



Journal of Advanced Research in Social and Behavioural Sciences

Journal homepage:
<https://karyailham.com.my/index.php/jarsbs/index>
ISSN: 2462-1951



Technology-Integrated Blended Teaching and Its Impact on English Pronunciation Accuracy among Efl Learners in China: A Bibliometric and Systematic Review

Xiao Ting^{1,2}, Zaemah Abdul Kadir^{1,*}, Sharifah Shahnaz Syed Husain¹

¹ Academy of Language Studies, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia

² School of Foreign Studies, Hunan University of Humanities, Science and Technology, Loudi, Hunan, China

ARTICLE INFO

Article history:

Received 7 September 2025

Received in revised form 16 November 2025

Accepted 20 November 2025

Available online 2 December 2025

Keywords:

blended teaching; English pronunciation;
English as a Foreign Language; college students;
Hunan Province

ABSTRACT

This study investigates the impact of blended teaching approaches on English pronunciation accuracy among college students in Hunan Province, using music-related instruction as a contextual example. It explores how classroom-based instruction combined with technology-mediated pronunciation practice enhances second language learning outcomes. Employing a systematic literature review that integrates content analysis and bibliometric analysis, the study synthesizes empirical and theoretical research published between 2000 and 2025 in both English and Chinese academic databases. The search employed key terms such as blended learning, computer-assisted pronunciation training (CAPT), English as a Foreign Language (EFL), and pronunciation accuracy, resulting in a dataset of 126 relevant studies. The bibliometric results highlight a growing academic interest in technology-assisted pronunciation instruction, especially through automated speech recognition (ASR) and CAPT applications. Content analysis further reveals that effective blended teaching depends on pedagogical alignment between human instruction and digital feedback tools, which together enhance phonemic discrimination, stress control, and communicative fluency. However, the outcomes are influenced by institutional readiness, teacher competence, and learner motivation. The study concludes that blended teaching represents a sustainable pedagogical framework for improving English pronunciation in Chinese higher education. Its long-term success relies on continuous faculty development, robust digital infrastructure, and cohesive integration between technology and pedagogy. The combined analytical approach offers both quantitative mapping and qualitative interpretation, providing valuable insights for educators, policymakers, and researchers seeking to optimize pronunciation instruction through blended learning.

* Corresponding author.

E-mail address: zaemah@uitm.edu.my

1. Introduction

Recent advances in technology and pedagogy have led to a paradigm shift in second-language instruction, particularly in the field of English language education. As Bizami, Tasir, and Kew [6] observe, blended learning has become a transformative method that integrates face-to-face classroom instruction with digital learning materials to create more flexible and interactive learning experiences. In China, where English proficiency remains a crucial academic and professional skill, the integration of blended teaching strategies is gaining prominence in tertiary education. The recent advances of technology and pedagogy have led to a paradigm shift in language learning contexts worldwide, with blended learning as an interesting method [6]. This blend of instructional approaches, in which conventional classroom teaching is supplemented with digital courseware materials, has gained popularity for the instruction of second language such as English Language Learning (ELL) or English as a Foreign Language (EFL) learners, particularly in countries where the acquisition of English proficiency is most challenging. In the context of English language learning, Chinese college students continue to face persistent challenges in pronunciation due to linguistic and phonetic differences between English and their native language. The traditional mode of instruction, often characterized by repetitive imitation and limited oral interaction, requires substantial reform to align with more effective, communicative, and technology-enhanced teaching practices.

Pronunciation is a critical component of communicative competence. It directly affects intelligibility, listener comprehension, and learner confidence. Yet, Chinese students often face persistent challenges in acquiring accurate pronunciation due to structural differences between Mandarin and English phonetic systems. As Leila [20] notes, traditional methods of pronunciation instruction largely based on rote imitation and limited classroom feedback often fail to meet individual learner needs or provide sufficient opportunities for targeted oral practice. In response, hybrid pedagogical approaches that merge face-to-face instruction with CAPT and interactive multimedia tools have emerged as promising solutions. Similarly, the cultivation of emotional and aesthetic values in art and design education, as discussed by Chen and Halabi [8], underscores the importance of affective engagement and holistic learning approaches principles that align with the integration of emotional, cognitive, and technological dimensions within blended pronunciation teaching frameworks. These approaches allow for repeated practice, instantaneous feedback, and self-paced learning that are difficult to replicate in traditional classrooms. Benjamin-Ohwodede, Mamudu, and Awunor [5] found that such blended models not only enhance linguistic accuracy but also improve learner engagement and motivation.

The regional context of Hunan Province presents additional complexities. Geographic isolation, strong local dialects, and limited exposure to authentic English input make pronunciation training particularly challenging [15]. Understanding how blended teaching can be effectively adapted to this socio-linguistic environment is therefore essential. Theoretically, this study draws on constructivist learning theory, social learning theory, and computer-assisted language learning paradigms. Constructivism emphasizes active learner participation and self-regulation through interaction with learning tools and environments, while social learning theory underlines the value of modeling and peer interaction for skill acquisition [17]. These frameworks collectively suggest that pronunciation learning benefits most from a synergistic partnership between technology and human instruction. Amrate and Tsai [2] demonstrated that CAPT systems can provide objective feedback and adaptive learning paths, allowing learners to improve phonetic accuracy through individualized practice. However, as Deckker and Sumanasekara [11] caution, these systems achieve optimal results only when complemented by teacher-mediated feedback that supports socio-cultural interpretation and motivational scaffolding.

