DEEP LEARNING TECHNOLOGIES IN ENHANCING SPEAKING PROFICIENCY: A LITERATURE-BASED ANALYSIS OF A NEW PEDAGOGICAL PARADIGM IN TEFL

Rampeng[1], A. Rizal[2], Dewi Nopitasari[3], Ulfah Syam[4]

[1234]Universitas Bosowa Makassar; Indonesia

Correspondence E-mail; rampeng@universitasbosowa.ac.id

Abstract	This study explores how deep learning (DL) technologies, particularly those integrated into tools such as automatic speech recognition (ASR), text-to-speech (TTS), and AI-powered pronunciation applications, can transform speaking instruction by introducing a new pedagogical paradigm. Through a qualitative, literature-based methodology, this research analyzes thirty peer-reviewed studies published between 2020 and 2025. The findings reveal that DL tools significantly enhance learners' pronunciation, rhythm, and fluency through real-time feedback and adaptive learning paths. In addition, these tools contribute to affective gains by reducing speaking anxiety and increasing learner motivation. However, challenges such as algorithmic bias, limited effectiveness in pragmatic discourse, and a lack of teacher training remain significant barriers to full integration. The discussion highlights the shifting role of TEFL instructors from content deliverers to facilitators of technology-enhanced learning environments. The study also emphasizes the importance of considering ethical implications and ensuring data privacy when implementing DL applications. Recommendations are provided for educators, institutions, and researchers to foster effective, equitable, and sustainable use of DL in language education. By bridging the gap between AI technologies and pedagogical practice, this paper proposes a forward-looking framework that positions deep learning not merely as a tool but as a transformative force in TEFL speaking instruction.
Keywords	Deep Learning, Speaking Proficiency, TEFL, Artificial Intelligence in Education, Automatic Speech Recognition, Language Pedagogy.

INTRODUCTION

Speaking proficiency is one of the most important components of communicative competence in the context of Teaching English as a Foreign Language (TEFL). These skills play an important role in learners' ability to express ideas, interact meaningfully, and participate in real communication situations. However, in practice, speaking ability is still one of the most challenging aspects for English learners as a foreign language. Various global studies show that the majority of EFL learners have difficulty in developing fluency, accuracy, and confidence when speaking. The EF English Proficiency Index (EF EPI) 2023 report shows that most countries in Asia, including Indonesia, fall into the "low proficiency" category, with speaking skills being the lowest-performing indicators among other language skills. These difficulties are exacerbated by factors such as a lack of opportunities to communicate authentically, speaking anxiety, limited personal feedback, and the dominance of conventional teaching methods that are not always responsive to individual needs.

The empirical data further reinforce the urgency of this issue. Liu and Jackson's (2008) study showed that more than 60% of EFL students in Asia face high levels of speech anxiety, so they tend to avoid verbal tasks that demand spontaneity. Research by Toubot and Sailin (2021) also found that affective barriers, such as the fear of making mistakes and the embarrassment of speaking in front of peers, significantly contribute to low speaking performance. In addition, speaking and learning are often hindered by traditional teaching methods that still rely on memorizing dialogue, structured exercises, and limited peer-to-peer interaction. Methods like this do not always provide sufficient training frequency, fail to accommodate personalization needs, and do not offer adaptive reps to develop fluency and accuracy continuously (Richards, 2008; Nation & Newton, 2009). This condition demands a new approach that is more adaptive, data-driven, and technology-supported to provide real-time feedback and unlimited training opportunities.

Advances in Artificial Intelligence (AI) technology, especially deep learning (DL), have opened up new opportunities in language learning innovation. Deep learning, which is a branch of machine learning, is capable of extracting complex patterns from large amounts of data and has been widely used in natural language processing (NLP), automatic speech recognition (ASR), and text-to-speech (TTS). The integration of these models into learning platforms—such as ELSA Speak, Duolingo, Google Speech API, and Rosetta Stone Tru Accent—allows learners to receive automated feedback on segmental aspects, suprasegmental aspects, and speech fluency. Wang and Vásquez (2022) demonstrate that DL-based applications can enhance EFL learners' phonological awareness

by providing customized visual and acoustic feedback. Zhang et al. (2021) also found that DL-based speech recognition systems can improve students' pronunciation accuracy through consistent and standardized correction.

