International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

INTERNATIONAL JOURNAL OF MODERN EDUCATION (IJMOE)
www.ijmoe.com

GLOBAL ACADEMIC EXCELLENCE

FROM CLASSROOM TO CAREER: LEVERAGING TEXT-TO-SPEECH TECHNOLOGY TO ENHANCE ESL PRONUNCIATION, VOCABULARY, AND EMPLOYABILITY IN HIGHER EDUCATION

Latisha Asmaak Shafie[1], Nor Azira Mohd Radzi[2*], Razlina Razali[3], Nor Alifah Rosaidi[4], Raisa Rastom[5], Hafizah Hajimia[6].

[1] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email:ciklatisha@uitm.edu.my
[2] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email: norazira202@uitm.edu.my
[3] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email: razlinarazali@uitm.edu.my
[4] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email:alifah.rosaidi@uitm.edu.my
[5] Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia
Email:raisa@uitm.edu.my
[6] School of Education, Universiti Utara Malaysia, Sintok, Kedah, Malaysia
Email:hafizah.hajimia@uum.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 27.03.2025
Revised date: 14.04.2025
Accepted date: 15.05.2025
Published date: 05.06.2025

To cite this document:

Shafie, L. A., Radzi, N. A. M., Razali, R., Rosaidi, N. A., Rastom, R., & Hajimia, H. (2025). From Classroom To Career: Leveraging Text-To-Speech Technology To Enhance ESL Pronunciation, Vocabulary, And Employability In Higher Education.

Abstract:

English language proficiency is a critical determinant of employability in globalized workplaces, yet many university students struggle to achieve competency in vocabulary and pronunciation. This study investigates the integration of Text-to-Speech (TTS) tools into English language pedagogy to address these gaps. Employing a qualitative case study design, the research developed and implemented TTS-supported lesson plans focusing on 1,000 employability-focused vocabulary items. Data were collected via classroom observations, lecturer journals, and semi-structured interviews with five lecturers and their students. Thematic analysis revealed five key outcomes: (1) enhanced vocabulary retention through consistent pronunciation models, (2) improved pronunciation accuracy via native-like intonation and stress patterns, (3) increased learner autonomy through self-regulated practice, (4) heightened engagement via interactive TTS features, and (5) strengthened employability skills such as critical thinking and communication confidence. The findings underscore TTS tools as transformative artefacts within Vygotsky's sociocultural framework, bridging pedagogical gaps and fostering self-

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

International Journal of Modern Education, 7 (25), 120-132.

DOI: 10.35631/IJMOE.725009.

regulation. Recommendations include institutional adoption of TTS tools, educator training, and longitudinal studies to assess long-term impacts. This innovation offers a scalable solution to align language education with global workforce demands.

Keywords:

Vocabulary Acquisition, Pronunciation Proficiency, TTS Tools, ESL Learners, Sociocultural Theory, Employability

Introduction

In an era where English proficiency is a gateway to global career opportunities, Malaysian graduates face persistent challenges in meeting employer expectations. National surveys highlight that 81% of employers cite poor communication skills as a hiring barrier (World Bank, 2014), with deficient English proficiency ranking among the top reasons for graduate unemployability (Malaysian Employers Federation, 2016). Despite policy initiatives like Malaysia's English Language Roadmap 2015–2025, students still struggle with vocabulary gaps and mispronunciation, which hinder workplace communication and global competitiveness (Zainuddin et al., 2019). Vocabulary deficits and mispronunciation persist as significant challenges for students, substantially hindering their capacity for effective workplace communication and consequently diminishing their global competitiveness. Such linguistic deficiencies may precipitate misunderstandings, misinterpretations, and a lack of precision in professional contexts, ultimately impeding career progression and restricting opportunities for international collaboration.

Research attributes these challenges to inadequate pedagogical focus on pronunciation (Lin & Ooi, 2023) and limited access to authentic language models (Levis, 2018). The existing body of literature has shown that text-to-speech (TTS) tools can facilitate students' language skills, such as pronunciation, reading comprehension, vocabulary and writing (Al-Jarf, 2022; Perez et al., 2021). While Computer-Assisted Pronunciation Training (CAPT) tools like TTS have shown promise in addressing these issues (Al-Jarf, 2022; Fitria, 2022), their integration into the Malaysian higher education curricula remains underexplored. This study bridges this gap by examining how TTS tools, when embedded into lesson plans, enhance vocabulary and pronunciation mastery while fostering employability skills.