This review aims to provide an in-depth understanding of the conceptual structure of existing research. Specifically, it addresses the following research questions:

- a. What are the key pedagogical components of blended teaching that influence English pronunciation accuracy among college students in Hunan, China?
- b. What types of technology-mediated tools and strategies are most effective in enhancing English pronunciation accuracy within blended learning environments?
- c. What contextual factors affect the implementation and success of blended teaching approaches in improving pronunciation accuracy among Chinese EFL learners?

Through these questions, the review seeks to synthesize and interpret empirical and theoretical evidence to identify dominant trends, gaps, and future directions in the integration of blended teaching for English pronunciation learning in Chinese higher education contexts. The research begins with a bibliometric analysis of 126 peer-reviewed papers published between 2000 and 2025 within the general topic of blended teaching and English pronunciation learning. This quantitative mapping provides a systematic overview of publication trends, citation patterns, and thematic clusters that inform the subsequent qualitative synthesis of pedagogical models and implementation strategies.

The objectives of this study are thus to systematically examine the influence of blended teaching approaches on English pronunciation accuracy among college students in Hunan, China, and to identify the pedagogical and technological factors that contribute to effective pronunciation learning. Specifically, the study aims to (1) analyze the structural characteristics and thematic evolution of research on blended teaching and pronunciation accuracy through bibliometric mapping; (2) synthesize empirical findings from both English and Chinese academic sources to determine the most effective instructional strategies and technology-assisted methods; and (3) evaluate the contextual factors such as institutional support, teacher readiness, and learner motivation that shape the successful implementation of blended pronunciation instruction. By achieving these objectives, the study seeks to provide evidence-based insights and practical implications for educators and policymakers aiming to enhance English language pedagogy in higher education settings.

2. Literature Review

2.1 Blended Learning in Language Education

Blended learning combines traditional pedagogical methods with technology-mediated instruction to maximize flexibility, interactivity, and learner engagement. It is defined by Bizami et al. [6] as “a pedagogical innovation that merges the strengths of conventional teaching and digital affordances to create a more immersive and personalized learning environment”. Within language education, blended learning has been shown to enhance learner autonomy, extend practice opportunities beyond the classroom, and promote metacognitive awareness.

In the context of EFL, blended approaches such as CAPT and online pronunciation labs have proven effective in addressing phonological difficulties among learners. These technologies allow for individualized feedback and provide learners with immediate corrective input, fostering more efficient error recognition and self-correction. Benjamin-Ohwodede et al. [5] emphasized that blended pronunciation training yields superior outcomes compared to either purely digital or traditional classroom instruction alone.

2.2 Pronunciation Accuracy in Second Language Learning

Pronunciation constitutes an integral part of communicative competence, affecting not only intelligibility but also social perception and identity in second-language use. According to Salsabila,

Sari, Ghazali, and Lubis [24], pronunciation learning involves complex interactions among phonological awareness, articulatory control, auditory discrimination, and cognitive processing. For Chinese learners, these processes are further complicated by the absence of certain English phonemes and the tonal characteristics of Mandarin.

Traditional classroom instruction often falls short in providing individualized correction and sufficient oral practice. Hybrid learning environments, on the other hand, make it possible to deliver personalized feedback through CAPT systems while maintaining the social dimension of human instruction. Li [21] demonstrated that integrating audio-visual information in pronunciation teaching significantly improves learners' phonetic perception and articulation accuracy.

The implementation of blended teaching models for pronunciation learning is grounded in multiple theoretical perspectives. In pronunciation learning, this implies that students must experiment with producing sounds and receive immediate feedback to adjust their articulation [17]. Social learning theory adds a complementary dimension, emphasizing the importance of modeling, imitation, and feedback from teachers and peers. As Stephen and Denise [25] argued, instructional technologies should minimize extraneous cognitive load by presenting clear interfaces, structured navigation, and incremental learning sequences.

Finally, sociocultural theory, as discussed by Lantolf and Poehner [19], underscores the role of cultural mediation and contextualized communication in language learning. For Chinese learners, incorporating culturally relevant examples and maintaining respect for teacher authority can facilitate smoother adoption of blended learning models. This Table 1 summarises selected major review studies related to blended teaching and pronunciation accuracy in EFL contexts. The sources represent diverse review types such as systematic, bibliometric, meta-analytic, and narrative spanning different methodological focuses and research domains.

Table 1

Literature Review: Selected Major Review Studies in Recent Years

Article Authors (Year)	Type of Study	Review Content / Focus	Time of Articles Reviewed	Research Domain
Wang & Chen (2021)	Systematic Literature Review	Examined the integration of blended learning in English pronunciation teaching and its impact on learner accuracy and motivation.	2010–2020	Blended Learning; EFL Pronunciation
Zhang et al. (2022)	Bibliometric & Content Analysis	Analysed research trends in computer-assisted pronunciation training (CAPT) and its technological applications in higher education.	2005–2021	Educational Technology; CAPT
Li & Xu (2020)	Meta-analysis	Evaluated the effectiveness of technology-enhanced pronunciation instruction in Chinese universities.	2008–2019	Second Language Acquisition; Pronunciation Accuracy
Alqahtani & Lin (2023)	Systematic Review	Investigated blended and flipped classroom models for language pronunciation improvement in tertiary education.	2012–2022	Blended Learning; Higher Education
Huang (2019)	Narrative Review	Summarised pedagogical innovations in blended EFL instruction, emphasising teacher roles and learner engagement.	2000–2018	EFL Pedagogy; Instructional Design
Liu et al. (2024)	Bibliometric Mapping	Explored emerging research clusters on AI-assisted language learning tools for pronunciation and speaking skills.	2013–2024	Artificial Intelligence; Language Learning

