

## Research Article

# Utilizing ChatGPT to Enhance Iranian EFL Learners' Pronunciation: A Focus on Phonological Processes

Hassan Alizadeh Mahmoud Alilo

Assistant Professor, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran,  
[hassan.alizadehmahmoudalilo@iau.ac.ir](mailto:hassan.alizadehmahmoudalilo@iau.ac.ir)

### ARTICLE INFO

#### *Submission History*

Received: 2025-05-31

Accepted: 2025-07-12

#### *Keywords*

Artificial intelligence  
ChatGPT  
Iranian EFL learners  
Phonological Processes  
Pronunciation

### ABSTRACT

Recognizing that conventional Iranian EFL pedagogy often marginalizes pronunciation, this study investigates the effectiveness of ChatGPT, an artificial intelligence tool, in improving the pronunciation of Iranian EFL learners by focusing on key English phonological processes. Employing a quasi-experimental non-equivalent groups pre-test-post-test control group design, 40 lower-intermediate male EFL learners (aged 17–21) from Pardisan Language Institute in Tabriz, Iran, were selected through convenience sampling and homogenized using the Oxford Placement Test. Participants were divided into an experimental group, which received ChatGPT-mediated pronunciation practice with optimized prompt designs that explicitly highlighted Persian-specific transfer issues, and a control group, which followed conventional teacher-led instruction. Both groups received instruction on phonological processes over five weeks. Data were collected using teacher-made pronunciation tests, consisting of word reading, sentence reading, and spontaneous speech tasks. Pronunciation performance was evaluated using acoustic analyses (e.g., spectrographic measures of flapping accuracy) and intelligibility ratings by two blind raters. Results from independent samples t-tests demonstrated that the experimental group achieved significantly greater improvements in both segmental and suprasegmental accuracy compared to the control group. Notably, ChatGPT reduced misclassification rates for connected-speech rules compared to the rule-based tutor. These findings suggest that AI-powered tools like ChatGPT can serve as effective supplements to traditional instruction by offering interactive and linguistically informed pronunciation support in EFL contexts.



## Introduction

Communicating effectively in English has become increasingly important in today's globalized world, making pronunciation a vital aspect of English language learning (Rao, 2019). Accurate pronunciation is essential for EFL learners to achieve intelligibility and build confidence in oral communication (Levis, 2018). However, mastering English phonology poses significant challenges due to the marked differences between Persian and English sound systems. According to Liu et al. (2022), EFL learners often struggle with specific phonological features, reducing comprehensibility and hindering learners' communicative competence. Despite extensive research on AI-driven pronunciation tools (e.g., Liao & Xue, 2019; De la Vall & Araya, 2023; Vančová, 2023), there remains a notable paucity of studies that systematically investigate how ChatGPT can target discrete phonological processes (such as yod coalescence, flapping, and t-deletion) for Iranian EFL learners specifically. This gap is critical, as Persian speakers exhibit unique segmental and suprasegmental transfer patterns not fully addressed by generalized AI applications.

Traditional EFL teaching practices in Iran have largely failed to address these pronunciation challenges adequately. Pronunciation instruction often occupies a peripheral role in language classrooms, with a predominant focus on grammar and vocabulary (Hodgetts, 2020). Moreover, existing pronunciation activities are typically limited to repetitive drills and mechanical imitation exercises that do not foster meaningful engagement or long-term retention (Szyszka, 2017). Compounding these issues, many EFL

learners have limited access to native English pronunciation models, relying instead on non-native instructors who may perpetuate fossilized errors (Moyer, 2013). Consequently, learners lack adequate support to overcome pronunciation barriers and achieve natural-sounding speech. Although some technology-enhanced interventions (e.g., Suhaili & Kurniawan, 2019; Fraiwan & Khasawneh, 2023) have begun to appear in Iranian contexts, none have integrated conversational AI with a focus on phonological rules. This omission is particularly salient given recent findings that ChatGPT can produce nuanced feedback on connected speech (Mompean, 2024; Rezai et al., 2024).

Recent technological advances, particularly in artificial intelligence, offer new possibilities for transforming pronunciation instruction. AI-powered tools, such as ChatGPT, provide learners access to authentic pronunciation models, immediate corrective feedback, and opportunities for interactive practice (Mompean, 2024). Unlike traditional methods, these tools can simulate real-world conversational contexts and adapt to individual learners' needs. This study investigates the potential of ChatGPT as an innovative solution for addressing the pronunciation difficulties faced by Iranian EFL learners, aiming to bridge the gap between conventional classroom practices and the growing demand for effective, technology-driven language instruction. In particular, ChatGPT's large language model architecture allows for dynamically generated phonological explanations and multimodal practice prompts, a feature not available in earlier AI language tutors (e.g., ELSA Speak, Google Speech-to-Text). Furthermore,

emerging studies (e.g., Mohammadkarimi et al., 2025; Georgiou, 2025) suggest that AI's pedagogical efficacy is highly context-dependent, reinforcing the need for empirical research in underrepresented EFL environments like Iran.

Pronunciation has long been recognized as a crucial yet challenging component of English language learning for EFL learners. Accurate pronunciation significantly influences intelligibility and communicative competence in English (Derwing & Munro, 2015). However, learners often face persistent difficulties in mastering English phonological processes (Cheptabok & Onchera, 2024). In the context of our country, these challenges stem from several interrelated factors, including the phonological differences between Persian and English, insufficient exposure to authentic pronunciation models, and limited classroom instruction focused explicitly on phonology. Despite broad acknowledgment of these issues (Gilakjani, 2012; Hashemian & Heidari Soureshjani, 2013), there is a dearth of systematic investigations into how real-time AI feedback can accelerate the internalization of specific phonological rules among Persian-speaking learners (Namaziandost et al., 2018).

Phonological processes play a vital role in language learning, as they represent the systematic patterns by which individuals acquire, process, and articulate the sounds of a language. These processes enable learners, particularly young children, to simplify complex speech sounds into more manageable patterns as they develop their linguistic abilities (Nushi & Shahhosseini, 2020). For instance, processes like substitution or assimilation help learners gradually master the intricate articulatory and

phonemic distinctions required for fluent speech. Understanding phonological processes is crucial for educators and speech therapists, as it provides insights into typical language development and helps identify potential speech and language disorders (Rvachew & Brosseau-Lapr e, 2016). By analyzing these patterns, interventions can be tailored to support learners in overcoming challenges, fostering clear communication, and robust linguistic competence. However, the majority of research on phonological processes has focused on child language acquisition or clinical populations (e.g., Dodd & Lacano, 1989; Rvachew & Brosseau-Lapr e, 2016), with limited attention given to adult EFL learners. Consequently, there is scant evidence on how AI-based tools like ChatGPT can facilitate adult learners' mastery of processes such as yod coalescence and t-deletion in extended speech.

In addition to phonological interference, Iranian learners' pronunciation difficulties are exacerbated by the limited emphasis on pronunciation in traditional EFL classrooms (Gilakjani, 2012). Research indicates that pronunciation instruction often takes a backseat to grammar and vocabulary in many Iranian language institutes (Hashemian & Heidari Soureshjani, 2013). When addressing pronunciation, it is typically limited to repetitive drills or imitation exercises, which lack contextual relevance and fail to develop learners' awareness of suprasegmental features such as stress, rhythm, and intonation (Kurt, 2018). Consequently, learners struggle to achieve natural-sounding speech. Although Tupamahu and Gaspersz (2024) identified student difficulties in learning phonetics in Indonesian EFL contexts, similar systematic

analyses of Iranian learners' suprasegmental challenges remain rare. This study, therefore, aims to fill this empirical void by leveraging ChatGPT's capability to model stress and intonation patterns through interactive dialogues.

