various speaking aspects, including oral fluency, grammatical range and accuracy, pronunciation, oral rhythm, idea organization skills, reading aloud skills, and presentation skills. However, the study

relied on self-reports rather than measurable speech analytics, hindering assessment of specific skill improvements. This study aims to fill this gap by analyzing learners' speech output across four key speaking subskills (GA, LR, PN, and FC).

Abdulhussein Dakhil, Karimi, Abbas Ubeid Al-Jashami and Ghabanchi (2025) identified that the AI tool ELSA improved Iraqi EFL learners' speaking skills (grammar, vocabulary, intonation, fluency) and communication willingness, though not pronunciation. However, feedback generated by AI tools may occasionally be inaccurate, which could negatively impact learners' speaking development (Mompean, 2024). This study aims to address the limitations of previous research by integrating corpora with AI tools. This method establishes an interactive oral ability learning framework while ensuring feedback reliability, leveraging both corpus data and AI feedback to optimize learning outcomes.

Challenges have been identified when students use AI tools in the educational context. Prompt engineering is an essential component, which could help boost effective interaction skills with AI tools (Walter, 2024). The development of critical digital literacy (CDL) assumes heightened importance (Satar, Hauck & Bilki, 2023). The formulation of precise prompts is recognized as fundamental for obtaining optimal outputs from generative language models (Huang, 2023). The current study will use corpus data to enhance learners' ability to develop prompts.

1.3. Speaking instruction on learning management system (LMS) platforms

The utilization of web-based learning tools within educational settings has become a prevalent approach due to its adaptability, convenience, self-paced nature, and cost-effectiveness. A majority of learners prefer web-based learning over traditional classroom methods (e.g. Mendis & Dharmawan, 2019). The effectiveness of web-based learning (e.g. LMS), however, depends on the level of interaction involved (e.g. Woo & Reeves, 2008).

Mendis and Dharmawan (2019) investigated an LMS system, Canvas, with high school learners' English oral ability learning. Successful interactions occurred on Canvas LMS, so there is a compelling need to conduct English speaking lessons on this LMS. Suartama and Dewa (2014) proposed that activities could be designed on the Moodle LMS, including instruction videos, discussion forums, chat, and quizzes. However, limited previous studies have provided suggestions on how to make use of the functions (e.g. discussion forum and quiz function) on the LMS platforms to develop an interactive course and observe students' learning process. This study will address this gap by introducing a self-directed corpus-based and AI-integrated English speaking course on Canvas and evaluating its effectiveness.

2. Research gaps and research questions (RQs)

Most approaches to learning oral communication provide learners with restricted feedback opportunities confined mostly to classroom settings. AI tools show promise in personalized and real-time feedback. However, prior studies on AI-powered tools for pronunciation training relied on self-reported data rather than empirical analysis of speech outputs. In addition, limited studies exist to synergize corpus data with AI's adaptive feedback for targeted English speaking improvement. Based on these research gaps, two RQs were drafted:

1. To what extent does the online corpus-based and AI-integrated language learning approach facilitate and enhance English speaking development among learners?
2. What are learners' attitudes and evaluations of the online corpus-based and AI-integrated language learning approach?

Table 1. The functions of corpus and AI tools in our approach

Tool type	Function
Corpus	- Notice common features in learners' speech - Compare features across different proficiency levels - Develop critical digital literacy and enhance prompts development ability

AI	Generative AI
	   - Give prompts to generate ideas for speaking tasks
	   - Give prompts to receive personalized feedback on grammatical range and accuracy, lexical resource, pronunciation, and fluency and coherence
	   - Give verbal prompts for interactive speaking practice
	
	Text-to-speech AI
	   - Generate customized sample recordings for imitation specifically on pronunciation and fluency and coherence

3. Methods

3.1. Participants

Sixty-two participants who were university students in Hong Kong (HK) with Chinese as their first language (L1) were enrolled in the online training. Among them, 20 participants engaged in a comprehensive study that encompassed a pre-test, a six-session training, a self-study session, a post-test, and a semi-structured interview.

3.2. Tests

During both tests, each participant received an IELTS speaking part 2 topic. We prepared six different topics for the two tests (three for each). These six topics were not included in the training sessions. A topic was randomly assigned to each participant. One minute was given to prepare the topic before participants delivered a two-minute speech on the assigned topic.

3.3. Training details

In the first session, we included a brief introduction of the spoken corpus, AI tools used in our training, and the corpus-based and AI-integrated oral ability learning procedure. In sessions 2 to 5, participants were instructed on how to use corpora and AI tools to build up ideas and practise the four subskills (GA, LR, PN, and FC). The functions of corpora and AI tools in our approach are shown in Table 1. Corpus-based practice is the core and could raise learners' awareness of their English speaking proficiency, enhance prompts development ability, cross-check feedback accuracy, and perform self-reflection and self-assessment by comparing their own speech with corpus data. Learners use generative AI tools to generate ideas of conversation topics, receive personalized feedback, and communicate. Text-to-speech AI tools provide learners with tailor-made native speaker samples for imitation. The last session included an AI package for English oral ability learning and hands-on activities on how to use the AI tools in this package.

    All training sessions took place on Canvas. Each session included an instructional video along with a quiz to evaluate participants' comprehension of the video content and an online discussion forum, giving participants the opportunity to communicate with one another, share their previous experiences in learning English oral communication, share the prompts and feedback they received on each subskill by using generative AI (e.g. Poe and Copilot), and express their opinions on the AI package. In the self-study session, all participants prepared a 250-word script based on an IELTS speaking part 2 topic, showcased the prompts used and responses received when interacting with AI tools, showcased how pauses were added to the script by using

the text-to-speech AI tool, prepared a two-minute recording with the revised script, and shared their thoughts about using corpora and AI tools to learn spoken English.