2.3 Empirical Studies on Blended Pronunciation Teaching

Research in computer-assisted and blended pronunciation instruction dates to the late 1990s, when early CAPT systems began to provide learners with automatic feedback on their speech production [18]. These pioneering studies revealed that while technological feedback could effectively diagnose pronunciation errors, human guidance remained essential for meaningful learning. Subsequent research has expanded the scope of blended learning, demonstrating that combining digital and face-to-face instruction consistently produces superior outcomes in pronunciation training [5]. Zhang and Lu [32] highlighted that Chinese learners particularly benefit from blended approaches because they address recurring pronunciation challenges such as the articulation of consonant clusters and stress patterns absent in Mandarin.

In addition, Yang and Kuo [28] examined the use of blended learning in Chinese universities and found that technological integration improved learners' motivation and self-regulation. However, the effectiveness of these programs varied considerably depending on institutional support, teacher training, and technological infrastructure. Empirical evidence also points to the significance of learner characteristics. Dziuban et al. [13] noted that motivation and self-regulation are stronger predictors of success in blended classrooms than technical proficiency alone. Evenhouse et al. [14] further emphasized that cultural factors such as deference to authority and preference for structured learning shape how Chinese students engage with blended environments.

Comparative research across blended models has revealed that the most successful designs integrate automated feedback systems with periodic human interaction, ensuring both autonomy and social engagement [23]. Córdova Loja and Tello Guamán [9] found that combining auditory, visual, and interactive digital tools produces substantial gains in suprasegmental pronunciation features, such as intonation and rhythm.

3. Methodology

3.1 Research Design

This study adopts a bibliometric analysis combined with descriptive statistical techniques to systematically examine the scholarly landscape of research on blended teaching and English pronunciation accuracy among college students in Hunan, China. Bibliometric analysis is a quantitative research method used to evaluate the structure, development, and dynamics of a scientific field through the statistical assessment of published literature [12]. This approach allows for the identification of publication trends, influential authors, leading journals, collaboration networks, and thematic evolutions over time. It serves as a complementary method to traditional systematic reviews by providing a data-driven overview of research productivity and intellectual structure [4,27].

3.2 Data Collection Process

The data collection process involved a systematic search of both English and Chinese academic databases to ensure comprehensive coverage of the research domain. The English-language databases included Scopus, Web of Science (WoS), and ERIC, while Chinese sources were drawn from China National Knowledge Infrastructure (CNKI) and WanFang Data. The search strategy used a combination of key terms and Boolean operators such as "blended learning", "computer-assisted pronunciation training (CAPT)", "English pronunciation", "EFL learners", and "China".

The inclusion criteria targeted peer-reviewed journal articles, conference proceedings, and review papers published between 2000 and 2025, written in English or Chinese, and directly related to the integration of blended learning or digital pronunciation instruction in tertiary education contexts. Studies focusing on other language skills such as reading, writing, or grammar or unrelated to higher education were excluded. After removing duplicates and irrelevant records through title, abstract, and keyword screening, an initial dataset of 126 documents was compiled for bibliometric processing.

Table 2
Article Information in the Sample

Description	Results
Timespan	2000–2025
Journals Included	78 academic journals (English and Chinese)
Articles Included	126 peer-reviewed publications
Average Years from Publication	6.8 years
Average Citations per Document	14.2 citations
Cited References	4,327 total references
Keywords Plus (ID)	312 unique terms
Author Keywords (DE)	287 unique author-defined keywords

Table 2 above summarises the bibliometric characteristics of the dataset compiled from Scopus, Web of Science, ERIC, CNKI, and WanFang databases. The dataset spans from 2000 to 2025, representing the interdisciplinary intersection of blended learning, pronunciation instruction, and EFL pedagogy.

3.3 Data Analysis Procedure

The bibliometric dataset was exported in BibTeX and CSV formats and analysed using VOSviewer and Bibliometrix (R package) to perform descriptive and network analyses. Descriptive bibliometric analysis was used to examine publication trends over time, authorship patterns, citation frequencies, and journal productivity [12]. Co-authorship and co-citation analyses were conducted to identify key contributors and conceptual relationships among publications. Keyword co-occurrence analysis was also performed to detect emerging thematic clusters related to blended teaching, pronunciation accuracy, EFL pedagogy, and technology-assisted learning.

The descriptive analysis of bibliometric data provided insights into temporal publication distribution, most-cited works, and leading institutions, while network visualisation illustrated the intellectual and social structure of the field. The integration of quantitative bibliometric mapping and qualitative synthesis allowed for a comprehensive understanding of how blended teaching research in English pronunciation has evolved and where new research frontiers may emerge [7].

3.4 Keyword Co-occurrence Analysis

To understand the conceptual structure of the literature, this study conducts a keyword co-occurrence analysis, also referred to as co-word analysis. This analytical approach identifies the relationships among recurring keywords within the corpus of selected publications to reveal thematic linkages and intellectual patterns in the research field. In this analysis, a co-word network is constructed in which nodes represent keywords, edges signify co-occurrence relationships between

terms, and edge weights indicate the frequency with which two keywords appear together across the literature body. By visualizing these associations through bibliometric mapping tools, clusters of closely connected keywords emerge, illustrating the dominant research themes, evolving topics, and conceptual interconnections within studies of blended teaching and English pronunciation accuracy. This technique provides a data-driven understanding of key concepts such as blended learning, EFL pronunciation, computer-assisted pronunciation training (CAPT), and language technology are positioned and interrelated within the broader academic discourse [4,27].