Additionally, several empirical studies have reinforced the effectiveness of deep learning-based approaches in developing speaking skills. Wang (2023) reported a significant improvement in phonological accuracy of students who used CNN-based pronunciation-training applications compared to the control group. Lin and Lin (2022) show that the use of LSTM-based feedback systems can improve speaking fluency through intensive repetition-based training. A Sung (2021) study highlights that adaptive feedback in DL-based mobile apps has a positive impact on student confidence, especially for those who experience speech anxiety. On the other hand, Chen and Zou (2021) found that integrating DL-based ASR in language teaching improved the quality of students' verbal responses, as the system can simulate an immersive learning environment. They demonstrate that DL technology has significant potential to complement speech teaching, which has traditionally relied heavily on face-to-face interaction and subjective assessment from teachers.

Nonetheless, the pedagogical integration of deep learning in TEFL speaking teaching is still not fully developed. Most AI research in the context of language learning remains general and does not specifically address the unique capacity of deep learning in supporting speech learning. The study of Lim and Griffiths (2022), for example, states that although many AI applications are used in language learning, the understanding of how DL can be systematically integrated within communicative teaching frameworks is still limited. In addition, DL still faces limitations in evaluating high-level speaking competencies, such as pragmatism, interactional strategies, cultural sensitivity, and negotiation of meaning skills. These dimensions are highly contextual and laden with social values, so they still require the presence of teachers to provide nuanced and contextual assessments.

This is where the novelty of this article lies. Unlike previous research that has focused more on the common use of AI or mobile-assisted language learning applications, this article explicitly highlights the role of deep learning in developing speaking skills. It presents a critical analysis of how the technology can be integrated into the pedagogical framework of modern TEFL. This article not only identifies relevant DL tools and models but also synthesizes empirical findings to extract learning principles that can be applied in the classroom. In addition, this article offers an integrative framework that connects deep learning affordances to speaking learning goals, communicative

approaches, and individual learning needs, thus making conceptual contributions that have not been widely discussed in the literature. This approach provides practical guidance for teachers seeking to integrate DL tools effectively into their speaking instruction.

Thus, this article aims to provide a comprehensive overview of the potential of deep learning to enhance speaking skills in the context of TEFL, offering a new pedagogical paradigm that combines advanced technologies with communicative teaching practices. In particular, this article identifies DL tools relevant to speech learning, analyzes the empirical evidence on their effectiveness, as well as proposes a DL-based pedagogical framework that can be implemented in modern EFL classrooms. Through a cross-disciplinary synthesis between applied linguistics, educational technology, and instructional design, this article is expected to provide a new direction in the development of technology-based speaking skills and enrich the research discourse in the field of language learning technology. It also emphasizes the alignment of technological affordances with learners' individual needs and classroom goals.

METHOD

This study employs a qualitative library research design with an exploratory orientation, focusing on synthesizing theoretical insights and empirical findings related to the use of deep learning technologies in TEFL speaking instruction. As a literature-based qualitative inquiry, the research aims to map the current landscape of deep learning applications in speaking pedagogy and identify conceptual and practical gaps that may support the formulation of a new pedagogical paradigm. The data were drawn from peer-reviewed journal articles, academic reports, and conference papers published between 2020 and 2025, selected from reputable databases such as Scopus, Web of Science, ERIC, ScienceDirect, and Google Scholar. The inclusion criteria emphasized studies conducted in TEFL/ESL contexts that explicitly address speaking instruction and incorporate deep learning tools such as ASR, TTS, NLP, pronunciation-training algorithms, or AI-driven feedback systems.

The analysis employed Braun and Clarke's (2006) thematic analysis framework, which comprises familiarization with the data, coding, theme identification, review, and refinement. NVivo 12 software was used to support the coding process, enabling the visualization and organization of recurring linguistic, pedagogical, and technological patterns. Themes were categorized into three core areas: (1) the effectiveness of deep learning tools in enhancing speaking

outcomes, (2) learner and teacher perceptions of AI-enhanced speaking instruction, and (3) pedagogical and ethical implications of DL integration. The rigor of the analysis was strengthened through peer debriefing with TEFL and educational technology specialists, as well as the maintenance of audit trails to ensure transparency in article selection and interpretation. Triangulation was accomplished by comparing diverse methodological approaches and cross-validating findings across multiple sources.

