

# Enhancing English Pronunciation with ChatGPT: An Effective AI Tutor for Transforming Language Instruction and Learning

Hanadi S. Mirza\*

Lebanese University, Faculty of Education Beirut, Lebanon

\*Corresponding author: Hanadi S. Mirza, Lebanese University, Faculty of Education Beirut, Lebanon.

Submitted: 04 September 2025 Accepted: 10 September 2025 Published: 17 September 2025

doi <https://doi.org/10.63620/MKPJSSHR.2025.1043>

**Citation:** Mirza, H. (2025). *Enhancing English Pronunciation with ChatGPT: An Effective AI Tutor for Transforming Language Instruction and Learning*. *J of Soc Sci & Hum Res*, 2(5), 01-10.

## Abstract

Artificial Intelligence (AI) has emerged as a transformative tool in language instruction, particularly for enhancing oral fluency and pronunciation of non-native English speakers. This action research investigates the extent to which ChatGPT, as an AI-powered language tutor, can improve the pronunciation of learners of English as a foreign language (EFL)—specifically, for senior student-teachers majoring in Early Childhood Education (ECE), whose profession requires frequent read-aloud sessions for young learners. A cohort of 11 female ECE student-teachers used ChatGPT's voice mode and the SpeechAnalyzer app to practice English sounds, stress patterns, and intonation. Participants audio-recorded three stories across pretest, posttest, and delayed posttest phases. These stories were generated by ChatGPT, and each included multi-syllabic words, interdental sounds, definite and indefinite articles, statements and questions. Additionally, a survey was administered to explore participants' perceptions of the advantages, challenges, and effectiveness of these AI tools. Findings indicated that ChatGPT significantly improved participants' pronunciation, stress accuracy, and learner autonomy by offering individualized feedback and a flexible learning experience. Speech Analyzer complemented this process by identifying misread and omitted words and by providing visual cues for improvement. The pronunciation scores increased from a pretest mean of 5.0 to a delayed posttest mean of 7.67. Despite these benefits, challenges included difficulties in effectively prompting ChatGPT, interpreting AI generated feedback, capturing the phonetic nuances, and addressing ethical concerns. Recommendations for future AI integration in pronunciation training are also discussed.

**Keywords:** ChatGPT, Artificial Intelligence, Pronunciation, EFL Learners, Higher Education, Learner Autonomy, Individualized Feedback.

## Introduction

For many university students speaking English as a Foreign Language (EFL), poor pronunciation remains a persistent barrier to effective oral communication. Despite years of formal instruction, numerous non-native university students often struggle to produce English sounds clearly and accurately, which can negatively affect their intelligibility, self-confidence, and willingness to engage in spoken interactions. This major difficulty can influence the success of communication in academic, professional, and social contexts. Clear and accurate pronunciation enables listeners to understand messages with minimal effort, and it plays a crucial role in supporting both the speaker's credibility and the listener's comprehension [1, 2]. Moreover, pronunciation is perceived as a marginalized component of the

language curriculum in higher education. Teachers often prioritize grammar, vocabulary, and the four language skills (reading, writing, listening, and speaking) while pronunciation receives little explicit focus. One reason for this neglect is the misconception that pronunciation will develop naturally as students gain more exposure to English, or that it is less important than the structural aspects of language. Additionally, instructors may feel underprepared to teach pronunciation, particularly the suprasegmental features, due to limited training in this area during their own education [3].

To develop adults' pronunciation skill, both explicit instruction and guided practice are required [4]. In the broader field of language learning, strategy-based instruction has gained recogni-

tion for its effectiveness in improving various language skills; however, pronunciation learning strategies (PLS) have not been as widely adopted or integrated into classroom routines compared to strategies for vocabulary, reading, or grammar [3]. Still, PLS proved effective when continuously practiced, such as self-recording, interacting with native speakers, and focusing on word stress and intonation. Incorporating both segmental and suprasegmental features can help EFL learners develop their English pronunciation skill [5]. In contrast, and over the past two decades, the growing integration of technology in language education has opened new avenues for pronunciation instruction. Computer-assisted pronunciation training (CAPT) tools and artificial intelligence (AI)-powered applications have become increasingly popular, offering learners opportunities for individualized, flexible, and repeated pronunciation practice. Studies have demonstrated that technological tools, including automated feedback systems and speech visualization software, can enhance learners' awareness of their pronunciation and contribute to noticeable improvements in accuracy and fluency [6]. However, several studies highlighted the impact of AI-tools on improving the segmental features of the English pronunciation more than its suprasegmental features [7].

### **Pronunciation Features**

Pronunciation is composed of two primary components: segmental and suprasegmental features. Segmental features refer to the individual sounds of a language, namely consonants and vowels, which are the smallest units of sound or phoneme. As explained by Sharma (2021), the English consonants constitute 24 phonemes while the vowels are 20 phonemes. These 44 sounds in total are directly linked to the alphabetic system of English whereby the spelling of various consonants and vowels can produce different sounds. A special attention to vowels should be made seen their variations in length and quality. Failing to pronounce a vowel accurately can lead to misunderstanding. Consequently, students are often trained to recognize and reproduce sounds that are closely tied to spelling patterns, making segmental features more accessible for teachers to present and easier for learners to notice and correct. In contrast, suprasegmental features — including stress, intonation, rhythm, and connected speech patterns — operate at a broader level across words, phrases, and entire utterances. These features are less visible in written text but are crucial for conveying meaning, emotion, and discourse-level coherence in spoken English. Despite the importance of these features, suprasegmentals are frequently overlooked in both teacher training and language instruction [8]. Although the Common European Framework of Reference for Languages (CEFR) provides detailed descriptions of language competencies at different proficiency levels (Council of Europe, 2001), there is a lack of comprehensive guidance on the phonological features that learners are expected to master at each level. Consequently, this leaves pronunciation development largely to the discretion of teachers and textbook writers. As Galaczi et al. (2011) have noted, this gap has led to limited systematic attention to pronunciation in both teaching practice and language assessment.