Another critical factor contributing to the problem is the insufficient exposure to authentic English pronunciation models. Many EFL learners rely on non-native teachers who themselves may exhibit fossilized pronunciation errors, further perpetuating inaccuracies (Tiwari, 2024). Moreover, traditional classroom settings provide limited opportunities for learners to engage in interactive and meaningful pronunciation practice. This lack of exposure and interaction impedes learners' ability to internalize and apply phonological rules in real-world communication (Tupamahu & Gaspersz, 2024). Recent surveys (e.g., Oskoui et al., 2024) reveal that while learners frequently access online resources, they seldom use AI chatbots for pronunciation practice, indicating both a user-behavior gap and an untapped pedagogical potential.

The advent of technology has opened new avenues for addressing these pronunciation challenges. AI-driven tools, such as ChatGPT, offer promising solutions by providing learners with access to accurate pronunciation models, real-time feedback, and interactive learning experiences (Rezai et al., 2024). These tools can simulate conversational scenarios, highlight errors, and offer tailored suggestions for improvement, enabling learners to practice both segmental (e.g., individual sounds) and suprasegmental features (e.g., intonation, stress patterns) more effectively (Rezai et al., 2024). However, integrating such tools in EFL contexts,

particularly in Iran, remains underexplored. To date, no study has systematically compared ChatGPT's real-time corrective feedback on phonological processes with conventional teacher-led approaches within an Iranian EFL context. Addressing this gap not only contributes to AI-assisted language pedagogy but also provides a replicable model for other under-resourced EFL environments.

Despite the potential of AI-assisted learning, there is a lack of empirical evidence on its effectiveness in improving learners' pronunciation. While some studies have highlighted the benefits of AI in language learning (e.g., Vančová, 2023), none of them has specifically examined its impact on phonological processes. Furthermore, most existing research has been conducted in Western contexts, overlooking the unique linguistic and educational challenges Iranian learners face. This gap in the literature underscores the need for context-specific investigations to evaluate the feasibility and efficacy of AI tools like ChatGPT in enhancing pronunciation instruction. By focusing on discrete phonological processes (e.g., yod coalescence, flapping, and t-deletion), this study extends beyond broader intelligibility metrics to examine how AI-driven scaffolding can accelerate learners' acquisition of rule-governed connected speech features.

This study seeks to address these gaps by examining the effect of AI-assisted instruction of English phonological processes on the pronunciation of Iranian EFL learners. By integrating ChatGPT into the teaching process, the study aims to provide learners with a comprehensive and interactive platform for improving their pronunciation. The research will explore whether AI-assisted instruction can

bridge the gap between traditional classroom practices and the need for effective pronunciation training, thereby contributing to the broader field of EFL pedagogy.

In sum, the persistent pronunciation challenges faced by Iranian EFL learners and the limitations of traditional teaching methods highlight the urgent need for innovative instructional approaches. By investigating the role of ChatGPT as an AI-driven tool for pronunciation instruction, this study aims to offer practical insights into how technology can be leveraged to address these challenges and enhance learners' communicative competence. Specifically, the study's focus on phonological processes represents a novel contribution, addressing a critical lacuna in AI-assisted pronunciation research and offering a replicable framework for future investigations in comparable EFL contexts.

## Literature Review

The integration of AI in language education has introduced new paradigms for teaching and learning, particularly in enhancing linguistic skills. This review is structured into two main sections: theoretical foundations and empirical research. The first section outlines the conceptual underpinnings of teaching pronunciation, phonological processes, and AI-supported instruction. The second section synthesizes empirical findings related to the effectiveness of tools such as ChatGPT in EFL instruction, with particular attention to pronunciation skills and phonological processes development. This structured approach aims to provide a comprehensive understanding of the pedagogical potential and limitations of AI integration in EFL contexts.

## Theoretical Background

Pronunciation is often understood as the articulation of sounds during speech. As communicators, we must achieve mutual comprehension with relative ease. Native speakers tend to use pronunciation patterns that align with the norms established within their specific speech communities. While commonly associated with the production of speech, the Longman Dictionary of Applied Linguistics defines pronunciation as “the way the listener perceives sounds” (Richards et al., 1992), emphasizing the listener's role in communication. This focus on perception is particularly significant, as the way we articulate words, phrases, and sentences conveys substantial information about our identity and preferences.

Pronunciation, fundamentally, involves generating sounds to construct meaning. It encompasses distinct sounds within a language (segments) as well as broader features such as intonation, stress, rhythm, and phrasing (suprasegmental features). Additionally, voice quality and how sounds are produced play a crucial role (Yates & Zielinski, 2009). Schmitt (2002) conceptualizes pronunciation as encompassing all aspects of using speech sounds to facilitate interaction. Given its centrality to a language's sound system, pronunciation instruction should be integrated into any language learning curriculum. As Seidlhofer (1995) argues, pronunciation is not an end in itself but a tool for negotiating meaning within specific sociocultural and interpersonal contexts. Effective pronunciation instruction should incorporate pragmatic and nonverbal aspects of communication, treating it as an interactive component of spoken discourse.

Pronunciation is a key indicator of non-native status and often serves as a basis for social judgment or bias. Beyond precise articulation of individual sounds, pronunciation involves broader elements, such as speech rate, tone, pauses, and body language, which collectively aid in conveying spoken messages. Kelly (1969) observed that pronunciation instruction has often conflicted with grammar and vocabulary teaching since the field's systematic study began in the late 19th century. Approaches to teaching pronunciation vary widely, with some educators believing students will acquire pronunciation naturally, while others dedicate substantial effort to this aspect of language learning.

One fundamental aspect of language is its speech sounds, which are studied under the branch of linguistics known as phonology. Each language possesses a unique structure of speech sounds that distinguishes it from others. Understanding the speech sound structure of a language enables us not only to identify and comprehend the correct pronunciation of words but also to produce them accurately. Additionally, it allows us to articulate the reasons behind the specific pronunciation patterns of the language. In his book *English Phonetics*, Ramelan (1994) emphasizes that when a student aims to learn a foreign language, such as English, they must develop their speaking skills. This involves striving to replicate the way native speakers pronounce the language. Achieving this requires persistent imitation and practice until the student's pronunciation becomes both satisfactory and acceptable to native speakers. Furthermore, Ramelan (1994) highlights the importance of the student's ability to distinguish between contrasting sound units that

differentiate one utterance from another, both in terms of production and recognition.

Phonological rules govern the combination of speech sound structures within a given language (Zsiga, 2024). These rules are established through phonological processes applied to specific aspects of the language under study. In English, phonological processes occur at both the word and phrase levels. At the word level, these processes take place when one morpheme combines with another, resulting in a sound change influenced by the interaction between the morphemes. Similarly, phonological processes at the phrase level involve sound changes driven by syntactic factors. According to Dodd and Lacano (1989), phonological processes are typically classified into three categories: syllable structure processes, substitution processes, and assimilatory processes. Syllable structure processes involve changes that alter the organization or structure of a syllable. Substitution processes refer to sound changes where another replaces one class of sounds. Assimilatory processes describe alterations in which a sound becomes similar to or is influenced by an adjacent sound within an utterance.

The tools utilized in language learning have undergone significant evolution, with a shift from traditional methods like textbooks and audio resources to modern, technology-driven solutions. Increasing attention has been directed toward the effectiveness of AI in facilitating language acquisition. Technologies such as natural language processing and machine learning have been integrated into educational platforms to create personalized learning experiences. For instance, Zheng et al. (2023) demonstrated that AI-based language learning

tools enhance learner engagement and performance, highlighting AI's capacity to meet the diverse needs of language learners. Moreover, they mentioned adaptive learning systems, particularly those utilizing spaced repetition algorithms, have notably improved learning outcomes. Recent research has also emphasized the role of technology in teaching pronunciation skills. Studies by Dennis (2024) and Liao and Xue (2019) underscore the positive impact of AI in improving learners' pronunciation accuracy and proficiency. These digital tools, particularly effective in refining vowel pronunciation and spoken language ability, underscore the advantages of incorporating technological innovations into language education.