Despite its systematic nature, this study carries several limitations inherent to qualitative library research. First, the analysis relies on the availability and accessibility of published studies, which may exclude emerging or unpublished findings that could provide additional nuance. Second, the rapid evolution of deep learning technologies means that the literature may not fully capture the most recent algorithmic developments or commercial applications released after the study's cut-off period. Third, variations in research contexts and methodologies across the selected studies may influence comparability, limiting the generalizability of the synthesized themes. Nonetheless, by combining rigorous selection criteria, structured thematic analysis, and expert consultation, this study offers a credible and comprehensive overview of the current state of deep learning-supported speaking instruction in TEFL.

FINDINGS AND DISCUSSION

Findings

This section presents key findings from 30 studies published between 2020 and 2025 concerning the role of deep learning (DL) technologies in supporting speaking proficiency in TEFL classrooms. The findings are organized into three major themes: the improvement of speaking competencies, the affective impact on learners, and the pedagogical implications for classroom practice.

Deep Learning Enhances Core Speaking Competencies

Deep learning tools have significantly advanced learners' pronunciation, fluency, and rhythm in speaking tasks. Tools such as ELSA Speak, Google ASR, and other speech evaluation engines provide real-time corrective feedback on segmental features (such as consonants and vowels) and suprasegmental elements (including pitch, stress, and intonation). Several studies report that consistent engagement with these tools helps learners identify fossilized errors and develop self-monitoring strategies that are often absent in traditional teacher-centered classrooms

(Wang & Vásquez, 2022; Ngo et al., 2023). This immediate feedback allows learners to adjust their speech in real time, enhancing the accuracy of their output. The visual and auditory cues provided by these systems improve learners' awareness of subtle sound distinctions. Consequently, learners can internalize correct pronunciation patterns more effectively than through traditional instruction alone.

Additionally, learners who use DL technologies consistently report an increase in oral fluency and coherence over time. Unlike rigid textbook-driven practice, these applications adapt to each learner's pace and provide multiple opportunities to repeat, correct, and refine utterances. For instance, Chen et al. (2023) demonstrated that learners who used AI feedback systems three times a week for a month outperformed their peers in terms of speech rate and articulation clarity in classroom presentations. The adaptive nature of DL feedback supports long-term internalization of speech patterns. This personalized adaptation ensures that learners focus on their specific weaknesses rather than following a generic curriculum. It also encourages a growth mindset, as learners can progressively improve without fear of immediate evaluation from peers. Learners can focus intensively on their individual weaknesses, creating a more personalized learning experience. This repeated, adaptive practice strengthens memory retention and automates correct speech patterns. In turn, learners gain confidence in using English spontaneously in real-life situations.

Importantly, DL technologies serve as powerful tools in expanding speaking opportunities outside the classroom. In large or resource-limited TEFL environments, DL systems simulate one-on-one instruction, offering continuous and individualized practice. Learners can revisit problematic areas and track their improvement without needing the constant presence of a teacher. This accessibility empowers learners to take ownership of their speaking development, bridging the gap between in-class learning and autonomous language practice. The flexibility of these systems enables learners to practice anytime and anywhere, overcoming constraints of time and location. Digital records of learners' performance provide valuable feedback for both students and instructors to monitor progress. Moreover, the integration of gamification and adaptive challenges increases learner motivation and engagement over the long term.

DL Tools Reduce Speaking Anxiety and Improve Learner Motivation

Deep learning (DL) technologies play a significant role in reducing speaking-related anxiety among EFL learners. Many learners feel intimidated by face-to-face speaking tasks due to the fear of judgment or making mistakes, but AI-powered systems such as chatbots and speech apps create

a non-threatening environment where learners can rehearse and experiment with language at their own pace. Lee & Warschauer (2020) found that students practicing with chatbots demonstrated lower anxiety and greater classroom participation after only four weeks of regular use. Moreover, gamification elements embedded in many DL platforms—such as speech-based games, real-time scoring, badges, and progress graphs—maintain learner motivation by incentivizing practice and providing immediate recognition for improvement. Studies by Lu & Wang (2023) and Hwang et al. (2021) confirm that learners who receive gamified feedback are more likely to complete speaking tasks and return to the platform consistently.