### **Advantages of AI Tools for Pronunciation Improvement**

AI-based tools, especially chatbots, offer accessible, personalized, and interactive pronunciation training [9, 10]. Prior studies confirm that AI-driven tools enhance oral fluency and pronun-

ciation, provide immediate corrective feedback, and promote self-directed practice [8-13]. Among these tools, ChatGPT stands out for its potential to enhance pronunciation through conversational engagement [10]. Aryanti and Santosa (2024) further noted that AI applications can reduce learners' anxiety by creating a safe practice environment. Fatima (2024) emphasized the flexibility AI offers, enabling learners to practice pronunciation anytime and anywhere. Sukma and Ardini (2023) emphasized that such tools promote learner autonomy and encourage self-directed practice. They also help learners retain pronunciation skills better than conventional methods due to their interactive approach and rich feedback [14]. Mohammadkarimi (2024) confirmed similar perceptions, noting that students found AI tools effective for improving pronunciation accuracy. According to Mohammadkarimi (2024), AI-powered tools can save class time by offering individualized feedback using speech recognition, real-time assessments, and adaptive learning strategies. These tools identify specific pronunciation errors and provide immediate correction, addressing gaps that traditional classrooms may overlook [14]. Hence, the use of AI tools in class can reduce teachers' workload by automating pronunciation error detection, freeing educators to focus on higher-order instructional goals [15].

### **Challenges of AI Tools in Pronunciation Training**

Despite their benefits, AI tools often focus on segmental pronunciation. The emphasis is on individual sounds with less focus on suprasegmental features such as stress and intonation [8]. Some AI tools fail to capture subtle pronunciation variations, making it hard for learners to make precise improvements [10]. Another challenge lies in the difficulty of interpreting AI generated feedback. Learners might misunderstand AI feedback mainly when it lacks detailed explanations [16]. Also, speech recognition errors can lead to incorrect feedback which can frustrate learners [18]. Generally, AI models are often trained on native English speakers, making them less effective for non-native English learners with diverse linguistic backgrounds [15].

Language proficiency and ethical concerns pose additional challenges. The usefulness of AI-tools can vary according to learners' different proficiency levels. Beginner learners, for instance, benefited the most, as AI tools provide foundational pronunciation support [15]. However, advanced learners found AI tools less effective for fine-tuning pronunciation, as most applications focus on basic phonetic corrections rather than advanced phonological training [8]. Integrating AI tools in pronunciation practice presents some ethical concerns. Data security and potential misuse of personal speech data have been raised [10]. Cost and accessibility also limit the widespread adoption of AI tools [13]. While many students find AI tools beneficial, they often emphasize the complementary role of human interaction in achieving optimal results in language learning [15-17]. In addition, Yang and Kyun (2022) stressed the importance of teacher's intervention when learning a new language.

### **Needed AI Development for Pronunciation Learning**

Current AI powered pronunciation tools need further development. Such tools should incorporate enhanced feedback for suprasegmental features such as stress, intonation, and rhythm training to provide a more holistic pronunciation improvement experience [7]. AI tools, moreover, need more diverse training

datasets to improve speech recognition accuracy, which include non-native English speakers to reduce pronunciation bias [18]. It's also recommended to combine AI pronunciation tools with human instruction to enhance pronunciation learning by allowing teachers to intervene when necessary [16]. AI-based pronunciation apps should be integrated into structured curricula to ensure systematic learning rather than relying solely on independent practice [14]. Additionally, governments and institutions should work on making AI tools more accessible and affordable for learners from different socioeconomic backgrounds [15]. Clear ethical guidelines for data privacy should also be established to prevent misuse of personal speech data collected by AI applications [10].

### Purpose of the Study and Research Questions

Pronunciation is essential for EFL learners, especially future teachers who must model accurate language use in the classroom. Student-teachers majoring in Early Childhood Education (ECE) frequently struggle with pronunciation, limiting their confidence during read-aloud activities. This study investigates whether AI tools, namely ChatGPT and SpeechAnalyzer, can address this challenge by enhancing pronunciation and promoting learner autonomy. This approach helps bridge the gap between traditional teaching methods and modern technological advancements, ensuring that future teachers are well-equipped to foster accurate pronunciation in their students. Moreover, this research was designed to answer the following questions:

RQ1: To what extent can ChatGPT and SpeechAnalyzer improve future ECE teachers' English pronunciation?

RQ2: How do ECE student-teachers perceive the role of these AI tools in enhancing their English pronunciation?

### Methodology

This study adopts a qualitative research design namely action research. It aims to explore practical ways to improve university students' English pronunciation using ChatGPT and SpeechAnalyzer as AI-based pronunciation tools. Action research, as discussed by Creswell (2015), is a reflective and participatory approach to research that seeks to address specific issues within educational settings. It focuses on the detailed exploration of a problem within its real-life context.

### Participants and Site

Eleven female student-teachers (N= 11) majoring in Early Childhood Education took part in this study. They were enrolled in the 5th semester in the Faculty of Education at the Lebanese University. Their age ranged between 21 and 26 years old. They were attending a 40-hour course entitled "Teaching Speaking and Writing Skills in Primary classes" delivered by the researcher during the fall semester of the academic year 2024-2025. The participants utilized ChatGPT and SpeechAnalyzer to enhance their English pronunciation integrated into the speaking project required for the course.

### Methods of Collecting and Analyzing Data

Data were collected using a) audio-recordings, b) ChatGPT as an AI tutor to practice pronunciation outside class, c) SpeechAnalyzer to detect mispronounced words and speed delivery, and d) a survey including open-ended and closed-ended questions to explore participants' views on the use of these 2 AI tools in their pronunciation practice. Participants recorded 3 stories across

pretest, posttest, and delayed posttest phases. These stories were generated by ChatGPT at the grade 2 level and were around 110 words each including problematic sounds, multi-syllabic words, and declarative and interrogative sentences. Participants filled the survey at the end of the semester. The data resulting from the survey open-ended questions were analyzed using thematic coding. Descriptive statistics helped analyze missing/misread words, pauses, and the closed-ended questions. Being the researcher and instructor, I protected my participants privacy by ensuring anonymity and confidentiality.

### Study Procedures

The procedures of this study were designed to systematically guide ECE student-teachers through a structured process aimed at improving their English pronunciation. The process was organized into six key stages.