Theoretical Framework

Guided by Vygotsky's sociocultural theory (SCT), this study positions TTS tools as **artefacts** facilitating **object regulation** (technology-mediated learning) and **other regulations** (lecturer feedback). Vygotsky's sociocultural theory (SCT) offers a theoretical framework for elucidating the mediatory influence of social interaction and cultural artefacts on learning and developmental processes. Within the purview of this investigation, Text-to-Speech (TTS) tools are conceptualized as artefacts integral to the modulation of the learning trajectory. These tools engender two fundamental modalities of regulation: object regulation and other regulation.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

Object regulation pertains to the technological mediation of learning through the alteration of learners' engagement and manipulation of information. TTS tools, through the transformation of textual content into auditory form, are posited to enhance comprehension, attentional focus, and engagement, thus facilitating object regulation. Other regulation encompasses the role of social interaction and feedback in the instructional paradigm. In the present context, lecturer feedback serves as a mechanism of other regulation, directing and scaffolding learners' utilisation of TTS tools and their broader learning experience.

The incorporation of TTS tools into the pedagogical environment, in conjunction with the provision of judicious feedback, enables lecturers to harness both object and other regulations to cultivate a more conducive and efficacious learning milieu for students. This methodological approach is consonant with Vygotsky's accentuation of the socio-cultural determinants of learning and underscores the potential of technology to mediate and augment the educational process. SCT posits that learners achieve self-regulation through scaffolded interactions with tools and social guidance (Lantolf et al., 2015). TTS tools provide immediate auditory feedback, enabling students to internalize pronunciation norms and vocabulary usage independently—a process which is critical for workplace readiness.

Literature Review

The Pronunciation Gap in ESL Education

Pronunciation remains a critical yet underprioritized component of ESL curricula, particularly in higher education. One of the main reasons that contribute to this is that pronunciation is often treated as an as-needed response to students' performance rather than being part of the main curriculum (Pennington, 2021). Research highlights its marginalization in favour of macro-skills like reading and writing, despite its direct impact on intelligibility and professional communication (Metruk, 2024). For instance, Malaysian educators cite insufficient phonetic training and time constraints as systemic barriers to effective pronunciation instruction, leading to learners' deficiencies in stress patterns and intonation (Wahid & Sulong, 2013; Nguyen & Newton, 2020). These gaps perpetuate communication challenges in workplace settings, where unclear pronunciation can hinder employability (Metruk,2024; Dja'far & Hamidah,2024).

A synthesis of recent studies underscores the limited exposure to diverse speech models in traditional classrooms, which restricts learners' ability to adapt to global accents. For example, Al-Shami's (2024) mixed-method study on High-Variability Phonetic Training (HVPT) integrated with TTS technology revealed that learners who were exposed to varied voices showed significant improvements in comprehensibility and accentedness compared to those using single-voice models. This aligns with findings from Dja'far and Hamidah (2024), who demonstrated that AI-based speech recognition tools like Google Read Along improved pronunciation accuracy by 23% in higher education cohorts, emphasizing the need for scalable, technology-driven solutions.

TTS Technology as a Pedagogical Catalyst

Text-to-speech (TTS) tools such as NaturalReader and Amazon Polly are revolutionizing pronunciation training by providing authentic, customizable speech models. These tools address the limitations of traditional instruction by offering learners immediate feedback, self-paced practice, and exposure to prosodic features like word stress and intonation. For instance,

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

Sain and Cobar (2023) found that TTS applications significantly enhanced learners' ability to master word stress patterns through repetitive, interactive drills (Fitria, 2022).

Recent empirical studies highlight TTS's dual role in fostering phonological awareness and learner autonomy. Al-Shami's (2024) research on TTS-integrated HVPT demonstrated that learners using varied voice models achieved measurable gains in both segmental (e.g., past-ed allomorphs) and suprasegmental (e.g., fluency) features. Similarly, Dja'far and Hamidah (2024) reported a high N-Gain score (0.7646) in pronunciation proficiency among students using TTS tools, outperforming conventional methods. These findings are corroborated by systematic reviews indicating that AI-driven tools like Elsa Speak and EnglishCentral enhance pronunciation through real-time feedback and adaptive learning pathways (Peña-Acuña, 2024).