3.4. The spoken corpus used in the training

The English Speech Corpus with Different Proficiency Levels (Chen, 2022; hereafter the Spoken Corpus; http://corpus.eduhk.hk/english_speech_corpus/) consists of 91 sets of spoken data (78 sets of spontaneous speeches and 13 sets of classroom presentations). Among the 78 sets of spontaneous speeches, 48 sets were collected from learners in Chinese Mainland and HK, while 30 sets were extracted from publicly accessible IELTS speaking test recordings. Each spontaneous speech was annotated comprehensively according to four criteria adopted from the IELTS speaking test: GA, LR, PN, and FC. Key indicators annotated in the Spoken Corpus appear in the Appendix (see supplementary materials).

The main aim of using the Spoken Corpus data is to raise learners' linguistic awareness of features that influence their speaking performance. Learners followed the following procedure: identifying the common features produced by speakers with lower proficiency level (e.g. band 5.5 or lower), performing self-reflection and self-checking on whether learners produce these features in their speech, identifying the common features produced by speakers with higher proficiency level (e.g. band 7.0 or higher), and comparing the features produced by the lower proficiency level with the higher level to identify the features that influence speaking performance.

3.5. Data analysis

For the pre- and post-tests, 20 features were analyzed by listener judgement (five for each subskill). For GA, we counted the number of sentences, the frequency of simple sentences, complex sentences, complex structures, and grammatical errors that the participants produced. For LR, we counted the frequency of inappropriate word choice, appropriate and inappropriate idioms, and appropriate and inappropriate collocations in participants' speech. In terms of PN, we counted the frequency of consonant errors, vowel errors, schwa insertion at the end of a word, mispronunciation, and C-V linking. For FC, we calculated speech length, speech rate, the frequency of silent pausing, filled pausing, and discourse markers and connectives.

Participants' prompts submitted to the discussion forum during the four after-training activities were further categorized based on the degree to which participants followed the sample prompt provided during training.

The interviews were conducted based on the themes related to how participants use the corpus-based and AI-integrated approach, as well as the evaluation of this approach and the LMS.

4. Results

In order to answer the first RQ, participants' pre- and post-test performances on the four subskills (GA, LR, PN, and FC), participants' self-report on how they used corpus and AI technologies to learn each subskill, and the prompts that they used for feedback were analyzed.

4.1. Pre- and post-test results

Participants' performances in the pre- and post-tests (Tables 2 and 3) were compared. For the positive features, the frequency of participants' production on these features in the post-test is higher than those in the pre-test, which indicates that participants attempted to use more positive features in their speech after training. Paired samples t-test results revealed that the number of sentences, the frequency of complex sentences, appropriate collocations, and C-V linking were significantly higher in the post-test compared to those in the pre-test (with p values equal to .030*,

Table 2. Frequency of positive features by participants in both tests

Subskills	Positive features	Pre-test		Post-test				
		M	SD	M	SD	t	p	Cohen's d
GA	Number of sentences	10.05	3.62	12.95	3.28	−2.350	**.030***	−0.504
	Simple sentences (frequency)	5.10	3.14	6.40	2.54	−1.542	.140	−0.331
	Complex sentences (frequency)	4.10	1.65	5.85	3.77	−2.150	**.045***	−0.461
	Complex structures (frequency)	0.25	0.55	0.50	0.83	−1.314	.204	−0.282
LR	Appropriate idioms (frequency)	0.00	0.00	0.70	2.00	−1.563	.135	−0.336
	Appropriate collocations (frequency)	13.20	4.24	21.10	6.80	−4.602	**<.001***	−0.988
PN	C-V linking (frequency)	5.50	2.78	11.50	7.04	−3.988	**<.001***	−0.856
FC	Speech length (second)	106.79	27.37	122.60	30.89	−2.032	.056	−0.436
	Speech rate (words per minute)	105.00	12.77	115.93	19.86	−2.618	**.017***	−0.562
	Discourse markers and connectives (frequency)	18.25	6.36	20.20	5.37	−1.185	.251	−0.254

Note. GA = grammatical accuracy; LR = lexical resource; PN = pronunciation; FC = fluency and coherence.

Table 3. Frequency of negative features by participants in both tests

Subskills	Negative features	Pre-test		Post-test				
		M	SD	M	SD	t	p	Cohen's d
GA	Grammatical errors (frequency)	7.55	5.38	3.95	2.84	3.174	**.005****	0.681
LR	Inappropriate word choice (frequency)	1.65	1.50	1.05	1.28	1.318	.203	0.283
	Inappropriate idioms (frequency)	0.00	0.00	0.45	2.01	−1.000	.330	−0.215
	Inappropriate collocations (frequency)	1.55	2.01	0.75	0.79	1.598	.126	0.343
PN	Consonant errors (frequency)	15.35	6.15	11.50	6.31	2.329	**.031***	0.500
	Vowel errors (frequency)	7.40	4.28	4.90	3.16	2.828	**.011***	0.607
	Schwa insertion (frequency)	1.45	2.33	0.70	1.59	1.994	.061	0.428
	Mispronunciation (frequency)	0.80	1.67	1.55	2.64	−1.598	.127	−0.343
FC	Silent pausing (frequency)	41.40	11.93	36.10	16.03	1.484	.154	0.319
	Filled pausing (frequency)	7.05	6.10	3.20	4.63	2.952	**.008****	0.634

Note. GA = grammatical accuracy; LR = lexical resource; PN = pronunciation; FC = fluency and coherence.

.045*, <.001***, and <.001*** respectively). Participants' mean speech rate also increased significantly from 105.00 words per minute in the pre-test to 115.93 words per minute in the post-test ($p = .017*$).