Beyond motivation and comfort, DL tools also foster a growth mindset among EFL learners. Because these systems enable repeated practice and visible performance tracking, learners start to perceive speaking proficiency as a skill that can be developed gradually. This perception shift—from fixed ability to trainable competence—boosts learners' intrinsic motivation and increases their willingness to speak both inside and outside the classroom. In addition, the ongoing feedback reinforces learners' self-efficacy and encourages continuous reflection on their speaking progress. Consequently, DL technologies not only reduce anxiety and enhance engagement but also promote a learner-centered approach that supports autonomy, self-efficacy, and sustained improvement in speaking skills. Furthermore, the integration of DL tools encourages learners to set personal goals and take ownership of their learning journey, fostering long-term engagement and persistence.

Pedagogical Shifts and Challenges in Integrating DL into TEFL Classrooms

The implementation of DL tools in speaking instruction has reshaped teacher roles and classroom dynamics. Teachers now function more as facilitators and instructional designers, selecting appropriate tools, guiding students through AI-based feedback, and scaffolding speaking activities around DL-generated insights. Instead of offering all corrective input themselves, teachers help learners interpret automated feedback and plan their learning trajectories accordingly (Godwin-Jones, 2020). This shift allows teachers to focus on higher-order pedagogical tasks, such as fostering critical thinking and communicative competence, and encourages collaborative learning as students engage with both technology and peers to co-construct knowledge. DL systems are also still evolving in their ability to evaluate discourse-level features, such as turn-taking, repair strategies, and interactional competence, which are critical components of communicative speaking that go beyond pronunciation and fluency (Liu et al., 2022).

Ethical considerations emerge as another crucial factor in DL integration. Concerns regarding data privacy, voice recording storage, and algorithmic fairness—especially for speakers with regional or non-standard accents—must be addressed to ensure equitable use in global TEFL contexts. Without clear institutional policies and transparent algorithmic design, DL tools may inadvertently reinforce linguistic biases. Institutions should establish clear guidelines for data handling, provide training on ethical AI use, and promote inclusive algorithmic design. Moreover, fostering awareness and critical thinking among learners about AI-mediated learning is essential to ensure responsible and reflective use of DL tools. A balanced integration strategy is needed—one that enhances speaking instruction while upholding learner rights and maintaining the human touch in language education.

Table 1. Synthesis Table of DL Tools in TEFL Speaking Studies

Author & Year	DL Tools Used	Focus	Main Findings
Wang & Vásquez (2022)	ELSA Speak	Pronunciation, Stress, Fluency	Significant improvements in speaking proficiency
Ngo et al. (2023)	ASR Tools (Umum)	Pronunciation Accuracy, Motivation	Moderate positive effects on pronunciation accuracy and motivation
Lee & Warschauer (2020)	AI Chatbots	Speaking Anxiety, Willingness to Communicate	Reduction of anxiety, enhanced classroom participation
Huang et al. (2023)	AI Feedback Tools	Coherence, Complexity, Lexical Resource	Improvements in speaking performance and self-regulation
Chen et al. (2023)	AI Feedback Systems	Speech Rate, Articulation Clarity	Enhanced fluency and articulation clarity in presentations
Lu & Wang (2023)	AI Pronunciation Feedback	Speaking Motivation	Increased learner motivation and engagement through gamification
Liu et al. (2022)	DL Personalized Systems	Speaking Development, Personalization	Personalized pathways improved learners' speaking outcomes
Godwin-Jones (2020)	AI General Tools	Teacher Training & Integration	Teachers need clear frameworks for AI tool integration
Tsai (2021)	AI Feedback Tools	Learner Perception & Performance	Learners perceive AI feedback as beneficial

			but require support
Xie et al. (2020)	ASR Systems	Bias & Fairness in ASR	ASR systems show bias against non-standard accents

The table summarizes various studies that demonstrate the positive contribution of deep learning technology to improving speaking skills in the context of TEFL. In general, various DL tools—from ELSA Speak, AI-based chatbots, to ASR systems—have been shown to improve key aspects such as pronunciation, fluency, articulation clarity, and learning engagement and motivation. Additionally, some studies have shown that the use of DL can reduce speech anxiety and support more personalized learning. However, some studies have also highlighted challenges such as accent bias, teacher training needs, and ethical issues in the integration of these technologies. Overall, these findings confirm that DL has great potential in strengthening speech learning, although its implementation needs to be done carefully and in a targeted manner. Future research and practice should focus on developing guidelines and strategies to optimize DL integration while addressing its limitations.