Vocabulary Mastery and Its Role in Employability

Vocabulary proficiency is a linchpin of professional communication, directly influencing employability outcomes. Studies show that students with robust vocabularies excel in job interviews and workplace interactions, as they can articulate ideas clearly and contextually (Bjørge et al., 2020). However, traditional vocabulary instruction often fails to contextualize terms within industry-specific scenarios, limiting their applicability (Şahan & Şahan, 2023).

Emerging research advocates for integrating TTS tools into vocabulary training to bridge this gap. For example, AI-driven platforms like Lingvist and Busuu adapt content to learners' progress, embedding terms within professional dialogues and simulations. A quasi-experimental study by Liao et al. (2023) on generative AI in ESL curricula found that tools like ChatGPT improved contextual vocabulary retention by 34% through scenario-based practice, such as mock negotiations and client interactions. Additionally, Huang et al. (2016) demonstrated that speech-to-text recognition (STR) apps enhance vocabulary acquisition by aligning spoken practice with written reinforcement, fostering both fluency and accuracy.

Synthesis and Future Directions

While TTS technologies show promise, gaps persist in their application to higher education and employability contexts. Current research predominantly focuses on basic literacy, with limited exploration of advanced professional communication skills. For instance, longitudinal studies are needed to assess the sustainability of TTS-driven improvements in pronunciation and vocabulary. Furthermore, equitable access to these technologies remains a concern, as socio-economic disparities often limit adoption in resource-constrained institutions.

Future studies should investigate hybrid models that combine TTS with human mentoring to address nuanced communication challenges (e.g., cultural pragmatics). Additionally, curriculum designers must prioritize industry-aligned scenarios, such as using TTS to simulate job interviews or client presentations, to enhance employability outcomes.

Methodology

The current research endeavour employs a qualitative case study approach to delve into the efficacy of Text-to-Speech (TTS) tools as aids in bolstering English vocabulary acquisition and refining pronunciation prowess within the realm of higher education. This methodological approach is underpinned by a theoretical framework that draws upon established principles from the domains of language acquisition and educational technology, with a pronounced emphasis on the pragmatic facets of implementation and evaluation.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

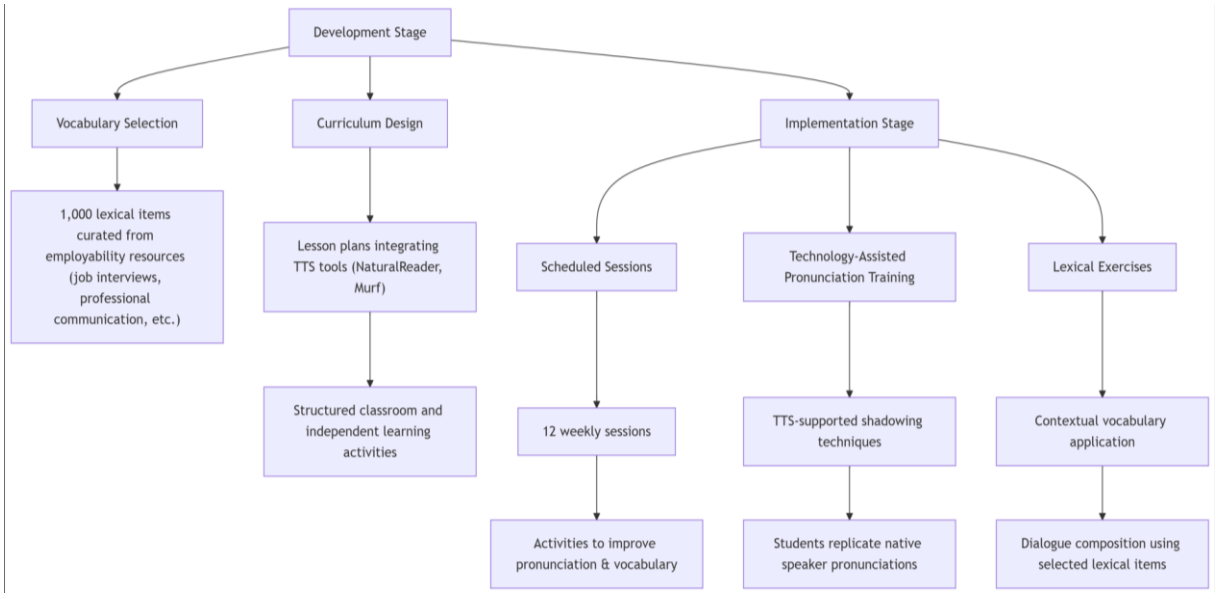
In the forthcoming discourse, a comprehensive exposition of the research design will be delineated, and substantiated by pertinent scholarly literature. This elaboration will encompass a meticulous elucidation of the research objectives, the rationale for the selection of the qualitative case study methodology, the specific TTS tools that will be employed, the participant selection criteria, the data collection instruments and procedures, the data analysis techniques, and the anticipated outcomes of the study. Additionally, the ethical considerations inherent in the research process will be explicitly addressed, ensuring the protection of participant rights and the maintenance of research integrity.

