Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

ASR Using Speechnotes for EFL Learners: A Study of the Effects on English Pronunciation and Prosody Skills

Fahad Aljabr[1]

Abstract

Artificial Intelligence has captured the language classrooms all over the world like never before, albeit not without many debates and discussions on the pros and cons. However, there exists some kind of consensus on the many opportunities that AI powered applications offer to the learners. This study concerns itself with the role of Automatic Speech Recognition technology alongside peer feedback on the English language (EFL) pronunciation and speaking skills of adult Saudi learners. The study uses a mixed-methods approach and an exploratory sequential design to examine whether and what is the impact of an AI application on the EFL output and whether such technology can be integrated as a learning aid in the Saudi higher education realm. The participants are 24 intermediate proficiency Saudi EFL learners at University of Hai enrolled in the sophomore year of the English Speaking Skills Course. Prior to the study, they had only received conventional instruction in pronunciation before being subject to the Speechnotes ASR application. The impact of the intervention on pronunciation and general speaking proficiency of prosodic components was studied and judged in the specially designed tasks and the standardized IELTS English Speaking Skills component. Qualitative data in the form of semi structured interviews with the sample helped triangulate the information to answer the main point of inquiry in this study: Can an ASR tech tool be effectively used as a supportive pronunciation and general speaking skills learning aid in an EFL class of university learners.

Keywords: *ASR, Speechnotes, EFL pronunciation, university learners, speaking skill development.*

Introduction

AI and its tools have captured the imagination of our race like never before, on the sidelines, this has generated a plethora of debates, doubts, and dichotomies, especially in the education sector as we argue if and how far the change will impact learners' knowledge and skill acquisition, affect teachers' training, recruitment, and pedagogies, and whether technology will (or can) replace humans.

Automatic speech recognition (ASR) technology is an exceptionally promising auxiliary tool for EFL students looking to enhance their spoken English proficiency. In the post-Covid times, new research is focusing on this upcoming educational aid in a big way. For instance, the main goal of a study by Liu et al. (2022) was to find out how students feel about ASR technology once it is used in traditional educational settings. The study sample comprised 249 English majors from a Northeastern Chinese institution, and they were split into two groups: a control group of 124 students and an experimental group of 125 students. During oral practice, participants in two groups utilized ASR technology; in contrast to the control group, the experimental group also included teacher instruction. Based on the ASR technology's scores, the teacher could provide more thorough speaking education. Participants were required to answer a survey relevant to the ASR technology and the need to learn reflective journaling while this experiment was going on. The study's findings demonstrated that both educators and individuals had favorable and contented opinions about the potential of ASR in oral training. They also thought the technology could fulfil many of their needs, including a scoring system that would enable them to more easily comprehend the actual speaking level.

Not only as a learning tool, the development of automatic speech recognition (ASR) technology has made it possible to diagnose learners' speech issues using ASR-based speech evaluation. In the meantime, there are more options for improving pronunciation with ASR-based pronunciation training. Wenqi and Moonyoung (2021) conducted a research to find out how well automatic speech recognition (ASR) identifies spoken English mistakes and to find out how teachers and students feel about using ASR technological advances as a learning aid and for pronunciation evaluation. ASR-assessed and human-

[1] English Language Department, College of Arts, University of Ha'il, Ha'il 81481, Saudi Arabia; f.aljaber@uoh.edu.sa

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

assessed read-aloud assessments were administered to five Chinese EFL students. The extent of overlapping areas was ascertained by comparing the sorts of pronunciation errors identified by the two tests. The results show that, at the segmental level, there were crossovers between robot and human assessment. Furthermore, it was discovered that the ASR device satisfied the diverse pronunciation learning requirements of the students. The study's implications shed light on the use of ASR technology to support instruction and evaluation of English pronunciation. This study is especially relevant to the current inquiry into the efficacy of Speechnotes in enhancing English pronunciation of Saudi EFL learners.