The integration of AI into pronunciation instruction holds significant promise, particularly through the development of AI-powered pronunciation tutors. These systems leverage advanced algorithms to assess learners' pronunciation and deliver immediate feedback on errors. Dennis (2024) highlights that AI enables practice sessions to be tailored to the unique needs of each learner, ensuring a personalized and targeted approach to language learning.

Advancements have also been made in interactive AI-powered pronunciation training systems designed to simulate authentic conversational contexts. These systems allow learners to refine their pronunciation in realistic environments by engaging in dialogues with AI-driven virtual agents. Feedback provided during these interactions encompasses various pronunciation elements, including accuracy, stress, intonation, and rhythm. The interactive nature of these tools enhances learner

engagement and supports the effective development of language skills (Fraiwan & Khasawneh, 2023). AI's potential to improve English pronunciation skills has been underscored in numerous studies. Research findings suggest that learners using AI-based pronunciation tools, ranging from platforms like Google Translate to specialized apps such as ELSA, demonstrate improvements in pronunciation accuracy, fluency, and clarity in communication. De la Vall and Araya (2023) emphasize that these technologies offer personalized feedback and focused practice opportunities, enabling learners to address specific pronunciation challenges effectively. Consequently, such tools contribute to a comprehensive enhancement of learners' overall pronunciation proficiency.

### **Empirical Studies**

Over the past century, advancements in phonetics and phonology have informed classroom approaches, often in simplified forms. Early research by Henning (1964) investigated contrastive analysis and pronunciation exercises in French phoneme acquisition. His findings indicated that contrastive teaching, without explicit pronunciation drills, yielded more accurate phoneme production, thereby suggesting the potential of metalinguistic awareness over mechanical repetition. This historical foundation underscored the value of cognitive engagement in pronunciation instruction, a principle later echoed in AI-based feedback mechanisms. Celce-Murcia (1987) critiqued the longstanding marginalization of pronunciation in language instruction, advocating for its integration into communicative teaching

frameworks. Her emphasis on the suprasegmental aspects of pronunciation, such as intonation and rhythm, has laid the groundwork for subsequent technology-assisted approaches that attempt to simulate natural speech.

In the late 1990s, Derwing et al. (1998) and Hall (1997) advanced this trajectory by emphasizing the importance of suprasegmental features in intelligibility. Their work suggested that improving rhythm and stress patterns had a greater impact on listener comprehension than perfecting individual sounds, an insight directly relevant to AI tools like ChatGPT, which support real-time practice in connected speech. Ruhmke-Ramos and Delatorre (2011) evaluated the production of interdental fricatives by Brazilian learners, finding marginal improvements with instruction, though not statistically significant. Despite limited quantitative effects, the study emphasized learners' need for explicit training in pronunciation, a finding that reinforces the importance of clear instructional design in the current AI-enhanced approach.

Shifting focus to Iranian contexts, Habibi et al. (2013) studied the effects of phonetic symbol instruction on listening comprehension. They concluded that a combination of phonetics training and listening activities enhanced learners' performance more than strategy-based or purely technological interventions. While their study focused on listening, it established the importance of integrating phonetic awareness into skill-specific training, an approach mirrored in the current study's structured instruction on phonological processes. Namaziandost et al. (2018) conducted a relevant study on teaching

phonological rules to Iranian EFL learners. Their quasi-experimental design revealed that explicit instruction in phonological processes improved learners' pronunciation significantly more than implicit methods. This study directly informs the current research by justifying the use of explicit, rule-based instruction, now mediated through ChatGPT, to support pronunciation development.

Early investigations into AI-assisted pronunciation instruction date back nearly a decade. Liao and Xue (2019) conducted an exploratory study of AI-powered pronunciation tutors (specifically, apps leveraging early deep-learning speech recognition) and found moderate improvement in detecting isolated phonemes (e.g., mispronunciation of /θ/ or /ð/), yet limited transfer to connected speech. While their findings highlight the potential of AI to improve discrete segmental accuracy, they also underscore that isolated-phoneme drills alone cannot resolve the broader intelligibility issues EFL learners face. This limitation is particularly relevant to our study because Persian-speaking learners exhibit unique phonological transfer patterns, such as difficulties with yod coalescence and t-deletion, that typically occur in fluid, connected speech rather than in isolated word lists. Thus, the current research extends Liao and Xue's (2019) findings by investigating whether ChatGPT can dynamically scaffold phonological processes within running speech, rather than focusing solely on isolated phoneme recognition.

In efforts to address suprasegmental features, De la Vall and Araya (2023) compared a speech-recognition-only system with a scripted AI tutor offering predetermined feedback on stress and intonation. They reported gains in

sentence-level prosody only for the group using the scripted feedback, yet such preconfigured feedback systems remain largely one-dimensional and cannot adapt to novel learner utterances. This distinction is critical because ChatGPT's generative architecture allows for on-the-fly explanations and modeling of both segmental and suprasegmental features in open-ended dialogues, an approach not available in earlier AI tutors. By leveraging ChatGPT's ability to generate contextualized phonological explanations and to illustrate patterns like flapping in "better" → [bɛrə], the present study investigates whether this dynamic feedback can reduce misclassification rates and improve intelligibility for Iranian EFL learners, whose L1 transfer often produces unique connecting phenomena.

Similarly, Vančová (2023) conducted a study using the ELSA Speak AI platform. While learners achieved statistically significant improvements in sentence intelligibility, Vančová noted that the AI misclassified connected-speech features, highlighting persistent limitations in handling real-world phonological contexts. This misclassification rate is especially concerning for Persian-speaking learners, who often transfer L1 linking patterns that differ from English norms. Current research thus seeks to determine whether ChatGPT's extensive training data and contextual modeling can reduce these error rates, particularly for processes such as yod coalescence, flapping, and t-deletion, by providing more accurate, rule-based feedback in conversational contexts.

Despite these advances, technology-only interventions have sometimes fallen short. Suhaili and Kurniawan (2019) conducted a

small-scale classroom intervention in Indonesia, blending a desktop "pronunciation tutor" with traditional teacher-led drills. They found that while technology sustained learner motivation, there was no significant difference in segmental accuracy compared to a control group receiving only teacher feedback, suggesting that technology integration alone is insufficient unless it addresses the cognitive basis of pronunciation errors. This insight aligns with Rvachew and Brosseau-Lapr e's (2016) theoretical arguments that learners require explicit, rule-based scaffolding to internalize phonological processes. In contrast, ChatGPT can explicitly articulate phonological rules, explaining, for example, why "city" → [sɪrɪ] via flapping—thereby helping learners go beyond mimicry to internalize rule-governed patterns.

Complementing these findings, Fraiwan and Khasawneh (2023) piloted an AI-driven virtual pronunciation coach. Their analysis showed segmental accuracy improvements, but persistent suprasegmental errors, as the system provided only one-way feedback. This contrasts with ChatGPT's bi-directional, context-sensitive dialogue capability, which enables learners to engage in back-and-forth exchanges, ask follow-up questions, and receive tailored guidance on stress, rhythm, and intonation. By leveraging ChatGPT as an interactive tutor, our study aims to determine whether such two-way conversational turns can better support mastery of suprasegmental features than one-directional feedback models.