The frequency of the negative features by participants decreased from pre-test to post-test except for the frequency of inappropriate idioms and mispronunciation. Participants' speech performance improved after training by avoiding using negative features in their speech. Paired samples t-test results showed that participants made fewer grammatical errors (pre-test: 7.55, post-test: 3.95, $p = .005**$), consonant errors (pre-test: 15.35, post-test: 11.50, $p = .031*$), and

vowel errors (pre-test: 7.40, post-test: 4.90, $p = .011^{\star}$) after training. Participants also produced fewer filled pauses in the post-test ($M = 3.20$) than in the pre-test ($M = 7.05$) with $p = .008^{\star\star}$.

4.2. Experience of learning the four subskills

Participants described how they applied corpora and AI tools for their speaking development, as well as how they integrated these two tools to enhance each speaking subskill. In the following sections, we detail how participants improved GA, LR, PN, and FC using corpora and AI tools.

4.2.1. GA learning

Fourteen participants found AI tools beneficial for identifying and correcting grammatical errors. Eight further stated that AI tools were useful in enhancing grammatical range, helping them construct more varied and complex structures:

> *Participant (P)16: AI corrects grammatical errors in my sentences and highlights areas where I've made mistakes, such as issues with verb tenses.*

> *P9: I use AI to help me adjust my grammar, such as increasing complexity.*

Only two participants found AI useful for simplifying overly complex sentences, making their speech more conversational and suitable for natural spoken English:

> *P11: AI can help me make my sentences more in line with everyday spoken English.*

Five participants confirmed that the corpus served a supporting role by offering authentic examples of grammar in context, especially from higher-band learners:

> *P1: I learn complex sentence structures like adverbial and attributive clauses by observing high-scoring answers in the corpus.*

4.2.2. LR learning

Twelve participants emphasized that AI tools helped them correct lexical errors (e.g. inappropriate collocations) and improve word choice by suggesting more natural or topic-specific expressions:

> *P12: AI tools like Poe help improve vocabulary by replacing unnatural phrases with more natural or appropriate ones.*

> *P3: AI corrects inappropriate collocations.*

Ten participants pointed out that the corpus was especially valued for exposing learners to vocabulary used in different band-level responses, helping them discover less common words, collocations, and contextual usage:

> *P1: By using a corpus, we can observe how high-band learners express themselves and then take note of their vocabulary for learning purposes.*

Importantly, two participants emphasized combining both tools using AI to analyze their word usage and the corpus to refine it with authentic examples, enabling targeted vocabulary development:

P15: I use an AI tool to analyze my spoken responses and get feedback on my lexical resource. Then, I consult corpus examples to make targeted improvements accordingly.

4.2.3. PN learning

Eight participants highlighted that the AI tools were helpful in identifying pronunciation issues, such as misplaced stress and mispronunciations:

P11: I read a passage to the AI, which identifies problems in my pronunciation, such as incorrect word stress or mispronunciations.

In addition, nine participants stated frequent use of AI-generated audio, especially from text-to-speech AI tools like Murf, for modelling native pronunciation. The ability to modify accent and speed for targeted practice was a popular function. Participants also emphasized that slowing down and imitating AI-generated speech helped improve their pronunciation step by step:

P6: In terms of pronunciation, I find Murf more helpful. For example, it allows you to set nationality and accent, so I can listen to the generated audio and imitate it.

The corpus was also beneficial for pronunciation practice, particularly because it offered authentic recordings of high-band learners (reported by six participants). Participants favoured corpus-based audio over AI-generated audio due to its greater naturalness:

P14: I think corpus has an advantage when it comes to pronunciation. For example, I can listen to recordings of high-band speakers and imitate their pronunciation.

Three participants described the combination of corpora and AI tools facilitated targeted improvement in pronunciation by receiving feedback from AI on pronunciation and using corpus to make targeted improvements:

P15: I input my spoken content into an AI tool, and it provides feedback on my pronunciation. Then I use corpus to make targeted improvements.

4.2.4. FC learning

Twelve participants reported that AI tools helped add discourse markers and connectives, thus improving logical flow:

P13: AI tools add discourse markers and connectives to improve sentence flow and coherence.

In terms of fluency, six participants shared that text-to-speech AI tools helped reduce fillers and unnecessary pauses and improved their awareness of speech pacing. Participants also noted that they could compare different types of pauses and adjust their speech accordingly:

P3: AI helps me remove redundant content and improves flow through better pauses and connectives.

P4: I think it's helpful, since we add different types of pauses. AI reads them differently, so we can compare how a weak pause compared to a stronger one and determine which is more effective.

The corpus further supported FC by providing authentic examples from spoken performances in IELTS practice tests by learners of different proficiency levels. Five participants observed from corpus data that high-band learners minimized excessive pauses and used clearer logical flow:

P14. In terms of fluency, I find the corpus especially helpful. I can listen to high-band speakers and try to imitate their fluency.

Another eight participants used examples in the corpus to identify common coherence issues in lower-band responses and to improve their own speech:

P1: I observed the differences between high- and low-band responses in the corpus to learn how to express myself more coherently.

Three participants emphasized that using both tools helped them diagnose and address fluency issues more effectively, creating a complementary and iterative learning process:

P15: I input my speech into AI for feedback on coherence, then use the corpus to make targeted improvements.

4.3. Prompt analysis

Participants reported the prompts that they gave to generative AI tools to request sample answers or feedback on GA, LR, PN, and FC, as well as the feedback that they received from the generative AI tools. A sample prompt was shared with them for each skill. For example, the sample prompt for PN is, "*Can you help me identify my pronunciation mistakes on English vowels in the following speech?*" Across the four sessions, 110 prompts were considered valid for analysis.

