IJoLLT Vol. 8, No. 2, September 2025
eISSN: 2637-0484

Enhancing Pronunciation Skills in ESL Learners Through ChatGPT: The Efficacy of an AI-Driven Learning Tool

PRISCILLA SHAK (Corresponding author)
[0000-0001-5034-2808]
Centre for the Promotion of Knowledge and Language Learning
Universiti Malaysia Sabah
Sabah, Malaysia
pshak@ums.edu.my

CHANG SIEW LEE
[0000-0003-4923-0221]
Centre for the Promotion of Knowledge and Language Learning
Universiti Malaysia Sabah
Sabah, Malaysia
cslee@ums.edu.my

DARON BENJAMIN LOO
[0000-0001-9203-3608]
Centre for the Promotion of Knowledge and Language Learning
Universiti Malaysia Sabah
Sabah, Malaysia
daronloo@ums.edu.my

ABSTRACT

Pronunciation is often taken as a gauge of language proficiency and to judge overall English ability—sometimes inaccurately. As ESL learners reach tertiary-level education, pronunciation training often diminishes in priority, as they are assumed to have mastered basic pronunciation skills. Unfortunately, this may not be the case, and many pronunciation errors persist into adulthood. If this remains unaddressed, they can become fossilised, making them more challenging to correct later. In higher education settings, English language courses rarely focus on pronunciation training, particularly in the context of correcting erroneous pronunciation. This neglect can lead to communication issues and negatively impact individuals' growth and others' perception of them in professional environments. Despite the potential of AI technology to transform language learning, there is a notable gap in the literature regarding the application of AI tools for pronunciation enhancement. This study aims to bridge this gap by investigating the use of ChatGPT for individualised pronunciation learning. It positions the study as a critical step toward enabling ESL learners to leverage the use of AI-driven approaches as potential solutions for pronunciation enhancement. Conducted among undergraduates taking English courses in a public university in Malaysia, this study adopts a quasi-experimental mixed-method design involving pre-test, post-test, and vocabulary identification. Findings indicated that ChatGPT has great potential for applications in pronunciation learning and can offer individualised self-accessed support for ESL students to improve their pronunciation.

Keywords: AI-driven approach, ChatGPT, individualised, pronunciation learning, and self-accessed learning.

Received: 22 May 2025
Accepted: 10 October 2025
Published: 17 October 2025

To cite this article: Shak, P., Lee, C. S., & Loo, D. B. (2025). Enhancing Pronunciation Skills in ESL Learners Through ChatGPT: The Efficacy of an AI-Driven Learning Tool. *International Journal of Language, Literacy and Translation 8*(2), 63-81. https://doi.org/10.36777/ijollt2025.8.2.135

To link to this article: https://doi.org/10.36777/ijollt2025.8.2.135

INTRODUCTION

In English as a Second Language (ESL) contexts, pronunciation remains one of the most crucial yet frequently neglected aspects, especially at the tertiary level in Malaysia. Despite its importance in ensuring intelligibility and enhancing confidence, emphasis on pronunciation often receives limited instructional time and resources. If left unchecked, pronunciation errors persist, and learners continue to face challenges, not only academically but also potentially undermining their employability. Hence, there is a growing need for innovative approaches that could supplement classroom teaching towards a more learner-centred pedagogy. Recent advances in Artificial Intelligence (AI) have given new avenues for the integration of AI tools for pedagogical support in language instruction. AI-driven tools such as ChatGPT are capable of providing instant, adaptive, and personalised feedback to address learners' pronunciation questions and concerns. Nevertheless, there are still limited empirical studies that examine the pedagogical effectiveness of AI tools for pronunciation instruction. Addressing this gap is essential to exploring ways in which such tools can be utilised effectively for pronunciation teaching and learning.

Various studies have reported how students learning English find pronunciation challenging (e.g., Almusharraf, 2022; Liang & Fryer, 2024). However, intelligible and correct pronunciations are important to ensure that meaning can be conveyed successfully. In contexts where English is spoken as a second or foreign language, developing acceptable pronunciation and speaking skills is a process that involves not only becoming aware of sounds and sound combinations but also developing multicultural competence (Low, 2021). This renders the learning process one that is complex, which, for students, may be cognitively and affectively draining (Suzukida, 2021). In the Malaysian classrooms, the teaching of pronunciation may have been neglected. This may be observed in the scarcity of current research regarding the pedagogy of pronunciation, which may also be indirectly reflected through recent research reporting on other aspects and skills of the English language, e.g., research on Malaysian students' English vocabulary size by Lew et al., 2025; a study on English grammar learning challenges by Md Zolkapli et al., 2025. The minimal teaching of pronunciation may be due to teachers' lack of confidence or skills, and the pressure to prepare students to sit and pass national exams (Ahmad Shah, 2024).

Nonetheless, the teaching of pronunciation may now be supported with artificial intelligence (AI) tools being integrated into English language classrooms across different levels and contexts. These tools include automated speech recognition systems that may be accessible through computers or mobile phones, which may offer immediate feedback and instruction (Evers & Chen, 2020; Bashori et al., 2024). Furthermore, AI tools may support classroom activities by providing individualised assistance to students beyond the formal school setting. Such support may help students in terms of improving pronunciation and spoken communication, as well as developing independent and self-regulation strategies (Jin et al., 2023). This study focuses on the use of ChatGPT for pronunciation enhancement due to its cost efficiency, ability to provide instant feedback, phonetic support, and individualised learning, which are features that are essential for ESL learners. This study focuses on the use of ChatGPT for pronunciation enhancement due to its cost efficiency, ability to provide instant feedback, phonetic support, and

individualised learning, which are features that are essential for ESL learners. Against this backdrop, this study aims to examine how ChatGPT may be helpful to Malaysian university students in improving their English pronunciation. The research questions that guide this study are; i) to what extent does ChatGPT-assisted learning improve the pronunciation accuracy of Malaysian undergraduate ESL learners, and ii) which types of pronunciation errors are corrected, and which persist after ChatGPT-assisted intervention? Findings from this study would help determine the feasibility of using AI tools to assist Malaysian university students in improving their pronunciation of English.