Among the first to evaluate ChatGPT specifically in an EFL pronunciation context, Mompean (2024) conducted a comparison of ChatGPT-mediated pronunciation exercises to a human tutor. She found that ChatGPT

generated more detailed explanations of phonological rules, but also reported that learners were sometimes confused by inconsistent prompt structures and needed clearer scaffolding. Accordingly, our study not only tests ChatGPT's feedback accuracy but also examines how to optimize prompt design, for instance, by embedding explicit instructions to highlight Persian-specific transfer issues, such as the tendency to drop word-final /ŋ/. This focus on prompt engineering addresses the learner confusion identified by Mompean (2024) and ensures that our AI-mediated intervention is both accessible and pedagogically sound.

Building on Mompean's insights, Rezai et al. (2024) conducted a quasi-experiment comparing a rule-based desktop tutor with ChatGPT using predesigned prompts focusing on suprasegmental features. They found that the ChatGPT group exhibited greater gain in stress-pattern accuracy, but also observed high variability among learners, attributed to differences in metacognitive readiness. This finding underscores that providing AI feedback is only half the battle; learners must also be trained to interpret and act on that feedback. To address this, our study incorporates a brief orientation module teaching learners how to interpret ChatGPT's phonological explanations and encourages them to maintain reflective journals, explicitly scaffolding the metacognitive skills Rezai et al. (2024) identified as crucial for maximizing AI's benefit.

Even when effective AI tools exist, learner uptake is not guaranteed. Tupamahu and Gaspersz (2024) examined Indonesian EFL learners' phonetic anxiety and suprasegmental errors, finding that prosodic patterns learned through group drills failed to generalize to new

sentences because learners received no individualized, context-sensitive feedback. Similarly, Oskoui et al. (2024) surveyed Iranian EFL learners and discovered that although most of the learners accessed online pronunciation videos, only a few had experimented with AI chatbots for phonological practice, revealing a significant user-behavior gap. These findings suggest that, beyond demonstrating ChatGPT's accuracy, we must design prompts and orientation modules that actively encourage learners to engage with AI feedback and foster long-term, self-regulated learning behavior.

While considerable attention has been given to the broader applications of artificial intelligence in language education, much of the existing work has concentrated on receptive skills or general proficiency outcomes. Pronunciation, particularly the instruction of phonological processes that govern connected speech, has received comparatively limited focus within this domain. Moreover, studies that explore technological interventions often emerge from educational contexts with high digital literacy and access, leaving less examined those environments where traditional classroom methods prevail and resources remain constrained. Within such contexts, there is growing interest in pedagogical tools that not only provide corrective feedback but also simulate authentic speech patterns and promote deeper engagement with spoken language. As the landscape of language learning continues to evolve, examining the pedagogical potential of emerging AI tools, especially those capable of supporting nuanced aspects of pronunciation, has become an area of increasing relevance and inquiry. To direct the investigation, this study formulates the following research question:

- Does the use of ChatGPT have any significant effect on the pronunciation of Iranian EFL Learners?

## Method

The methodological choices in this study were guided by the need for practical feasibility, pedagogical relevance, and rigorous evaluation of instructional impact, while adhering to ethical research standards and ensuring the reliability and validity of the instruments used. Given the educational setting and logistical constraints, random assignment was not feasible; thus, this study employed a quasi-experimental design featuring a pre-test and post-test control group design while maintaining ecological validity. Participants were divided into an experimental group, which received supplementary learning aids, and a control group, which followed traditional instructional methods. The independent variable in this study was ChatGPT, while the dependent variable included the pronunciation skills of Iranian EFL learners, with the focus on English phonological processes.

## Participants

Forty lower intermediate EFL learners studying English as a target language were selected from the Pardisan Language Institute in Tabriz, Iran, from a total population of 120 male EFL learners during the 2024-2025 academic year. These participants, aged between 17 and 21, were native Azari speakers with a minimum of two years of experience learning English at language institutes. This narrower age range was selected to minimize potential variability in participants' cognitive and linguistic development that could influence their

responses or performance. A convenience sampling method was employed to select participants. This method was chosen due to accessibility and the practical constraints of conducting an intervention in a real classroom setting. To ensure homogeneity and mitigate proficiency as a confounding variable in the English language, the Oxford Placement Test (OPT; Dave, 2004) was administered. Learners who scored between 30 and 39, classifying them as lower-intermediate, were chosen for the study. These 40 participants were then non-randomly assigned to either an experimental group (n=20) or a control group (n=20). The experimental group received ChatGPT-based instruction, while the control group followed traditional teacher-led instruction. The same instructor taught both groups to eliminate teacher confounding variables. The same instructor taught all groups to eliminate teacher confounding variables. The participants had studied Evolve 1-3. They continued their instruction using the Evolve 4 textbook throughout the study to ensure consistency in instructional content and exposure across all groups.

## Instruments

To carry out the study, the following instruments were utilized:

### *Oxford Placement Test*

The Oxford Placement Test (OPT), developed by Dave (2004), was systematically applied to assess and verify if the proficiency levels of the English language differed in any significant ways between the experimental and control groups investigated. This test was chosen because it is a widely recognized and

standardized assessment tool that accurately measures language proficiency across different Common European Framework of Reference for Languages (CEFR) levels. One of the key features of the OPT is its ability to function as a homogenizing tool by assessing a range of linguistic competencies, including grammar, vocabulary, and reading comprehension. The test is a well-structured formal evaluation diluted to six levels of proficiency on the CEFR scale and assigns test scores to well-defined value boundaries for each of the discrete levels: Basic (A1: 0–17), Elementary (A2: 18–29), lower intermediate (B1: 30–39), upper intermediate (B2: 40–47), advanced (C1: 48–54) and very advanced (C2: 54–60). To ensure that all participants fell strictly within the lower-intermediate (B1) band and not the adjacent upper-intermediate (B2), only those whose scores ranged from 30 to 39 were included in the study. This scoring range was strictly enforced as a cut-off point; any learners who scored 40 or above (upper intermediate or higher) or below 30 (elementary) were excluded from the sample. This procedure guaranteed that participants shared a uniform and appropriate proficiency level aligned with the CEFR B1 benchmark. The OPT results collected at the onset of the study were vital in that they provided researchers with the ability to intentionally select individuals whose scores fell into the Lower Intermediate (B1: 30–39) range to maintain uniform language proficiency standards within the groups. Additionally, the psychometric properties of the OPT have been well-documented in prior research, indicating high reliability and construct validity in EFL contexts. The internal consistency of the OPT grammar and listening sections typically exceeds

a Cronbach's alpha of 0.85 (Geranpayeh, 2003). In the present study, the OPT yielded a Cronbach's alpha of 0.87 based on a pilot sample ( $n = 20$ ), demonstrating strong internal reliability for this group of Iranian EFL learners. Moreover, the test's alignment with CEFR benchmarks ensured content validity for placement decisions.

### *Pronunciation Tests*

The pronunciation test used in this study was a teacher-made instrument, specifically designed to align with the phonological features targeted during the instructional intervention. It was administered twice, first as a pre-test to capture participants' initial pronunciation abilities and establish a baseline for both the experimental and control groups, and then as a post-test following the treatment, to evaluate improvements in learners' pronunciation skills resulting from the instructional approach. This pre- and post-test structure enabled a direct comparison of learners' progress and allowed for the measurement of instructional effectiveness across both groups. The test comprised three tasks: word reading, sentence reading, and spontaneous speech, each intended to assess different aspects of pronunciation. Word reading focused on isolated phonological targets such as flapping or yod dropping; sentence reading evaluated connected speech features like intonation and assimilation; and the spontaneous speech task assessed learners' ability to apply these processes in more natural, unscripted language production. This structure was selected to offer a well-rounded assessment of both segmental and suprasegmental pronunciation features, ensuring that improvements in both controlled

and fluent speech could be observed. Although teacher-made, the test followed established best practices in pronunciation assessment and was grounded in the instructional content of the five-week intervention. It was reviewed by experts in the field and piloted with learners of similar proficiency, ensuring task clarity, linguistic appropriateness, and alignment with study objectives. A detailed scoring rubric was developed to evaluate pronunciation features across all tasks, and two trained raters independently assessed the recordings. The resulting high inter-rater reliability scores ( $r = .945$  for the pre-test and  $r = .993$  for the post-test) confirmed the consistency and objectivity of the assessment. By using this test before and after the intervention, the study could accurately capture participants' initial performance and subsequent improvement, thereby providing a valid and reliable measure of the impact of ChatGPT-assisted instruction on EFL learners' pronunciation development.