#### Stage 1: Pretest (Story 1)

At the beginning of the semester (Week 1), participants were asked to audio-record Story 1, titled "Playground". This initial recording served as the baseline to assess their existing pronunciation skills. The recordings were analyzed using SpeechAnalyzer application (app) which helped identify errors such as mispronounced words, omitted words, and pauses across all submissions.

#### Stage 2: Targeted Feedback

After analyzing the pretest recordings, the researcher/instructor provided individualized and group feedback. Emphasis was placed on common problem areas such as the interdental sounds /ð/ and /θ/, the articles "a," "an," and "the," the schwa sound /ə/ in multi-syllabic words. These sounds were difficult for these non-native speakers whose first language is Arabic. Appropriate word and sentence-level stress patterns and proper intonation in statements and questions were also challenging.

The difficult sounds encountered by the participants are briefly described below. This study addressed the interdental fricative sounds, which are a well-documented source of pronunciation difficulty for EFL learners. These include the voiceless interdental sound /θ/ — as found in initial positions in words like *think* and *thought* — and the voiced interdental sound /ð/, present in words like *they* and *that*. Many learners substitute these sounds with /s/, /z/, /t/, or /d/ due to the absence of interdental articulation in their native spoken language [19]. The second focus was on the definite and indefinite articles. Learners were trained to distinguish and produce the two context-dependent pronunciations of the article *the*. When the word following the *the* begins with a consonant sound, *the* is pronounced with a schwa /ðə/ (e.g., *the table*, *the book*). When the following word begins with a vowel sound, *the* is pronounced with a long vowel /ði:/ (e.g., *the apple*, *the umbrella*). This subtle distinction is often overlooked by learners due to the wrong application of this pronunciation rule. Additionally, the study emphasized the proper articulation of the indefinite articles "a" and "an". These articles are usually pronounced in rapid speech with the schwa sound /ə/ (e.g., *a car*, *a toy*) or /æn/ before vowel-initial words (*an elephant*, *an orange*). Learners tend to overemphasize or misplace stress on these articles. Another critical focus was the correct production of the schwa /ə/ in multi-syllabic words. In spoken English, the schwa is typically found in unstressed syllables, especially in the

final syllable of words such as teacher, nation, and animal. The schwa sound is frequently mispronounced or replaced by a more distinct vowel by learners, negatively affecting intelligibility.

### Stage 3: ChatGPT Integration

In this stage, students were guided on how to use ChatGPT voice mode for pronunciation practice. Learners were encouraged to select problematic words, phrases, and sentences from Story 1 and engage in targeted practice using specific question prompts such as:

1. Where does the stress fall in these words?
2. How should these words or phrases be pronounced correctly?
3. How are words with interdental sounds articulated?
4. How should the articles "a," "an," and "the" be pronounced in context?
5. Can you model the correct rising or falling tone for this sentence or question?

This AI-assisted stage practicing with ChatGPT was designed to help learners receive immediate feedback and refine their pronunciation through repeated practice outside the classroom.

### Stage 4: Posttest (Story 2)

In Week 5, participants completed a posttest by audio-recording a second story titled "The Missing Doll" after practicing with ChatGPT. The recordings were once again analyzed using SpeechAnalyzer to measure changes in three key areas: speech rate, the frequency of missing words, and the accuracy of word pronunciation. Students also submitted screenshots of their ChatGPT conversations and their performance analysis from

SpeechAnalyzer as documentation of their learning process in addition to their audio-recording of Story 2.

### Stage 5: Delayed Posttest (Story 3)

In Week 11, a delayed posttest was performed using a third story titled "Sandy's Big Adventure." This step aimed to assess the long-term retention and transferability of pronunciation improvements. Participants were instructed to practice the same way as in Step 3 using ChatGPT, followed by recording their reading of the story and submitting both the audio file and supporting screenshots from ChatGPT and SpeechAnalyzer.

### Stage 6: Feedback and Reflection

Towards the end of the semester, participants were asked to complete a feedback survey. The survey consisted of two sections: the first included 12 statements rated on a 5-point Likert scale measuring their perceptions of using ChatGPT as a pronunciation tutor. The second section comprised open-ended questions, inviting learners to reflect on the advantages, challenges, and need for improving the integration of AI tools, specifically ChatGPT and SpeechAnalyzer, in enhancing their English pronunciation and ultimately their read aloud sessions.

## Results

The findings resulting from practicing with ChatGPT and SpeechAnalyzer are detailed below.

### ChatGPT as an Effective AI-Based Pronunciation Tool

The questionnaire used was adapted from Mohammadkarimi (2024) while specifying ChatGPT as the AI tool in this study.

**Table 1:** Rating the Effectiveness of ChatGPT in Enhancing English Pronunciation

|    | Statement   | Average |
|----|---|---------|
| 1  | ChatGPT has significantly improved my English pronunciation.  | 4.6     |
| 2  | My pronunciation accuracy improved since using ChatGPT.   | 4.6     |
| 3  | ChatGPT helped me identify/correct specific pronunciation errors  | 4.4     |
| 4  | ChatGPT has increased my confidence in speaking English.  | 3.3     |
| 5  | ChatGPT provided immediate/beneficial feedback on my pronunciation.   | 4.6     |
| 6  | ChatGPT is motivating for practicing English pronunciation.   | 4.5     |
| 7  | ChatGPT made learning pronunciation more enjoyable.   | 4.2     |
| 8  | Interpreting ChatGPT's feedback accurately is sometimes challenging.  | 2.6     |
| 9  | ChatGPT doesn't always capture the nuances of my pronunciation accurately.  | 3.5     |
| 10 | My current English proficiency level affects how I perceive the effectiveness of ChatGPT as AI- based pronunciation tool. | 4.1     |
| 11 | ChatGPT is user-friendly if I'm familiar with technology.   | 4.4     |
| 12 | ChatGPT positively impacted my overall English learning experience.   | 4.1     |