Research Design

A qualitative case study was conducted at Universiti Teknologi MARA, Perlis Branch, to examine the implementation of text-to-speech (TTS) tools within English language instruction. The study population comprised five faculty members and 15 undergraduate students from five distinct academic courses. The process is being simplified in Figure 1 below.

Intervention Design

Figure 1: Intervention Design



1. Development Stage:

- **Vocabulary Selection:** A selection of 1,000 lexical items was methodically curated and collated from a diverse range of materials pertaining to employability skills. These materials comprised resources on job interviews, professional communication, and other pertinent domains.
- **Curriculum Design:** Detailed lesson plans were formulated to integrate text-to-speech (TTS) technologies, including NaturalReader and Murf, into both classroom instruction and independent learning exercises. These plans were designed to offer a structured and stimulating educational experience for students.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

2. Implementation Stage:

- **Scheduled Sessions:** Instructors conducted 12 weekly sessions that incorporated a variety of activities intended to augment students' pronunciation and vocabulary proficiency.
- **Technology-Assisted Pronunciation Training:** The sessions included pronunciation training supported by TTS tools, such as shadowing techniques, wherein students listened to and replicated native speaker pronunciations.
- **Lexical Exercises:** Furthermore, the sessions incorporated lexical exercises that emphasized contextual application, such as the composition of dialogues incorporating the designated vocabulary items.

Data Collection and Analysis

Data were collected from classroom observations, lecturers' journals and semi-structured interviews with five lecturers and 15 students at the UiTM Perlis Branch. Thematic analysis was used to analyse data to identify the themes and to make sense of the meanings and contexts (Braun & Clarke, 2013). To begin with, classroom observations were conducted to record students' interactions with TTS tools. Simultaneously, lecturers recorded their reflective accounts of pedagogical challenges and successes in their lecturers' journals. After that, semi-structured interviews were conducted to explore post-intervention insights from five lecturers and 15 students. Data were collected inductively using the manual process.

Alignment with Research Objectives

The current study aimed to explore the effectiveness of Text-to-Speech (TTS) tools in enhancing vocabulary and pronunciation among ESL learners in higher education. A qualitative case study was ideal because it allowed for an in-depth investigation of a specific phenomenon (the use of TTS tools) within a real-world context (university ESL classrooms). Not only that, it was focused on understanding the experiences, perceptions, and outcomes of both lecturers and students, which aligned with the study's goal of evaluating the impact of TTS tools on learning outcomes and employability skills (Liu, 2021).

A corpus of 1,000 professional vocabulary terms was curated, focusing on employability skills (e.g., business communication, technical jargon) derived from industry reports and academic literature. Lesson plans were designed to incorporate TTS tools, such as Natural Reader, which allows customization of pronunciation models (e.g., British/American accents) and playback speed to accommodate diverse learner needs. The design drew inspiration from studies demonstrating TTS's efficacy in providing consistent auditory input for pronunciation practice. For instance, Natural Reader's ability to generate MP3 files enabled students to access pronunciation models outside the classroom, fostering self-paced learning.