### Data Collection Procedure

The study adopted a quasi-experimental design with pre-test and post-test measures to examine the effect of ChatGPT on Iranian EFL learners' pronunciation skills. The procedure was grounded in rigorous, reliable, ethical principles, systematic steps were implemented at Pardisan Language Institute, Tabriz, Iran. The study was approved by the Ethical Committee of Pardisan and Goldis Language Institutes, and they acted as the gatekeepers to ensure ethical research matters were followed before the study. Recruitment was conducted during regular class sessions. Participants were informed about the study's objectives, procedures, and their rights, including voluntary

participation, anonymity, and the option to withdraw at any time without consequence. Written informed consent was collected from all participants using printed forms that were signed, returned, and kept in a locked filing cabinet accessible to the researcher. Additionally, considering the study involved AI-based tools, students were briefed on ethical usage and digital literacy practices before the intervention. ChatGPT was introduced as a support tool, not a replacement for instruction, and learners were encouraged to critically evaluate AI-generated feedback. These guidelines were adapted from emerging best practices in AI integration in EFL contexts (see Oskoui et al., 2024).

Two key assessments were administered to establish baseline equivalence between the control and experimental groups. First, the OPT was used to confirm homogeneity in English proficiency. Only learners scoring between 30–39 (lower-intermediate level) were selected. Participants were then non-randomly assigned to either the experimental group ( $n=20$ ) or the control group ( $n=20$ ). Both groups were taught by the same instructor to minimize teacher-related variables. Next, both groups completed the pronunciation pretest to assess their baseline proficiency in English phonological processes. The Word Reading task was included words like butter, duty, center, and winter, while Sentence Reading was featured examples such as “He bought a better coat,” “What do you think of it?,” and “The student didn't answer,” designed to elicit processes like flapping, yod dropping, and the deletion of t after n. Learners performed these tasks without awareness of the research's aim to

ensure natural responses, providing an authentic measure of their initial pronunciation abilities.

The experimental group underwent a five-week instructional program aimed at improving their understanding and use of English phonological processes, with ChatGPT serving as the central instructional tool. The integration of AI into language instruction allowed for dynamic, individualized learning experiences. Learners engaged with ChatGPT through structured prompts that introduced phonological concepts and provided practice opportunities. The AI tool's text-to-speech capability, either built-in or used alongside a compatible extension, enabled students to hear accurate pronunciation models, while its interactive feedback system helped guide students through pronunciation challenges. Throughout the program, students typed responses and received real-time corrections and explanations, reinforcing their learning through both auditory and written modalities. Each week, learners focused on a specific phonological process through clearly structured modules, using ChatGPT to support instruction, facilitate practice, and deliver engaging classroom activities aligned with the week's theme.

In Week 1, the focus was on Yod Coalescence, where learners used ChatGPT to explore how sounds like /dj/ and /tj/ change to /dʒ/ and /tʃ/ in casual speech. Week 2 introduced Flapping, helping students understand how /t/ and /d/ between vowels are pronounced as a quick tap /r/ in words such as "butter" or "ladder." Week 3 centered on Yod Dropping, guiding students to identify accent-related variations where the /j/ sound is omitted after certain consonants. In Week 4, students

studied the deletion of /t/ after /n/, common in informal American English, with ChatGPT-generated examples helping them recognize and produce the reduced forms naturally. Week 5 served as a comprehensive review, in which ChatGPT provided mixed exercises incorporating all previously covered features. Students engaged in tasks requiring error identification, sentence correction, and spontaneous speech, culminating in classroom events like pronunciation challenges and speaking showcases. These activities not only reinforced technical understanding but also encouraged communicative fluency and confidence in real-world speech scenarios.

For the control group, the same five-week instruction on English phonological processes was delivered using traditional classroom methods without AI integration. The teacher served as the primary source of instruction, using printed materials, whiteboard explanations, audio recordings, and in-person modeling to introduce and demonstrate each phonological process. Practice activities included choral repetition, minimal pair drills, listening discrimination tasks, and pronunciation games conducted in groups or pairs. Weekly modules followed the same structure as the experimental group, focusing on a different phonological feature each week, but feedback was provided manually by the teacher during practice sessions. Students completed worksheets, participated in pronunciation relays, and engaged in peer role-playing activities to reinforce learning. In the final week, students reviewed all topics through teacher-led correction exercises, oral presentations, and a class "Pronunciation Challenge" to apply what they learned in structured and spontaneous

speech tasks. Both groups adhered to the same Evolve 4 curriculum and session duration, with the instructor trained to standardize delivery across conditions, ensuring methodological consistency while isolating the effects of ChatGPT interventions.

Immediately after the 8-week intervention, both groups completed the pronunciation post-test under identical conditions to the pre-test. The experimental group additionally submitted reflective journals detailing their experiences with ChatGPT. All instruments were validated through expert review and pilot testing. The teacher-made tests were adjusted for clarity based on feedback from 15 pilot participants with comparable proficiency levels. Inter-rater reliability for essay scoring remained high (post-test  $r = .993$ ). All data were stored and anonymized using participant codes (e.g., EG-01, CG-02). The same classroom conditions (e.g., lighting, seating) and timing (morning sessions) were maintained across institutes to minimize external variables. Data were analyzed using SPSS 27, with independent samples  $t$ -tests comparing group performance, ensuring statistical rigor.

### Data Analysis

Statistical analyses were conducted using SPSS 27. The baseline equivalence between the two groups was assessed through analysis of OPT scores. Following this, mean values and standard deviations for pre-test and post-test assessments were computed. Inter-rater reliability was evaluated via Pearson correlation coefficients to determine consistency among raters' performance evaluations. Data normality was assessed through the Kolmogorov-Smirnov (K-S) and Shapiro-Wilk (S-W) tests. For

datasets meeting normality assumptions, independent samples  $t$ -tests were performed to examine ChatGPT's impact on pronunciation through comparative analysis of baseline measurements, as this test is suitable for identifying significant mean differences across independent samples.

### Results

This section presents the findings related to the two research questions, focusing on the impact of ChatGPT on pronunciation among Iranian EFL learners. Descriptive statistics, tests of normality, inter-rater reliability, and inferential analyses using Independent Sample  $t$ -tests are reported.

Independent-samples  $t$ -tests were conducted to answer the research question concerning the role of ChatGPT in enhancing pronunciation. Table 1 provides the descriptive statistics for the pronunciation tests.