The questionnaire contained 12 statements evaluating the participants' views on the effectiveness of ChatGPT as an AI- pronunciation tutor in enhancing the participants' English pronunciation, confidence, engagement, and overall learning experience. Each statement was rated on a 5-point Likert scale ranging from Strongly Agree to Strongly Disagree. The average ratings for each statement are displayed here after. Based on the results in Table 1, respondents strongly agreed (4.6/5) that ChatGPT significantly improved their pronunciation. They reported a noticeable enhancement in their overall pronunciation accuracy (4.6/5). ChatGPT was generally considered user-friendly, de-

pending on the users' familiarity with technology (4.4/5). This AI tool was perceived as highly engaging and motivating for pronunciation practice with an average of 4.5/5. Also, ChatGPT was found effective in helping users identify and correct pronunciation mistakes with an average of 4.4/5. The participants highly valued the immediate feedback provided by ChatGPT (4.6/5). Most participants found that ChatGPT made learning pronunciation more enjoyable (4.2/5). However, the impact on users' confidence in speaking English was moderate (3.3/5), indicating room for improvement. Some users struggled to accurately interpret the feedback provided by ChatGPT (2.6/5). There were



some concerns that ChatGPT does not always capture pronunciation nuances effectively (3.5/5). Finally, participants' perception of ChatGPT's effectiveness was influenced by their English proficiency level (4.1/5), and most of them (4.1/5) agreed that using ChatGPT positively contributed to their overall English learning experience.

### Sample Conversations with ChatGPT

Below are sample conversations between some participants and ChatGPT voice mode showing how learners were practicing stress, difficult sounds, and intonation using Story 2: The Missing Doll (see Table 2).

**User:** Where is the stress in the following words: missing, terrible?

**ChatGPT:** Sure! Here's where the stress falls in those words:

**Missing:** Stress on the first syllable - MISsing. Terrible: Stress on the 1st syllable - TERrible /

**User:** How do you say the in the missing and the in the dollhouse?

**ChatGPT\*:** The word "the" can be pronounced in two ways depending on the following sound:

1. "The" (/ðə/): Used before words that begin with a consonant sound. For example, "the missing" (since "missing" starts with a consonant sound, /m/).
2. "The" (/ði:/): Used before words that begin with a vowel sound. "The dollhouse" (since "dollhouse" starts with a vowel sound, /dɔ:/).

**User:** How do you say: a terrible mess? As (A)?

**ChatGPT\*:** You should say "a terrible mess" with the article "a" pronounced as the alphabet "A" because "terrible" starts with a consonant sound (/t/).

**User:** Why? Isn't "a" a schwa here?

**ChatGPT:** Yes, exactly! In "a terrible mess," the article "a" is pronounced with a schwa sound, /ə/, because it comes before a

word that starts with a consonant sound /t/.

**User:** What is the intonation in the sentences: She could not find her. Did you see my doll?

**ChatGPT:** For these sentences "She could not find her" you should use a falling intonation. Start with a normal pitch and let your voice drop at the end to show certainty and finality. So, it would sound like: She could not find her. For the sentence "Did you see my doll?" use a rising intonation at the end to show you're asking a question.

Prompting ChatGPT effectively is crucial to get accurate results and comprehensible feedback. However, in 2 of the chats, ChatGPT responded with inaccurate explanations of the rules for pronouncing the definite article "the" and the indefinite article "a". ChatGPT\* said that "the" is pronounced as /ði:/ before the word "doll" which starts with a vowel though the words "missing" and dollhouse" both start with a consonant /m/ and /d/ respectively. Similarly, ChatGPT\* agreed first with the participant that the indefinite article "aa" is pronounced as the alphabet letter "A" in the phrase "a terrible mess". When the participant inquired whether the article "a" should be a schwa, ChatGPT confirmed that "a terrible mess" starts with a schwa sound /ə/. Regarding the localization of the stress in the provided disyllabic words and the rising or falling intonation in the given short sentences and question, ChatGPT could provide detailed explanations. In other conversations, participants copied and pasted the whole story and asked ChatGPT to read it aloud for them before they recorded it themselves.

### SpeechAnalyzer: Missing and Misread Words and Pauses

This section shows one of the participant's read aloud performance of Story 2 using SpeechAnalyzer. (X) represents the missing words and the misread words are underlined.

**Table 2:** Transcription of Story 2 Using SpeechAnalyzer App

|  |
|--|
| <p><b>Story 2:</b> The Missing Doll Jane was very sad. She could not find her doll. Where is she? Did you see my doll? Her mom said, "Look in your room. Maybe your doll is there." Jane looked in her room. It was a terrible mess. She started to clean her room. She put all the books on the shelf. She put her blankets on her bed. She picked up her clothes off the floor. "I found her", Jane said happily. "My doll was hiding under my clothes. I think she was playing hide and seek." Jane put her dear doll in the pink dollhouse. Her room was clean, and now it was time to play.</p> |
| <p><b>Transcription of Story 2:</b> the missing doll Jane was very sad she could not find her dog where is she did you see my doll her mom said look in your room maybe your dolls there Jean looked in her room it was X terrible mess X started to clean her room she put all the books on the shelf she put hair blankets on her bed she picked up her clothes off the floor I found her Jean sad happily my daughter was hiding under my clothes I think she was playing hide and seek X put hair dear doll in the pink doghouse her room was clean X now it was time to pea lay</p>             |

Table 2 displays the original text "The Missing Doll" and contrasted it with the speech-to-text provided by SpeechAnalyzer. The (X) show that the words were omitted such as "she, a Jane". More words were misread and hence misspelled by this AI tool. The following words "Jane, doll, said, dollhouse, her, doll is, play" were misread respectively as "Jean, dog or daughter, sad,

doghouse, hair, dolls, pea lay".

The most frequent missing (Ms) and misread words obtained by the SpeechAnalyzer app for all recordings of the 3 stories were displayed in Table 3 and summarized below.