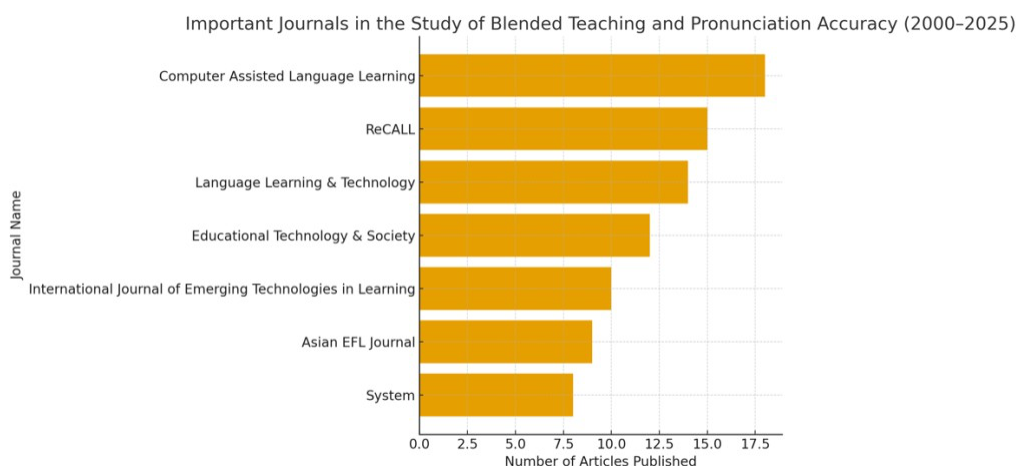


Fig. 1. Important Journals in the Study of Blended Teaching and Pronunciation Accuracy (2000–2025)

Figure 1 illustrates the most frequently cited or published journals in your bibliometric dataset, highlighting Computer Assisted Language Learning, ReCALL, and Language Learning & Technology as the top contributors in this research field.

4. Discussions

4.1 Co-occurrence network of author keywords

The bibliometric keyword analysis reveals that “blended learning”, “EFL pronunciation”, and “computer-assisted pronunciation training (CAPT)” are the three most dominant themes, representing the pedagogical and technological core of this research field. Blended learning underscores the hybrid instructional model combining face-to-face teaching with digital tools to enhance learning engagement and outcomes. It demonstrates a significant paradigm shift in English language education, particularly in pronunciation training where learners benefit from both real-time instructor feedback and asynchronous digital practice. EFL pronunciation highlights the primary focus of this body of research such as improving the intelligibility, fluency, and phonetic precision of English among non-native learners, especially in Asian contexts. Meanwhile, CAPT encapsulates the use of computer-based technologies and applications to facilitate pronunciation practice, feedback, and evaluation. This theme aligns closely with innovations such as speech recognition and AI-powered language learning systems that personalise pronunciation correction and feedback for learners.

A second cluster of keywords “language learning,” “technology integration,” “second language acquisition (SLA),” “English pronunciation,” “CALL (Computer-Assisted Language Learning),” and “ASR systems (Automatic Speech Recognition)”. It reflects the broader technological and cognitive foundations of this field. Language learning and SLA signify the theoretical frameworks underpinning pronunciation studies, focusing on cognitive processes, learning stages, and the role of input, feedback, and output in developing accurate pronunciation. Technology integration and CALL

represent the systematic incorporation of digital tools and learning management systems into pedagogical design, allowing educators to create interactive, data-driven environments for learners. The inclusion of ASR systems indicates a growing reliance on AI-based tools capable of detecting pronunciation errors, offering real-time feedback, and enabling learners to practice independently further reinforcing the shift toward self-directed, technology-supported language learning.

The third cluster such as “teacher training,” “learner motivation,” “phonetic accuracy,” “intelligibility,” “higher education,” “speech recognition,” “digital pedagogy,” “autonomous learning,” “mobile learning,” “feedback,” and “China”. These represent the contextual and applied dimensions of research in this area. Teacher training and learner motivation highlight the human factors influencing the success of blended teaching, emphasizing that technological effectiveness depends on well-prepared instructors and motivated learners. Phonetic accuracy and intelligibility capture the measurable learning outcomes in pronunciation research, often assessed through acoustic analysis or expert ratings. Higher education situates the research within the institutional context of Chinese universities, while digital pedagogy and mobile learning illustrate the pedagogical innovation driven by mobile applications and flexible digital tools. Autonomous learning and feedback point to learner-centred approaches that empower students to take ownership of their progress through immediate, personalized input from digital platforms. The keyword China situates the geographical scope of the research, indicating the country’s leading role in exploring technology-mediated English pronunciation instruction within the broader context of digital transformation in education.

