Utilizing AI-Powered Speech Recognition Technology In Saudi EFL Learners Speech

Abdullah Nijr Alotaibi
Department of English, College of Education, Majmaah University, Al-Majmaah, 11952, Saudi Arabia

Abstract
This study explored the way Saudi EFL students speak English after using speech recognition software operated by artificial intelligence. To obtain data about students' experience and observable improvement, researchers used questionnaires and pre- and post-tests to assess improvement. During the span of eight weeks, 45 Saudi intermediate students worked with Speechling, a computer-based voice recognition system. Speechling gave students immediate feedback about their pronunciation of English and a structured approach for practicing pronunciation. In order to get accurate information about students' attitudes towards the technology and how they responded to it, researchers administered them surveys along with spoken tests. The tests measured how much pronunciation and fluency had improved and the questionnaires asked about feelings towards. In addition to testing by speech, researchers gave students questionnaires to get detailed information on what they felt about the technology and how they responded to it. The tests measured how much pronunciation and fluency had improved and the surveys on attitudes on and feelings towards the technology. The participants' fluency and confidence in English speech were improved after going through AI-SRT. Their average pronunciation score was much better from 7.2 to 14.6 (Cohen's d = 2.4). The same outcome was observed in the fluency tests: participants' rhythm was improved and fluency was higher with reduced hesitations. They sounded more natural overall. Most of the participants had a positive attitude towards using AI-SRT. 78% of the respondents liked getting instant feedback and almost 82% reported it had helped their standard classroom learning. Understanding of the connection between language and AI learning is made possible through this initiative. A perfect example of how this approach of teaching improves speaking ability and pronunciation is the utilization of speech recognition software by English language learners of Arabic origin.
Keywords: AI-SRT, Speechling, AI-SRT intervention, pronunciation, fluency, speech recognition

INTRODUCTION

English instruction is no longer written. Everything is evolving due to technology, and significant shifts in teaching are disrupting traditional methods. For the record, the primary reason that transformed schooling is artificial intelligence. Among the most pervasive elements of school curricula that have been implemented worldwide is English as a Foreign Language (EFL). Students learn differently when their English language teachers leverage AI tools to assist them in teaching. Emergent needs of the students can be addressed in tailoring the learning process. This is aligned with the objective of contemporary education, which is to make learning smart, productive, and technology-enabled for every student.
English Arabic native speakers also have particular challenges unique to them compared to other language learners. These are not merely related to vocabulary; according to the research referenced by Alotaibi (2021, 2022) and Altakhaineh et al. (2024), conventional approaches have not considered the unique complexity between English and Arabic phonemes. However, with the pace at which artificial intelligence is developing, there is a

lot of optimism. The future challenges can be met through the application of AI-SRT, or artificial intelligence-based speech recognition technology. As AI-SRT has the ability to apply personalized pronunciation practice, appropriate feedback, and individualized guessing of their level using learning, it provides native Arabic speakers with multiple language improvement strategies (Huang et al., 2022).

Exploring AI-SRT's possible value to students of Arabic-English automatically raises questions that are outside the purview of a study of educational methods. It raises questions about the international adoption of technology, the position of language in technology application, and whether technologies developed for the purpose of serving global educational requirements can be locally used. Examining Arab-speaking students' experience with AI gives us the potential to develop the tools with potential value to students (the hope being that they have an impact). It also positions itself within the wider debate surrounding AI and learning. This study is an investigation into just how effective AI-SRT is within a context of Arabic-speaking English students. It addresses the impact of language on learners' response to technology, their preparedness to utilize it, and consequently their outcomes. At its heart, such research aims to identify true insights based on evidence that will enable the creation of language learning supports for particular language communities. Concurrently, it also seeks to make a contribution towards our understanding of how language influences interactions with technology in the process of learning a new language.

This research is mainly about two questions that quantify the efficiency of AI-SRT for English language learners of Arabic. The study begins to examine whether using AI-SRT is helpful to the performance of Arabic learners under pronunciation and speaking skill categories, that is, to enhance in the difficult areas which can be trained because of frequent switching from Arabic to English. Second, the study seeks to know how students perceive the AI-SRT. What do they like? What are they worried about? And how do they want to integrate technology with traditional pedagogy in the teaching of lessons? The study also raises concerns regarding the use of AI-based language instruments in other settings, e.g., in schools. Hence, with findings - against the shortcomings in real-world implementations of AI-SRT for Arabic-English learning - this study is a first step for schools and educators in finding out how to incorporate new technology in their pedagogy without sacrificing the quality of good teaching or the language gap.

REVIEW OF LITERATURE

AI-SRT has been able to revolutionize the field of computer-assisted language learning. Beforehand, language learners could not receive support to improve pronunciation and speaking skills one on one, but it is very much possible now (Crompton et al., 2024). State-of-the-art AI-SRT systems are not just bells and whistles; they apply machine learning, neural networks, and natural language processing to hear a speech, identify communicative breakdowns, and provide genuinely meaningful feedback (Nguyen, 2024). These technologies have developed much over the last decade. Contemporary systems demonstrate a range of capabilities such as speech recognition, pronunciation assessment, and providing immediate feedback with excellent precision, and this is done with outstanding accuracy and even when the language is not known or the speaker comes from another geographical region (Ngueajio & Washington, 2022; Rogers et al., 2020).