**Table 3:** Frequent Missing and Misread Words in the Recordings of the 3 Stories

| Playground    |    |         | The Missing Doll |    |         | Sandy's Big Adventure |    |         |
|---------------|----|---------|------------------|----|---------|-----------------------|----|---------|
| Words/Phrases | Ms | Misread | Words/Phrases    | Ms | Misread | Words/Phrases         | Ms | Misread |
| Molly         | 5  | Money   | Jane             | 9  | -       | Sandy's               | 3  | Sindy   |

|           |   |         |           |    |          |           |   |        |
|-----------|---|---------|-----------|----|----------|-----------|---|--------|
| Counted   | 3 | -       | Dollhouse | 3  | Doghouse | Burrow    | 8 | Borrow |
| To 100    | 4 | 200     | Doll      | 18 | -        | Hill      | 3 | -      |
| A turn    | 3 | -       | Mess      | 3  | -        | Owl       | 6 | Our    |
| The swing | 4 | -       | Jane said | 2  | Jean sad | Grandpa's | 8 | -      |
| Said      | - | Sad     | My doll   | -  | My dog   | A note    | - | A not  |
| Were      | - | They're | Her       | -  | Hair     | That read | - | Thread |
|           |   |         | -         | -  | -        | Saw       | - | So     |

According to Table 3, it was generally observed that proper nouns such as “Molly,” “Jane,” “Sandy,” and “Grandpas” were often misread or omitted. In the recordings of the story Playground, “the swing” and “to 100” were missing each 4 times and “could” and “a turn” 3 times. The proper noun “Molly” was omitted 5 times. Moreover, the following words “Molly”, “were”, “to 100”, and “said” were sometimes misread as “money”, “they’re”, “200”, and “sad” respectively. In the recordings of the second story The Missing Doll, the following words were omitted: “Jane” (9 times), “dollhouse” (3 times), “doll” (18 times), “Jane said” (2 times), and “mess” (3 times). “My doll” was sometimes misread as “my dog”, “Jane said” as “Jean sad”, and “her” as “hair”. The words “doll” and “dollhouse” were especially problematic. When checking the recordings of the third story Sandy’s big Adventure, “Sandy’s” was missing 3 times, “hill” 3 times, and “grandpa’s” (8 times). The word “burrow” was missing 8 times and was misread sometimes as “borrow”. Also, “owl” was missing 6 times and was misread at times as “our”. “A note”, “that read”, and “saw” were often misread as “a not”, “thread”, and “so” respectively.

Besides missing and misread words, the average pauses were also provided for the 3 stories. According to the SpeechAnalyzer app, the base line for the pauses is 0.4. In the recordings of the Playground Story, the pause average was 0.29 which means that participants were somehow rushing while reading the story aloud. In the recordings of Story 2, the participants were reading aloud slightly more slowly with a pause average of 0.32. the recordings of the third story showed that the participants were reading with a moderate speed whereby the average was 0.38, which was close to the baseline. This is due to participants’ familiarity with technology and the procedures of practicing with ChatGPT and SpeechAnalyzer.

Based on the data provided by SpeechAnalyzer and which were shown in Table 2 and Table 3, the student-teachers were able to locate there missing and misread words by simply comparing their transcribed text to the original story. They were also able to monitor the speed with which they were reading the story aloud by checking the average pauses. However, it's noteworthy mentioning that the speech or audio transcribed by the Speech

Analyzer app does not use any uppercase letters except for the proper noun Jane to indicate the beginning of sentences nor any punctuation marks to signal the end of various sentences and questions. Therefore, participants were not able to evaluate their intonation or rhythm seen that there are no visual cues provided by this tool.

### ChatGPT and SpeechAnalyzer as Effective AI Tools

Before using ChatGPT and SpeechAnalyzer, the participants rated their pronunciation skills at an average score of 5 on a scale from 1-10 whereby 1 is very poor and 10 is excellent. After practicing with these AI tools, the average score increased to 7.67, indicating significant improvement in their English pronunciation. Additionally, the participants answered open-ended questions related to the advantages, challenges, effectiveness, and recommendations for using ChatGPT and SpeechAnalyzer to improve their English Pronunciation. The results were displayed in Table 4 and summarized below.

When asked about the advantages of using AI Tools for Pronunciation Practice, participants identified several key benefits of using ChatGPT and SpeechAnalyzer. The participants verbatim below reflected the benefits of ChatGPT in improving their English pronunciation. ChatGPT “provided instant feedback on pronunciation and stress patterns”. Also, “it provides a clarification of pronunciation rules, such as vowel sounds, stress patterns, or the difference between similar sounds.” “ChatGPT helped me realize the many mistakes I made when I try to pronounce different words.” “I didn’t realize that we were pronouncing even the simplest words, like ‘the,’ incorrectly until we took this course.” “I really enjoyed listening to ChatGPT and learned how to pronounce every word in a professional way... to help improve my speaking skills as a teacher”. As for the benefits of using SpeechAnalyzer, the participants’ answers were reported as follows. It provides visual and written feedback. SpeechAnalyzer “Provided visual feedback on speech patterns such as: if I am saying the word correctly.” It also helps improve speech clarity and speed. “It helped me slow my pace in reading since I talk fast and showed me which words were not pronounced properly.” These advantages indicated that both tools serve as valuable resources for learners looking to refine their pronunciation skills.

**Table 4:** Views on the Usage of ChatGPT and SpeechAnalyzer

|            | ChatGPT  | Speech Analyzer  |
|------------|--|--|
| Advantages | <ol style="list-style-type: none"> <li>1. Provides instant feedback</li> <li>2. Clarifies pronunciation rules</li> <li>3. Helps recognizing mistakes</li> <li>4. Supports teachers and students</li> </ol> | <ol style="list-style-type: none"> <li>1. Provides visual and written feedback</li> <li>2. Helps improve speech clarity and speed</li> <li>3. Enhances pronunciation accuracy</li> </ol> |

|                 |   |  |
|-----------------|---|--|
| Challenges      | <ol style="list-style-type: none"> <li>1. Sometimes lacks detailed pronunciation explanations</li> <li>2. Has contextual limitations</li> <li>3. Requires precise prompting</li> </ol>              | <ol style="list-style-type: none"> <li>1. Couldn't be downloaded by 1 participant</li> <li>2. Struggles with accent recognition</li> </ol> |
| Effectiveness   | <ol style="list-style-type: none"> <li>1. Helpful AI tools but not entirely sufficient</li> <li>2. Need for diverse learning resources</li> <li>3. Stress the importance of practice</li> </ol>     |  |
| Recommendations | <ol style="list-style-type: none"> <li>1. Useful for teachers and student-teachers</li> <li>2. Not very suitable for young kids</li> <li>3. Must be integrated with other learning tools</li> </ol> |  |

Despite the benefits mentioned earlier, participants reported several challenges when using the tools. Two participants mentioned that they found ChatGPT challenging in its inability to provide detailed pronunciation explanations. "Sometimes, stress and intonation needed more detailed explanations." It also requires precise prompting. "I have to ask the question in a very specific and detailed way to get the information I need about pronouncing different words." As for the use of SpeechAnalyzer, one participant reported having technical issues in downloading this AI tool on her laptop and mobile phone.