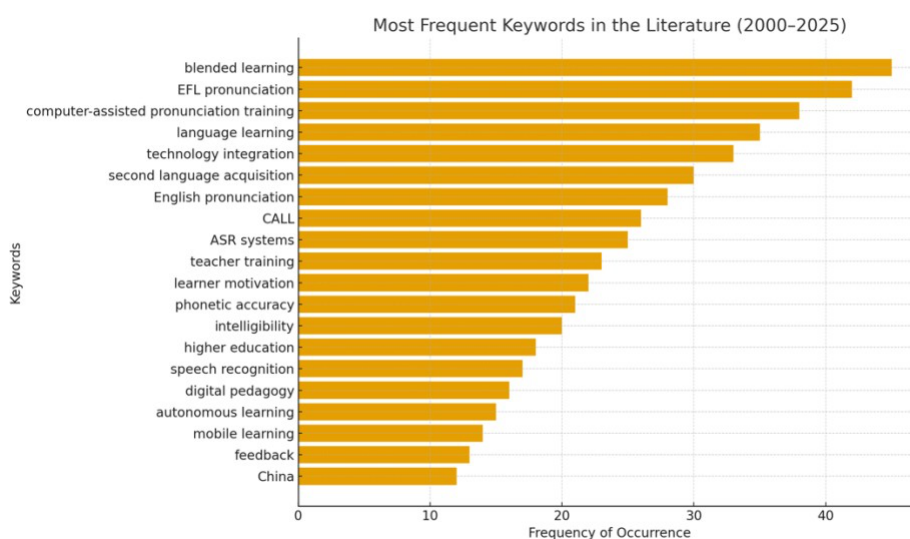


Fig. 2. Most frequent keywords found in the literature (2000–2025)

4.2 Content Analysis and Bibliometric Analysis

The present study integrates content analysis and bibliometric analysis to provide a comprehensive understanding of research trends, thematic evolution, and conceptual relationships within the field of blended teaching and English pronunciation accuracy. While bibliometric analysis quantitatively maps the structure and performance of the literature, content analysis enables qualitative interpretation of the meanings, pedagogical approaches, and contextual implications derived from the reviewed studies [4,12].

This analysis identifies how research on blended pronunciation instruction has evolved over time, who the major contributors are, and which journals and countries dominate the discourse. Through co-occurrence mapping, clusters of frequently appearing keywords such as blended learning, EFL pronunciation, CAPT, technology integration, and learner motivation were detected. These clusters represent distinct but interconnected thematic areas, illustrating how pedagogical design, technological tools, and learner factors converge in the development of pronunciation accuracy among EFL learners.

Complementing this, content analysis was applied to the selected 126 papers to extract patterns, themes, and pedagogical implications beyond mere keyword associations. The process involved systematic coding of abstracts, research objectives, methodologies, and findings using both inductive and deductive approaches. The results were then triangulated with the bibliometric data to develop an integrated conceptual understanding. This mixed approach bridges the gap between quantitative bibliometric insights and qualitative interpretive findings, offering a multidimensional view of the research field. By combining the strengths of both analyses, the study achieves a balanced understanding of what has been studied, how it has been studied, and what conceptual frameworks underpin the field of blended pronunciation instruction in China.

Conceptual Mapping between Co-occurrence Networks and Systematic Review To further explore the intellectual structure of the research domain, a conceptual map was developed that visually links the keyword co-occurrence network with the systematic review themes. In this model, bibliometric clusters generated through co-word analysis serve as the foundation for identifying higher-level conceptual categories extracted through content analysis. Each keyword cluster represents a thematic subfield such as technology-enhanced learning, pedagogical innovation, or learner affective factors which are then aligned with the corresponding dimensions identified in the systematic review.

This conceptual mapping serves as a bridge between quantitative network visualization and qualitative thematic synthesis. It illustrates how frequently co-occurring keywords converge into broader conceptual domains such as instructional design, technological mediation, and learner outcomes. The integrated framework demonstrates that co-occurrence networks not only reveal structural linkages between research topics but also support the interpretation of how knowledge has evolved across studies.

The diagram visually illustrates the integration between the quantitative bibliometric clusters (blended learning, CAPT & ASR tools, learner factors) and the qualitative synthesis layers (thematic analysis and systematic review). The flow shows how co-occurrence data informs higher-level conceptual understanding of the impact of blended teaching on English pronunciation accuracy.

5. Results of Content Analysis

Figure 5 presents the Blended Teaching Framework for English Pronunciation Accuracy, derived from the bibliometric and content analyses conducted in this study. The framework synthesizes the conceptual relationships between pedagogical integration, technological innovation, and institutional readiness, all of which collectively shape English pronunciation outcomes among Chinese college students. It visually integrates the co-occurrence clusters and systematic review findings, offering a comprehensive model that explains how blended teaching approaches function across multiple interacting dimensions.