The analysis of prompts (Table 4) revealed notable differences in how participants utilized the sample prompt structures provided during training. Participants demonstrated more customization in idea generation. The most frequent type of prompt involved customizing the sample to target a higher IELTS band score (41.38%). In contrast, prompts that strictly followed the sample were relatively rare (10.34%). Other recurring types of modification included simplified prompts (13.79%), prompts without a specified band score (13.79%), requests for a time limit (10.34%), and the inclusion of additional contextual details (10.34%); 62.06% of the occurrences involved some form of prompt customization. One participant's prompt is as follows:

P10: Please act as an IELTS candidate with a good command of English and answer the questions in part 2 of the Speaking test. Based on the background information given by me, give a reference answer (100s) that can get 7 points [band 7] according to the speaking test standards. The sentences and words used in the answer are colloquial and the content is easy to understand. And please try to reuse your used sentences, words, time, names, in order to improve the reusability, easy for me to remember.

For GA and LR, around 90.00% of prompts adhered closely to the sample format. Participants predominantly targeted lexical range (40.48%) and grammatical accuracy (30.95%), with fewer prompts addressing grammatical range (9.52%) or lexical accuracy (7.14%). Only five participants requested additional refinement for naturalness or alignment with a target band level. The heavy reliance on structured prompts in this session indicates a strong preference for explicit scaffolding. The sample prompts included (1) "*Can you help me to correct my grammar mistakes for my following speech?*" (for GA accuracy); and (2) "*Can you help me to rewrite the following speech by using five paraphrases, five less common words, and five idioms?*" (for LR range).

Table 4. Categorization of valid prompts

Categorization	Occurrence	Percentage
Idea generation		
Strictly following sample prompt	3	10.34%
Following sample prompt, but targeting higher-band score (8.5: 1; 8.0: 8; 7.5: 2; 7.0 and above: 1)	12	41.38%
Following sample prompt, without specifying band score	4	13.79%
Simplified prompt	4	13.79%
Additional request for setting time limit	3	10.34%
Others (e.g. including other details in the prompt)	3	10.34%
GA & LR		
Grammatical accuracy	13	30.95%
Grammatical range	4	9.52%
Lexical accuracy	3	7.14%
Lexical range	17	40.48%
Additional request for target band level	3	7.14%
Additional request for refinement of speech (e.g. reduce the use of unusual words and idioms)	2	4.76%
PN		
Consonant	17	24.29%
Vowel	13	18.57%
Word stress	14	20.00%
Intonation	13	18.57%
C-V linking	2	2.86%
Additional request for fluency	9	12.86%
Others (e.g. request for the International Phonetic Alphabet, PN in general)	2	2.86%
FC		
Discourse markers and connectives	15	40.54%
Filled pause	6	16.22%
Silent pause	4	10.81%
Fluency and coherence in general	12	32.43%

Note. GA = grammatical accuracy; LR = lexical resource; PN = pronunciation; FC = fluency and coherence.

PN and FC learning showed a more balanced distribution between sample-based and customized prompts. PN demonstrated a more distributed pattern of focus across various pronunciation features. Participants most frequently targeted consonants (24.29%) and vowels (18.57%), word stress (20.00%), and intonation (18.57%). Interestingly, 12.86% of prompts included explicit requests related to fluency. Participants tended to associate pronunciation with fluency rather than viewing it in isolation. FC also exhibited a balanced distribution of prompt types, with 32.43% addressing FC in general terms and 40.54% explicitly focusing on discourse markers and connectives. A smaller proportion of prompts incorporated filled pauses (16.22%) and silent pauses (10.81%).

4.4. Evaluation of corpus-based and AI-integrated oral ability training

Participants' interview data were used to answer RQ2. Particular attention was given to how the approach contributed to creating an interactive and engaging learning environment and what participants perceived as its major strengths in facilitating self-directed speaking practice.

Analysis of interviews identified AI's dialogic real-time responses as central to perceived interactivity (Table 5). Thirteen participants described how generative AI tools enabled dialogue-like interaction, which allowed them to refine their responses through multiple rounds of feedback, providing examples and suggestions. Three participants also reported that feedback from AI tools created an interactive environment, which enhanced participants' motivation and activated engagement. However, five participants found the interaction less engaging, describing it as lacking authentic conversational dynamics.

Personalized and targeted feedback provided by AI tools emerged as a prominent strength (Table 5), with 12 participants reporting using the AI tools to address specific weaknesses in GA, LR, PN, or FC. Participants also highlighted the convenience and accessibility of the tools (P16), as well as their ability to tailor AI-generated audio to individual preferences, such as accent and speed.

Regarding the corpus, participants considered it particularly useful for offering localized, exam-relevant examples, especially those from Chinese learners, which they found relatable and accessible. Additionally, six participants reported that the corpus helped them identify and avoid common errors made by lower-band learners, particularly in pronunciation and coherence.

Furthermore, five participants highlighted that the combined use of corpora and AI tools contributed to a systematized and efficient learning process. One participant described it as a "closed loop" (P5), in which feedback, examples, and revision were closely linked. Although one participant (P15) pointed out that AI feedback may not always be fully accurate, they still found it immediate and useful, emphasizing its overall practical value.

4.5. Evaluation of the Canvas LMS

All 20 participants highlighted the flexibility of learning on Canvas as a major benefit. The ability to access course materials anytime and from any location was frequently mentioned. Seven participants emphasized the benefits of self-paced learning, including features like video replay and speed control. However, four participants reported difficulties maintaining concentration in the online environment, considering a lack of real-time engagement and external distractions as key issues. In addition, 11 participants expressed concerns about the limited interaction and feedback available in the online mode compared to face-to-face classes.

For the effectiveness of the discussion forum on Canvas, 16 participants reported that reading peers' posts helped them reflect on their own responses. Additionally, three of them mentioned that peers' posts also inspired how they wrote prompts for the interactive AI tools. Three participants suggested that requiring students to respond to peers' posts could increase interaction and engagement. Regarding the effectiveness of quizzes, 14 participants reported that quizzes were effective in consolidating knowledge and reinforcing key concepts. Seven participants noted that quizzes helped them identify gaps in their understanding and reflect on their learning. Examples of participants' reports also have been listed in Table 6.