Despite the massive contribution of research to second and foreign language learning, pronunciation and training in prosody have taken a back seat. The reason for this may be the belief that understanding or comprehension of the message is not critically dependent on pronunciation (Al-Ahdal, 2020; Crowther et al., 2015; Gilakjani & Ahmadi, 2011; Ngo et al., 2023). In addition to this lackadaisical attitude to pronunciation, teachers may feel ill-equipped for pronunciation training of learners due to their own lack of training in the same (Evers & Chen, 2021) or lack of knowledge in the wherewithal (Couper, 2017). However, in foreign language learning milieus, learner needs and satisfaction spectrum as in Saudi Arabia place high weightage on native -like pronunciation for reasons as diverse as social value, perceptions of future needs such as foreign education plans, and even occupational mandates (as concluded from this study). Thus, pronunciation and prosody training cannot be given less emphasis especially in the Saudi context.

Literature Review

AI bots in language learning have evolved massively in the recent times, and especially since the laudable reception of ChatGPT an year ago. The popularity of this Generative model spread like wild fire inspiring an uptake in ASR research since then. Evers and Chen (2021) examined the effects of automatic speech recognition (ASR) technology and visual/verbal learning approaches on the development of adult learners of English as a second language over the course of a 12-week pronunciation course. The teacher corrected and provided feedback on the adult learners' pronunciation in the control group (n = 28); dictation ASR was used in conjunction with peer correction in experimental group 1 (n = 33); and dictation ASR was used alone in experimental group 2 (n = 31). Two-way ANCOVA was used to examine their pre- and post-tests on pronunciation in reading activities and live conversation while accounting for their learning styles. The findings imply that learning styles significantly impacted each group's pronunciation performance on the reading task. Learners with a visual style did better than learners with verbal style in the task given related to reading. The combination of ASR and peer correction resulted in the highest improvement in both reading tasks and live conversations.

Jiang et al. (2023) looked at how Chinese students' willingness to communicate (WTC) in spoken English and the evolution of their interconnected qualities in a switched EFL setting were affected using automatic speech recognition (ASR) technology. The 14-week quasi-experiment involved 160 undergraduate students from a Chinese institution. Both sets of students received instruction in a reversed order. During their pre-class self-study, the individuals in the treatment team had to practice their oral communication skills using the ASR technological devices, whereas the control group did not use the ASR equipment. According to the findings, students' WTC with the teacher and class and WTC with non-Chinese significantly differed between groups because of the ASR-based oral practice, suggesting that the ASR technology may help to improvise Chinese learners WTC in spoken language. On other hand on interactional part there was not a significant difference between both the teams. Though findings showed a limited role of ASR devices on Chinese learners' interactional features. In this study they studied impact if ASR devices on WTC and peer interaction had impact on divided teams.

The effect of using automatic speech recognition technology (ASR) with others to correct them on the development of second language (L2) articulation and ability to speak among English as a Foreign Language (EFL) learners will be investigated using an explanatory sequential design. In contrast to traditional teacher-led feedback and instruction, Sun (2023) research sought to determine whether this method could be a useful tool for improving L2 speaking and articulation. The results of the research imply that using ASR

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

technology in conjunction with constructive criticism can be a powerful strategy for improving EFL learners' L2 speaking and pronunciation abilities. The possibility of integrating ASR technology into language learning settings is demonstrated by the EG's superior performance over the CG in speaking and pronunciation tests. Furthermore, the encouraging comments from the EG participants highlight the importance of utilizing ASR technology as a helpful resource in instructional settings.

Chen's (2023) meta-analytic study investigates the overall impact of automatic speech recognition (ASR) on the pronunciation skills of ESL/EFL students. Results from 15 research, totaling 38 effect sizes, that were discovered between 2008 and 2021 were meta-analyzed. ASR has an instrument broad impact size, according to the meta-analysis's findings (g = 0.69). (1) ASR has a large effect on segmental pronunciation but a small effect on suprasegmental pronunciation; (2) substance to extended treatment duration of ASR shows improved educational outcomes, but short duration offers no differential effect compared to a non-ASR condition; (3) practicing pronunciation with peers; and (4) moderator analyses indicate that ASR with explicit corrective feedback is largely effective, while ASR with indirect feedback (e.g., ASR dictation) is moderately effective(.5) For adolescent (i.e., 18 years of age and up) and advanced English speakers, ASR is generally advantageous. All things considered; ASR is a useful tool that is suggested to support the improvement of L2 students' pronunciation.