LITERATURE REVIEW

AI AS SUPPORT FOR PRONUNCIATION DEVELOPMENT

With the proliferation of the internet, computers, and mobile phones, students everywhere are given access to various digital and AI tools. For language learners, the emergence and integration of AI tools have left a positive impact, as they have been deemed helpful to those of varying proficiency levels (Evers & Chen, 2020). In the teaching of pronunciation, AI tools are a valuable support to the teaching and learning activities in the classroom. Studies have reported how these tools have helped students gain an awareness of pronunciation mistakes (Evers & Chen, 2020; Bashori et al., 2024). Studies have also reported that using AI tools not only improves students' proficiency in the target language but also their learning motivation and strategies (Liu et al., 2025). Nuñez et al.'s (2025) study involving 49 EFL learners utilising ChatGPT for speaking skills learning indicated that, based on their post-test results, the use of ChatGPT led to significant improvement in their pronunciation. Given ChatGPT's accessibility, students may also seek help with pronunciation beyond class time. The application can provide various language models to students (Barrot, 2023) and help guide students by providing language input (Huang et al., 2022). As such, students do not need to overtly rely on their teachers for help; instead, consistent and accessible help may be rendered to them through these tools (Yahaya et al., 2021). This includes the provision of immediate and specific feedback through AI tools (Fatima, 2024). Lara et al. (2024) demonstrated this through the development of an AI tool that records and transcribes a student's speech, which is then compared against a target pronunciation benchmark. Feedback that includes phonemes, intonation, and stress points is then provided to the students for independent revision (see also Bashori et al., 2024). Nonetheless, some findings have indicated that AI tools may be more useful in providing feedback on segmental features, such as consonant and vowel sounds, rather than suprasegmental features such as stress and intonation (Senowarsito & Ardini, 2023).

AI tools may be supported with other sources of feedback, such as from peers (Evers & Chen, 2020) and teachers (Liu et al., 2025). Feedback from AI may be sought beyond the classroom and may be curated according to the students' needs (Zhao, 2024; Mohammadkarimi, 2024). By allowing students to work with AI tools outside of a formal classroom setting, they are being encouraged to be autonomous learners without having to rely solely on the instruction of their teachers (Saraini et al., 2022; Senowarsito & Ardini, 2023). This may further encourage students

to develop effective metacognitive strategies that would help them self-correct errors in pronunciation and address issues affecting speaking (Shafiee Rad & Roohani, 2024; Zhao, 2024).

Speaking is a cognitively heavy process for students as they will need to manage vocabulary, grammar, and the organisation of information, besides attending to accurate and comprehensible pronunciation (Fatima, 2024). Moreover, there are other social stressors linked with speaking, such as accent and speech style (Zhai & Razali, 2022). When expected to speak spontaneously, students who are still developing fluency and pronunciation may hesitate and pause, which may lead to lower marks (Yahaya et al., 2021; Evers & Chen, 2020). Other negative emotions that students may face when asked to speak or produce sounds in front of their peers may be embarrassment and a heightened level of self-consciousness (Muniandy & Selvanathan, 2024). This may inhibit students' willingness to participate and even speak out in class (Yang et al., 2024). Anxiety that they experience from having to speak may affect students' confidence to use other language skills (Yahaya et al., 2021; Kho & Ting, 2024). When using ChatGPT, anxiety may be minimised, and the fear of judgement may be eliminated. Students can also do repeated practice according to their own needs. This creates a safe space for students to practise speaking skills and pronunciation (Hu & Škultéty, 2024). Being able to practise in this manner helps students develop long-term retention of sounds and even vocabulary. This may reinforce instructions already given to students in the classroom setting (Kazu & Kuvvetli, 2023). Essel et al. (2024) investigated the cognitive effects of using ChatGPT for individual language practice on university learners and discovered increased self-efficacy and improved reflective thinking, which are important skills for tertiary-level students to develop.

It should not be assumed that learners would know how to use AI tools; moreover, there may be sounds that are difficult to articulate due to subtle pronunciation features (Mohammadkarimi, 2024). The use of AI tools should also be regulated or checked by teachers, as students may develop an over-reliance on them if left unsupervised (Pangestu & Suwartono, 2024). Furthermore, students may not necessarily receive feedback about pronunciation positively. This would depend on students' learning goals; if they viewed pronunciation as an important aspect of good communication, they would then appreciate feedback given to them (Liu et al., 2025).

CHALLENGES IN TEACHING PRONUNCIATION AND USING AI TOOLS: THE MALAYSIAN CONTEXT

In the Malaysian classroom, pronunciation or speaking activities may not be a regular teaching and learning activity. More time is instead spent on teaching students to master grammar and syntax (Asikin & Ibrahim, 2020). Other classroom challenges are teachers' lack of training in teaching pronunciation and limited resources and facilities (Mohd Asikin & Ibrahim, 2020). Furthermore, the challenges faced by Malaysian students may be compounded by the preference for native-speaker accents. Despite a recognition of a Malaysian English accent, the preoccupation with native-speaker accents would inhibit students from communicating openly in English (Abu Hassan et al., 2021). Moreover, Malaysian students have been reported to be hyper-aware when having to speak in English, especially when giving a presentation (Kho & Ting, 2024). Some related difficulties include the inability to speak spontaneously, hesitant speech, and inappropriate and long pauses (Chuah & Ch'ng, 2023). In terms of pronunciation difficulties or

challenges, studies from the context of Malaysia have reported factors such as interference from students' first language, which may be Malay or other local languages (Hamzah et al., 2015; Author et al., 2016). This has resulted in Malaysian students finding certain sounds in English difficult to pronounce, such as differentiating long and short vowel sounds, diphthongs (Abu Hassan et al., 2021), consonant sounds such as plosives, fricatives, and affricates, and silent letters and grammatical endings, such as the past tense "-ed" sound (Author 1 et al., 2016).