The development of AI-SRT is a very logical evolution of language learning. It basically takes all the theory that is behind second language acquisition and places it within a platform that provides automatic feedback, corrections, and guides learners towards independent learning (Dimitriadou & Lanitis, 2023). A recent research article demonstrates

that the pronunciation of learners improves optimally in computer-assisted pronunciation training when feedback that they receive is timely, targeted, and consistent. These are exactly the domains where CALL systems excel (Chapelle, 2001; Chapelle & Sauro, 2017; Zhao et al. 2021). Human teachers can make every moment of instructional learning accurate, but they sometimes miss very subtle nuances of pronunciation mistakes and neglect to provide identical feedback to various students in an inconsistent manner. In contrast, AI offers constantly available, unbiased, and consistent pronunciation analysis and correction for every learner (Yalcin & Korkmazgil, 2021; Zhang et al., 2024).

In acknowledgment of the increasing trends in experimental study in the effectiveness of AI-SRT in the language-learned domain, there has been a surge of experimental studies within AI-SRT's efficacy. These have exemplified many variables including the enhancement of pronunciation, establishment of speaking skill, and learner engagement in various language contexts. In fact, even recent studies outrightly assert that AI-SRT appears to be an effective intervention in various contexts of languages as well as numerous learner populations (Al-Mamary et al., 2024; Mohammadkarimi, 2024; Du & Daniel, 2024). Xu et al. (2025) were exhaustive in their review of literature of AI-SRT studies, and categorized a sequence of studies showing a substantial impact on learners' pronunciation accuracy, speaking fluency, and learner confidence on a range of language pairs. Similarly, Mingyan et al. (2025) investigated the effectiveness of AI feedback interventions on second language pronunciation. Their research showed drastic modifications of accent and segmental and suprasegmental features of speech production for their learners.

Most of the positive effect for which AI-SRT have been widely shown in numerous studies, but such challenges and restraints outlined, in particular, the ones, which are stated in the articles by Abbas et al. (2024) and Perkins et al. (2024), may hinder the proper implementation of such technologies into real-world educational settings. Incorrect feedback due to low speech recognition accuracy can be a cause of a learning experience being harmful. For instance, if the user has a heavy accent and is a non-native speaker, and the speech recognition does not accurately understand the input, then the feedback received by the user will also be incorrect. Then in such a situation, the user may be tempted to reproduce his/her mistakes repeatedly. Moreover, differences between learners to whom AI-SRT is addressed pose challenges in solving the AI-SRT problem. The differences devoured by technological expertise, preference regarding learning, and reaction to programmed feedback among learners (Hubbard, 2013). While on the one hand, there are people who would shy away with human interaction and feedback, on the other hand, there could be others who would view AI-generated feedback as being more objective and less stress inducing to their egos.

It is education that unmasks and resolves these personal differences that is the primary cause for AI-SRT interventions. Another challenge posed by the advent of AI-SRT is the lack of ease of effective incorporation of AI-SRT into the current curricula and pedagogies. Effective incorporation can only be possible when there exists a blueprint on how the AI tools can be put to use to complement and enhance the current teaching practices without replacing them. This may entail massive training of teachers, curriculum restructuring, and institution-level facilitation for effective rollout and long-term achievement.

METHODOLOGY

The research took on a mixed-methods design. Quantitative experiments were combined with qualitative questionnaires' analysis. The design helped in describing the effectiveness of the AI-SRT program and the learners' feelings towards it. The following questions are the research questions:

Question 1: What is the scope of the AI-SRT to enhance the ability of Arabic-native-language learners to speak and pronounce English in a better way? In this case, emphasis is placed on the problems caused by the differences between Arabic and English.
Question 2: What are the attitudes of Arabic learners toward the AI-SRT program? This encompassed their attitude toward the use of technology, their fears regarding language appropriateness, and their preferred teaching methods.

Participants

The study involved 45 Saudi students taking English language courses at a locally based university. In order to enjoy an excellent mix of ability levels and backgrounds, they were chosen in random base. They had to fulfill conditions, including being native speakers of Arabic, possessing mid-level English, and being prepared to engage in an eight-week intervention with frequent training sessions. There were 23 males and 22 females in the age range of 19-28 years (M = 23.4, SD = 2.8). They had all attended at least six years of English instruction in schools, and they had experience in learning language from 6 to 15 years (M = 9.2, SD = 2.1). A survey was used to assess experience with prior technology, and 89% reported having used language learning apps, while 12% reported having used AI-based pronunciation training software. This breakdown made possible the examination of technology acceptance among learners as a function of their AI experience while ensuring that most participants possessed the basic technology skills required for the intervention. The research framework is demonstrated in Figure 1.