Discussion

The findings of this study underscore the transformative potential of deep learning technologies in enhancing speaking proficiency among EFL learners. In general, the evidence indicates that DL-powered applications not only function as complementary learning tools but also signal a paradigmatic shift in pedagogical design, learner autonomy, and teacher roles in TEFL classrooms. Nevertheless, while most studies report substantial benefits, a critical reading reveals varying levels of effectiveness depending on learner context, tool design, and institutional readiness. Therefore, successful integration of DL technologies requires a holistic approach that combines careful selection of appropriate tools, alignment with curriculum goals, professional development for teachers, and continuous monitoring of learner progress to ensure both equitable access and meaningful learning outcomes.

Interpreting the Impact on Speaking Skills

Improvements in pronunciation, fluency, and speaking confidence reported across the reviewed studies align with several foundational theories in second language acquisition. Krashen's Affective Filter Hypothesis suggests that anxiety reduction facilitates greater language intake; DL-powered tools provide non-threatening feedback environments that help lower affective barriers (Krashen, 1985; Lee & Warschauer, 2020). Similarly, Swain's Output Hypothesis emphasizes the

importance of pushed output in developing linguistic accuracy; DL applications provide iterative practice spaces that encourage learners to refine articulation and syntactic control (Swain, 1995). Additionally, DL-based feedback resonates with Vygotsky's Zone of Proximal Development (ZPD), where learners rely on external scaffolding to perform beyond their current competence. Real-time corrective feedback, personalized difficulty adjustment, and performance tracking represent forms of AI-mediated scaffolding that extend the learner's ZPD in productive language skills.

However, not all studies demonstrate equally strong gains. Xie et al. (2020), for example, caution that ASR-based feedback may misjudge learners with non-standard accents, resulting in inconsistent scoring. Likewise, several studies note that DL tools tend to focus disproportionately on segmental features (e.g., phoneme accuracy) rather than communicative competence, reducing their ability to capture broader discourse-level speaking skills. These inconsistencies suggest that DL tools are not universally effective, and their benefits must be examined in relation to the learner's linguistic background, proficiency level, and frequency of exposure. Therefore, careful implementation, ongoing teacher guidance, and complementary classroom activities are essential to maximize the pedagogical potential of DL technologies in speaking instruction.

Pedagogical Realignment and Teacher Roles

The introduction of DL technologies has catalyzed a notable shift in teacher roles. Educators are increasingly serving as facilitators, mediators, and curators of technology-enhanced instructional environments, rather than as the sole sources of linguistic input. This transition is supported by frameworks such as the Technological Pedagogical Content Knowledge (TPACK) model, which underscores the need for teachers to harmonize technology, pedagogy, and content knowledge in meaningful ways (Mishra & Koehler, 2006). Several studies show that teachers who effectively integrate DL tools can design richer, more interactive speaking tasks that combine AI-assisted feedback with communicative classroom practices (Godwin-Jones, 2020; Hinojo-Lucena et al., 2020). Nevertheless, a major theme emerging from the literature is the lack of sufficient professional development. Tsai (2021) and Bai & Guo (2021) reveal that teachers often rely on trial-and-error due to insufficient guidance from institutions, resulting in suboptimal classroom integration.

Some studies further challenge overly optimistic portrayals of DL efficacy. For example, teacher–student interaction studies suggest that excessive reliance on technology may reduce opportunities for human negotiation of meaning, an essential component of interactionist SLA theories (Long, 1996). Moreover, students may develop a false sense of proficiency if automated

feedback is not supplemented with real-time conversational practice. In addition, the lack of social cues and immediate clarification in AI-mediated interactions can limit the development of pragmatic competence and collaborative problem-solving skills. This suggests that while DL systems may enhance individual practice, they cannot fully replace dialogic, socially constructed learning opportunities, which are central to communicative competence. Therefore, effective TEFL instruction requires integrating DL tools in ways that complement, rather than replace, meaningful human interaction.