Using technology to teach pronunciation has been reported to have a positive impact on Malaysian students, especially when done with active learning strategies (Zhang & Al-Saqqaf, 2025). This may be observed in the study of Chuah and Ch'ng (2023), where students were found to experience improvement in their pronunciation when doing voice-over challenges on TikTok. The type of digital or AI tool chosen is an important consideration, as it would have a bearing on the motivation and consistency of learners in improving their pronunciation (John & Lo, 2024). This was also observed in the study of Chandran et al. (2024), where school students were taught English pronunciation along with English Language Speech Assistant (ELSA), an AI-powered tool that supports pronunciation improvement. Their study found that students developed positive learning behaviours, which enhanced their motivation to improve their pronunciation. Previous studies on the use of the ChatGPT Voice mode for language learning mainly revolved around its use in two-way communication with the learners, where learners interacted with ChatGPT Voice to practise their speaking or communication skills. There is scant research on the use of ChatGPT for targeted pronunciation practice.

METHODOLOGY

This study utilised a mixed-method design in its methodology, combining both quantitative and qualitative approaches. Such a design allows for a more comprehensive understanding of the research issue (Creswell & Clark, 2017; Creswell & Creswell, 2023; Maxwell, 2019). A quasi-experimental design was undertaken to measure the effects of using ChatGPT, and the word identification task was conducted to explore participants' recognition of unfamiliar vocabulary. Adopting this design enables direct observation of the pre-post measurement of a controlled intervention (Creswell & Creswell, 2023). Purposive sampling was utilised in recruiting Malaysian ESL students from a public university as participants. This sampling procedure ensures that the data obtained enable the in-depth exploration of the topic from a representative target group (Mackey & Gass, 2021). The data collection methods involved the use of pre-test, post-test, and identification of unfamiliar words.

PARTICIPANTS

First and second-year undergraduates taking university English courses were selected through purposive sampling for this study. Based on their classroom assessment performance, those with weak pronunciation were identified and invited to participate in the study by their English course instructors. A total of sixteen student participants were recruited. They were considered as students with low proficiency in English based on their Malaysian University English Test

(MUET) results, a test used to measure the English language proficiency of pre-university Malaysian students (Malaysian Examination Council, 2019).

INSTRUMENTS

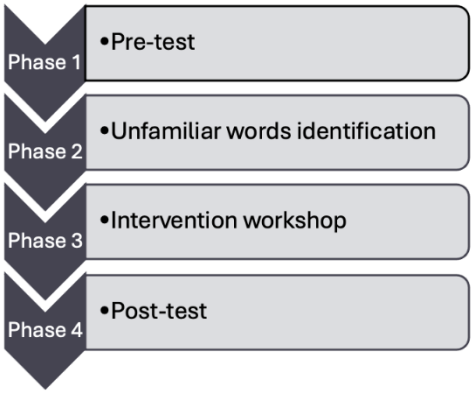
For this study, three short reading texts of between 113 and 217 words each were utilised (Appendix A). The texts were used for pre-reading and post-reading tests and unfamiliar words identification. The first text (Text A) was adopted from the Handbook of the International Phonetic Association (IPA) (International Phonetic Association, 1999), entitled *The North Wind and the Sun* which has been used widely in past pronunciation research (Deterding, 2006; Author 1 et al., 2016). The second reading text (Text B), entitled *The Boy Who Cried Wolf* was adapted from Deterding (2006), who adapted the reading text from Aesop's fable. It was deemed to be more suitable, as there was less content word repetition and it contained a wider range of English sounds compared to *The North Wind and the Sun* (Deterding, 2006). The final text (Text C), *The Legend of Mahsuri*, was specifically devised by the main researcher of this study based on a well-known local Malaysian legend typically taught at the primary school level and incorporates commonly mispronounced sounds by Malaysian ESL learners based on an earlier study (Author 1 et al., 2016) with a similar group of ESL undergraduates. Comparatively, *The North Wind and the Sun* is more compact in nature with 113 words, *The Boy Who Cried Wolf* consists of 217 words, while *The Legend of Mahsuri* contains 199 words.

PROCEDURES

The study was conducted in four phases as depicted in Figure 1.

Figure 1

Phases of the study



During the pre-test phase, participants were asked to read aloud the three texts to identify their pronunciation errors. The read-aloud sessions were conducted online using the MS Teams platform and were video recorded by the researchers. These sessions were conducted before the introduction of an AI tool for pronunciation practice. The pre-test provided data on students'

pronunciation errors, which were later used to compile a list of customised words for the students to learn how to pronounce using ChatGPT. After completing their pre-test, in the second phase of the study, the participants were then asked to underline words they were unfamiliar with on another copy of the same reading text. Doing so allowed the researchers to obtain data on words unfamiliar to the students. The third phase was the implementation of the 'Say It Right with AI' pronunciation workshop, where all the participants attended a full-day workshop where they learnt to use an AI tool, namely ChatGPT, to practise and improve their pronunciation. During the workshop, they were given explanations and instructions on how to use ChatGPT for pronunciation practice. There were also demonstrations on how to interact with ChatGPT using the application's voice mode. In addition, the participants were also given a set of suggested prompts to guide them to ask relevant questions when prompting ChatGPT. Once they were familiar with the voice mode and prompts, the participants were given a list containing words that they had mispronounced for them to practise. This list of mispronounced words was extracted from their pre-test and provided them with an individualised list of sounds for each of them to focus on. At the final phase of the study, participants were required to complete a post-test with the researchers. During this session, the participants performed the post-reading tests where they read the same three texts and were again video recorded.

For data analysis, video recordings of the pre- and post-tests were qualitatively analysed by the researchers. Based on their reading of the three texts, each participant's pronunciation errors before and after the intervention were identified. Doing this entailed several rounds of replaying the video recordings to ensure that all pronunciation errors were correctly identified. In this study, a pronunciation is considered erroneous when it deviates from standard British English (received pronunciation) and general American English. Hence, they included errors due to dialect and accent variations. Subsequently, the errors were systematically listed, transcribed using broad transcription, and categorised based on the error types. The participants' mispronunciations during pre- and post-tests were also tabulated and later quantitatively analysed. This allowed the researchers to identify and compare low- and high-occurring errors. In addition, a list of unfamiliar words was compiled for each participant to enable comparison with pronunciation errors and elucidate findings.