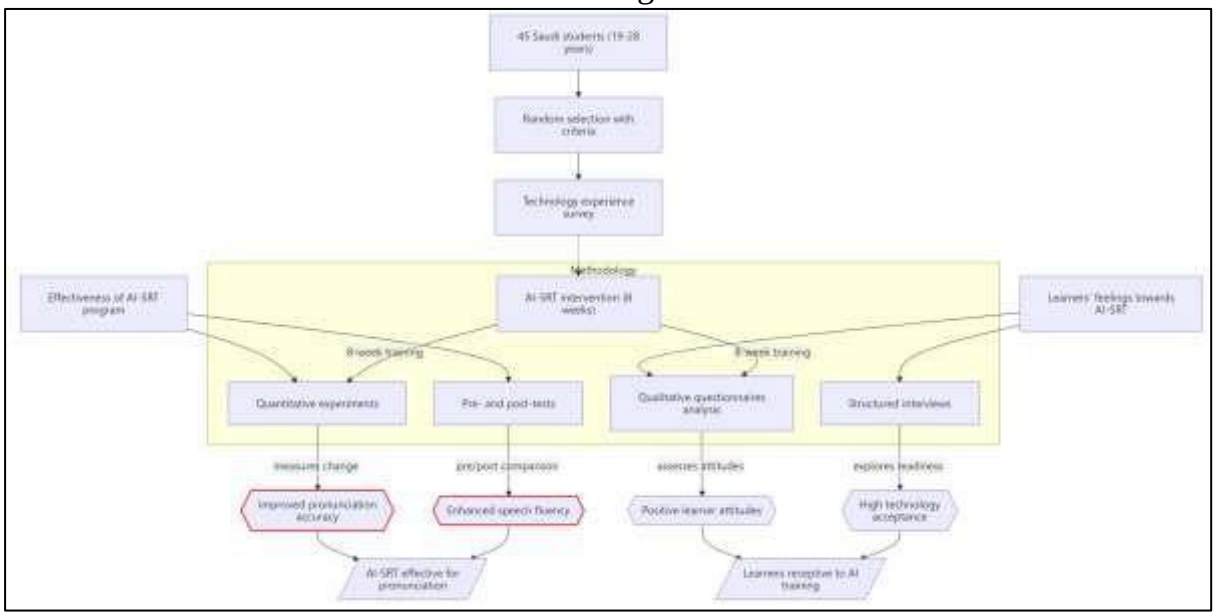

Figure 1: Conceptual Framework of the Data Collection and Analysis

Instruments and Procedure

The study employed several tests to investigate different phenomena of pronunciation development and development of speaking skills and attitudes of learners. The quantitative section of the research was based on the pretest and post-test model of study in exploring the growth in terms of pronunciation accuracy and speech fluency due to the AI-SRT treatment. The qualitative stage implied structured questionnaires and semi-structured interviews as the tools of assessment of the attitudes of learners, the readiness to technology, and the interaction with the application.

AI-SRT System and Intervention Design

The intervention was a speech recognition system based on AI titled Speechling. Speechling uses machine learning algorithms to analyze the speech patterns, detect pronunciation errors, and give detailed feedback regarding the segmental and

suprasegmental properties of speech. The eight-week intervention was structured to advance gradually through increasingly difficult pronunciation problems in a coherent sequence. It began with the production of sounds and moved on to connected speech and dialogues. The participants were required to perform at least 30 minutes of training daily for five days a week. They were free to schedule their training sessions at their convenience and in accordance with their preference.

Data Collection and Analysis

In three times the study data was gathered: pre-program, during the 8-week AI-SRT intervention, and the week after the program ended. The pre-intervention data collection was done a week before the program, and the post-intervention data collection was done the week after. Intervention monitoring was achieved through weekly progress assessment and technology use recording throughout the 8-week AI-SRT program. The analysis plan incorporated both quantitative and qualitative methodology to the highest extent possible to answer the research questions. Quantitative analysis was aimed at measuring change from pre-test to post-test in speaking fluency and pronunciation accuracy. Qualitative analysis was searching for patterns in learners' acceptance and perceptions of technology.

RESULTS

Phonetic errors instances pointed to consonant cluster issues (M = 6.8 out of a potential 20 points, SD = 2.1), absence of English vowel contrast in Arabic (M = 7.4 out of a potential 20 points, SD = 1.9), and also to speech orintonation traits like word stress patterns and intonation (M = 8.2 out of a potential 20 points, SD = 2.3). These issues were initially inscribed, which assisted in defining the intervention areas as well as establishing a quantitative measure for performance.

The analysis of the change in the accuracy of pronunciation from pre-test to post-test indicated that across all dimensions tested, there was significant improvement. The effect sizes that were calculated indicated that the improvement was practical in significance.

Global pronunciation accuracy for all subjects was high, with the pre-test average being 7.2 and standard deviation of 2.8. The mean value of the post-test was 14.6 and the standard deviation was 3.1. The average change was thus equal to 7.4, $t(44) = 12.8$, $p < .001$, Cohen's $d = 2.47$. The result showed a large effect size for the average participant equivalent to the change of nearly 2.5 standard deviations. Thus, this finding shows that the change is remarkably large from the point of practical significance. The production of consonant clusters also increased significantly, with a pre-test mean value of 6.8 and standard deviation of 2.1 (Table 1, 2, and 3).

The post-test mean was 15.2 with standard deviation 2.7, $t(44) = 15.3$, $p < .001$, Cohen's $d = 3.42$. The pattern of change in pronunciation accuracy shows that the AI-SRT system has been a very effective tool in solving the most difficult English pronunciation problem for Arabic-speaking learners. The detailed analysis of the clusters obtained has presented the evidence that there has been a reduction in the onset clusters like /st/, /sp/, and /sk/ whereas the coda clusters in the pre-test period remained constant.