Focus on Contextual Understanding

The study was situated in a specific educational setting (Universiti Teknologi MARA), and the qualitative case study methodology was well-suited for examining how TTS tools were integrated into lesson plans and how they influenced teaching and learning practices. This approach has provided rich, contextual insights that quantitative methods might overlook.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

The five lecturers integrated the TTS-supported lesson plans into their courses, blending in-class activities (e.g., listening exercises, pronunciation drills) with self-directed tasks. The tools were used to:

- Model native-like pronunciation: TTS provided standardized auditory examples, reducing reliance on variable instructor accents.
- Facilitate contextual learning: Vocabulary was embedded in dialogues and scenarios mirroring professional settings, aligning with task-based language teaching principles.

In addition, the lecturers received training on optimizing TTS features, such as adjusting speech speed and accent preferences, to address individual student needs.

Data Collection Methods

The use of a qualitative case study methodology supports the use of multiple data collection methods, which are essential for capturing diverse perspectives and triangulating findings. Data were collected from classroom observations, lecturers' journals and semi-structured interviews. Thematic analysis was used to analyse data to identify the themes and to make sense of the meanings and contexts (Braun & Clarke, 2013). To begin with, classroom observations were conducted to record students' interactions with TTS tools. Simultaneously, lecturers recorded their reflective accounts of pedagogical challenges and successes in their lecturers' journals. After that, semi-structured interviews were conducted to explore post-intervention insights from five lecturers and 15 students. Data were collected inductively using a manual process.

Data Analysis

Thematic analysis was used to identify recurring patterns, such as "enhanced self-regulation" and "accent adaptability." Transcripts were coded manually, with triangulation across data sources to ensure validity. This approach aligns with qualitative frameworks in TTS research, which emphasize learner experiences and pedagogical adaptability.

Emphasis on Thematic Analysis

The study used thematic analysis to identify patterns and themes related to the impact of TTS tools on vocabulary acquisition, pronunciation proficiency, and employability skills. This analytical approach is well-suited to qualitative case studies, as it allows researchers to organize and interpret complex data in a way that highlights key insights and practical implications.

Practical Implications

The qualitative case study methodology employed in the current study enables the researchers to provide actionable recommendations for educators and policymakers. By focusing on a specific context, the study can offer tailored insights into how TTS tools can be effectively integrated into ESL curricula to enhance learning outcomes and prepare students for the workforce (Fitria, 2022).

The qualitative case study methodology is the most suitable approach for this research because it aligns with the study's objectives, provides rich contextual insights, and supports the use of multiple data collection methods. By focusing on a specific educational setting and exploring the experiences of both lecturers and students, this methodology enables the researchers to generate meaningful findings and practical recommendations for enhancing ESL education through TTS tools (Cardoso et al., 2015).

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

Results And Discussion

The investigation revealed significant progress in learners' linguistic proficiency and their readiness for professional integration. These advancements were characterized by five principal themes derived from the empirical data. These themes presumably comprise particular domains in which learners exhibited marked development, such as improved communicative competence, heightened self-assurance in occupational environments, and a more profound comprehension of the skills and expertise necessary for successful employment. The results indicate that the pedagogical interventions or approaches implemented in the study exerted a favourable influence on learners' linguistic capabilities and general employability, thereby underscoring their applicability for wider deployment in educational and vocational training settings. :

Enhanced Vocabulary Acquisition

The results indicate that students exhibited higher retention rates for professional vocabulary compared to traditional methods as shown in Table 1. TTS tools provided repetitive, context-rich auditory input, reinforcing word meanings and usage. Lecturers noted that features like adjustable playback speed allowed students to process complex terms at their own pace, a finding consistent with studies on TTS's role in reducing cognitive load. Students exhibited improved retention through auditory reinforcement. Similar to the findings of Abdel-Reheem Amin (2024), lecturers noted that TTS tools contextualized vocabulary in professional scenarios (e.g., "negotiation," "deadline"), aiding nuanced understanding.

Table 1: Enhanced Vocabulary Acquisition

Students	Lecturers
I can learn different accents on my own. I learned many new words. I learn to listen carefully to the individual sounds and stressed patterns.	The students are exposed to different pronunciations. The students are exposed to many new words.