Validity was established through the use of well-known Aesop's fables widely used by IPA and linguists for pronunciation investigation and a passage based on a Malaysian folk tale designed to include common phonological difficulties encountered by Malaysian ESL learners, such as the pronunciation of fricatives, consonant clusters, and past-tense -ed endings. Thus, ensuring that the instruments are aligned with the objectives of this study to accurately identify and measure the intended construct. Reliability was maintained through standardised administration procedures following a list of established step-by-step procedures for the three reading texts for all participants. Meanwhile, rubrics for the error analysis adhered to a predetermined threshold (as discussed earlier) for a uniform categorisation of errors.

Data triangulation was employed in this study to strengthen the trustworthiness and credibility of the findings. For this, the pre- and post-test scores and pronunciation errors were examined. The test scores yielded quantitative findings that offered insights into the participants' performance before and after the ChatGPT-assisted intervention. Meanwhile, the pronunciation error analysis offered qualitative data of the error patterns by pinpointing phonemes and word

forms that either improved or remained problematic. This methodological triangulation provided a more comprehensive understanding of the impact of the intervention on the participants' pronunciation development.

FINDINGS AND DISCUSSION

PRONUNCIATION ERRORS

The data collected from the pre-test revealed a total of 38 categories of pronunciation errors and seven (7) whole-word errors made by the participants. Analysis of the data determined that students have problems in pronouncing fricatives [/ð/(front), /ð/(mid), /θ/(front), /θ/(end), /z/], plosives (/t/, /d/, /g/), diphthongs (/aɪ/, /eɪ/, /əʊ/, /aʊ/, /eə/, /æ/), pure long vowels (/ɑː/, /ɜː/, /ɔː/, /uː/), sub-pure long vowel (/juː/), pure short vowels [/ə/(front), /ə/(mid), /ɪ/(front), /ɪ/(mid), /ɪ/(end), /ɒ/, /e/, /ʌ/, /i/], sub-pure short vowels (/əs/, /ər/, /ɪŋ/), consonant clusters (/lf/, /lt/), affricative (/dʒ/), sibilant (/s/), and –ed endings (/d/, /ɪd/, /t/). These findings are consistent with those observed in past studies from the Malaysian context, which have reported pronunciation difficulties with long and short vowel sounds, diphthongs (Abu Hassan et al., 2021), plosives, fricatives, affricates, and the past tense –ed sound (Author 1 et al., 2016).

In addition to the segmental errors, the seven (7) whole-word mispronunciations noted include wrapped (/ræpt/), obliged (/əblaɪdʒd/), echoed (/ekəʊd/), pierced (/pɪrst/), shepherd (/ʃepəd/), famine (/fæmɪn/), and remorse (/rɪmɔːs/). The whole-word errors were associated with phoneme substitution, syllable insertion, syllable substitution, phoneme insertion, syllable deletion, and consonant cluster simplification. Previous studies conducted in Malaysia (Hamzah et al., 2015; Author et al., 2016) have highlighted mother tongue interferences as a contributing factor to students' pronunciation difficulties, which may stem from Malay or other local languages. Therefore, the errors observed in the participants can be partly attributed to mother tongue influence, where certain phonological features of their native languages may shape how whole words were mispronounced.

CORRECTED ERRORS

Results obtained from the post-test revealed that the participants showed notable improvements in their pronunciation. The participants initially made seven (7) whole-word errors in the pre-test, but these errors were successfully corrected following the ChatGPT-assisted workshop session, with no further whole-word errors being observed. In addition, 11 categories of segmental sound errors were successfully corrected. The corrected segmental pronunciation errors included two (2) fricatives [/ð/(mid), /z/], one (1) diphthong (/aɪ/), two (2) pure long vowels (/ɑː/, /ɜː/), two (2) pure short vowels [/ə/(front), /ɪ/(front)], one (1) sub-pure short vowel (/əs/), two (2) consonant clusters (/lf/, /lt/), and one (1) affricative (/dʒ/). These findings indicate a positive impact of ChatGPT in correcting pronunciation errors, both at the word and phoneme levels. These findings suggest that ChatGPT could support participants in developing greater control over their pronunciation.

These improvements are attributed to ChatGPT's voice-enabled mode, which provides precise phoneme-level feedback, whole-word pronunciation guidance, and interactive pronunciation practice through prompts provided by the participants themselves. In the voice-enabled mode, ChatGPT allows users to listen to accurate pronunciation models, ask for explanations and elaboration, request phoneme-level and/or whole-word level pronunciation guidance, and listen to slower repetitions of sound to facilitate their learning of the targeted sounds, leading to greater control over their pronunciation. This makes ChatGPT voice-enabled mode comparable to the two (2) automatic speech recognition-based systems, I Love Indonesia (ILI) and NovoLearning (NOVO), discussed in Bashori et al. (2024).

PERSISTENT ERRORS

Despite these improvements, 27 errors remained persistent even after the ChatGPT-assisted session. These included difficulties in producing specific phonemes and the mispronunciation of -ed endings. These included three (3) fricatives [/ð/(front), /θ/(front), /θ/(end)], three (3) plosives (/t/, /d/, /g/), five (5) diphthongs (/eɪ/, /əʊ/, /aʊ/, /eə/, /æ/), two (2) pure long vowels (/ɔː/, /uː/), one (1) sub-pure long vowel (/juː/), seven (7) pure short vowels [/ə/(mid), /ɪ/(mid), /ɪ/(end), /ɒ/, /e/, /ʌ/, /i/], two (2) sub-pure short vowels (/ər/, /ɪŋ/), one (1) sibilant (/s/), and three (3) –ed endings (/d/, /ɪd/, /t/). The higher number of persistent errors in vowel categories, particularly with diphthongs and pure short vowels, suggests that participants faced ongoing challenges in perceiving and producing subtle vowel contrasts. These errors may stem from difficulty in distinguishing between similar vowel sounds influenced by their mother tongue. This reflects previous observations by Abu Hassan et al. (2021) regarding Malaysian students' difficulties in producing diphthongs and differentiating between long and short vowels. Furthermore, difficulties in perceiving the English regular past tense -ed endings are common among second language learners, regardless of their proficiency level (Abu Hassan et al., 2021). Moreover, challenges in producing the -ed endings suggest that learners struggle not only with understanding but also with accurately pronouncing past tense forms in English. These pronunciation errors may arise from the complexity of the –ed suffix, which requires the learners to understand its different pronunciations based on the preceding sound in the verb. This type of error is further elaborated in the subsequent section.

