AI-Driven Listening Hub: Enhancing ESL Listening Skills Using ChatGPT & TTS

[*1]Rafidah Amat, [2]Nur Syazwanie Mansor, [3]Nor Asni Syahriza Abu Hassan, [4]Zawani Badri, [5]Mas Aida Abd Rahim, [6]Muhammad Shyazzwan Ibrahim Brian

[1,2,3,4,5,6]Academy of Language Studies, Universiti Teknologi MARA, Kedah Branch, Kedah

***Corresponding Author**

DOI: https://dx.doi.org/10.47772/IJRISS.2025.925ILEIID000037

Received: 23 September 2025; Accepted: 30 September 2025; Published: 05 November 2025

ABSTRACT

The invention aims to generate customised listening practices for ESL learners. The artificial intelligence namely ChatGPT and Text-To-Speech website were selected to create a list of listening practices. With the emerging trends of the influences of artificial intelligence, it is undeniably in need of intellectual usage. The innovation lies in its adaptability and accessibility where it enables educators to prompt ChatGPT to produce an exclusive script-based content and the level of proficiency and later these scripts will be converted into audio using TTS tools. These listening practices offer a wide range of resources for non-native speakers to practice their listening skills. Apart from that, the privilege to use these artificial intelligence tools enable the educators to cater their listening practices to suit students' level of proficiency. The key features and functionalities of these products are, it offers educators to create listening practices based on their students; proficiency, the listening practices can be catered according to the students' CEFR level as required by the university or institution. Ever since the demand of self-regulated and scaffolding listening practices is growing, this tool is beneficial to higher education settings. In fact, it promotes inclusivity and cost-effectiveness where it empowers educators to create targeted, engaging materials without relying solely on commercial platforms. Thus, this innovation redefines how educators' developing and deliver the listening practices, aligning with modern pedagogical needs in language education.

Keywords: listening practices, listening skills, ChatGPT, Text-To-Speech

INTRODUCTION

Listening comprehension, still, is among the most difficult skills for ESL students, especially in the tertiary settings such as universities in which the proficiency requirement is in accordance with the internationally recognized benchmarks e.g. the CEFR. With the swift development of AI, educators now have the chance to develop responsive and learner-focused listening activities. In particular, we develop an AI-Driven listening hub, which merges the use of both ChatGPT and Text-to-Speech (TTS) in order to make personalised and accessible listening materials available in an affordable manner. Using AI, the hub produces text-based scripts specific to learners' proficiency levels, which are then transformed into audio by TTS tools in order to foster inclusive learning and learner agency (Kohnke et al., 2023). Thus, this paper provides a detailed overview of the system design from generating ChatPGT script to generating the audio from TTS.

Problem Statement

Common listening materials are mostly based on commercial platforms which cannot adapt to learners' differences. Pre-recorded listening software tends to be inflexible, expensive and not tailored to individual proficiency levels. ESL Tertiary students find it difficult to access self-regulated and scaffolded practices in keeping with their tertiary expectations. Teachers, however, are limited in their ability to source diverse, authentic content that is relevant to a specific student level without the added burdens of the cost of purchasing and the time it takes to source authentic listening content - such as these guided listening tasks. This calls for

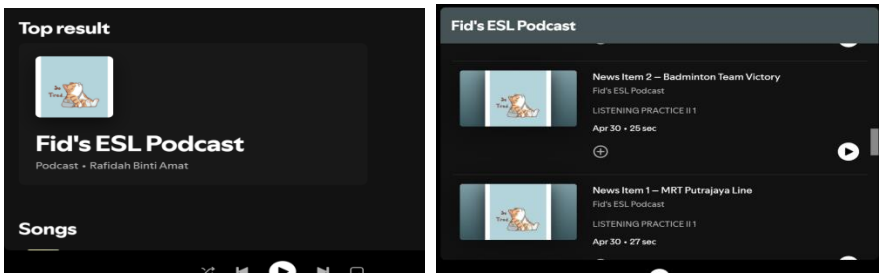
creative solutions that would strike a balance between flexibility, affordability, and pedagogical congruence, resonating with recent voices advocating for AI-mediated listening practices (Chou&Lee, 2021).

Objectives

1. To design and develop an AI-Driven Listening Hub that integrates ChatGPT and TTS technologies for ESL listening practices.
2. To generate customised listening scripts based on learners' CEFR proficiency levels.
3. To convert the generated scripts into audio resources using TTS tools for listening practice.
4. To evaluate the pedagogical effectiveness of the AI-generated practices in promoting listening comprehension and learner autonomy.
5. To explore the potential of the hub for implementation in higher education institutions.