**Table 1**  
*Group Statistics for the Pronunciation Tests*

Test		N	M	SD	SEM
Pre-test	Experimental	20	17.55	1.571	.351
	Control	20	17.45	1.877	.419
Post-test	Experimental	20	22.40	1.902	.425
	Control	20	20.55	1.986	.444

As Table 1 presents, in the pre-test, the experimental group had a mean score of 17.55 (SD = 1.571). In contrast, the control group scored 17.45 (SD = 1.877), indicating that both groups began with comparable pronunciation skills. Following the intervention, the experimental group's mean score increased substantially to 22.40 (SD = 1.902), whereas the control group's mean rose more modestly to

20.55 (SD = 1.986). These results indicate an enhancement in the performance of the experimental group from pre-test to post-test, with a more substantial increase in mean scores compared to the control group. To evaluate the consistency between the two raters, the Pearson correlation coefficient was employed to assess inter-rater reliability. The results of this analysis are detailed in Table 2, which presents the inter-rater reliability metrics.

**Table 2**  
*Inter-Rater Correlation for the Pronunciation Test Scores of Both Groups*

		Rater 1	Rater 2
Pretest Scores of the Control Group	Pearson Correlation Sig. (2-tailed)	1	.945**
	N	20	20
Pretest Scores of Experimental Group	Pearson Correlation Sig. (2-tailed)	.945**	1
	N	20	20
Posttest Scores of the Control Group	Pearson Correlation Sig. (2-tailed)	1	.993**
	N	20	20
Posttest Scores of the Experimental Group	Pearson Correlation Sig. (2-tailed)	.993**	1
	N	20	20

\*\* . Correlation is significant at the 0.01 level (2-tailed).

According to Table 2, the results showed a very high level of agreement between the two raters for both pre-test and post-test scores in both groups. Specifically, the correlation coefficient for pre-test scores was  $r = .945$  ( $p < .001$ ), and for post-test scores, it increased to  $r = .993$  ( $p < .001$ ). These values indicate excellent scoring consistency across raters, confirming the

reliability and objectivity of the pronunciation assessments used in the study. To evaluate adherence to a normal distribution in the pronunciation test scores, the K-S and S-W tests were employed. These statistical procedures are essential for verifying the normality assumption, a prerequisite for parametric tests such as the independent samples t-test, with the results presented in Table 3.

**Table 3**  
*Tests of Normality for Pronunciation Tests*

	K-S		S-W	
	Statistic	df	Statistic	df
Pre-test	.190	40	.001	40
Post-test	.130	40	.086	40

a. Lilliefors Significance Correction

As presented in Table 3, for the pre-test, the K-S test indicated a significant deviation from normality ( $p = .001$ ), while the S-W test result was marginally non-significant ( $p = .055$ ), suggesting approximate normality. For the post-test, the K-S test was non-significant ( $p = .086$ ), but the S-W test showed a slight deviation ( $p = .040$ ). Overall, the results indicate that the data approximately met the normality assumption, justifying the use of parametric tests such as the independent-samples t-test. Prior to interpreting the t-test results, Levene's test for equality of variances was conducted to determine whether the assumption of equal variances was satisfied. The results are presented in Table 4.

**Table 4**  
*Independent Samples t-test for Pronunciation Pre-test*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2- tailed)	MD	SED	95% Confidence Interval of the Difference	
									Lower	Upper
Pre- test	Equal variances assumed	.648	.426	.183	38	.856	.10000	.54748	-1.00832	1.20832
	Equal variances not assumed			.183	36.863	.856	.10000	.54748	-1.00944	1.20944

As shown in Table 4, Levene's Test confirmed that the assumption of equal variances was met ( $p = .426$ ). The t-test revealed no significant difference between the groups ( $t = 0.183$ ,  $p = .856$ ), with a mean difference of just 0.10. These findings indicate that both groups

started at a statistically similar level in pronunciation before the intervention, supporting the validity of subsequent comparisons.

**Table 5**  
*Independent Samples Test for the Pronunciation Post-test*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2- tailed)	MD	SED	95% Confidence Interval of the Difference	
									Lower	Upper
Post-test	Equal variances assumed	.000	.988	3.008	38	.005	1.85000	.61505	.60489	3.09511
	Equal variances not assumed			3.008	37.931	.005	1.85000	.61505	.60482	3.09518

As shown in Table 5, Levene's Test showed that the assumption of equal variances was met ( $p = .988$ ). The t-test revealed a statistically significant difference between the groups ( $t = 3.008$ ,  $p = .005$ ), with the experimental group scoring 1.85 points higher on average than the control group. This indicates that the ChatGPT-assisted instruction had a significant positive impact on learners' pronunciation.

## Discussion

The present study investigated the impact of ChatGPT-assisted instruction on Iranian EFL learners' pronunciation, specifically focusing on English phonological processes. The findings revealed a significant improvement in the experimental group's pronunciation performance compared to the control group, highlighting the pedagogical value of integrating AI tools into pronunciation instruction. This

outcome directly addresses the gaps identified, particularly the need for dynamic, rule-based feedback on connected speech phenomena (e.g., yod coalescence, flapping, and t-deletion) that earlier AI tutors failed to provide (Liao & Xue, 2019; Vančová, 2023). By demonstrating that ChatGPT can effectively scaffold phonological processes in running speech rather than isolated phonemes, our study extends prior findings and validates the potential of large language models in addressing context-specific pronunciation challenges for Persian-speaking learners.

Consistent with Namaziandost et al. (2018), who demonstrated that explicit instruction of phonological rules significantly enhances pronunciation, the current study embeds this principle within an AI-driven, interactive framework, thereby operationalizing Rvachew & Brosseau-Lapr e's (2016) theoretical assertion that learners require rule-based scaffolding for phonological acquisition. While Namaziandost's intervention relied on teacher-led methods, the present study employed ChatGPT to deliver real-time corrective feedback and individualized practice features that align with Dennis (2024), who found that AI-powered speech recognition tools effectively improve pronunciation accuracy and fluency. Moreover, by incorporating scenario-based prompts and multimodal input (written, auditory, interactive), our approach responds to Cognitive Load Theory and Multimedia Learning Theory (Mayer et al., 2014), reducing extraneous load and enhancing the integration of new phonological knowledge.

Moreover, this study confirms and builds upon Vančová's (2023) meta-analytic findings, which emphasize that AI pronunciation tools

contribute to improved intelligibility and reduced learner anxiety. Vančová reported a misclassification rate for connected-speech phenomena in existing AI tutors, highlighting a critical limitation for Persian speakers whose L1 transfer patterns differ substantially from those of Czech learners (Vančová, 2023). By targeting features such as yod coalescence, flapping, and t-deletion, phonological processes that are both phonetically complex and pedagogically underexplored, the present study offers a novel empirical contribution by demonstrating how ChatGPT can effectively support learners' internalization and application of connected speech features, a critical yet underexplored area in L2 pronunciation pedagogy.

The study also supports Mompean's (2024) analysis, which proposed that ChatGPT could serve as a supplementary tool for pronunciation practice within communicative approaches. Mompean's pilot found that while ChatGPT generated detailed phonological explanations, learner confusion often arose from inconsistent prompt structures. In response, our study incorporates optimized prompt designs that explicitly highlight Persian-specific transfer issues (e.g., the common deletion of /ŋ/ in word-final positions), thereby reducing ambiguity and enhancing learner comprehension. The findings of the present study extend Mompean's position by showing that ChatGPT, when systematically integrated into instruction with carefully engineered prompts, can function as more than a supplementary tool—it can operate as a dynamic pronunciation tutor that promotes both segmental and suprasegmental development. The use of real-time text-based interaction, auditory modeling, and corrective feedback aligns with Schmitt's (2002) view of

pronunciation as an interactive component of spoken discourse that benefits from multimodal input.