According to some participants, these 2 AI tools were highly effective for Pronunciation Improvement yet not entirely sufficient. "While ChatGPT and SpeechAnalyzer were considered very helpful, they were not enough alone. Pronunciation improvement also depends on real-life practice." "I also learned a lot of English from interviews, podcasts, songs, and videos I watch and listen to online." Others emphasized the importance of practice. "I recorded the story several times until I was satisfied with the results of SpeechAnalyzer." After using ChatGPT and SpeechAnalyzer, practice and repetition are the most important factors to master pronunciation. At the end of the survey, most participants highly recommended ChatGPT and SpeechAnalyzer to both in-service and pre-service teachers aiming at perfecting their English pronunciation. Some added that teachers and even older students can use these AI tools to "make sure they are

pronouncing words correctly, but they also need to read, watch English content, and engage in real-life communication." Some specified that such AI tools can be difficult to use by young elementary kids; however, "K-6 teachers can use these tools in their lesson plans" to accurately speak English in class and read stories aloud to their students.

### Difficult English Sounds

The difficult sounds mentioned earlier in Stage 2 i.e. the interdental sounds, the articles the, a, an, schwa, and proper intonation were not shown as missing or misread words when using the SpeechAnalyzer app. This is due mainly to 2 factors: a) the participants could clearly pronounce these sounds and b) SpeechAnalyzer as a speech-to-text recognition tool was able to recognize various sounds regardless the accurate pronunciation according to the English language. For example, in "the/za/zi/ze" playground were all transcribed correctly as /the/. The main aim of SpeechAnalyzer as public speaking enhancement tool was to help speakers' talks be comprehensible and fluent. Therefore, the progress in pronouncing these difficult sounds accurately was evaluated by the researcher. These sounds were identified in the pretest using Story 1 and are listed in Table 5, which showed the number of participants who mispronounced these sounds in the recordings of the stories "Playground" and "Sandy's Big Adventure". (See Appendix A)

**Table 5:** Mispronounced Difficult Sounds in Story 1 and Story 3

| Difficult Sounds | Pretest Story 1 | Delayed Posttest Story 3 |
|------------------|-----------------|--------------------------|
| THis -THink      | 4               | 5                        |
| a/an             | All (11)        | All (11)                 |
| The + C          | 8               | 6                        |
| The + V          | 3               | 6                        |
| Schwa            | 1               | 1                        |
| Stress           | 4               | 2                        |
| Her              | 4               | 3                        |

According to Table 5, there were slight improvements in the articulation of some problematic sounds by some participants. However, certain errors remained despite the pronunciation practice with ChatGPT, the AI tutor. Four (4) student-teachers continued to struggle with interdental sounds  $\theta$  and  $\delta$ . All participants mispronounced the indefinite articles  $\text{ə}$  a and  $\text{æn}$  an in the recordings of both stories 1 and 3. Moreover, five (5) participants frequently mispronounced "the" / $\delta\text{ə}$ / before consonants in Story 1 while 3 additional student-teachers mispronounced "the" / $\delta\text{i}$ / before vowels in Story 3. Regarding stress patterns, the number of incorrect stress placements decreased from 4 errors

in Story 1 to 2 errors in Story 3, likely due to repeated practice with ChatGPT. Although participants engaged in pronunciation practice with ChatGPT, many of these errors persisted from the pretest to the delayed posttest, as shown in Table 5. These challenges can be attributed to the limitations of ChatGPT which Lacks automatic phonetic distinction. ChatGPT does not recognize subtle phonetic variations, such as those between  $\delta\text{ə}$ ,  $\delta\text{i}$ , and  $\text{zə}$  unless participants explicitly prompted it to explain phonological rules and practiced pronunciation through repetition. Guided input is needed to internalize these distinctions. Moreover, SpeechAnalyzer accurately transcribed different variations

of ðə and ði as long as the pronunciation was clear and audible. However, these nuances had to be identified by the researcher, and the associated spelling and pronunciation rules were reinforced in class. While AI tools facilitated pronunciation practice, these findings underscore the need for more refined AI feedback mechanisms and explicit instruction to help learners overcome persistent phonetic challenges.

## Discussion and Conclusion

The findings of the study are discussed in this final section and the 2 research questions are answered below.

**RQ1: To what extent can ChatGPT and SpeechAnalyzer Improve Future ECE Teachers' English Pronunciation?**

The results in Table 1 indicate that ChatGPT was highly effective in enhancing English pronunciation, particularly in pronunciation improvement (4.6), accuracy enhancement (4.6), and immediate feedback (4.6). Additionally, the tool was well-received for its ability to engage and motivate learners (4.5). However, certain areas require improvement, especially in boosting users' confidence in speaking English (3.3) and ensuring clearer interpretation of feedback (2.6). Concerns regarding ChatGPT's ability to accurately capture pronunciation nuances (3.5) suggest that AI-based pronunciation tools need further refinement to address this limitation. Moreover, the results displayed in Table 3 showed the frequency of words omitted or misinterpreted when using SpeechAnalyzer to process audio recordings of the three stories. Proper nouns such as Molly, Sandy, and Jane were frequently omitted or sometimes wrongly substituted. Additionally, some action verbs—said, were, and saw—were often misread as sad, they're, and saw, respectively. Several words posed pronunciation challenges for participants, leading to recognition errors in SpeechAnalyzer. For instance, doll, burrow, and her were often mispronounced and subsequently transcribed as dog, borrow, and hair. Furthermore, participants who initially read Story 1 at a relatively fast rate (pauses = 0.29) demonstrated improved pacing in Story 3, achieving a more moderate reading speed with an average pause duration of 0.38. This suggests that, over time, learners developed greater control over their speech rhythm, possibly due to practice and increased familiarity with the AI tools.