PRODUCT DESCRIPTION & METHODOLOGY

Figure 1 The AI-Driven Listening Hub



Product Description

Figure 1 presents the AI-Driven Listening Hub screenshots. The hub is an online platform created to provide personalised listening activities for ESL learners, particularly within the Malaysian context. The hub generates content through ChatGPT, where educators can specify themes, CEFR levels, and vocabulary focus to produce relevant scripts. These scripts are then converted into audio using a text-to-speech (TTS) system, resulting in natural-sounding listening materials that align with learners' proficiency levels. For instance, the prompts used to generate text scripts are described based on learners' proficiency level, a specific context and choosing accent that best fit learners' comprehension ability. This ensures more natural delivery as well as assist learners in comprehending the contents. The completed scripts and audio files are stored within the hub, offering educators an organised resource for lesson delivery and practice.

METHODOLOGY

Figure 2 The Development Process

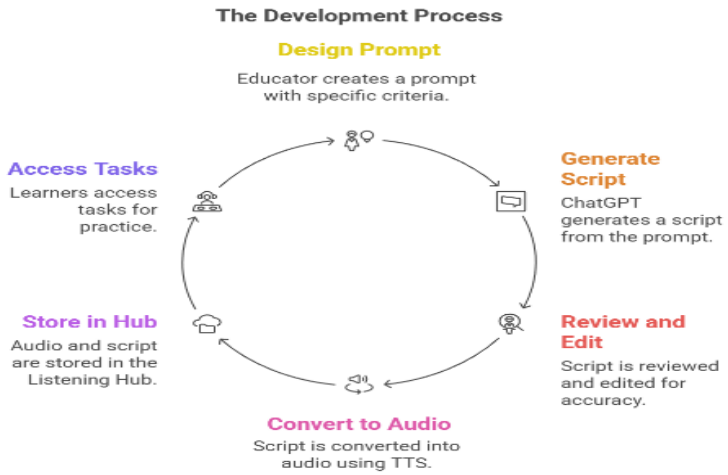


Figure 2 describes the development process of AI-Driven Listening Hub. The process started with Design Prompt to include all the necessary items in generating the required text. Then, the researcher generated the text using the prompt. The scripts were then be analysed and assessed by two ESL (English as a Second Language) lecturers to ensure its relevance to the CEFR B1 level. Once it completed, the text was uploaded to the TTS website to generate the audio. The audio was produced based on Malaysian or Singaporean accent. Some of these websites do not provide Malaysian accent but only with Singaporean accent. Thus, the researcher selected either Malaysian or Singaporean accent options. The methodology of this study is designed to explore ESL learners' perspectives on the AI-Driven Listening Hub. The development cycle consists of two interconnected stages: (1) design and development of the hub, and (2) evaluation through learners' perceptions of its usability, engagement, and pedagogical value.

Phase 1: Design and Development

Figure 3 The ChatGPT Prompts

General Template Prompt:

"Generate a [number of words] listening script for ESL learners at CEFR [level]. The topic is [theme]. Use [target vocabulary words] within the script. The context should be [academic/social/professional]. Ensure the script is clear, level-appropriate, and suitable for conversion into audio using a text-to-speech tool."

Example Prompt 1 (CEFR B1 – Academic):

"Generate a 250-word listening script for ESL learners at CEFR B1. The topic is 'Environmental Sustainability in Universities'. Include the vocabulary: recycle, energy, waste, environment. The context is academic. The script should sound like a short university lecture and be suitable for TTS conversion."

Figure 4 The Script

Script: *Time Management for University Students*

Time management is an important skill for students in higher education. When you learn how to organise your time well, you can reduce stress and become more successful in your studies. Many students often feel that there are not enough hours in the day. However, the real problem is not the lack of time, but how the time is used.