Table 1: Pre-Test and Post-Test Performance Across Pronunciation Measures

Pronunciation Measure	Pre-Test Mean (M)	Post-Test Mean (M)	Mean Difference
Global Pronunciation Accuracy	7.2	14.6	+7.4
Consonant Cluster Production	6.8	15.2	+8.4
English Vowel Contrast	7.4	12.8	+5.4

Suprasegmental Features (Stress & Intonation)	8.2	13.9	+5.7
Rate of Speech (WPM)	98.3	127.4	+29.1

A significant and quite strong increase in production of English vowel contrast can be observed. Pre-test mean score has increased from 7.4 (SD = 1.9) to 12.8 (SD = 2.4) in a post-test with the following result for the statistics: t (44) = 11.2, p < .001, Cohen's d = 2.51. The relatively smaller size of effect for vowel production compared to consonant clusters may suggest that it is more challenging for English learners to separate vowels from consonants. Beyond this, it also suggests that more practice is needed to master eventually such subtle acoustic details. Suprasegmental feature generation, which encompasses both word stress and intonation patterns, was also clearly boosted.
Pre-test mean score was 8.2 (SD = 2.3) and post-test mean was 13.9 (SD = 2.8), respectively, t (44) = 10.7, p < .001, Cohen's d = 2.26. The improvements in the prosodic features were justified due to the complex nature of stress patterns in English and their very importance in general clarity and "sounding more like a native speaker.". Besides, the fluency measures of speech were significantly improved in quite several different ways. Thus, the AI-SRT intervention not only improved the pronunciation accuracy but also the verbal communicative function.
The rate of speech was also significantly improved. The pre-test mean was 98.3 words per minute (SD = 18.7) and for the post-test, it was 127.4 words per minute (SD = 21.2). Thus, the mean words per minute were enhanced by 29.1 with t (44) = 8.9, p < 0.001, Cohen's d = 1.45.

Table 2: Inferential Statistics for Pronunciation-Related Measures

Measure	**t-value**	**p-value**	**Effect Size (Cohen's d)**	**Interpretation**
Global Pronunciation Accuracy	12.8	< .001	2.47	Very large effect
Consonant Cluster Production	15.3	< .001	3.42	Very large effect
English Vowel Contrast	11.2	< .001	2.51	Very large effect
Suprasegmental Features	10.7	< .001	2.26	Very large effect
Rate of Speech	8.9	< .001	1.45	Large effect

The researcher observed much heterogeneity when he looked at differences in learning outcomes separately. With the character of the participants, efficacy for them of the intervention was predicted. By conducting a multiple regression, the researcher pinpointed these three most significant factors that were accountable to a great degree for overall pronunciation improvement: baseline English ability (β = .34, p < .01), technological self-efficacy (β = .28, p < .05), and number of practices done daily (β = .41, p < .001). Here, the model accounted for 47% of improvement score variation in pronunciation (R^2 = .47, F (3,41) = 12.1, p < .001). Improvement was weakly negatively correlated with age (r = –.23, p = .12), such that the younger subjects improved a little more, but this was not statistically significant. Gender achievement differences were close to zero and not statistically significant (t (43) = 1.2, p = .24), thus the AI-SRT intervention was equally effective for male and female participants.

Table 3: Error Patterns and Predictors of Pronunciation Improvement

A. Phonetic Error Patterns (Pre-Test)

Error Type	Pre-Test Mean (out of 20)	Standard Deviation (SD)
Consonant Cluster Issues	6.8	2.1
Absence of English Vowel Contrast in Arabic	7.4	1.9
Word Stress & Intonation Errors	8.2	2.3

B. Multiple Regression Analysis

Predictor Variable	Standardized Beta (β)	p-value
Baseline English Ability	.34	< .01
Technological Self-Efficacy	.28	< .05
Number of Daily Practices	.41	< .001
Model Fit (R^2)	0.47	< .001

A survey of familiarity with technology found that 89% of users were familiar with language learning applications, but just 12% were aware of AI-assisted pronunciation practice tools. Such an allocation of information permitted it to investigate technology acceptance across levels of experience with AI, while at the same time the vast majority of the participants were screened for the usual tech competence necessary for their adequate utilization of the intervention. Technology acceptance questionnaire results were indicative of the general positive predisposition of the participants towards the AI-SRT intervention. The mean scores were above the scale's neutral point on each of the dimensions measured.

The best rating was on perceived usefulness (M = 4.2 on a 5-point scale, SD = 0.7), as 84% of the participants strongly agreed or agreed that the AI-SRT system helped them improve their English pronunciation. Participants also rated the system as easy to use (M = 3.9, SD = 0.8); 78% of the participants either agreed or strongly agreed that they could use the system and gain advantages from it. Besides that, there was also strong behavioral intention to continue using AI-SRT technology (M = 4.0, SD = 0.9) since 76% of the participants mentioned that they would continue practicing pronunciation using the same tools after a study had finished. High behavioral intention in such a manner is an indication of the perceived value by the intervention and readiness to use similar technologies again in the future.

Thematic analysis of responses to open-ended questionnaires and interview transcripts indicated several key themes related to participants' experience of the AI-SRT intervention. The highest-rated theme was the positive experience of immediate and consistent feedback. Participants indicated that receiving an immediate and objective reaction to their pronunciation attempt was very important for them. Two of the most common quotes were: "The instant feedback made me aware of precisely what I did incorrectly" and "I was able to practice at any time and get consistent feedback, unlike human instructors whose responses would vary."

The second general theme was the accessibility of the AI-SRT system. Students enjoyed the aspect of being able to practice pronunciation at their own speed and on their own time, without the restrictions of classroom timetables or teacher availability. Statements like "I was able to practice whenever I had the chance, even late at night" and "The ability to repeat exercises as many times as needed was very useful" were the articulation of this theme.