Improved Pronunciation Skills

The findings indicate that students reduced phonetic errors after using TTS models as illustrated in Table 2. The tools' native-speaker simulations helped learners grasp intonation and stress patterns, addressing common L2 challenges such as vowel length and consonant clusters. For example, mimicking British English RP (Received Pronunciation) via TTS improved clarity in professional presentations. TTS models reduced fossilized errors (e.g., mispronouncing "strategy" as /stra-tuh-jee/). Students reported increased confidence in oral presentations, aligning with findings by Al-Jarf (2022). Sain and Cobar (2023) also discovered that the repetition technique is one of the main reasons students become confident with their pronunciation. Bonces (2019) also cited accurate pronunciation as one of the learners' developments in using TTS for oral reading fluency besides linking sounds and reading timing.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

Table 2:Improved Pronunciation Skills

Students	Lecturers
The natural-sounding voices are not scary. I learn the tone and intonation of words in sentences. I mimic the TTS voice as closely as I can.	They now know where to practice their pronunciation skills. The students know how to pronounce their words better. Different tones/intonation will give different meanings. It trains the mouth's muscles and auditory perception.

Enhanced Learner Autonomy

The findings suggest that students utilised TTS-generated MP3 files for independent practice, reflecting principles of self-regulated learning as indicated in Table 3. One participant noted, "I could practice anytime, without feeling judged," highlighting TTS's role in reducing classroom anxiety—a phenomenon observed in prior TTS studies. In agreement with van Compernolle and Williams (2013), TTS-enabled self-practice outside class hours empowered students to address individual weaknesses, aligning with SCT's emphasis on self-regulation. Students become confident to learn on their own without teachers and classmates (Sain & Cobar, 2023). The use of TTS as an out-of-class pedagogical tool could increase in-class time, thus more focus could be placed on other important tasks such as feedback and real-life communication (Van Lieshout & Cardoso, 2022).

Table 3: Enhanced Learner Autonomy

Students	Lecturers
I could practice anytime, without feeling judged. Practice makes perfect. I learned how to break the words into syllables.	They are more independent and learn to regulate their learning. They can try to learn from mistakes. They can use the phonetic spellings and can break up words into "chunks".

Enriched Learning Engagement

The research highlights gamified TTS activities (e.g., accent imitation challenges) increased participation as depicted in Table 4. The interactive nature of tools like Natural Reader aligned with findings that multimodal inputs (audio + text) boost motivation. Lecturers also reported higher attendance in TTS-integrated sessions. The use of TTS increased students' enthusiasm

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

for learning the targeted language (Widyana et al.,2022). Interactive features like voice customization (e.g., British vs. American accents) heightened interest. One student remarked, "Imitating the TTS voice made practice feel like a game." The finding is similar to the study conducted by Fatima (2024), who discovered that students were more engaged in improving their pronunciation when they listened to a native-like English accent.

Table 4: Enriched Learning Engagement

Students	Lecturers
Imitating the TTS voice made practice feel like a game. I love the practices. I learned individual sounds and how they form words.	My students are more involved. The students do not feel pressured. The practices have customised features.

Development of Employability Skills

The results reveal that by practising job interview simulations and technical presentations with TTS, students gained confidence in professional communication as reflected in Table 5. This aligns with research linking pronunciation accuracy to career success in globalized industries. Additionally, problem-solving skills emerged as students troubleshoot software issues independently. The findings are consistent with the study by Şahan and Şahan (2023), which found that simulations of workplace scenarios (e.g., mock client calls) using TTS tools enhanced problem-solving and critical thinking skills prioritized by employers.\

Table 5: Development of Employability Skills

Students	Lecturers
I now know many words about jobs. I discover new things about my job.	My students are exposed to workplace vocabulary. I understand many words about work (jargon).