The findings mentioned above highlight the valuable role that AI tools—specifically ChatGPT and SpeechAnalyzer—played in enhancing the English pronunciation skills of ECE student-teachers. These results are consistent with the growing body of literature that recognizes AI as an effective, accessible, and individualized resource for language learning, especially in improving pronunciation [12-16]. The observed increase in participants' pronunciation scores from a pretest mean of 5 to a delayed posttest mean of 7.67 reinforces previous research on AI-driven pronunciation tools. Similar to findings reported by Humardhiana (2022) and Aryanti & Santosa (2024), this study demonstrated that AI-powered feedback—both auditory through ChatGPT and visual through SpeechAnalyzer—supported learners in identifying errors, adjusting pronunciation, and improving accuracy over time. Additionally, the participants' feedback echoed Mohammadkarimi's (2024) observations that learners value AI for its instant feedback, especially in contexts where traditional classroom feedback is limited by time and teacher availability.

**RQ2: How do ECE Student-Teachers Perceive the Role of These AI Tools in Enhancing their English Pronunciation?**

All participants acknowledged that ChatGPT and SpeechAnalyzer contributed to improving their English pronunciation. ChatGPT mainly helped them identify accurate pronunciation, stress patterns, and appropriate intonation, while SpeechAnalyzer provided feedback on mispronounced words that were either omitted or replaced with unintended alternatives. These AI tools allowed participants to practice at their own pace, receive individualized feedback, and develop greater learning autonomy.

However, some participants encountered difficulties in effectively prompting ChatGPT and interpreting the feedback generated by both AI tools. Despite these challenges, all future early childhood education teachers in the study recommended the use of ChatGPT and SpeechAnalyzer for both practicing teachers and other student-teachers to enhance their own pronunciation as well as that of their EFL students. Participants also expressed that while these AI tools could be highly beneficial for older students, young elementary learners might struggle with using them independently and correctly interpreting the provided feedback. Respondents also emphasized the importance of supplementing AI-based pronunciation practice with other resources such as podcasts and videos in addition to engaging in real-life communication.

These findings highlight both the strengths and limitations of AI-assisted pronunciation tools. While ChatGPT and SpeechAnalyzer effectively support pronunciation practice, they require further refinement to enhance accuracy, particularly in recognizing nuanced speech patterns mainly in intonation and providing clearer feedback. Much like the findings of Sukma and Ardini (2023), who emphasized the link between AI tools and learner autonomy, this study found that participants were empowered to self-direct their pronunciation practice by prompting ChatGPT for targeted feedback on specific words, phrases, and stress patterns. This autonomous, self-paced learning aligns with the flexibility and motivation benefits highlighted in studies by Fatima (2024) and Gutiérrez (2023).

However, according to earlier research, AI tools—especially ChatGPT—show limitations in addressing suprasegmental features such as intonation and stress [7, 8]. The participants could improve their pronunciation of English segments as aligned with Sharma (2021); however, certain features like the correct articulation of articles (“a,” “an,” “the”) and intonation patterns remained problematic. This is mainly due to the challenges noted by Aryanti and Santosa (2024) and Gutiérrez (2023), who similarly reported that AI tools struggle to capture the full range of natural speech variation. Emphasizing the teaching of both segmental and suprasegmental features can better develop the English pronunciation in EFL classrooms as mentioned by Fang (2012). Moreover, participants' difficulties in interpreting AI feedback parallel the issues highlighted by Mohammadkarimi (2024) and Fatima (2024), who stressed that learners sometimes find AI-generated suggestions lacking in detailed phonetic explanation or context, which can slow progress. The study also echoed Noviyanti's (2020) and Lazzat's (2024) observations that while AI enhances pronunciation learning, it cannot fully replace the human element in language instruction, especially for mastering nuances and real-world speech contexts. Furthermore,



the ethical and technical challenges raised by the participants—such as the need for precise prompting and concerns over AI’s limited contextual understanding—align with earlier findings [6]. These researchers emphasized the importance of addressing data privacy, speech bias, and accessibility to ensure equitable AI integration in language education. Finally, the participants’ recommendation to blend AI-based practice with traditional methods and real-life exposure reflects the broader consensus in the literature that combining AI tools with human instruction can ensure a more balanced and comprehensive learning experience [13–20]. Moreover, integrating ChatGPT and SpeechAnalyzer in pronunciation practice as part of the speaking project in the “Teaching Speaking and Writing in Primary Classes” course highlighted the importance of using PBL to improve EFL speaking and strengthen communication skills, as mentioned by John and Levshits (2024) and Dijo et al. (2024).

## Conclusion

This study highlights the potential of ChatGPT as a valuable AI-based tool for pronunciation practice, offering personalized, flexible, and engaging learning experiences. The effectiveness of ChatGPT in improving pronunciation was influenced by the participants’ ability to formulate specific and detailed prompts. When used in conjunction with SpeechAnalyzer, both AI tools significantly enhanced pronunciation skills, as evidenced by the increase in the pretest mean score from  $M = 5$  to  $M = 7.67$  in the delayed posttest. While these 2 AI-tools significantly enhance pronunciation accuracy and learner engagement, challenges remain in areas such as suprasegmental training, feedback interpretation, and speech recognition accuracy [21]. In brief, ChatGPT and SpeechAnalyzer have served as effective supplements for pronunciation training, offering learner autonomy, real-time feedback, and engaging practice. However, their optimal use requires human guidance, diversified resources, and continued development to address nuanced pronunciation features and ethical concerns. Governments and educational institutions need to integrate AI-based pronunciation tools in EFL curricula to improve learners’ communication skills. Moreover, they should make AI-tools accessible to learners in various socio-economic classes and protect the privacy of users’ speech data used by speech recognition apps.

## Recommendations for Future Research

Future AI developments should focus on enhancing contextual understanding, improving AI-human integration, and addressing ethical and cost-related issues to make AI-based pronunciation training more effective and inclusive. Additionally, personalizing feedback based on individual proficiency levels could provide a more tailored learning experience. Future research should also explore hybrid approaches that integrate AI-driven pronunciation training with human instruction to optimize learning outcomes [22,23]. Therefore, both educators and learners need to develop their digital literacy skills to maximize the benefits of AI tools while mitigating potential challenges. The study results cannot be generalized seen the small number of participants. Hence, more studies are needed to fill the gap related to pronunciation teaching in K-6 EFL classes seen that pronunciation instruction is generally overlooked and mainly left to the concerns of individual teachers.