5. Discussion

5.1. Effectiveness of corpus-based and AI-integrated training on speaking performance

Speech task results in the pre- and post-tests showed positive effects of the online corpus-based and AI-integrated speaking course on learners' English oral ability. This finding is consistent with previous studies (e.g. Fathi *et al.*, 2024; Gablasova & Bottini, 2022; Mendis & Dharmawan, 2019;

Table 5. Participants' evaluation of the corpus-based and AI-integrated oral ability training

Theme	No. of participants (out of 20)	Participants' reports
Interactive and engaging learning environment		
Real-time feedback, examples, and suggestions from generative AI tools enable back-and-forth interaction	13	P18: AI tools create an interactive environment by giving real-time feedback based on my instructions or audio input.
AI enhances motivation	2	P18: The AI helps create an interactive environment and provides stimuli or interest to practice speaking.
AI activates engagement	1	P9: It really activates my engagement because I found that the AI's responses are closely related to how I ask the questions. So, I have to really think.
There is a lack of natural or realistic interaction	5	P1: I think there's still a bit of a lack of interactivity. Without real human conversation, communicating with AI feels a little stiff.
Strengths of the corpus-based and AI-integrated approach		
AI tools provide targeted and personalized feedback	12	P1: The advantage is that each AI tool offers targeted feedback. For example, it can give you ideas, help you generate text, and provide audio for practice. It's highly focused. You can choose specific tools based on your weak areas to improve more effectively.
Text-to-speech AI tools offer customizable audio samples	2	P4: When it comes to AI, one of its strengths is the ability to customize it based on your own needs. You can choose from different nationalities for the voice, adjust the speaking speed, or even change the persona. I think that's one of its biggest advantages.
AI tools are convenient	1	P16: It's pretty convenient. Yeah, because AI is developing so quickly now, and everyone has access to it – on their phones, on their computers – so it's just really convenient to practice speaking in a conversational way.
Corpus provides localized, exam-relevant examples	8	P4: The corpus [contains speech responses] from various speakers, and [some] speakers are not foreigners but students from Hong Kong and mainland China. I think this makes it more relevant to our English pronunciation and accents.
Corpus helps identify and avoid common errors	6	P7: Helps identify the characteristics of candidates at different band levels (e.g. Band 5, 6, 7, etc.) and relate them to personal performance. Provides insight into common mistakes and characteristics of certain band levels, which helps avoid similar errors.
This approach provides instructional materials and sample prompts	1	P2: From the start, it gives you a reference model. Based on that, you can give AI specific instructions and get tailored modifications.
Systematized and efficient learning process	5	P5: This method feels like a closed loop. Other apps only do part of the job, but this combines everything into one system. P18: The combination of corpus and AI tools also makes the feedback more accurate and precise compared to just using apps. Corpus serves as a supplement to AI tools, providing resourceful materials, and making the AI-generated feedback more natural and accurate.
Inaccurate but helpful	1	P15: One advantage of AI is that it provides instant feedback. While it may not be 100% accurate, it's still really helpful.

Table 6. Participants' evaluation on the use of Canvas learning management system

Theme	No. of participants (out of 20)	Participants' reports
Overall online learning experience		
Flexibility and convenience	20	P1: Canvas offers flexible timing and allows repeated video watching.
Self-paced learning (replay, speed control)	7	P3: Canvas videos can be replayed anytime. P12: Pre-recorded videos allow speed adjustments like 1.5x or 2x.
Focus reduction in an online setting	4	P18: I find it harder to concentrate when studying online due to distractions at home, such as messages on my phone or interruptions from family members.
Lack of interaction or feedback	11	P5: Canvas lacks timely feedback, unlike face-to-face classes where I can ask questions immediately.
Learning experience through discussion forum		
Knowledge gained from peers' responses (*Prompt creation)	16	P1: By observing other students' answers, I can compare mine and learn better ways to express [myself]. P19*: If I'm not sure how to prompt, I learn by checking others' prompts. That's also a kind of interaction.
Suggestions on setting response mandatory to enhance interaction	3	P13: Adding a requirement to give peer feedback might make it more interactive.
Learning experience through recap quiz		
Assistance in reinforcing learning	14	P4: Quizzes help consolidate learning because they are based on video content and provide auto-grading.
Understanding assessment	7	P8: I check whether I've understood the session through whether I answered correctly.

Zou *et al.*, 2023), indicating the effectiveness of corpus and AI technologies, as well as LMS platforms, in speaking practice. For participants' performance on the features (both positive and negative) of the four subskills, the findings are new. Participants in this study attempted to use more positive features in their post-test compared to their production in the pre-test. Participants' performances of five among the 10 positive features (number of sentences, complex sentences, appropriate collocations, C-V linking, and speech rate) improved significantly from pre-test to post-test. For the other five positive features (simple sentences, complex structures, appropriate idioms, speech length, and discourse markers and connectives), participants attempted to produce these features more frequently in the post-test, but the production frequency improvement was insignificant. The possible reason could be attributed to the relatively short overall training duration. The entire training programme in this study had only six sessions. Nevertheless, initial evidence of improvement was observed. Participants had awareness of the positive features that may influence oral proficiency, but they may require additional time for practice.

For the negative features, participants tried to avoid producing eight out of the 10 negative features in the post-test. Their performance of four features among the eight showed a significant decrease for grammatical errors, consonant errors, vowel errors, and filled pausing. No significant decrease in the performances of inappropriate word choice, inappropriate collocations, schwa insertion, and silent pausing could also be attributed to the relatively short duration of the training. The frequency of two negative features (inappropriate idioms and mispronunciation) by the participants in the post-test increased slightly. After the training, participants had an

awareness for using more idioms, producing longer speech, and using more complex words. In the post-test, they attempted to use more idioms in their speech, but their proficiency levels were not so high that they were familiar with the usage of some idioms. They made errors when using these idioms. The possible reason for producing more mispronounced words could be participants' production of longer and more complicated utterances in the post-test. Inevitably, they made pronunciation mistakes with some complex words.