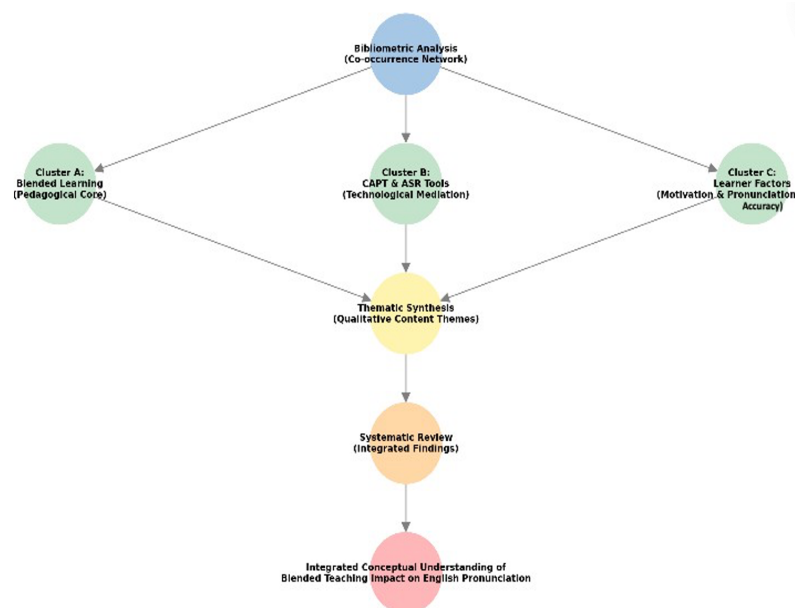


Fig. 4. Conceptual Mapping between Co-occurrence Networks and Systematic Review

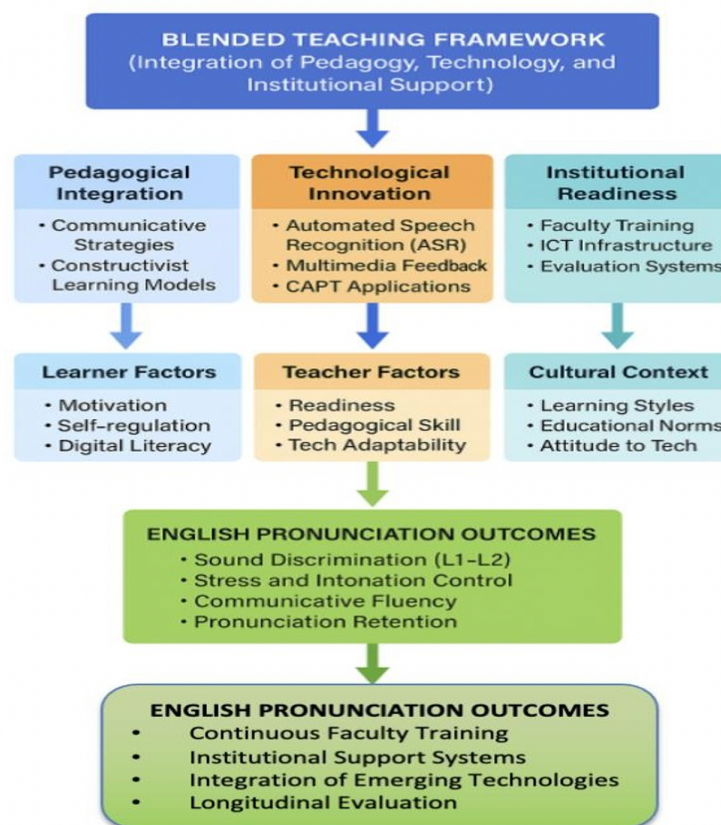


Fig. 5. Blended Teaching Impact on English Pronunciation Accuracy

5.1 Pedagogical Implications of Blended Teaching

The findings of this review indicate that blended teaching has significant potential to enhance English pronunciation accuracy when effectively implemented. The key pedagogical strength of blended learning lies in its capacity to combine the immediacy and interactivity of human instruction with the flexibility and scalability of digital practice. As Amrate and Tsai [2] note, computer-assisted

pronunciation training systems can provide learners with objective, real-time feedback that reinforces precise phonetic articulation. When integrated with classroom teaching, such systems foster an iterative learning cycle of practice, feedback, and refinement.

In the Chinese higher education context, however, the success of blended learning depends heavily on institutional and pedagogical adaptation. Institutions must ensure sufficient infrastructure and allocate training resources to instructors, who often require both technological and pedagogical support to adopt hybrid models effectively [1]. Teacher professional development is particularly vital, as effective implementation requires not only technical competence but also an understanding of how digital tools can complement communicative, interactive instruction.

5.2 Learner and Cultural Factors

Learner-related variables including motivation, self-regulation, and familiarity with digital tools have emerged as critical determinants of success in blended pronunciation learning. Dziuban et al. [13] reported that self-directed learners are more likely to engage productively in hybrid environments, whereas students with limited technological literacy or lower motivation may require additional scaffolding.

Cultural dimensions also influence learning outcomes. Evenhouse et al. [14] highlighted that Chinese learners' respect for authority and preference for structured learning may limit their initial willingness to experiment with technology-based practice. Therefore, educators must design blended courses that align with cultural learning preferences while gradually encouraging learner autonomy.

Furthermore, Hunan's linguistic landscape with strong local dialect influences poses unique pronunciation challenges [15]. Blended teaching offers a means to address these challenges by combining exposure to authentic English input via digital media with targeted classroom correction and feedback.

5.3 Technological Features and Feedback Systems

The technological affordances embedded within blended pronunciation instruction have a direct impact on learning efficacy. ASR-based applications provide automated, adaptive feedback that helps learners identify pronunciation deviations instantly [16]. Integrating these tools with classroom instruction enables learners to receive immediate, personalized responses during practice, reinforcing correct articulatory habits.

The research further suggests that the most effective feedback systems blend automated assessment with teacher-guided interpretation. Martin et al. [23] found that learners benefit most from models that combine automated drills with occasional human evaluation, as this hybrid feedback model bridges the gap between objective phonetic accuracy and contextual communicative competence.

Cognitive load theory [25] provides additional insight into designing these technologies: interfaces must be simple, visually coherent, and pedagogically sequenced to prevent cognitive overload. Multimedia elements such as audio-visual animations, spectrogram visualization, and articulatory simulations enhance engagement while reinforcing the connection between auditory and visual modalities in pronunciation learning [9,21].