In general, incorporating ChatGPT in pronunciation correction has yielded encouraging outcomes, particularly in addressing whole-word pronunciation errors and certain phonemes [/ð/(mid), /z/, /aɪ/, /ɑː/, /ɜː/, /ə/(front), /ɪ/(front), /əs/, /lf/, /lt/, and /dʒ/]. Although the persistent errors outnumber the corrected errors, the improvements suggest meaningful progress. This evidence supports the feasibility of employing AI tools such as ChatGPT to assist Malaysian university students in enhancing their English pronunciation, further suggesting the potential of ChatGPT as a valuable tool to supplement pronunciation instruction in the classroom and for individualised self-accessed support for ESL students.

On the other hand, a limitation in the use of ChatGPT was observed during the workshop. Three (3) participants encountered instances where ChatGPT misinterpreted their speech input and produced responses that differed from their intended meanings. As a result, for users with

weaker pronunciation, this limitation might prevent them from effectively using ChatGPT's voice mode. If the system fails to accurately recognise speech input, it could lead to frustrating interactions and a decline in user satisfaction. This limitation is consistent with the findings of Mohammadkarimi (2024), which emphasised users' concern about whether the AI tool is sensitive enough to accurately understand or assess their pronunciation, especially in interpreting subtle pronunciation differences. Therefore, this calls for refining speech recognition capabilities or algorithms to more effectively accommodate a range of pronunciation styles, which would significantly boost ChatGPT's effectiveness and user satisfaction.

ERRORS OF PAST TENSE –ED FORM

As found in a previous study conducted in a Malaysian ESL classroom (Author 1 et al., 2016), students made frequent pronunciation errors of the –ed form (past tense). Similar findings were also found in Nyarks and Inyang's (2024) study, where their bilingual Nigerian respondents tended to ignore past tense suffixes, and in Marchena et al.'s (2020) study, where the pronunciation of the –ed form is considered one of the hardest to master by Spanish EFL learners. Listed in Table 3 are some of the frequently mispronounced –ed forms (past tense). In the case of the past tense –ed form, mispronunciation can affect intelligibility, especially when the speech context is not clear or obvious. The –ed has three (3) allomorphs, as shown in Table 3. Mispronouncing these (e.g., saying *"oblig-id"* instead of *"obliged" or "linger-id"* instead of *"lingered"*) may lead to confusion, especially when speaking fast or when the context is unclear.

Table 3

Pronunciation errors of the past-tense -ed form

Pronunciation errors of the past-tense –ed form	*Allomorphs for –ed form*
used, agreed, considered, obliged, tried, cried, stayed, lingered, stabbed, accused, sentenced, surrounded, flowed, pierced, returned, filled, revealed, confessed, echoed, judged	/d/
wrapped, rushed, convinced escaped, confessed	/t/
succeeded, kind-hearted	/ɪd/

Past studies distinguish between intelligibility, how much can be understood, and comprehensibility, ease of understanding (Hu et al., 2022; Isaac & Trofimovich, 2012; Levis, 2018). While mispronouncing –ed endings can make speech slightly harder to understand, it does

not significantly impede the listener's ability to grasp the intended message as the listener gradually gets accustomed to the speaker's speech patterns. However, it may cause a delay in meaning processing or misunderstandings, especially for listeners unfamiliar with the context. As mentioned above, while the use of ChatGPT may be effective in correcting mispronunciation, some persistent errors, such as the use of the past tense –ed form, may continue to prove challenging for students to overcome and may require a longer time for them to adapt to the different allomorphs' pronunciation.

ERRORS DUE TO UNFAMILIAR WORD

Mispronunciation was also attributed to unfamiliar words in previous studies (Altiar & Suwartono, 2020; Fikriyana et al., 2023; White et al., 2013). As the participants had identified unfamiliar words before undergoing the pronunciation intervention, the researchers were able to identify these unfamiliar words, which were shown to lead to mispronunciation. Due to technical constraints, this section limits its discussion to unfamiliar words that were identified by more than 60% of the participants. One of the most frequently (86.7%) identified as unfamiliar word was 'obliged'. Data from the post-test indicated that while most of the participants were able to correct their mispronunciation of this word, there were still 26.7% who continued to mispronounce it. This may be attributed to the fact that the word 'obliged' is in the past-tense form, which, as highlighted in the previous section, is a common cause of mispronunciation for ESL students. Another word identified by 66.7% of the participants as unfamiliar was 'shepherd'. Upon learning the pronunciation and meaning of this word through ChatGPT, only one participant mispronounced this word during the post-test, from the initial 6 participants who mispronounced the word during the pre-test.

On the other hand, there were also instances of participants indicating a word as unfamiliar, but they were able to pronounce the word correctly. One such example is the word 'flocks', where a majority, 80% of the participants, identified it as unfamiliar, but only 20% mispronounced this word, and this was mainly attributed to the dropping of the –s sound (plural form) at the end of the word. One possible reason for the participants' ability to pronounce the word 'flocks' may be due to the similarity of this word to the common noun 'locks'. Hence, the participants were able to pronounce this word due to familiarity with a similar-sounding word (Uchihara et al., 2022).

Other words that were identified as unfamiliar by 60% of the participants are 'cloak', 'lush', and 'disputing'. Despite this, only 33.3% mispronounced 'cloak' initially, and only one participant mispronounced the word during the post-test. Similarly, only 20% of the participants mispronounced 'lush', and only one was still unable to pronounce the word correctly during the post-test. As for 'disputing', despite being identified by many as unfamiliar, only 20% mispronounced it during the pretest, and they were all able to pronounce the word correctly during the post-test. All these instances indicate that with the help of ChatGPT, the pronunciation of unfamiliar words can mostly be corrected. Where mispronunciation still occurred, it may be attributed to the lack of practice time or the lack of attention paid by the participants to say the word correctly.

The researchers, when implementing the pronunciation intervention, have noticed that participants with weaker pronunciation tended to require guidance and support, especially in terms

of determining the accuracy of their pronunciation, as ChatGPT was unable to discern the accuracy of their pronunciation. In some cases, mispronunciations were accepted by ChatGPT as "correct" or as "a variety of the pronunciation" of the word (ChatGPT's response after being asked if the pronunciation of a participant was correct). Hence, despite previous study findings that ChatGPT can be useful for students for speaking practice, its efficacy may be limited for non-native ESL learners who are looking for feedback and guidance to improve their pronunciation or articulation of specific phonemes (Zou et al., 2024).