Participants' PN improvement identified in this study is not in line with Abdulhussein Dakhil *et al*. (2025), in which learners' pronunciation performance remained unchanged. The current study enhanced participants' pronunciation performance following three steps (1. enhancing learners' perception, 2. providing feedback, and 3. using text-to-speech AI tools to improve learners' production). Participants first used learner corpus data to raise their awareness of the possible PN features that they may have and the PN features that may influence speaking performance. Participants then sought feedback on PN from generative AI tools (e.g. ChatGPT), followed by generating and imitating native speakers' samples by text-to-speech AI tools. This finding supported the pedagogical suggestions proposed by Chen and Tian (2025), indicating that using only one type of tool in language learning is not sufficient. Tools with different functions need to be combined, complement each other, to develop their strengths and compensate for their weaknesses.

5.2. Attitudes and evaluations of the corpus-based and AI-integrated oral ability training on Canvas

The current study confirmed that using corpus and AI technologies in pronunciation and fluency training helped create an interactive environment. This finding supported previous studies encouraging the application of AI to design interactive speaking activities and tasks (e.g. Fathi *et al*., 2024; Zou *et al*., 2023). Previous studies highlighted that the use of large language model AI tools helped create a highly interactive and engaging learning environment (e.g. Shafiee Rad & Roohani, 2024) by providing instant feedback and offering opportunities for personalized and interactive practice, which fosters interaction, engagement, and language development. In the current study, giving prompts to AI tools and receiving feedback from AI tools could help build up an interactive learning environment.

The effectiveness of corpus use in learning oral communication has also been confirmed in the current study by offering authentic models and examples being consistent with previous studies (e.g. Gablasova & Bottini, 2022). Participants also reported on the prompts development function when combining corpus use and AI technologies in speaking, which helped enhance CDL. Corpus data is considered a valuable resource for prompts development in a variety of domains, such as enhancing LR range. Using the feature list summarized from the corpus data supported learners to formulate effective prompts for the AI and receive more accurate and detailed feedback. This is the beauty of the combination of the two types of technologies. Participants' reports on the effectiveness of combining the two tools in speaking learning also supported the advice of integrating tools with different functions in teaching and learning by Chen and Tian (2025).

The corpus-based and AI-integrated training on Canvas LMS is flexible, convenient, and provides learners with self-paced learning, which is in line with Qaddumi and Smith (2024). We utilized the quiz function on Canvas LMS to verify participants' understanding of the key knowledge and help them reinforce learning in each session.

Interaction is the main factor in determining the success of a course on LMS (Woo & Reeves, 2008). Participants in this study confirmed interactions of the discussion forum on Canvas, which was the main interaction among peers. Participants could learn from others by browsing others' posts and confirmed the interactions on Canvas, which is consistent with Mendis and Dharmawan (2019). However, participants also reported a lack of interaction or feedback and proposed that teachers should make participation mandatory to enhance more peer interactions during self-directed online learning.
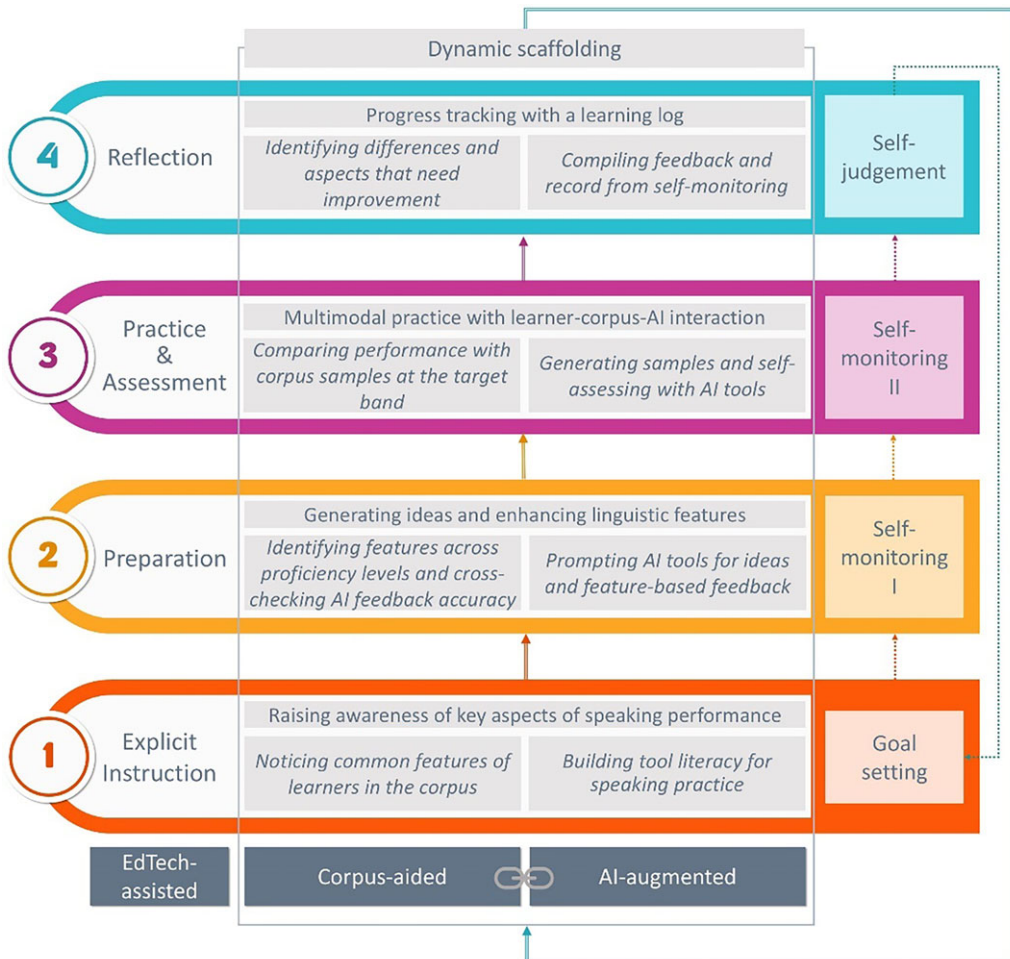
Figure 1. Corpus-based and AI-integrated oral ability learning framework.