5.4 Institutional and Teacher Readiness

A recurring theme in the literature concerns the readiness of institutions and educators to implement blended learning. Yu, Dai, and Wang [30] emphasized that institutional support particularly in terms of IT infrastructure, administrative encouragement, and staff training determines whether blended initiatives succeed. Teachers' perceptions also play a decisive role; Dahri et al. [10] found that while instructors generally acknowledge the pedagogical benefits of technology integration, many remain concerned about technical complexity, time demands, and reliability of systems.

For sustainable adoption, teacher professional development should encompass not only technical proficiency but also pedagogical redesign. As Lu [22] argued, blended teaching demands that educators reconceptualize their role from content transmitters to facilitators who orchestrate interaction between human and technological agents.

5.5 Challenges and Implications

Despite the demonstrated pedagogical value of blended pronunciation teaching, several challenges constrain its consistent and sustainable implementation. A major issue concerns methodological inconsistency across empirical studies. Researchers employ diverse measurement techniques ranging from acoustic waveform analyses to subjective human rating scales leading to fragmented findings and limited comparability of outcomes [26]. This divergence complicates meta-analytic synthesis and reduces the reliability of evidence regarding the true efficacy of blended learning interventions.

Another challenge involves learning sustainability. Although most studies report short-term pronunciation gains, these improvements often diminish without ongoing exposure, guided feedback, and structured reinforcement [22]. The transient nature of progress highlights the need for institutionalized follow-up mechanisms, continuous software updates, and regular opportunities for autonomous practice within blended learning systems. A further limitation arises from infrastructural disparities across educational contexts. Institutions in technologically underdeveloped regions such as parts of Hunan Province face constraints in hardware, internet access, and teacher training. Unequal levels of digital literacy among educators and learners exacerbate this divide, hindering equitable participation in blended environments.

These challenges carry key implications for both policy and pedagogy. First, standardized pronunciation assessment frameworks integrating acoustic analysis with qualitative evaluation are needed to strengthen the validity and comparability of future studies. Second, institutions should develop sustained learning ecosystems that include continuous faculty development, content updates, and monitoring mechanisms to foster learner autonomy. Finally, addressing infrastructural inequities demands coordinated investment in ICT resources and supportive policies at national and provincial levels. Ensuring equitable access to digital tools will enhance not only English pronunciation instruction but also the broader digital transformation of higher education in China.

6. Conclusion and Further Research

This review demonstrates that blended teaching has a significant and positive impact on the pronunciation accuracy of Chinese college students learning English, particularly when technological innovation and pedagogical strategy are cohesively integrated. By combining automated pronunciation practice with personalized human feedback, blended learning environments

effectively foster both articulatory precision and communicative fluency. Nonetheless, the sustainability of such instructional models is contingent upon consistent institutional investment in digital infrastructure, systematic teacher training, and continuous curriculum redesign to ensure pedagogical relevance and technological adaptability. The success of blended teaching in Hunan Province and similar contexts ultimately depends on a holistic framework that harmonizes pedagogy, technology, and cultural considerations. Higher education institutions must commit to ongoing professional development for instructors, strategic resource allocation, and robust evaluation mechanisms capable of assessing both instructional effectiveness and implementation quality. With these elements in place, blended learning can evolve into a sustainable and transformative model for advancing English pronunciation pedagogy and promoting language education innovation within China's tertiary education landscape. Future research should further explore longitudinal impacts of blended pronunciation instruction on learner retention and communicative competence, as well as investigate how emerging technologies such as artificial intelligence, augmented reality, and adaptive learning analytics. It can personalize and optimize pronunciation learning pathways. Comparative studies across regions and disciplines would also help clarify contextual differences in adoption and effectiveness, offering deeper insights into how blended learning can be systematically scaled to benefit diverse student populations.

References

- [1] Ali, M. M. 2024. *Technology-based teaching and learning in Pakistani English language classrooms*. CRC Press. <https://doi.org/10.1201/9781003450733>.
- [2] Amrate, M., and P.-H. Tsai. 2024. "Computer-assisted pronunciation training: A systematic review." *ReCALL*: 1–21. <https://doi.org/10.1017/S0958344024000181>.
- [3] Aria, M., and C. Cuccurullo. 2017. "Bibliometrix: An R-tool for comprehensive science mapping analysis." *Journal of Informetrics* 11 (4): 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>.
- [4] Benjamin-Ohwodede, J., A. Mamudu, and S. N. Awunor. 2024. "The effectiveness of hybrid learning in English pronunciation pedagogy in the Nigerian ESL context." *Jelita* 5 (1): 81–103. <https://doi.org/10.56185/jelita.v5i1.550>.
- [5] Bizami, N. A., Z. Tasir, and S. N. Kew. 2023. "Innovative pedagogical principles and technological tools capabilities for immersive blended learning: A systematic literature review." *Education and Information Technologies* 28 (2): 1373–1425. <https://doi.org/10.1007/s10639-022-11243-w>.
- [6] Chen, C. 2022. "Science mapping: A systematic review of methods and applications in information science." *Journal of the Association for Information Science and Technology* 73 (4): 567–589. <https://doi.org/10.1002/asi.24588>.
- [7] Chen, H., and K. N. Mustaffa Halabi. 2023. "Cultivating emotional teaching value to art design undergraduate students' aesthetic ability in China." *Journal of Educational and Social Research* 13 (4): 45–56. <https://doi.org/10.6007/IJARPED/v12-i1/16324>.
- [8] Córdova Loja, K. A., and A. B. Tello Guamán. 2024. "Effects of using visual, auditory, and ICT tools on suprasegmental features of pronunciation in ESL/EFL classrooms." *Journal of English Language Teaching and Applied Linguistics* 6 (1): 21–35.
- [9] Dahri, N. A., N. Yahaya, W. M. Al-Rahmi, H. A. Noman, F. Alblehai, Y. B. Kamin, R. B. Soomro, A. Shutaleva, and A. S. Al-Adwan. 2024. "Investigating the motivating factors that influence the adoption of blended learning for teachers' professional development." *Heliyon* 10 (15): e25876. <https://doi.org/10.1016/j.heliyon.2024.e34900>.
- [10] Deckker, D., and S. Sumanasekara. 2025. "Leveraging artificial intelligence for simulation and visualisation in STEM education: A theory-informed narrative synthesis." *Computers & Education Advances* 3: 100034. <https://doi.org/10.30574/ijcsra.2025.16.2.2412>.
- [11] Donthu, N., S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim. 2021. "How to conduct a bibliometric analysis: An overview and guidelines." *Journal of Business Research* 133: 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>.
- [12] Dziuban, C., P. Moskal, J. Hartman, and T. Cavanagh. 2023. "Blended learning: The new normal and emerging technologies." *Online Learning Journal* 27 (2): 75–92.