Synthesis And Implications

This study underscores TTS's potential as a scalable tool for addressing gaps in pronunciation training and vocabulary acquisition, particularly in resource-limited settings. Future research could explore AI-driven TTS adaptations for dialect-specific training or integrate speech recognition for real-time feedback.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

Conclusion and Recommendations

This study demonstrates that TTS tools when strategically integrated into curricula, address pronunciation and vocabulary gaps while fostering employability. Key recommendations include:

1. Institutional Adoption: Universities should invest in TTS licenses and training workshops.
2. Curriculum Design: Embed TTS tools across disciplines (e.g., business communication).
3. Future Research: Longitudinal studies to assess retention and workplace integration.

TTS tools offer a pathway to equitable, future-ready education by aligning pedagogical innovation with labour market needs. This study underscores the transformative potential of Text-to-Speech (TTS) tools in bridging pedagogical gaps in pronunciation and vocabulary while simultaneously fostering employability skills in higher education. By integrating TTS into curricula, institutions can create inclusive, technology-driven learning environments that align with labour market demands. The findings affirm that TTS-supported instruction not only addresses linguistic challenges but also cultivates critical soft skills—self-directed learning, problem-solving, and adaptability—identified as vital for modern workplaces.

To maximize the impact of TTS tools, the following evidence-based strategies are proposed:

Institutional Adoption and Training by investing in TTS Licenses

Universities should prioritize procuring versatile TTS platforms (e.g., Natural Reader, Speechify) that offer multilingual and accent customization features. These tools align with the Jisc Employability Toolkit's call for embedding technology into curriculum delivery.

Universities should conduct workshops for educators by providing training programs to equip lecturers with technical and pedagogical skills for TTS integration. Educator proficiency is critical to mitigating implementation barriers. At the same time, universities need to give access to infrastructure to ensure TTS tools are compatible with existing LMS platforms (e.g., Moodle, Canvas) to support seamless remote learning, as advocated by curriculum development tools like CurricuPlanner.

Curriculum Design and Cross-Disciplinary Integration

Embed TTS Across Disciplines: Design discipline-specific modules (e.g., engineering, healthcare) using tools like TeachFlow Pro or SyllaSoft, which support multimedia integration and real-time collaboration. For instance, business communication courses could simulate job interviews with TTS-generated dialogues. Furthermore, universities need to adopt universal Design for Learning (UDL) by integrating TTS with UDL principles to accommodate diverse learning styles. Zascavage and Winterman (2009) highlight how assistive technologies like TTS align with UDL frameworks to promote engagement. At the same time, develop interactive TTS activities (e.g., accent imitation challenges) to mirror the engagement strategies of tools like EduDesign Hub, which emphasize interactivity.

Future Research Directions

Longitudinal Studies are needed to track the long-term retention of pronunciation skills and workplace application of TTS-acquired vocabulary. Now, with AI-driven personalization, one must explore transformer-based models (e.g., LaBSE) to tailor TTS content to individual learner needs, as proposed in arXiv's curriculum recommendation paradigm (Zascavage & Winterman, 2009). In a way, this will help the cross-cultural adaptation to investigate TTS's

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009

efficacy in non-English contexts, addressing translation ambiguities through language-switching strategies.

By aligning TTS integration with institutional policies, curriculum innovation, and labour market needs, universities can create equitable, future-ready educational ecosystems. Collaborative efforts between educators, policymakers, and industry stakeholders will be essential to scale these practices globally. Technology's true value lies in its ability to empower learners—transforming classrooms into incubators for both linguistic excellence and professional resilience.

Acknowledgement
The authors wish to thank the Innovation in Teaching and Learning Grant (I-TALENT), UiTM Perlis, for funding this research.

References

Abdel-Reheem Amin, E. (2024). EFL students' perception of using AI text-to-speech apps in learning pronunciation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4746800

Aizawa, I., Rose, H., Thompson, G., & Curle, S. (2020). Beyond the threshold: Exploring English language proficiency, linguistic challenges, and academic language skills of Japanese students in an English medium instruction programme. *Language Teaching Research, Advance online publication*. https://doi.org/10.1177/1362168820965510

Al-Jarf, R. (2022). Text-to-speech software for promoting EFL freshman students' decoding skills and pronunciation accuracy. *Journal of Computer Science and Technology Studies*, *4*(2), 19–30.

Al-Shami, F. (2024). *Exploring high-variability phonetic training through TTS technology* (Master's thesis). Concordia University.