5.3. Pedagogical implication

The findings of this study offer significant pedagogical implications based on the development of a corpus-based and AI-integrated oral ability learning framework (as shown in Figure 1). Teachers could use the four-step framework to design training lessons and speaking tasks. Learners first analyze authentic samples from spoken corpora at various proficiency levels. This comparative analysis not only raises awareness of effective speaking features but also highlights common difficulties encountered by learners, enhancing learners' prompts development ability. Then, learners are encouraged to use generative AI tools to generate topic ideas and receive personalized feedback, focusing on the four key subskills (GA, LR, PN, and FC). Learners are also encouraged to use corpus data for cross-checking AI feedback accuracy. Text-to-speech AI tools provide audio models for learners to imitate, offering perceptual input that benefits PN, FC, and even GA and LR acquisition through multimodal practice. Learners then compare their own performance with corpus samples at their target band level as established benchmarks for self-monitoring. The framework concludes with compiling feedback and records from AI tools, comparing their

performance against the corpus feature list, identifying the aspects that need further improvement, and reflecting on their progress.

6. Conclusion

This study examined the effectiveness of the online corpus-based and AI-integrated approaches in facilitating and enhancing English speaking development. Quantifiable improvements were observed in several key linguistic features. After receiving the training, participants attempted to use more positive features, such as complex sentences, and avoid using negative features, such as grammatical errors. Learners expressed positive attitudes towards the online corpus-based and AI-integrated training, considering it as creating a more interactive and engaging speaking environment. The combination of corpora and AI tools was widely regarded as an effective approach for self-directed speaking practice. Learners highlighted the flexibility afforded by the Canvas LMS, along with the benefits of discussion forums and recap quizzes on this platform.

This study contributes to the field of intelligent computer-assisted language learning by highlighting the pedagogical potential of combining corpora with AI tools to improve learners' oral proficiency. One limitation of this study is the relatively short duration of the training, which may have hindered participants' speaking knowledge and tool consolidation. Future research should investigate the long-term effects of this approach. Another limitation is a lack of control group participants. For future studies, a control group using traditional speaking learning methods should be recruited.

References

Abdulhussein Dakhil, T., Karimi, F., Abbas Ubeid Al-Jashami, R. & Ghabanchi, Z. G. (2025) The effect of artificial intelligence (AI)-mediated speaking assessment on speaking performance and willingness to communicate of Iraqi EFL learners. *International Journal of Language Testing*, 15(2): 1–18. https://doi.org/10.22034/IJLT.2024.486564.1383

Afna, M. (2018) The correlation between vocabulary and speaking skill. *JL3T (Journal of Linguistics, Literature and Language Teaching)*, 4(1): 43–71. https://doi.org/10.32505/jl3t.v4i1.750

Ahn, T. & Lee, S.-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

Arkorful, V. & Abaidoo, N. (2015) The role of e-learning, advantages and disadvantages of its adoption in higher education. *International Journal of Instructional Technology and Distance Learning*, 12(1): 29–42.

Bahadorfar, M. & Omidvar, R. (2014) Technology in teaching speaking skill. *Acme International Journal of Multidisciplinary Research*, 2(4): 9–13.

Brown, H. D. & Abeywickrama, P. (2019) *Language assessment: Principles and classroom practices* (3rd ed.). Hoboken: Pearson Education.

Cahyono, B. Y., Fauziah, H., Santoso, D. R. & Wulandari, I. (2024) Revisiting technological tools used in EFL speaking classes during the COVID-19 pandemic: What are the implications for the post-pandemic? *The JALT CALL Journal*, 20(2): 1–21. https://doi.org/10.29140/jaltcall.v20n2.1054

Chau, K. G. (2021) The effect of ICT on learners' speaking skills development. *International Journal of TESOL & Education*, 1(1): 22–29.

Chen, H. C. (2022) *The English Speech Corpus with Different Proficiency Levels*. The Education University of Hong Kong. http://corpus.eduhk.hk/english_speech_corpus/

Chen, H. C. & Tian, J. X. (2022) Developing and evaluating a flipped corpus-aided English pronunciation teaching approach for pre-service teachers in Hong Kong. *Interactive Learning Environments*, 30(10): 1918–1931. https://doi.org/10.1080/10494820.2020.1753217

Chen, H. C. & Tian, J. X. (2025) CALL and speaking. In McCallum, L. & Tafazoli, D. (eds.), *The Palgrave encyclopedia of computer-assisted language learning*. Cham: Palgrave Macmillan, 1–11. https://doi.org/10.1007/978-3-031-51447-0_297-1

Criado, R. (2013) A critical review of the presentation-practice-production model (PPP) in foreign language teaching. In Monroy, R. (ed.), *Homenaje a Francisco Gutiérrez Díez*. Murcia: Edit.um, 97–115.