During the workshop, it was observed that using ChatGPT's standard voice mode was quite challenging for some of the participants, especially for those with weaker pronunciation. For instance, during their interaction with ChatGPT, the participants were asked to use the voice mode to listen to the correct pronunciation of the word. However, when the participants utilised the voice mode to prompt ChatGPT, the AI assistant at times provided irrelevant responses, which were not helpful and sometimes confusing for the participants. To overcome this, the participants were instructed to use typewritten prompts to ensure that ChatGPT could understand their instructions and provide the assistance required. As such, the voice mode was mainly useful for generating the audio samples of the sounds of the correct word pronunciation, but not for ascertaining the accuracy of pronunciation (Lima & Wallace, 2023).

While the use of ChatGPT appeared to be fairly intuitive after the learners were familiar with the functions and prompts, there were observable technical issues that they faced. As the participants were working with the free standard version of the application, there were some instances where the students could not obtain the right input from ChatGPT, and they resorted to the use of another application, Google Search, which incorporates AI capabilities for pronunciation guidance. This alluded to the notion that the effectiveness of ChatGPT can be further enhanced with the use of other AI tools for greater efficacy.

Overall, it can be surmised that the occurrence and persistence of pronunciation errors could be attributed to several linguistic and phonological factors. As exemplified earlier, one major cause is first language (L1) interference, where ESL learners transfer phonological features of their L1 to English sounds. For example, the substitution of dental fricatives (/θ/, /ð/) for plosives (/t/, /d/) as the learners are not familiar with such phonemes. Similarly, they also face difficulties with producing -ed endings (/d/, /t/, /ɪd/) due to a lack of equivalent forms in their L1. Another contributing factor is the lack of auditory discrimination training, where learners are not sensitive to the subtle distinctions between minimal pairs, such as /eɪ/ and /æ/, and short versus long vowels. On top of these, the longstanding incorrect articulations due to fossilised habits further perpetuate pronunciation errors.

CONCLUSION

In conclusion, incorporating AI technology, specifically ChatGPT, in pronunciation improvement is effective and highly promising. Findings of this study have indicated that many of the participants benefited from the AI intervention and showed improvement in their post-tests, as evidenced by the reduction in the number of mispronounced words and their improved abilities in correcting previously mispronounced words. In particular, ChatGPT is especially helpful in helping learners to correct whole-word mispronunciation, such as the pronunciation of unfamiliar

words. ChatGPT is also helpful in providing sample sounds for learners to imitate to improve their pronunciation. With customised prompts, learners were also able to prompt ChatGPT to provide them with a detailed breakdown of the sounds and phonemes, making it useful for clarifying the pronunciation of the targeted sounds or words. The targeted input given by ChatGPT is also especially beneficial for individual learners facing difficulties with certain sounds.

Several practical implications for English language pronunciation teaching and learning arise from this study. First, it has been demonstrated that ChatGPT can serve as a self-access tool to supplement pronunciation learning through individualised feedback and targeted self-paced practice, which is valuable in the Malaysian higher education settings, considering the lack of emphasis on pronunciation instruction. Second, the use of ChatGPT promotes learner autonomy and encourages learners to learn in a low-pressure environment without having to worry about what pronunciation mistakes they make in front of others. Third, integrating the use of ChatGPT supports the integration of AI technology for blended learning in the language classroom. Lastly, it underscores the importance of teacher training for them to fully leverage ChatGPT for effective pronunciation instruction.

Unlike previous studies that relied on the use of AI tools for conversation learning, this study demonstrates the distinct dynamic and adaptive pronunciation support that ChatGPT could render to learners. A key finding revealed ChatGPT's ability to help improve pronunciation accuracy by enhancing learners' metalinguistic awareness, leading to a clearer understanding of how specific English sounds are formed and differentiated, and in doing so, promotes self-correction. The study has also demonstrated that ChatGPT fosters self-directed learning and targeted practice, which greatly reduces the need for teacher-led instruction. Additionally, the study also identified segmental pronunciation errors that are more easily corrected (e.g., fricatives, diphthongs, and consonant clusters) and those that are more persistent, such as those involving -ed endings and vowel contrasts. These nuanced findings, although demonstrating the efficiency of ChatGPT in some aspects for pronunciation correction, indicated that teacher involvement and reinforcement are still necessary. As the ChatGPT intervention workshop was held for one day and the post-test was conducted on the same day, the long-term effect of the pronunciation improvement cannot be attested. Therefore, future research can investigate the long-term effects of AI pronunciation correction interventions and gauge their effectiveness over time.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Centre for the Promotion of Knowledge and Language Learning, Universiti Malaysia Sabah, for the funding of this study (GSM2411).

REFERENCES

Abu Hassan, N. A. S. Wan Ismail, W. N. F., Ahmad Zaidi, N. N. (2021). Malaysian pronunciation: i\Is it wrong to sound Malaysian? *Voice of Academia (VOA), 17*(2). pp. 239-242. https://voa.uitm.edu.my/v1/images/voa_issue/2021/vol17_no2_title20.pdf

Ahmad Shah, S. S. (2024). Integrating pronunciation: What challenges do Malaysian ESL teachers face? In W. A. Wan Aslynn, S. Pillai, & S. Y. Chew (Eds.), *Contemporary Malaysian English Pronunciation* (pp. 167-186). IIUM Press.

Almusharraf, A. (2022). EFL learners' confidence, attitudes, and practice towards learning pronunciation. *International Journal of Applied Linguistics, 32*(1), 126-141. https://doi.org/10.1111/ijal.12408

Altiar, V., & Suwartono, T. (2020). Mispronunciation of rarely-used English vocabulary by 8th semester students in ELT department. http://dx.doi.org/10.13140/RG.2.2.18475.87844

Author 1 et al. (2016).