Aladini (2023) sought to determine how artificial intelligence (AI) applications might enhance senior university students' logical reasoning and academic writing abilities. All of the participants (N=32) were senior English language teaching majors enrolled in two courses, Teaching English Language I and Teaching English Language II, at Dhofar University's Education department. The quasi-experimental method was used by the researcher. The data was gathered via a prep-posttest. Ten weeks were allotted for the trial (fall semester, 2022). The researcher employed artificial intelligence (AI) tools, such as Grammarly, Jasper, Quill Bot, Sudo write, and Chibi, which students utilized to complete coursework and other writing tasks. The findings demonstrated that the AI dictionary software and applications had a great impact in improvising learners academic writing ability and their logical thinking. The research has attempted to determine whether it is feasible to use AI in EFL/ESL courses in order to give a comprehensive picture of the current level of artificial intelligence in EFL/ESL instruction. It will also review previous studies that are pertinent to AI in EFL/ESL instruction, such as the use of natural language processing, virtual reality, self-regulated learning, intelligent teaching systems, and immersive virtual environments to teach English as a foreign language. Alhalangy and AbdAlgane's (2023) study used a questionnaire to collect data, which was then analyzed and discussed to reach conclusions. The findings showed that both teachers and students now bear the ethical burden of using AI in the classroom in the most effective manner possible.

383 out of 412 Chinese university EFL students participated in the current study and were split into the experimental and control groups in another study (Lingjie & Xiaojuan, 2025). Before and after a 12-session intervention, both groups took part in pre- and post-intervention tests. The experimental group's FLE, motivation, and engagement all showed statistically significant increases when a One-way Analysis of Covariance (ANCOVA) was conducted. Conversely, the control group showed a slight change in the previously indicated characteristics. AI solutions have the power to capture students' interest and, as a result, increase their desire to engage fully in the learning process. The study's findings indicate that integrating AI tools into the EFL classroom setting has important ramifications for both instructors and teacher educators.

The main goal of Marzuki (2023) was to investigate the variety of Artificial Intelligence (AI) writing tools that are currently accessible and evaluate how EFL teachers view their impact on students' writing, specifically regarding content and structure. The study was built using a case study design and a qualitative methodology. Semi-structured interviews were used to gather data regarding the variety of AI writing tools and how they affect students' writing. To shed light on the range of AI writing tools utilized in their classrooms, the study collected data from four EFL professors at three different Indonesian universities. QuillBot, WordTune, Jenni, Chat-GPT, Paperpal, Copy.ai, and Essay Writer were among these programs.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

Research Questions

The study aims to answer the following research questions:

1. What are the most problematic components in Saudi EFL learners' speaking of English as identified by Speechnotes?

2. What impact does an ASR tool like Speechnotes have on EFL learners' pronunciation in English?

3. How do Saudi EFL learners perceive the efficacy of Speechnotes in English pronunciation enhancement?

Method

In the first phase of the research, we collected quantitative data to follow the explanatory sequential mixed method of Creswell et al. (2004). The study was conducted in two stages: In the first phase, the aim was to compute the frequency of incorrect identification on Speechnotes. The content was read out by the participants in the Dictate function. In the second phase, the aim was to investigate the efficacy of an ASR application in enhancing the pronunciation and general speaking skills of adult Saudi EFL learners. This central idea germinated from earlier literature which established the importance of pronunciation accuracy in reaching fluency and general proficiency (e.g., Isaac, 2019). Moreover, there is hardly any language learning area in modern classrooms which is not touched by educational technology. The ASR technology first gained ground in the Saudi EFL scenario during the Covid times when online learning ushered in many of the tech tools that now pervade all classrooms. From an initial introduction to Zoom and Google Meet for learning which was even then mostly lecture based in Saudi Arabia, the sector has come a long way with parallel (multimodal) models of education in place where the conventional lecture method is supplemented with educational technology to address typical learning and learner needs.