Fathi, J., Rahimi, M. & Derakhshan, A. (2024) Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121: Article 103254. https://doi.org/10.1016/j.system.2024.103254

Feak, C. B., Reinhart, S. M. & Rohlck, T. N. (2018) *Academic interactions: Communicating on campus*. Ann Arbor: University of Michigan Press. https://doi.org/10.3998/mpub.288586

Gablasova, D. & Bottini, R. (2022) Spoken learner corpora for language teaching. In Jablonkai, R. R. & Csomay, E. (eds.), *The Routledge handbook of corpora and English language teaching and learning*. Abingdon: Routledge, 296–310. https://doi.org/10.4324/9781003002901

Huang, F. & Zou, B. (2024) English speaking with artificial intelligence (AI): The roles of enjoyment, willingness to communicate with AI, and innovativeness. *Computers in Human Behavior*, 159: Article 108355. https://doi.org/10.1016/j.chb.2024.108355

Huang, J. (2023) Engineering ChatGPT prompts for EFL writing classes. *International Journal of TESOL Studies*, 5(4): 73–79. https://doi.org/10.58304/ijts.20230405

Jao, C.-Y., Yeh, H.-C., Huang, W.-R. & Chen, N.-S. (2024) Using video dubbing to foster college students' English-speaking ability. *Computer Assisted Language Learning*, 37(4): 585–607. https://doi.org/10.1080/09588221.2022.2049824

Kachru, B. B. (ed.) (1992) *The other tongue: English across cultures* (2nd ed.). Urbana: University of Illinois Press.

Kohen-Vacs, D., Milrad, M., Ronen, M. & Jansen, M. (2016) Evaluation of enhanced educational experiences using interactive videos and web technologies: Pedagogical and architectural considerations. *Smart Learning Environments*, 3: Article 6. https://doi.org/10.1186/s40561-016-0029-z

Kostka, I. & Toncelli, R. (2023) Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *TESL-EJ*, 27(3): 1–19. https://doi.org/10.55593/ej.27107int

Luthfillah, D. A. & Fauzia, N. (2023) Using digital flashcards (DFs) as a learning media to enhance students' speaking skill. *Journal of English Teaching, Literature, and Applied Linguistics*, 7(2): 26–31. https://doi.org/10.30587/jetlal.v7i2.5941

Mendis, U. & Dharmawan, Y. Y. (2019) Understanding learner interaction in web based learning to improve English speaking skills in second grade of SMAN 4 senior high school Bandar Lampung: Using Canvas learning management system. *Journal of English Education Studies*, 2(1): 74–85. https://doi.org/10.30653/005.201921.38

Mompean, J. A. (2024) ChatGPT for L2 pronunciation teaching and learning. *ELT Journal*, 78(4): 423–434. https://doi.org/10.1093/elt/ccae050

Nunan, D. (2015) *Teaching English to speakers of other languages: An introduction*. New York: Routledge. https://doi.org/10.4324/9781315740553

Parveen, B. W. (2016) Use of technology in improving speaking skills. *Journal of English Language and Literature (JOELL)*, 3(2): 121–124.

Qaddumi, H. A. & Smith, M. (2024) Implementation of learning management systems (Moodle): Effects on students' language acquisition and attitudes towards learning English as a foreign language. *Trends in Higher Education*, 3(2): 260–272. https://doi.org/10.3390/higheredu3020016

Satar, M., Hauck, M. & Bilki, Z. (2023) Multimodal representation in virtual exchange: A social semiotic approach to critical digital literacy. *Language Learning & Technology*, 27(2): 72–96. https://doi.org/10.64152/10125/73504

Shafiee Rad, H. & Roohani, A. (2024) Fostering L2 learners' pronunciation and motivation via affordances of artificial intelligence. *Computers in the Schools*. Advance online publication. https://doi.org/10.1080/07380569.2024.2330427

Shin, J. L. K. & Yunus, M. M. (2021) A systematic review of e-learning in teaching and learning of speaking skills. *International Journal of Academic Research in Business and Social Sciences*, 11(4): 725–740. https://doi.org/10.6007/IJARBSS/v11-i4/9245

Suartama, K. & Dewa, K. S. (2014) *E-learning Berbasis MOODLE*. Yogyakarta: Graha Ilmu.

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

Talley, P. C. & Tu, H. (2014) Implicit and explicit teaching of English speaking in the EFL classroom. *International Journal of Humanities and Social Science*, 4(6): 38–46.

Tran, N., Hoang, D. T. N., Gillespie, R., Yen, T. T. H. & Phung, H. (2024) Enhancing EFL learners' speaking and listening skills through authentic online conversations with video conferencing tools. *Innovation in Language Learning and Teaching*. Advance online publication. https://doi.org/10.1080/17501229.2024.2334809

Walter, Y. (2024) Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21: Article 15. https://doi.org/10.1186/s41239-024-00448-3

Woo, Y. & Reeves, T. C. (2008) Interaction in asynchronous web-based learning environments: Strategies supported by educational research. *Journal of Asynchronous Learning Networks*, 12(3–4): 179–194. https://doi.org/10.24059/olj.v12i3-4.1691

Yan, X., Kim, H. R. & Kim, J. Y. (2021) Dimensionality of speech fluency: Examining the relationships among complexity, accuracy, and fluency (CAF) features of speaking performances on the Aptis test. *Language Testing*, 38(4): 485–510. https://doi.org/10.1177/0265532220951508

Yeh, H. C., Chang, W. Y., Chen, H. Y. & Heng, L. (2021) Effects of podcast-making on college students' English speaking skills in higher education. *Educational Technology Research & Development*, 69(5): 2845–2867. https://doi.org/10.1007/s11423-021-10026-3

Zou, B., Guan, X., Shao, Y. & Chen, P. (2023) Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. *Sustainability*, 15(4): Article 2872. https://doi.org/10.3390/su15042872

About the authors

Hsueh Chu Chen is an associate professor at the Education University of Hong Kong and has been investigating a wide range of issues in interlanguage phonetics and phonology, third language phonology, foreign accent and intelligibility, and computer-assisted/corpus-based pronunciation teaching and learning.

Xiaona Zhou is an EdD student specializing in phonetics and phonology. Her research focuses on integrating technology tools into pronunciation teaching and learning in diverse linguistic contexts.

Jing Xuan Tian is a PhD student majoring in phonetics and phonology with over 10 years of experience in language education and has also published articles in linguistics, multilingualism, computer-assisted language teaching and learning, and pronunciation teaching and learning.