That neutrality of AI feedback was also a source of concern for the participants. Some of them noted that they felt more relaxed in committing errors and experimenting with pronunciation if they were going to interact with the AI system rather than human instructors. Some of the comments were: "I didn't feel embarrassed when I made mistakes" and "AI doesn't judge me, so I felt free to try different pronunciations."

However, many points of the AI-SRT system were pointed out by the participants as needing changes. Some of them would like to have more contextual training sessions. They said that though the system performed well with isolated words and sentences, they would be of greater use to them if there were more interactive and realistic speaking conditions. Others proposed improving the visual feedback elements. They wanted the system to offer them more understandable phonetic comments. The majority of participants were worried about being overly dependent on technology. They saw the AI-SRT system as a helpful pronunciation aid but marked that the interaction and response from a human are still most beneficial in other areas of language acquisition. These participants stressed that there is a necessity for a mix of AI tools and conventional teaching approaches instead of the complete replacement of human instructors.

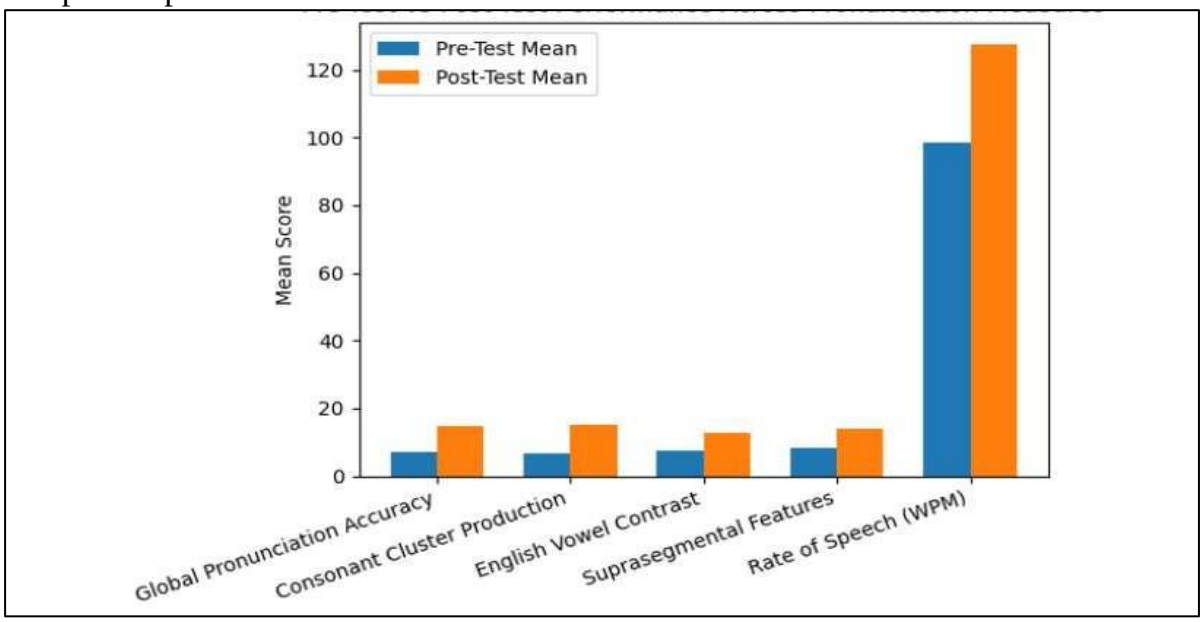

Figure 2: Pre-test and post-test mean scores across pronunciation measures

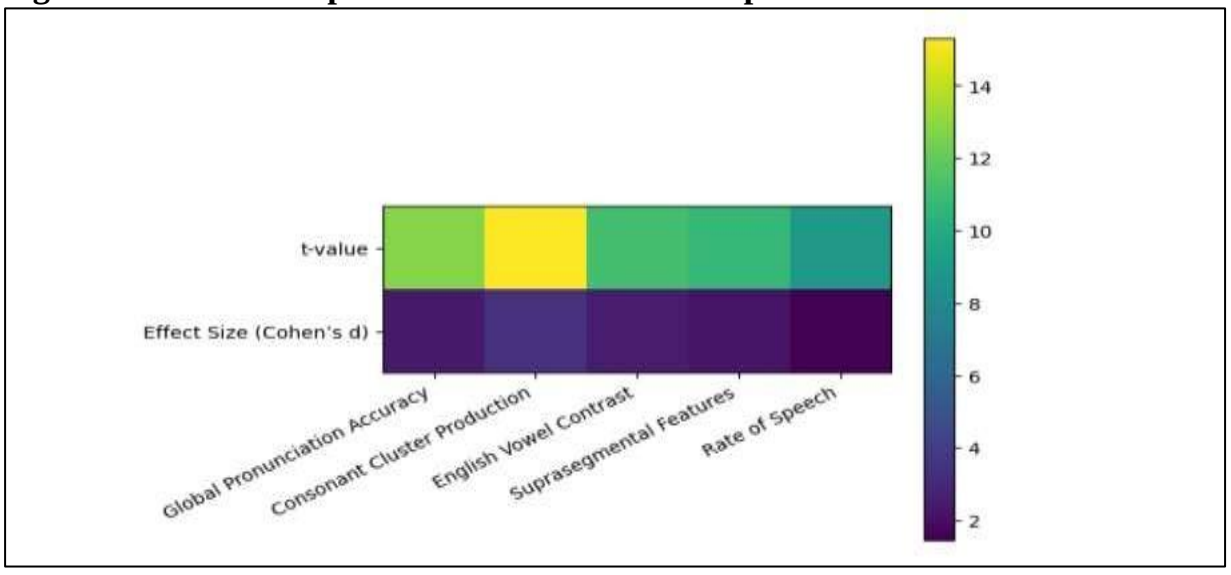

Figure 3: Demonstration of t-values and effect sizes (Cohen's d) across pronunciation-related measures