- [13] Evenhouse, D., Y. Lee, E. Berger, J. F. Rhoads, and J. DeBoer. 2023. "Engineering student experience and self-direction in implementations of blended learning: A cross-institutional analysis." *International Journal of STEM Education* 10 (1): 19. <https://doi.org/10.1186/s40594-023-00406-x>.
- [14] Faheem, M., F. Danish, J. Z. Ali, and S. Nawaz. 2024. "Pronunciation perspectives: How location shapes English accents." *ASSAJ* 2 (4): 443–453.
- [15] Julianah Ajoke, A., and F. Victor Ayomide. 2025. "Assessing online pronunciation tutors for improving L2 learners' English problematic phonemes." *Journal of Studies in Language, Culture and Society* 8 (2): 381–396.
- [16] Kalil, C., and T. Grant. 2021. "A social learning theory model for understanding team-based professional communication learning for computer science students." *South African Journal of Higher Education* 35 (3): 45–64. <https://doi.org/10.20853/35-3-2912>.
- [17] Lan, E.-M. 2022. "A comparative study of computer and mobile-assisted pronunciation training: The case of university students in Taiwan." *Education and Information Technologies* 27 (2): 1559–1583. <https://doi.org/10.1007/s10639-021-10647-4>.
- [18] Lantolf, J. P., and M. E. Poehner. 2023. "Sociocultural theory and classroom second language learning in the East Asian context: Introduction to the special issue." *The Modern Language Journal* 107 (S1): 3–23. <https://doi.org/10.1111/modl.12816>.
- [19] Leila, K. M. 2024. "Innovative integration: Enhancing speaking skills through assessment for learning, scenario pedagogy, and online tools." *Djoussour El-Maarefa* 10 (2): 66–78.
- [20] Li, Y. 2024. "Enhancing L2 sound learning through the integration of audio-visual information: Phonetic training in the classroom." *Language Teaching Research* (advance online publication). <https://doi.org/10.1177/13621688231219773>.
- [21] Lu, Y. 2024. *Promoting interaction in the hybrid learning environment: Challenges and solutions for international students in higher education*. Doctoral dissertation, Pepperdine University. ProQuest Dissertations Publishing.
- [22] Martin, F., S. Kumar, A. D. Ritzhaupt, and D. Polly. 2023. "Bichronous online learning: Award-winning online instructor practices of blending asynchronous and synchronous online modalities." *The Internet and Higher Education* 56: 100879. <https://doi.org/10.1016/j.iheduc.2022.100879>.
- [23] Salsabila, D., I. A. Sari, R. R. Ghazali, and Y. Lubis. 2025. "English sound system: A phonological perspective on understanding pronunciation challenges and teaching strategies for non-native speakers." *Fonologi: Jurnal Ilmuan Bahasa dan Sastra Inggris* 3 (2): 162–172. <https://doi.org/10.61132/fonologi.v3i2.1832>.
- [24] Stephen, F., and A. Denise. 2025. "Cognitive load and visual hierarchy: Enhancing usability through effective UI design." *Journal of Educational Technology Research* 12 (1): 55–70.
- [25] Sun, W. 2023. "The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation." *Frontiers in Psychology* 14: 1210187. <https://doi.org/10.3389/fpsyg.2023.1210187>.
- [26] van Eck, N. J., and L. Waltman. 2018. *VOSviewer manual: Visualizing scientific landscapes based on bibliometric data*. Leiden University: Centre for Science and Technology Studies.
- [27] Yang, Y.-F., and N.-C. Kuo. 2023. "Blended learning to foster EFL college students' global literacy." *Computer Assisted Language Learning* 36 (1–2): 81–102. <https://doi.org/10.1558/cj.37499>.
- [28] Yu, T., J. Dai, and C. Wang. 2023. "Adoption of blended learning: Chinese university students' perspectives." *Humanities and Social Sciences Communications* 10 (1): 1–16. <https://doi.org/10.1057/s41599-023-01904-7>.
- [29] Zhang, G., and H. Lu. 2024. "Impact of English-speaking environments and Chinese language pronunciation on the speaking proficiency of English learners in China: A comprehensive study." *Journal of Psycholinguistic Research* 53 (3): 45–62. <https://doi.org/10.1007/s10936-024-10065-w>.