Ethical, Technical, and Pedagogical Limitations

Despite the promising findings, the adoption of DL technologies raises several ethical, technical, and pedagogical concerns. Ethically, issues related to privacy, data ownership, and algorithmic transparency remain unresolved in many educational contexts. Learners' voice data stored on external servers can pose serious security risks unless robust institutional policies are enforced. Algorithmic bias is another concern, as DL models trained on limited datasets may marginalize learners with minority linguistic backgrounds (Xie et al., 2020). Furthermore, educators and institutions must consider informed consent and the potential psychological impact of constant monitoring on learners. Ethical guidelines should be clearly communicated to students to build trust in AI-assisted learning. In addition, collaboration with developers is needed to ensure that DL systems adhere to inclusive and transparent design principles.

Technically, DL systems remain more proficient at analyzing micro-level features, such as pronunciation accuracy, rhythm, and stress patterns. Still, they struggle to assess macro-level dimensions, including coherence, pragmatic appropriateness, and turn-taking strategies. From a pedagogical standpoint, this limitation supports the argument for hybrid approaches, allowing DL tools to handle repetitive diagnostic tasks. At the same time, teachers focus on the complex, socially embedded aspects of spoken interaction. Moreover, technical limitations may vary depending on the quality of the algorithms and the diversity of the training data. Continuous updates and calibration are necessary to maintain reliability across different learner populations. Finally, integration of multimodal data, such as gesture and facial expressions, could enhance the assessment of communicative competence beyond speech alone.

Some studies also raise concerns about learner overdependence on automated feedback. When learners rely heavily on DL scoring, they may become less inclined to engage in natural communication or real conversational tasks, limiting the internalization of authentic interactional

norms. This critique echoes warnings in cognitive load theory and autonomous learning research, suggesting that technology must be carefully scaffolded to avoid cognitive oversimplification or passive dependency. In addition, excessive reliance on AI feedback may reduce opportunities for peer-to-peer learning and collaborative problem-solving, which are crucial for developing pragmatic skills. Teachers should balance automated guidance with interactive tasks to maintain student engagement in real-world discourse. Lastly, cultivating learner metacognition about their speaking progress can prevent overreliance and promote reflective, self-directed practice.

Taken together, these insights suggest that while DL tools offer significant pedagogical advantages, their effectiveness is mediated by instructional quality, teacher preparedness, theoretical grounding, and ethical safeguards. To maximize long-term learning outcomes, TEFL classrooms should adopt a balanced integration strategy—one that combines DL-supported personalization with human-centered communicative pedagogy, ensuring that technology enhances rather than replaces meaningful language use. Such an approach ensures that learners benefit from adaptive feedback while still engaging in authentic social interaction. Teachers play a crucial role in interpreting AI-generated data, designing meaningful communicative tasks, and scaffolding higher-order language skills. Moreover, ongoing evaluation and refinement of both technological tools and instructional practices are essential to sustain effectiveness and equity in diverse learning contexts.

CONCLUSION

This study demonstrates that deep learning technologies are reshaping the teaching and learning of speaking skills in TEFL classrooms by providing personalized, data-driven feedback that significantly enhances learners' pronunciation, fluency, accuracy, and confidence. Through a synthesis of 30 studies, the findings affirm that DL-powered tools not only reduce speaking anxiety and promote learner autonomy but also shift instructional dynamics toward more learner-centered environments. At the same time, the role of teachers is evolving, requiring greater digital literacy and the ability to integrate AI tools meaningfully into pedagogy. Despite these advantages, key concerns—including data privacy, algorithmic fairness, and the limited capability of DL systems to assess discourse-level features—remain persistent challenges. These findings underscore that deep learning is not merely a technological supplement; it is a transformative pedagogical catalyst with implications that extend beyond classroom practice into broader curriculum development, teacher education, and policy-level reform.