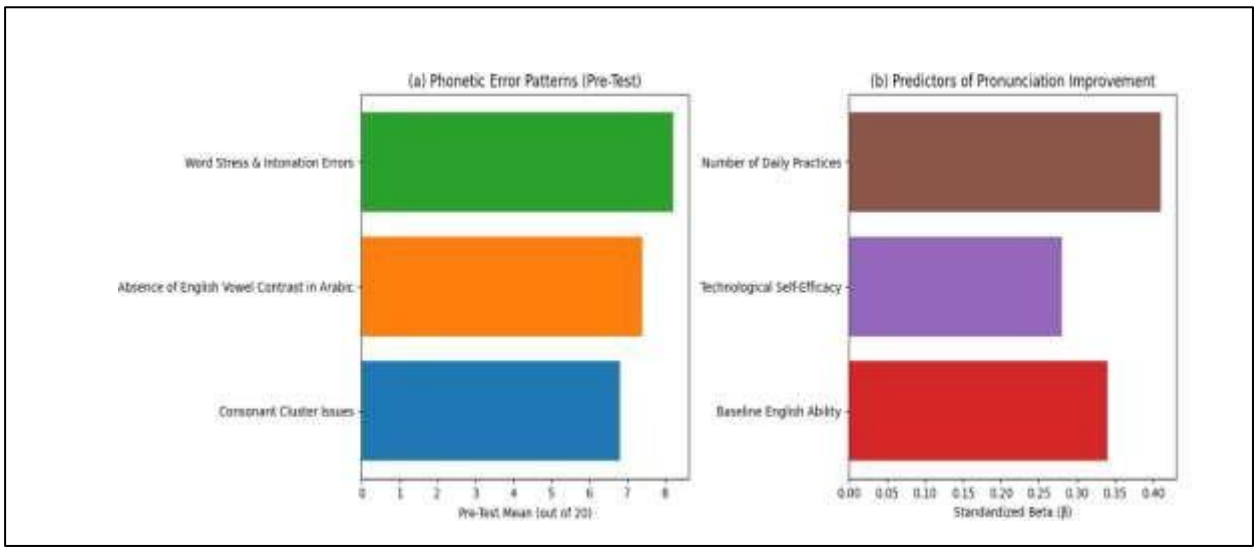

Figure 4: Comparison of (a) phonetic error patterns in the pre-test and (b) predictors of pronunciation improvement

Figures 2, 3, and 4 demonstrate the results of instructional selection that will be effective and the connections between the current factors that will cause pronunciation to improve. In Figure 2, a comparative analysis of pre-test and post-test mean scores in several pronunciation measures was undertaken, showing significant gains in pronunciation performance of learners in terms of global pronunciation accuracy, consonant cluster production, vowel contrast, suprasegmental features, and rate of speech, leading to an overall improvement in pronunciation performance of learners. The inferential statistical results are graphically illustrated in Figure 3, depicting t-values and effect sizes (Cohen d) of each of the measures associated with pronunciation, with statistically significant results and large to very large effect sizes, supporting the observed increases. Figure 4 offers a two-way comparison, with panel (a) presenting the most common patterns of phonetic errors observed at the stages of pre-testing, and panel (b) presenting the main predictors of pronunciation improvement calculated with the help of multiple regression analysis, revealing that the initial level of English skills, self-efficacy towards technologies, and the frequency of daily practice prove to be important predictors of pronunciation improvement.

DISCUSSION

The strong enhancement of the precision of pronunciation and speech fluency after this study is very convincing of the reality that AI-SRT is the best option when addressing the language problems of Arabic-speaking English learners. Besides, a very large effect size for all the domains evaluated with Cohen's d measures of 1.15 to 3.42 can be observed. This means that AI-SRT is a highly effective tool for pronunciation practice in an environment in which English learners are Arabic speakers. Notably, there were incredible changes in consonant cluster production (Cohen's d = 3.42), vowel contrasts (Cohen's d = 2.51), and stress and intonation patterns (Cohen's d = 2.26), the most common difficulties in Arabic learners.

Talking specifically regarding AI-SRT system, the performance of learners improved since it could provide immediate and direct feedback on such factors and thereby, they seem to have been able to overcome phonological transfer patterns better than if they were using traditional teaching. Finally, the theoretical implications of such findings go beyond applying teaching methods per se to questioning aspects like cross-linguistic transfer and how teaching pronunciation can be made to work. The outcome of AI-SRT treatment is

that one of the main reasons why immediate and constant feedback is very powerful in reducing the harmful consequences of phonological transfer is that the feedback stems from the most salient contrasts between the learners' first and target languages' phonetic systems.

The relationship between baseline English ability and gain ($r = .42$) was moderate, which would suggest that those with greater English ability at the beginning would derive more benefit from the AI-SRT interventions. Yet still, the approach is effective with varying levels of ability.