Barrot, J. S. (2023). ChatGPT as a language learning tool: An emerging technology report. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-023-09711-4

Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2024). I Can Speak: improving English pronunciation through automatic speech recognition-based language learning systems. *Innovation in Language Learning and Teaching, 18*(5), 443-461. https://doi.org/10.1080/17501229.2024.2315101

Chandran, M., Wang, S. L., Dzulkifly, S. B., Ng, T. F., & Ghoneim, A. S. (2024). Enhancing English pronunciation learning in primary education through natural language processing: A quantitative study. *Journal of ICT in Education, 11*(2), 1-17. https://doi.org/10.37134/jictie.vol11.2.1.2024

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research* (3rd ed.). Sage.

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). Sage.

Deterding, D. (2006). The North Wind versus a Wolf: Short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association, 36*(2), 187-196. https://doi.org/10.1017/S0025100306002544

Essel, H. B., Vlachopoulus, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language *models (LLMs). Computers and Education: Artificial Intelligence, 6,* 1-13. https://doi.org/10.1016/j.caeai.2023.100198

Evers, K., & Chen, S. (2022). Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults. *Computer Assisted Language Learning, 35*(8), 1869-1889. https://doi.org/10.1080/09588221.2020.1839504

Fatima, S. N. (2024). Artificial intelligence for pronunciation improvement in second language acquisition. *Journal for Research Scholars and Professionals of English Language Teaching, 8*(41). https://doi.org/10.54850/jrspelt.8.41.001

Fikriyana, H., Ni'mah, D., & Sholihah, F. A. (2023). The factors of English words mispronunciation encountered by EFL students. *Jurnal Penelitian, Pendidikan, dan Pembelajaran, 18*(21), 1-13. https://jim.unisma.ac.id/index.php/jp3/article/view/21511/16108

Hu, X., Kuo, L. J., & Dixon, L. Q. (2022). Effectiveness of interventions on improving English language learners' intelligibility and comprehensibility: A meta-analytic review. *Language Teaching Research, 0*(0). https://doi.org/10.1177/13621688221106832

Hu, W. C., & Škultéty, R. (2024). Unlocking the Learning Potential: ChatGPT as a Virtual Platform for Cross-Interaction in English Language Learning. *Engineering Proceedings, 74*(1). https://doi.org/10.3390/engproc2024074059

Huang, W., Hew, K., & Fryer, L. (2022). Chatbots for language learning—are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237–257. https://doi.org/10.1111/jcal.12610

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition, 34*, 475–505. http://doi.org/10.1017/S0272263112000150

Jin, S. H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education, 20*(1). https://doi.org/10.1186/s41239-023-00406-5

John, J., & Lo, Y. Y. (2024). Improving English speaking proficiency among non-English major learners in Malaysia through mobile language learning applications. *Pertanika Journal of Social Sciences & Humanities, 32*(4), 1599-1618. https://doi.org/10.47836/pjssh.32.4.17

Kazu, İ. Y., & Kuvvetli, M. (2023). A triangulation method on the effectiveness of digital game-based language learning for vocabulary acquisition. *Education and information technologies, 28*(10), 13541-13567. https://doi.org/10.1007/s10639-023-11756-y

Kho, M. G. W., & Ting, S. H. (2024). A Multidimensional Assessment of Malaysian Polytechnic Students' Communication Apprehension in Oral Presentations. *Asian Journal of Research in Education and Social Sciences, 6*(3), 343-360. https://doi.org/10.55057/ajress.2024.6.3.28

Lara, M. S., Subhashini, R., Shiny, C., Lawrance, J. C., Prema, S., & Muthuperumal, S. (2024). Constructing an AI-assisted pronunciation correction tool using speech recognition and phonetic analysis for ELL. *Proceedings of the 2024 10th International Conference on Communication and Signal Processing (ICCSP),* 1021-1026. https://doi.org/10.1109/ICCSP60870.2024.10543462

Lew, J. W., Łuniewska, M., Lee, S. T., & Yap, N. T. (2025). Assessing vocabulary size in Malaysian preschoolers: insights from the Malaysian English cross-linguistic lexical task and parents of multilingual children questionnaire. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2024.2446996

Levis J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press. https://doi.org/10.1017/9781108241564

Liang, L., & Fryer, L. K. (2024). Phonological instruction in East Asian EFL learning: A scoping review. *System, 123*. https://doi.org/10.1016/j.system.2024.103336

Lima, E. & Wallace, L. (2024). Friend or Foe: Considering What Pronunciation Teachers Can Do that ChatGPT Cannot. In D. J. Olson, J. L. Sturm, O. Dmitrieva, & J. M. Levis (eds), Proceedings of 14th Pronunciation in Second Language Learning and Teaching Conference, (pp. 1-11). Purdue University, September 2023. https://doi.org/10.31274/psllt.17652

Liu, C. C., Hwang, G. J., Yu, P., Tu, Y. F., & Wang, Y. (2025). Effects of an automated corrective feedback-based peer assessment approach on students' learning achievement, motivation, and self-regulated learning conceptions in foreign language pronunciation. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-025-10484-z

Low, E. L. (2021). EIL pronunciation research and practice: Issues, challenges, and future directions. *RELC Journal*, *52*(1), 22-34. https://doi.org/10.1177/0033688220987318

Mackey, A., & Gass, S. M. (2021). *Second language research: Methodology and design (3rd ed.)*. Routledge. https://doi.org/10.4324/9781003188414

Malaysian Examination Council. (2019). Malaysian University English Test: Regulations and test specifications.https://www.mpm.edu.my/images/dokumen/calon-peperiksaan/muet/regulation/Test_Specification_Regulation.pdf

Marchena, P. C., Vargas, S. S., & Sánchez, G. Z. (2020). Language teaching techniques for pronouncing the past tense phonemes /t/, /d/, and /ɪd/ in regular verbs when reading aloud. *Revista Ensayos Pedagógicos*, *XV*(2), 277-300. https://www.revistas.una.ac.cr/index.php/ensayospedagogicos/article/view/14725/20413

Maxwell, J. A. (2019). Distinguishing between quantitative and qualitative research: A response to Morgan. *Journal of Mixed Methods Research*, *13*(2), 132–137. https://doi.org/10.1177/1558689819828255