## References

1. Ali, A. Z. M., & Segaran, K. (2013). 3D talking-head mobile app: A conceptual framework for English pronunciation learning among non-native speakers. *English Language Teaching*, 6(8), 66–76. <https://doi.org/10.5539/elt.v6n8p66>
2. Galaczi, E., Post, B., Li, A., & Graham, C. (2011). Measuring L2 English phonological proficiency: Implications for language assessment. In *Proceedings of the 44th Annual Meeting of the British Association for Applied Linguistics: The Impact of Applied Linguistics* (pp. 67–72).
3. Eckstein, G. T. (2007). A correlation of pronunciation learning strategies with spontaneous English pronunciation of adult ESL learners (Unpublished master’s thesis). Brigham Young University, Rexburg, ID.
4. Fang, T., & Lin, C. C. (2012). Taiwan EFL learners’ pronunciation strategies in two learning contexts. *Journal of Language Teaching and Research*, 3(5), 888–897. <https://doi.org/10.4304/jltr.3.5.888-897>
5. Gilakjani, A. P. (2012). A study of factors affecting EFL learners’ English pronunciation learning and the strategy for instruction. *International Journal of Humanities and Social Science*, 2(3), 119–128.
6. Gilakjani, A. P., & Sabouri, N. B. (2014). Role of Iranian EFL teachers about using pronunciation software in the instruction of English pronunciation. *English Language Teaching*, 7(1), 139–149. <https://eric.ed.gov/?id=EJ1075692>
7. Sukma, S., & Ardini, S. N. (2023). The use of artificial intelligence to promote autonomous pronunciation learning: Segmental and suprasegmental features perspective. *IJEL-TAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 8(2), 133–148.
8. Humardhiana, A. (2022). AI-powered apps to enhance novice news readers’ English pronunciation. *ELT Echo: The Journal of English Language Teaching in Foreign Language Context*, 7(2), 172–184. <https://doi.org/10.24235/eltecho.v7i2.12503>
9. Sotomayor Cantos, K. F., Varas Giler, R. C., & Castro Magayanes, I. E. (2023). Artificial intelligence in language teaching and learning. *Ciencia Latina Revista Científica Multidisciplinar*, 7(4), 5629–5638.
10. Gutiérrez, L. (2023). Artificial intelligence in language education: Navigating the potential and challenges of chatbots and NLP. *Research Studies in English Language Teaching and Learning*, 1(3), 180–191. <https://doi.org/10.62583/rseltl.v1i3.44>
11. Dávila Macías, A. M., Armijos Solano, D. O., Palma Perero, L. M., Roca Panimboza, J. A., & Lucas Soledispa, C. J. (2024). The potential of artificial intelligence to improve speaking skills in a second language (English) fluently. *Ciencia Latina Revista Científica Multidisciplinar*, 8(3), 3826–3836. <https://doi.org/10.37811/cl-rcm.v8i3>
12. Abimanto, D., & Sumarsono, W. (2024). Improving English pronunciation with AI speech-recognition technology. *Acitya: Journal of Teaching and Education*, 6(1), 146–156. <https://doi.org/10.30650/ajte.v6i1.3810>
13. Aryanti, R. D., & Santosa, M. H. (2024). A systematic review on artificial intelligence applications for enhancing EFL students’ pronunciation skill. *The Art of Teaching English as a Foreign Language*, 5(1), 102–113. <https://doi.org/10.36663/tatefl.v5i1.718>
14. Vančová, H. (2023). AI and AI-powered tools for pronunci-

- ation training. *Journal of Language and Cultural Education*, 11(3), 12–24. <https://doi.org/10.2478/jolace-2023-002>
15. Noviyanti, S. D. (2020). Artificial intelligence (AI)-based pronunciation checker: An alternative for independent learning in pandemic situation. *ELT Echo: The Journal of English Language Teaching in Foreign Language Context*, 5(2), 162–171.
  16. Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skill. *LLT Journal: A Journal on Language and Language Teaching*, 27(1), 98–115. <https://doi.org/10.24071/lt.v27i1.8151>
  17. Lazzat, K. (2024). The evolution of language learning: Exploring AI's impact on teaching English as a second language. *Eurasian Science Review*, 2(4). <https://doi.org/10.63034/esr-42>
  18. 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>
  19. Mirza, H. S. (2015). ESL and EFL university students improve differently in pronunciation: The case of Lebanon. *Procedia - Social and Behavioral Sciences*, 199, 486–495. <https://doi.org/10.1016/j.sbspro.2015.07.533>
  20. Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, 38(5), 17–36. <https://doi.org/10.14742/ajet.7492>
  21. Creswell, J. W. (2015). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (5th ed.). Pearson Education.
  22. Little, D. (2006). *The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact*. *Language Teaching*, 39(3), 167–190. <https://doi.org/10.1017/S0261444806003557>
  23. Sharma, L. R. (2021). Significance of teaching the pronunciation of segmental and suprasegmental features of English. *Interdisciplinary Research in Education*, 6(2), 63–76. <https://doi.org/10.3126/ire.v6i2.47512>

## Appendix A: AI Generated Stories 1 and 3

### Story 1: Playground

Molly and her friend Bella were playing at the playground. They were supposed to be taking turns on the swing. Molly counted to 100 while Bella was on the swing. Bella didn't get off the swing. "Bella! I counted to 100, it's my turn!" said Molly.

"You counted too fast, it's still my turn," Bella said.

Molly was upset. She wanted a turn on the swing, too! What could she do? Molly could tell the teacher, but her teacher always wanted them to try to solve the problem first. What could she try?

She decided to try an "I message". That means to tell the other person how you're feeling.

### Story 3: Sandy's Big Adventure

Sandy was a brave and curious rabbit. One morning, she hopped out of her burrow to explore the forest. She found an old map near a tree and thought, "This is the treasure map from Grandpa's stories!"

She followed the map through a thick forest and over a rocky hill. On the way, Sandy saw an owl sitting on a branch. "Hello!" said the owl. "Are you looking for something?" Sandy replied, "Yes, I'm on an adventure!"

Finally, Sandy reached a hidden cave. Inside, she found a big box full of colorful marbles and a note that read, "The greatest treasure is bravery." Sandy smiled and carried the marbles home.