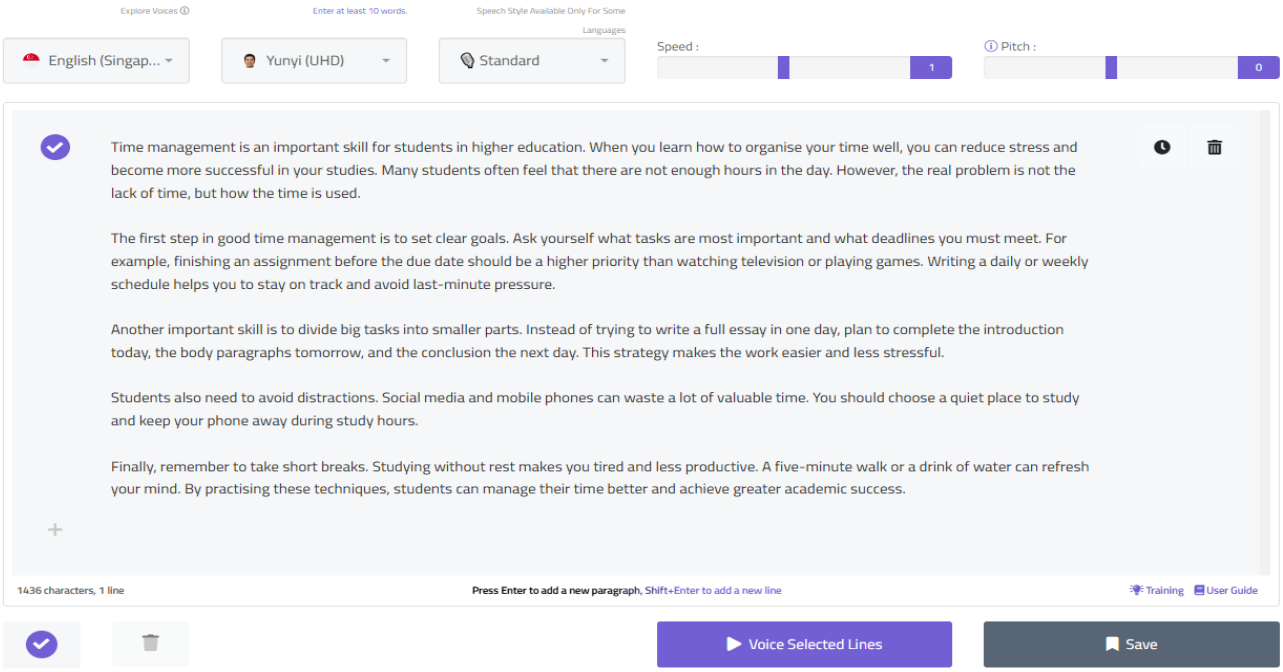
The first step in good time management is to set clear goals. Ask yourself what tasks are most important and what deadlines you must meet. For example, finishing an assignment before the due date should be a higher priority than watching television or playing games. Writing a daily or weekly schedule helps you to stay on track and avoid last-minute pressure.

Another important skill is to divide big tasks into smaller parts. Instead of trying to write a full essay in one day, plan to complete the introduction today, the body paragraphs tomorrow, and the conclusion the next day. This strategy makes the work easier and less stressful.

Students also need to avoid distractions. Social media and mobile phones can waste a lot of valuable time. You should choose a quiet place to study and keep your phone away during study hours.

Finally, remember to take short breaks. Studying without rest makes you tired and less productive. A five-minute walk or a drink of water can refresh your mind. By practising these techniques, students can manage their time better and achieve greater academic success.

Figure 5 The Script Conversion to TTS website



In this phase, educators prepared listening materials by generating scripts using ChatGPT shown in Figure 3. Prompts in Figure 4, were designed to specify CEFR levels, targeted vocabulary, and relevant contexts (academic, social, or professional). These scripts were then converted into audio files in Figure 5, through a text-to-speech (TTS) system. This TTS was selected based on Malaysian or Singaporean accent as to assist learners' familiarisation as well as better understanding of the contents delivered. The final materials (both scripts and audio) were integrated into the hub and organised for structured use by learners.

Phase 2: Learners' Perspectives

The second phase focused on collecting ESL learners' perceptions of the hub. A group of students from higher education institutions were introduced to the platform and assigned listening tasks based on the generated materials. Data collection employed surveys and open-ended questions to capture their views on usability, accessibility, and engagement. Emphasis was given to how learners experienced scaffolding, inclusivity, and autonomy within the platform. Qualitative feedback provided insights into the perceived effectiveness of the hub in supporting listening development and meeting their academic needs.

Potential Findings and Commercialisation

The AI-Driven Listening Hub will help students improve their listening skills as it will offer individualized and level-appropriate student tasks. It will also contribute to self-regulated learning and scaffolding, both of which will lead to learner autonomy and to academic achievement (Oxford, 2017). It is expected that learners will benefit from synthetic listening materials, notwithstanding that some challenges in terms of robotic tone and lack of prosody and technological issues continue to exist (Choi, 2022; Kang et al., 2009). In order to mitigate these issues, the audios can be altered according to its speed, tone as well as the pitch to lessen the robotic style of the audio which hinders learners' understanding.

Table 1 First Question of Open-ended

Themes	Examples of Responses
Technical Issues (internet, lagging, connection problems, downloading)	S17: "Must have strong Network connection." S20: "No Internet cause difficulty to me as a student."
Pronunciation & Accent Problems	S1: "AI cannot understand emotion." S10: "The intonations … are quite monotonous."