Md Zolkapli, R. B., Mohd Kenali, S. F., Abdul Hadi, N. F., Jaafar, A. J., Mohamad, H. A., Abd Rahman, A. L., & Shaharudin, N. A. D. (2025). Addressing English Grammar Learning Challenges Among Malaysian Islamic Studies Students. *Malaysian Journal of Social Sciences and Humanities*, *10*(1). https://doi.org/10.47405/mjssh.v10i1.3199

Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. *LLT Journal: A Journal on Language and Language Learning*, *27*(1), 98-115. https://doi.org/10.24071/llt.v27i1.8151

Muniandy, J., & Selvanathan, M. (2024). ChatGPT, a partnering tool to improve ESL learners' speaking skills: Case study in a Public University, Malaysia. *Teaching Public Administration*, 43(1), 4-20. https://doi.org/10.1177/01447394241230152

Nuñez, A. A. C., Nuñez, M. S. C., Pachay, J. F. Z., & Bosquez, A. M. C. (2025). Using ChatGPT voice to improve speaking skills in English language learners. Ciencia Latina Revista Cientifica Multidisciplinar, 9(1), 7143-7161. https://doi.org/10.37811/cl_rcm.v9i1.16390

Pangestu, H. P., & Suwartono, T. (2024). Exploring the role of artificial intelligence (AI) support in assisting students' English-speaking skills. *English Education, Linguistics, and Literature Journal*, *3*(2), 134-47. https://doi.org/10.5281/zenodo.13134979

Saraini, Miladiyenti, F., Rozi, F., Haslina, W., & Marzuki, D. (2022). Incorporating mobile-based artificial intelligence to English pronunciation learning in tertiary-level students: developing autonomous learning. *International Journal of Advanced Science Computing and Engineering*, *4*(3), 220-232. https://doi.org/10.62527/ijasce.4.3.92

Senowarsito & Ardini, S. A. (2023). The use of artificial intelligence to promote autonomous pronunciation learning: Segmental and suprasegmental features perspective. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *8*(2), 133-147. https://doi.org/10.21093/ijeltal.v8i2.1452

Suzukida, Y. (2021). The contribution of individual differences to L2 pronunciation learning: Insights from research and pedagogical implications. *RELC Journal*, *52*(1), 48-61. https://doi.org/10.1177/0033688220987655

Uchihara, T., Webb, S., Saito, K., & Trofimovich, P. (2022). Does mode of input affect how second language learners create form-meaning connections and pronounce second language words? *The Modern Language Journal*, *106*(2), 351-370. https://doi.org/10.1111/modl.12775

White, K. S., Yee, E., Blumstein, S. E., & Morgan, J. L. (2013). Adults show less sensitivity to phonetic detail in unfamiliar words too. *Journal of Memory and Language, 68*(4), 362-378. https://doi.org/10.1016/j.jml.2013.01.003

Yahaya, R. S., Madzlan, N. A., & Mastura Muhammad, M. (2021). ESL learners' obstacles in learning speaking skills. *The Malaysian English Language Teaching Association*, *50*(3), 173-181. https://doi.org/10.52696/WYZY7207

Yang, F., Li, K., & Li, R. (2024). AI in language education: Enhancing learners' speaking awareness through AI-supported training. *International Journal of Information and Education Technology*, *14*(6), 828-833. https://doi.org/10.18178/ijiet.2024.14.6.2108

Zhai, X., & Razali, A. B. (2022). International Chinese postgraduate students' adaptation strategies for oral English communication practices in Malaysian higher education institutions. *Education Research International*. https://doi.org/10.1155/2022/6439726

Zhao, Y. (2024). The exploration of artificial intelligence in pronunciation teaching. *Proceedings of the 1st International Conference on Data Science and Engineering (ICDSE 2024), 1,* 127-131. https://doi.org/10.5220/0012829100004547

Zou, B., Liviero, S., Ma, Q., Zhang, W., Du, Y., & Xing, P. (2024). Exploring EFL learners' perceived promise and limitations of using an artificial intelligence speech evaluation system for speaking practice. *System*, *126*. https://doi.org/10.1016/j.system.2024.103497

APPENDIX A

Text A

The North Wind and the Sun

The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shone out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

(Full English text as it occurs in the *Handbook of the IPA* (IPA 1999:39)

Text B

The Boy Who Cried Wolf

There was once a poor shepherd boy who used to watch his flocks in the fields next to a dark forest near the foot of a mountain. One hot afternoon, he thought up a good plan to get some company for himself and also have a little fun. Raising his fist in the air, he ran down to the village shouting 'Wolf, Wolf.' As soon as they heard him, the villagers all rushed from their homes, full of concern for his safety, and two of his cousins even stayed with him for a short while. This gave the boy so much pleasure that a few days later he tried exactly the same trick again, and once more he was successful. However, not long after, a wolf that had just escaped from the zoo was looking for a change from its usual diet of chicken and duck. So, overcoming its fear of being shot, it actually did come out from the forest and began to threaten the sheep. Racing down to the village, the boy of course cried out even louder than before. Unfortunately, as all of the villagers were convinced that he was trying to fool them a third time, they told him, 'Go away and don't bother us again.' And so the wolf had a feast.

(Deterding, 2006)

Text C

The Legend of Mahsuri

Mahsuri was a beautiful, kind-hearted woman who lived in Langkawi, an island surrounded by lush valleys and blue seas. Her grace brought admiration but also jealousy. A jealous villager falsely accused her of adultery. The villagers, quick to judge, believed the rumour without seeking the truth and sentenced her to death. On the day of her execution, Mahsuri stood with dignity and declared, "I am innocent. If I speak the truth, let the earth and sky bear witness to my words." Her voice echoed through the valleys, silencing the birds. She was tied to a tree and stabbed. As the dagger pierced her, white blood flowed from the wound, signifying her innocence, shocking the crowd. Realizing their grave mistake, the villagers were filled with guilt and remorse. Mahsuri's final words were haunting: "My spirit will stay until the truth is revealed." Her curse lingered, bringing storms, famine, and despair to the once-prosperous village. Years later, the truth came to light when the jealous villager confessed her lies. Though peace returned, the villagers could not erase the scar left by their false judgment. Mahsuri's tale teaches us to value truth and weigh our words, for they can wound like daggers.

(Author 1, 2025)