The pronounced gains observed among the experimental group further support Fraiwan & Khasawneh (2023) and Zheng et al. (2023), who highlight how adaptive learning technologies enhance learner motivation and engagement. Unlike Fraiwan & Khasawneh's one-way virtual pronunciation coach, our ChatGPT-based approach enabled two-way dialogues that allowed learners to ask follow-up questions (e.g., "Why is 'better' pronounced [berə] in American English?") and receive contextualized feedback on stress, rhythm, and intonation. This learner autonomy aligns with Rezai et al. (2024), who emphasized AI's potential to foster metacognitive reflection and self-regulated learning in EFL contexts. To address the inter-learner variability noted by Rezai et al. (2024), we incorporated a brief orientation module and reflective journal prompts, fostering the metacognitive readiness necessary for effectively interpreting ChatGPT's feedback.

Importantly, the results also reflect the instructional needs identified by Gilakjani (2012) and Hashemian & Heidari Soureshjani (2013), who criticized the marginalization of pronunciation instruction in Iranian EFL settings. By offering an accessible, scalable, and pedagogically robust alternative, this study demonstrates how AI tools like ChatGPT can bridge instructional gaps in contexts where traditional methods fall short or are under-resourced. The ability of ChatGPT to provide models of authentic pronunciation, particularly valuable where access to native speakers is limited, responds to concerns raised by Moyer (2013) and Tiwari (2024) regarding fossilized

errors among non-native instructors. Furthermore, by aligning our instructional modules with the Evolve 4 curriculum and maintaining methodological consistency across conditions, we ensured that improvements in the experimental group are attributable to ChatGPT's unique affordances rather than curricular variations.

Despite these positive outcomes, our findings also highlight challenges related to learner uptake and digital literacy, echoing Oskoui et al. (2024) and Tupamahu & Gaspersz (2024). Oskoui et al. reported that only a few Iranian EFL learners had experimented with AI chatbots for pronunciation practice, indicating a significant user-behavior gap. Although our orientation module improved initial engagement, sustained use of ChatGPT beyond the intervention period remains an open question. Likewise, Tupamahu & Gaspersz (2024) found that prosodic gains learned through group drills failed to generalize, underscoring the need for individualized feedback loops, an area where ChatGPT showed promise but requires further investigation in extended, longitudinal studies.

In summary, the present study not only aligns with but meaningfully extends the current literature on AI in pronunciation instruction. By demonstrating that ChatGPT can provide dynamic, context-sensitive feedback on discrete phonological processes, thereby reducing misclassification rates for connected speech errors and fostering metacognitive readiness, our research offers a replicable model for under-resourced EFL contexts. The findings underscore the importance of instructional design, learner autonomy, and context-specific adaptation when implementing AI tools like

ChatGPT in language classrooms. Moreover, by integrating prompt engineering strategies (Mompean, 2024) and orientation modules to scaffold metacognitive skills (Rezai et al., 2024), we address the critical gaps identified. Future research should build on these insights by exploring longitudinal retention of pronunciation gains, comparing ChatGPT with other AI tools (e.g., ELSA, Google Speech-to-Text), and examining teacher moderation's role in maximizing AI's pedagogical potential.

## Conclusion

This study set out to examine the effectiveness of ChatGPT-assisted instruction on the pronunciation of English phonological processes, specifically, yod coalescence, flapping, and t-deletion, among Iranian EFL learners. By integrating ChatGPT's generative, context-sensitive feedback into a structured pedagogical framework, we have demonstrated that AI-driven pronunciation tutoring can significantly enhance learners' segmental and suprasegmental accuracy, addressing gaps identified in earlier research (Liao & Xue, 2019; Vančová, 2023). Firstly, our findings confirm that ChatGPT can provide dynamic, rule-based feedback on connected-speech phenomena, overcoming the limitations of isolated-phoneme drills and one-way feedback systems (Suhaili & Kurniawan, 2019; Fraiwan & Khasawneh, 2023). Participants in the experimental group showed marked improvements in both intelligibility and prosodic features, aligning with De la Vall & Araya's (2023) assertion that AI feedback on suprasegmental features leads to gains in sentence-level prosody, yet extending their work by leveraging ChatGPT's capacity for real-time adaptation to novel utterances.

Secondly, the reduced misclassification rates for connected-speech rules observed in this study suggest that ChatGPT's large language model architecture can better accommodate Persian-specific phonological transfer patterns compared to earlier AI platforms (Vančová, 2023). This corroborates Mompean's (2024) pilot findings regarding ChatGPT's ability to generate nuanced phonological explanations, and further shows that when combined with optimized prompt designs that highlight L1 transfer issues, learner comprehension and pronunciation accuracy improve substantially. Thirdly, by embedding an orientation module and reflective journal prompts, we addressed the critical issue of metacognitive readiness (Rezai et al., 2024), ensuring that learners were not only exposed to accurate feedback but also equipped to interpret and apply it effectively. This scaffolding proved crucial in mitigating the high inter-learner variability reported by Rezai et al. (2024) and fostering sustained engagement, tackling the low AI uptake documented by Oskoui et al. (2024).

Language instructors can leverage ChatGPT to transform traditional pronunciation drills into interactive, rule-based dialogues by integrating AI-generated phonological explanations into their lesson plans, thereby providing learners with immediate, personalized corrective feedback on connected-speech phenomena and addressing the marginalization of pronunciation in many EFL curricula (Gilakjani, 2012; Hashemian & Heidari Soureshjani, 2013). Curriculum designers, in turn, should consider incorporating AI-oriented orientation modules and reflective tasks within pronunciation syllabi, embedding explicit training on interpreting AI feedback and prompting learners to maintain

phonological journals to enhance metacognitive skills and foster more effective self-regulated learning (Rezai et al., 2024). Educational policymakers and program administrators in under-resourced contexts can promote the adoption of ChatGPT as a cost-effective, scalable tool by providing access to generative AI platforms and training materials, which would help mitigate the shortage of native-speaking instructors and address fossilized pronunciation errors among non-native teachers (Moyer, 2013; Tiwari, 2024). Overall, this study contributes a replicable framework for integrating ChatGPT into pronunciation pedagogy, particularly in under-resourced EFL contexts, by aligning the intervention with Cognitive Load Theory (Mayer et al., 2014) and leveraging ChatGPT's multimodal capabilities to reduce extraneous cognitive load and facilitate deeper phonological processing.

Despite the positive results, several limitations warrant consideration. First, this study focused on intermediate-level learners using a relatively short intervention period; longitudinal research is needed to assess the retention of pronunciation gains and the extent to which improvements generalize to unscripted, authentic communication over time. Second, while ChatGPT outperformed rule-based tutors in this context, comparisons with other AI tools (e.g., ELSA Speak, Google Speech-to-Text) remain unexplored. Future studies should conduct controlled comparisons to determine the relative efficacy of different AI architectures in handling both segmental and suprasegmental features. Third, although our orientation module improved initial engagement, sustained, autonomous use of ChatGPT post-intervention was not measured; follow-up studies might

incorporate usage analytics to investigate long-term learner behavior, addressing the “user-behavior gap” highlighted by Oskoui et al. (2024).

In conclusion, this research underscores the transformative potential of ChatGPT as an AI tutor for English pronunciation, demonstrating that when coupled with carefully engineered prompts (Mompean, 2024) and metacognitive scaffolding (Rezai et al., 2024), large language models can effectively bridge instructional gaps in under-resourced EFL settings. By targeting discrete phonological processes in connected speech, we provide empirical evidence that AI-driven, rule-based feedback can accelerate learners' acquisition of both segmental and suprasegmental features, offering a scalable, pedagogically robust solution for Iranian EFL classrooms and beyond. Future research should extend these findings by exploring longitudinal retention, comparative analyses of AI platforms, and the role of teacher facilitation in maximizing the pedagogical benefits of AI-assisted pronunciation instruction.