Bjørge, A. K., Baciu, S. T., & Whittaker, S. (2020). Student perspectives on future careers: What is the role of languages? A Romanian case study. *Studia Universitatis Babeș-Bolyai Philologia*, 49–64.

Bonces, J. A. R. (2019). Implementing text-to-speech technology as a means of enhancing L2 reading fluency. *The International Education & Learning Review*, *1*(2), 2019. ISSN 2659-5915.

Braun, V., & Clarke, V. (2013). Successful qualitative research: A practical guide for beginners. Thousand Oaks, CA: SAGE.

Cardoso, W., Smith, G. F., & Fuentes, C. G. (2015). Evaluating text-to-speech synthesizers.

Cox, J. L., Henrichsen, L. E., Tanner, M. W., & McMurry, B. L. (2019). The needs analysis, design, development, and evaluation of the "English Pronunciation Guide: An ESL Teachers' guide to pronunciation teaching using online resources." *The Electronic Journal for English as a Second Language*.

Dja'far, V. H., & Hamidah, F. N. (2024). Improving English pronunciation via AI-based speech recognition. Ethical Lingua.

Fatima, S. (2024). Impact of accent and pronunciation on listening and speaking skills of university students. *International Journal of Contemporary Issues in Social Sciences*, *3*(2), 365–377.

Fitria, T. N. (2022). Utilizing text-to-speech technology: Natural Reader in teaching pronunciation. *JETLEE: Journal of English Language Teaching, Linguistics, and Literature*, *2*(2), 70–78.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume 7 Issue 25 (June 2025) PP. 120-132
DOI: 10.35631/IJMOE.725009
Gottardi, W., Almeida, J., & Tumolo, C. (2022). Automatic speech recognition and text-to-speech technologies for L2 pronunciation improvement. Texto Livre.

Levis, J. M. (2018). Intelligibility, oral communication, and the teaching of pronunciation. Cambridge, UK: Cambridge University Press.

Liu, C. (2021). Application of speech recognition technology in pronunciation correction of college oral English teaching. In V. Sugumaran, Z. Xu, & H. Zhou (Eds.), Application of intelligent systems in multi-modal information analytics (pp. 563–570). *Springer.* https://doi.org/10.1007/978-3-030-51556-0_76

Metruk, R. (2024). Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, *29*(13), 16255–16282. https://doi.org/10.1007/s10639-024-12453-0

Peña-Acuña, A., et al. (2024). Learning English as a second language with artificial intelligence for prospective teachers: A systematic review. *Frontiers in Education, 9*, 1490067. https://doi.org/10.3389/feduc.2024.1490067

Pennington, M. C. (2021). Teaching pronunciation: The state of the art 2021. *RELC Journal*, 52(1), 3–21.

Pérez, A., Diaz-Munio, G. G., Gimenez, A., Silvestre-Cerda, J. A., Sanchis, A., Civera, J., ... & Juan, A. (2021). Towards cross-lingual voice cloning in higher education. Engineering Applications of Artificial Intelligence, 105, 104413.

Şahan, Ö., & Sahan, K. (2023). A narrative inquiry into the emotional effects of English medium instruction, language learning, and career opportunities. *Linguistics and Education, 75*, 101149.

Sain, Y., & Cobar, J. (2023). Improving students' pronunciation in word stress through TTS (Text to Speech) feature: On a Google Translate application. IDEAS: *Journal on English Language Teaching and Learning, Linguistics and Literature, 11*, 837–844. https://doi.org/10.24256/ideas.v11i1.3321

Van Lieshout, C., & Cardoso, W. (2022). Google Translate as a tool for self-directed language learning. Language Learning & Technology, 26(1), 1–19.

Widyana, A., Jerusalem, M. I., & Yumechas, B. (2022). The application of text-to-speech technology in language learning. *In Advances in Social Science, Education and Humanities Research*. https://doi.org/10.2991/978-2-494069-91-6_14

Zainuddin, S. Z. B., Pillai, S., Dumanig, F. P., & Phillip, A. (2019). English language and graduate employability. Education + Training, 61(1), 79–93.

Zascavage, V., & Winterman, K. G. (2009). What middle school educators should know about assistive technology and universal design for learning. *Middle School Journal*, *40*(4), 46–52.