The technological self-efficacy component to play the central role in learning performance ($\beta = .28$) reveals that learners' confidence and familiarity with technology influence the effectiveness of such interventions to that degree. The implication of this discovery is that the proper AI-SRT implementation might just need adequate training along with the provision of tech support in particular to learners who have had very limited prior exposure to the application of AI tools. By providing initial orientation and technical support to enhance learners' interest and learning outcomes, schools implementing AI-SRT can optimize the deployment of the technology for learner achievement enhancement.

The absence of considerable differences in learning achievements by gender ($t(43) = 1.2$, $p = .24$) indicates that AI-SRT interventions are equally effective for male and female learners. It is a significant discovery for the management of equal access to learning opportunities for language through technology. This discovery differs from several previous findings that commented on gender contrasts in terms of using technology and might be dependent on the type of intervention, or sample, that was studied.

The different technology acceptance consequences that have been positively evaluated in this research tell a lot about the determinants that compel users to adopt and keep adopting AI-based language learning programs. The initial point that can be made from the statistics is the remarkably high mean score of 4.2 out of 5 for perceived usefulness and the mean score for the intention to keep using the tools as 4.0 out of 5. On the basis of these data, the conclusion can be made that the participants felt there were real benefits in the AI-SRT intervention. Then they were to apply the same techniques on their own after the experiment had been completed. It can be said that this high-point score of immediate feedback ($M=4.4$ out of 5) reflects the worth of that feature in user satisfaction and learning effectiveness. Instant and unemotional feedback on pronunciation is a remedy for the greatest disadvantage of traditional pronunciation instruction, which is the delay, inconsistency, or lack of feedback outside the classroom. Users' impressions derived from their feedback enable us to understand what leads to positive experiences with AI-SRT.

Most of the feedback emphasizes the benefit of judgment-free comments and postulate that AI systems might offer some psychological advantages over human learning in some cases. Most importantly, the students who get anxious while they are talking erroneously in front of the teacher or classmates can be supported in this manner. This finding is highly important for the creators of the language learning environment because it shows that AI instruments can be an excellent alternative for students with the problem of speaking fear or low self-esteem.

The pre-post-test single-group design of the AI-SRT intervention gives us clear evidence of improvement but does not allow us to establish cause and effect or compare it with other teaching methods. Hence, the results of this research must be scrutinized and new research planned accordingly. Future research needs to use controlled experimental methodologies with comparisons of AI-SRT interventions against standard pronunciation teaching or other technology-based methods.

Study Limitations

The research has yielded some optimistic results on the application of AI-SRT that enhances the English-speaking proficiency of Saudi EFL students; however, there are a few limitations that may impact the scope and generalizability of the findings. One limitation of this study is the sample size, with only 45 intermediate learners included, which may not be representative of the overall population of Saudi EFL students. The human factor is also limited in terms of conducting the study, which affects the reliability of the results and the diversity of the sample in terms of proficiency levels, regions, or age groups. Future research should aim to involve larger and more diverse groups of participants to generalize the results of the technique across various populations. Another limitation is the short duration of the intervention, which lasted only eight weeks. While improvements may have been observed during this period, it may not be sufficient to assess long-term retention of skills and continuous improvement in language proficiency.

To overcome these limitations, it would be better to implement an active intervention timeline and provide a follow-up program for the long-term implications of AI-SRT. Moreover, there was no control group in the study, which does not allow the author of the study to make causal assertions on the success of the intervention. Without a comparison group, it would be difficult to assess whether the improvement was directly caused by the intervention or by other factors such as regular classroom activities or external incentives. Future studies should involve control groups to provide a clear picture of the effectiveness of AI-SRT. In addition, the research emphasized pronunciation and fluency more than other fundamental elements of language acquisition, such as vocabulary, grammar, and communicative competence in different contexts. For a more balanced intervention, extensive analysis of AI-SRT effects on language proficiency should be conducted. Additional studies need to be done on various language constructions to have a more thorough understanding of the technology's effects. Another weakness was the self-report measures, which were dialectical, as the attitudes and experiences of the participants with the technology were measured using surveys and questionnaires. Although such methods are qualitative, they are vulnerable to biases or social desiring effects.

Recommendations and scope for future research

Longitudinal research would aid in finding out the long-term effects of AI-SRT on the process of language learning and study whether the positive gains in pronunciation, fluency, and general language proficiency are maintained after the course of the preliminary intervention. In addition, controlled group research design would be more adequate in providing evidence of the causal relationship between the effectiveness of AI-SRT and traditional teaching procedures. Future research should also incorporate a multidimensional, holistic approach because they focus on various aspects of language learning. The researchers can explore the characteristics of AI-SRT, including vocabulary, grammar, and listening comprehension, to gain a better understanding of the impacts of AI-SRT on the general proficiency of learners. Mixed-methods research, which combines both quantitative and qualitative methods of data collection, would provide a deeper description of the learning process. This approach could involve interviews, focus groups, or case studies to illuminate the nuances behind the experiences of learners that may not be easily quantified. Lastly, future studies may establish the efficiency of different AI-based language-learning tools outside of speech recognition. Another area that can be continued to be studied in the future is other applications, such as AI-driven conversational companions or adaptive learning, which should enable the creation of a more holistic perspective on how technology can positively impact language learning. Despite the fact that the opportunities of AI-Powered Speech Recognition Technology as a tool that can help EFL learners are described in the given study, the insights of the limitations condition

the further research that will reduce the knowledge about effective teaching language in the world that becomes more technology-centered.