	S29: "Unnatural accent."
Lack of Emotional / Natural Expression	S22: "…difficult to understand because of the tone sounds unnatural and it lacks emotions." S23: "AI didn't… pick up subtle cues like hesitation, sarcasm, or passive-aggressiveness."
Speed and Clarity Issues	S13: "its maybe speak fast." S22: "…sometimes the AI talks too fast or uses words I don't know." S43: "The speed level is very high may be difficult to some listeners."
Content Accuracy / Quality	S4: "AI-generated listening tools have errors sometimes." S41: "Accuracy Issues: AI may misinterpret accents, slang, or technical terms." S54: "misinformation."
Accessibility / Cost Barriers	S17: "Must have strong Network connection." S30: "need to purchase the pro."
No Challenges	S5: "Nothing." S40: "I think it's nothing." S48: "so far I didn't face any of it."

The findings show that the AI-Driven Listening Hub offers innovative opportunities for ESL listening practice, but learners also reported two main types of challenges. The first are technical issues, such as unstable internet connections, lagging systems, and subscription costs. The second are linguistic issues, including robotic or unclear pronunciation, unnatural accents, lack of emotion, and fast speech. These limitations are consistent with earlier studies that highlight weaknesses in speech synthesis, particularly in prosody and cultural nuance (Choi, 2022; Kang et al., 2009). At the same time, some learners indicated that they faced no difficulties, suggesting that AI-generated materials can work well for students with strong internet access or those more adaptable to synthetic voices. This shows that while AI-based tools are useful, learners value the authenticity of human voices, especially the natural rhythm, intonation, and emotional expression that support deeper listening comprehension (Susilo, 2023; Nainggolan & Hanifah, 2024). Overall, the results suggest that AI listening tools should be used as a complement rather than a replacement for traditional resources. A blended approach would balance the accessibility and flexibility of AI with the authenticity and cultural richness of human-based listening practices (Back & Kabulis, 2025; Jung, 2025).

NOVELTY AND RECOMMENDATIONS

The novelty of this innovation lies in its dual adaptability and inclusivity. Unlike static commercial resources, the hub empowers educators to create dynamic, script-based listening materials tailored to learners' proficiency levels. It integrates two key AI functionalities which are ChatGPT for generating authentic, context-rich scripts and TTS for producing natural-sounding audio resources This AI-Driven Listening Hub provides learners with scaffolding, learner autonomy as well as cost-effective access to listening practices. The hub is also accessible through mobile phones in which it provides convenience to the learners. Future research should examine how blending AI-generated content with human-delivered listening practices can maximise authenticity and pedagogical impact (Wang & Vasquez, 2023). It is recommended that the project undergo pilot testing within a higher education setting to validate its effectiveness. Furthermore, continuous refinement should be pursued through feedback loops involving educators and learners. Future developments may explore multilingual expansions, integration with mobile learning applications, and incorporation of analytics features to track learner progress. Last but not least, this listening hub can be expanded to include diverse ESL populations as well as the tracking to learners' progress.

ACKNOWLEDGEMENTS

The authors would like to extend their deepest appreciation to the Akademi Pengajian Bahasa (APB), Universiti Teknologi MARA (UiTM) Kedah Branch, for their continuous support and encouragement throughout the development of this project. Sincere thanks are also extended to the students involved in the

study, who feedback and engagement significantly informed the process of innovation and a special thanks to research team members who dedicated their best ideas, labour, and abilities to this project.

REFERENCES

1. Back, M., & Kabulis, K. (2025). Modality or Authenticity?: The Role of Authentic versus AI-Generated Texts in Language Learner Engagement. In Rethinking Language Education in the Age of Generative AI (pp. 68-89). Routledge.
2. Birtchnell, T. (2018). Listening without ears: Artificial intelligence in audio mastering. Big Data & Society, 5(2), 2053951718808553.
3. Jung, H. (2025). AI-Assisted Student-Generated Listening Materials in High School EFL: A Sociomaterial Perspective. Multimedia-Assisted Language Learning, 28(2), 55-76.
4. Nainggolan, E. E., & Hanifah, H. (2024). Listening Comprehension Problems Encountered by EFL Students at Coastal Area. Teaching and Learning Journal of Mandalika (Teacher) e-ISSN 2721-9666, 5(2), 208-217.
5. Raza, M. A., Khan, H., & Bukhari, S. (2024). Transforming EFL Listening Skills: The Power of AI Integration in Classrooms. Social Science Review Archives, 2(2), 2284-2295.
6. Susilo, J. (2023). An Analysis Of Students Difficulties In Using Authentic Recording In Listening Skill Of The Tenth Grade On SMK Citra Angkasa Bandar Lampung (Doctoral dissertation, IAIN Metro).
7. Trang, M. (2020). Understanding listening comprehension processing and challenges encountered: Research perspectives. International Journal of English Language and Literature Studies, 9(20), 63-75