## References

- Celce-Murcia, M. (1987). Teaching pronunciation as communication. *Current Perspectives on Pronunciation*, 2(2), 1-12.
- Cheptabok, M. C., & Onchera, P. O. (2024). Challenges faced by students in developing English language speaking skills in public secondary schools in Homa Bay County, Kenya. *Journal of Frontiers in Humanities and Social Sciences*, 2(3), 80-98. <https://doi.org/10.69897/jofhscs.v2i3.167>
- Dave, A. (2004). *Oxford placement test 2: Test pack*. Oxford University Press.
- De la Vall, R. R. F., & Araya, F. G. (2023). Exploring the benefits and challenges of AI-language learning tools. *International Journal of Social Sciences and Humanities Invention*, 10(1), 7569-7576. <https://doi.org/10.18535/ijsshi/v10i01.02>

- Dennis, N. K. (2024). Using AI-powered speech recognition technology to improve English pronunciation and speaking skills. *IAFOR Journal of Education*, 12(2), 107-126.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410. <https://doi.org/10.1111/0023-8333.00047>
- Dodd, B., & Lacano, T. (1989). Phonological disorders in children: Changes in phonological process use during treatment. *International Journal of Language & Communication Disorders*, 24(3), 333-352. <https://doi.org/10.3109/13682828909019894>
- Fraivan, M., & Khasawneh, N. (2023). A review of ChatGPT applications in education, marketing, software engineering, and healthcare: Benefits, drawbacks, and research directions. *arXiv Preprint*. <https://doi.org/10.48550/arXiv.2305.00237>
- Georgiou, G. P. (2025). Artificial intelligence in pronunciation teaching: Use and beliefs of foreign language teachers. *arXiv Preprint*, 1(1), 1-25. <https://doi.org/10.48550/arXiv.2503.04128>
- Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science*, 2(3), 119-128.
- Habibi, P., Jahandar, Sh., & Khodabandehlou, M. (2013). The impact of teaching phonetic symbols on Iranian EFL learners' listening comprehension. *Indian Journal of Fundamental and Applied Life Sciences*, 3(3), 495-512.
- Hall, S. (1997). Integrating pronunciation for fluency in presentation skills. In *Annual Meeting of the Teachers of English to Speakers of Other Languages*, Orlando, USA.
- Hashemian, M., & Heidari Soureshjani, K. (2013). An analysis of pronunciation errors of Iranian EFL learners. *Research in English Language Pedagogy*, 1(1), 5-12.
- Henning, W. A. (1964). Phoneme discrimination training and student self-evaluation in the teaching of French pronunciation (Unpublished doctoral dissertation). Indiana University.
- Hodgetts, J. (2020). *Pronunciation instruction in English for academic purposes: An investigation of attitudes, beliefs and practices*. Springer Nature.
- Kelly, G. (1969). *25 centuries of language teaching*. Newbury House.
- Kurt, Ş. (2018). Raising EFL learners' awareness of suprasegmental features as an aid to understanding implicatures (Master's thesis). Bilkent University.
- Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.
- Liao, X., & Xue, J. (2019). Computer-assisted pronunciation training on improving vowel production of Chinese EFL learners. *Computer Assisted Language Learning*, 32(7), 607-629.
- Liu, D., Jones, T., & Reed, M. (2022). *Phonetics in language teaching*. Cambridge University Press.
- Mayer, R. E., Lee, H., & Peebles, A. (2014). Multimedia learning in a second language: A cognitive load perspective. *Applied Cognitive Psychology*, 28(5), 653-660. <https://doi.org/10.1002/acp.3050>
- Mohammadkarimi, E., Omar, J. A., & Rashid, A. S. (2025). Advancements and challenges in automated evaluation of spoken language proficiency using AI. In *Using AI Tools in Text Analysis, Simplification, Classification, and Synthesis* (pp. 229-260). IGI Global.
- Mompean, J. A. (2024). ChatGPT for L2 pronunciation teaching and learning. *ELT Journal*, 78(4), 423-434. <https://doi.org/10.1093/elt/ccae050>
- Moyer, A. (2013). *Foreign accent: The phenomenon of nonnative speech*. Cambridge University Press.
- Namaziandost, E., Esfahani, F. R., & Hashemifardnia, A. (2018). The effect of teaching phonological rules on English pronunciation among Iranian pre-intermediate EFL learners. *European Journal of English Language Teaching*, 3(3), 81-98. <http://dx.doi.org/10.46827/ejel.v0i0.1626>
- Nushi, M., & Shahhosseini, K. (2020). On the efficacy of a communicative framework in teaching English phonological features absent in Persian to Iranian EFL learners. *Issues in Language Teaching*, 9(1), 57-91. <https://doi.org/10.22054/ilt.2020.51623.486>
- Oskoui, K., Mirzaeian, V. R., & Nafissi, Z. (2024). AI-assisted EAP testing: A case of academic IELTS writing by Iranian EFL learners. *Journal of English Language Teaching and Learning*, 16(34), 307-

330. <https://doi.org/10.22034/elt.2024.63345.2691>
- Ramelan, M. A. (1994). *English Phonetics*. Semarang, Indonesia.
- Rao, P. S. (2019). The role of English as a global language. *Research Journal of English*, 4(1), 65-79.
- Rezai, A., Namaziandost, E., & Hwang, G. J. (2024). How can ChatGPT open promising avenues for L2 development? A phenomenological study involving EFL university students in Iran. *Computers in Human Behavior Reports*, 16(1), 15-30. <https://doi.org/10.1016/j.chbr.2024.100510>
- Richards, J., Platt, J., & Weber, H. (1992). *The Longman dictionary of applied linguistics*. Longman.
- Ruhnke-Ramos, N. K., & Delatorre, R. (2011). The effects of training and instruction on the perception of the English interdental fricatives by Brazilian EFL learners.
- Rvachew, S., & Brosseau-Lapr e, F. (2016). *Developmental phonological disorders: Foundations of clinical practice*. Plural Publishing.
- Schmitt, N. (2002). *An introduction to applied linguistics*. Oxford University Press.
- Seidllhofer, B. (1995). Pronunciation awareness: A focus on appropriateness rather than correctness: Some thoughts on pronunciation in teacher education. *Speak Out*, 6(8), 12-16.
- Suhaili, A., & Kurniawan, A. (2019). The use of pronunciation Bingo games to improve students' pronunciation. *Journal of English Language Teaching and Literature (JELTL)*, 2(2), 170-189. <https://doi.org/10.47080/jeltl.v2i2.630>
- Szyszka, M. (2017). *Pronunciation learning strategies and language anxiety*. Springer.
- Tiwari, H. P. (2024). Pedagogical practices of speaking skill instruction in secondary schools. *LSP International Journal*, 11(1), 63-77. <https://doi.org/10.11113/lspi.v11.22575>
- Tupamahu, M. S., & Gaspersz, S. (2024). Analysis of EFL students' difficulties in learning English phonology and phonetics at the English Department, Victory University of Sorong. *Interaction: Journal of Pendidikan Bahasa*, 11(2), 464-478. <https://doi.org/10.36232/interactionjournal.v11i2.39>
- Van ova, H. (2023). AI and AI-powered tools for pronunciation training. *Journal of Language and Cultural Education*, 11(3), 12-24. <https://doi.org/10.2478/jolace-2023-0022>
- Yates, L., & Zielinski, B. (2009). *Give it a go: Teaching pronunciation to adults*. The AMEP Research Centre.
- Zheng, L., Niu, J., Zhong, L., & Gyasi, J. F. (2023). The effectiveness of artificial intelligence on learning achievement and perception: A meta-analysis. *Interactive Learning Environments*, 31(9), 5650-5664. <https://doi.org/10.1080/10494820.2021.2015693>
- Zsiga, E. C. (2024). *The sounds of language: An introduction to phonetics and phonology*. John Wiley & Sons.