To ensure the effective and ethical integration of deep learning in TEFL speaking instruction, several strategic directions must be pursued at various levels of the educational ecosystem. Teachers should continuously develop their technological and pedagogical competencies, employing DL tools as supportive resources while maintaining the human interaction essential for communicative and pragmatic development. Educational institutions must provide adequate infrastructure, digital training, and clear ethical guidelines to safeguard learner data and promote the responsible use of AI. Meanwhile, researchers and developers should collaborate to design more inclusive and culturally responsive DL systems capable of evaluating higher-level discourse skills while addressing biases that disadvantage learners with diverse linguistic profiles. Through coordinated efforts across teachers, institutions, researchers, and policymakers, deep learning technologies can be harnessed to build more equitable, scalable, and future-oriented models of language education.

REFERENCES

Bai, J., & Guo, H. (2021). Teachers' perceptions of AI-assisted English language teaching in China. *Journal of Educational Technology Development and Exchange*, 14(1), 1–16.

Chen, Y., Chen, Z., & Lin, C. (2023). Effects of AI-assisted pronunciation training on EFL learners' oral fluency. *Language Learning & Technology*, 27(1), 45–61.

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford University Press.

Godwin-Jones, R. (2020). Artificial intelligence and language learning: Ready for prime time? *Language Learning & Technology*, 24(3), 3–11.

Goh, C. C. M., & Burns, A. (2012). *Teaching Speaking: A Holistic Approach*. Cambridge University Press.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

He, Y., Zhang, H., & Li, W. (2023). Integrating deep learning for interactive EFL learning: A review of applications and implications. *Educational Technology Research and Development*, 71(2), 589–607.

Hinojo-Lucena, F. J., Aznar-Díaz, I., Cáceres-Reche, M. P., & Romero-Rodríguez, J. M. (2020). Artificial intelligence in education: A bibliometric study on its impact in the scientific literature. *Education Sciences*, 10(12), 1–12.

Huang, Y., Wang, L., & Chen, X. (2023). The effects of AI-based scaffolding on EFL learners' speaking performance and self-regulated learning. *Interactive Learning Environments*, 31(2), 1–15.

Jiang, Y., & Song, X. (2021). Intelligent language learning systems: Enhancing EFL speaking performance using neural networks. *ReCALL*, 33(2), 137–153.

Krashen, S. D. (1985). *The Input Hypothesis: Issues and Implications*. Longman.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.

Lee, J., & Warschauer, M. (2020). The effects of chatbot-based learning on EFL learners' speaking skills. *Language Learning & Technology*, 24(1), 1–15.

Li, W., & Zou, B. (2022). Review of AI-powered feedback systems in EFL speaking instruction. *Computer Assisted Language Learning*, 35(4), 463–479.

Lin, J. M. C., & Lan, Y. J. (2022). Exploring the effects of deep learning in mobile-based speaking activities. *British Journal of Educational Technology*, 53(6), 1452–1471.

Littlewood, W. (2004). The Task-Based Approach: Some Questions and Suggestions. *ELT Journal*, 58(4), 319–326.

Liu, Y., Chen, L., & Yao, Z. (2022). Deep learning integration in students' speaking development: A personalized approach. *Frontiers in Psychology*, 13, 929175.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 413–468). Academic Press.

Lu, C., & Wang, J. (2023). Enhancing speaking motivation through AI-based pronunciation feedback. *Computer Assisted Language Learning*, 36(1), 77–93.

Luo, H., & Li, Y. (2020). A study on speech anxiety and self-efficacy of EFL learners using AI chatbots. *Journal of Language Teaching and Research*, 11(3), 382–390.

Nation, P., & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking*. Routledge.

Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2023). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 35(1), 1–19.

Ortega, L. (2009). *Understanding Second Language Acquisition*. Routledge.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and Practice in Applied Linguistics* (pp. 125–144). Oxford University Press.

Tsai, Y. H. (2021). EFL learners' experiences with AI-based feedback: Perceptions and performance. *Educational Technology & Society*, 24(1), 95–108.

VanPatten, B. (2015). *While We're on the Topic: BVP on Language, Acquisition, and Classroom Practice*. ACTFL.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.

Wang, Y., & Vásquez, C. (2022). Efficacy of AI-powered pronunciation training on EFL learners' speaking proficiency. *Computer Assisted Language Learning*, 35(3), 1–20.

Xie, X., Liu, C., & He, Q. (2020). Fairness and bias in automatic speech recognition: An educational ethics perspective. *AI & Society*, 35(2), 375–388.