CONCLUSION

The study provides strong evidence of the efficacy of AI speech recognition technology in addressing pronunciation and speech issues directly targeted by Saudi English language learners. The substantial gains achieved in all sections where the impact was assessed with effect sizes ranging from 2.26 to 3.42, is a clear demonstration that AI-SRT is an effective educational device that can provide noteworthy learning achievements. The highly remarkable improvements in producing consonant clusters, recognizing English vowel contrasts, and learning suprasegmental features show that AI speech recognition technology can be the ideal way to overcome the long-time transfer problems that Arabic language learners have.

The research presents ideas that are not just limited to explicit instruction but also raise a query about the role of technology-mediated feedback in second language learning. The AI-SRT treatment to deep-rooted phonological tendencies has been successful and thereby it has established that problems of negative transfer could be addressed best by instant, consistent, and certain feedback. Negative transfer issues are those resulting from differences in native language phonological systems and the target language. Further, the general positive attitude of participants towards technology suggests that language learning software employing AI can be an effective means of engaging users and instruction. These results have enormous implications for schools looking to improve the instruction of English to Arabic learners.

The effectiveness of AI-SRT interventions coupled with high levels of acceptance and participation of learners indicate toward such strategies being a useful addition to the standard speech teaching methodology. On the other hand, discovery of technology self-efficacy being an important factor in determining outcomes indicates toward understanding the necessity of provision of training and assistance so as to maximize the effectiveness of these interventions. The next question needs to investigate how long pronunciation gains last as well as what the best methods of hybridizing AI with traditional teaching methods are.

References

1. Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education, 21*(1), 10. https://doi.org/10.1186/s41239-024-00444-7

2. Alotaibi, A. N. (2021). Phonological constraints on the utterance of L2 clusters by Saudi ESL learners. *Arab World English Journal, 12*(4), 551–559.

3. Alotaibi, A. (2022). Voicing contrast of L2 final stops: A case study on ESL learners from Saudi. *Journal of Language and Linguistic Studies, 18*(2), 1194–1207.

4. Al-Mamary, Y. H., Alfalah, A. A., Shamsuddin, A., & Abubakar, A. A. (2024). Artificial intelligence powering education: ChatGPT's impact on students' academic performance through the lens of technology-to-performance chain theory. *Journal of Applied Research in Higher Education, 16*(1), 23–45.

5. Altakhaineh, A. R. M., AL-Junaid, N. A., & Younes, A. S. (2024). Pronunciation and spelling accuracy in English words with initial and final consonant clusters by Arabic-speaking EFL learners. *Languages, 9*(12), 356.

6. Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge University Press.
7. Chapelle, C. A., & Sauro, S. (Eds.). (2017). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell.
8. Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, 55 (6), 2503-2529
9. Dimitriadou, E., & Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments, 10*, 12.
10. Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence, 6*, 100230.
11. Hubbard, P. (2013). Making a case for learner training in technology enhanced language learning environments. *CALICO Journal*, *30*(2), 163–178.
12. Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237–257.
13. Mingyan, M., Noordin, N., & Razali, A. B. (2025). Improving EFL speaking performance among undergraduate students with an AI-powered mobile app in after-class assignments: An empirical investigation. *Humanities and Social Sciences Communications, 12*(1), 370.
14. Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. *LLT Journal: A Journal on Language and Language Learning, 27*(1), 98–115. Sanata Dharma University.
15. Ngueajio, M. K., & Washington, G. (2022). Hey ASR system! Why aren't you more inclusive? Automatic speech recognition systems' bias and proposed bias mitigation techniques. A literature review. In *International Conference on Human-Computer Interaction* (pp. 421–440). Springer Nature Switzerland.
16. Nguyen, H. A. (2024). Harnessing AI-based tools for enhancing English speaking proficiency: Impacts, challenges, and long-term engagement. *International Journal of AI in Language Education, 1*(2), 18–29.
17. Perkins, M., Roe, J., Vu, B. H., Postma, D., Hickerson, D., McGaughran, J., & Khuat, H. Q. (2024). GenAI detection tools, adversarial techniques and implications for inclusivity in higher education. *arXiv*. https://doi.org/10.48550/arXiv.2403.19148
18. Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A primer on neural network models for natural language processing. *Journal of Artificial Intelligence Research, 57*, 615–732.
19. Xu, G., Yu, A., & Liu, L. (2025). A meta-analysis examining AI-assisted L2 learning. *International Review of Applied Linguistics in Language Teaching*. Advance online publication. https://doi.org/10.1515/iral-2024-0213
20. Yalcin, S., & Korkmazgil, S. (2021). The effectiveness of an AI-powered mobile application on the English pronunciation skills of EFL learners. *International Journal of Emerging Technologies in Learning (iJET), 16*(12), 135–148.
21. Zhang, J., Zhu, C., & Zhang, Z. (2024). AI-powered language learning: The role of NLP in grammar, spelling, and pronunciation feedback. *Applied and Computational Engineering, 102*(1), 18–23.
22. Zhao, Y., Chen, Y., & Ma, F. (2021). Effects of AI-assisted technology on second language pronunciation accuracy, fluency, and anxiety: A meta-analysis. *Computer Assisted Language Learning, 34*(7), 756–779.