Participants

We used the language laboratory at University of Hail, Saudi Arabia. This is a large, public funded university that offers diverse programs, especially in language and translation studies. The study participants were 24 intermediate proficiency Saudi EFL learners at enrolled at the University, randomly enlisted for the study. So far these participants had only received conventional instruction, and were subject to the Speechnotes ASR application with focus on English pronunciation and general speaking proficiency only during the study. All participants were native speakers of Arabic, males in the age range of 19-21 years at the time of the study. Since the IELTS at the beginning of proficiency determination procedure of the University had placed all the participants in the Intermediate speaking proficiency group, a similar test was conducted to find differences in the proficiency levels at the end of the study.

Findings and Results

Therefore, in this phase of the study we used Speechnotes (https://speechnotes.co/dictate/), a powerful speech-enabled online application with engaging cutting-edge speech-recognition technology. Its additional appeal lies in its no download, no install and even no registration policy, enabling users to start working right away. In this application, users start every note anew, thus stimulating the mind to crystallize new ideas. The user's focus is on the immediate content (a big plus for non-native language users whose imagination is wont to wander) as all but the immediate content fades out of sight. Lastly, the opportunity to speak at one's own pace, as fluently as one can, and in an uninterrupted flow, are other features that boost concentration. Figure 1 below captures a grab of a typical screen of Speechnotes.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

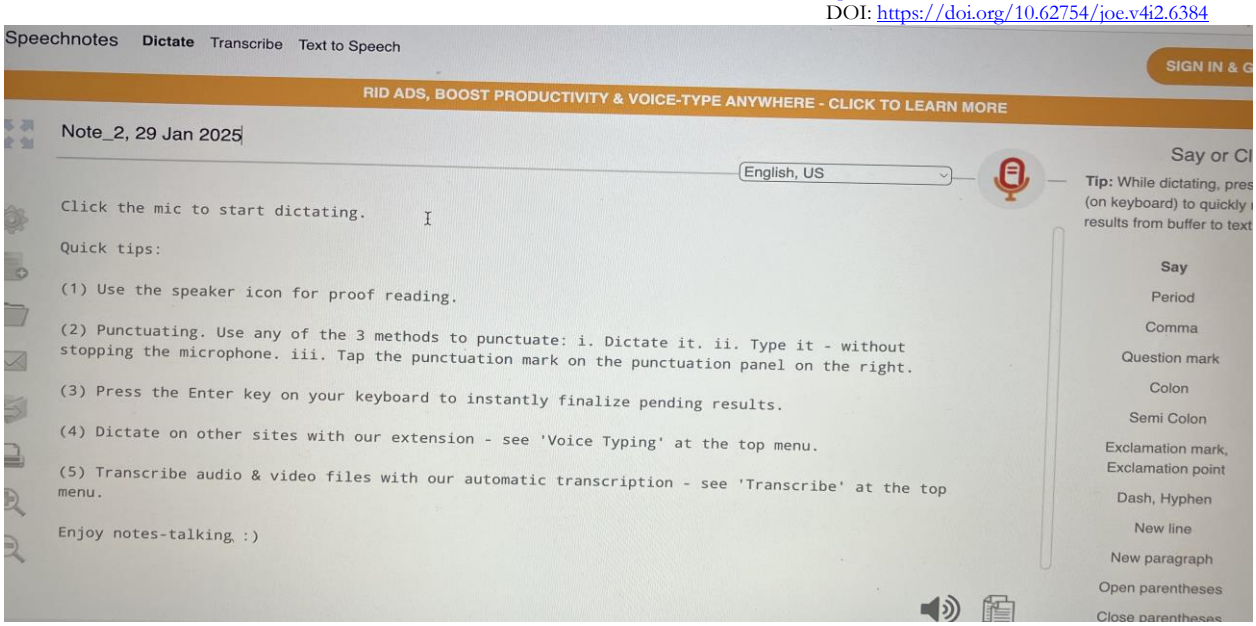

Figure 1: Screenshot grab of Speechnotes

The aim of the two groups during the four-week intervention period was to enhance English pronunciation and general speaking skills (prosodic features as in American English), with a target word bank of 28 problematic words as identified by the researcher in consultation with the EFL teachers at the University. Texts incorporating these words were then extracted from the study materials and a small corpus of lessons thus prepared for the study period. Standard practices were followed in the CG with the teacher inputs forming the only corrective or feedback measure. In the EG, however, as Task 1, the Dictate function of Speechnotes was used by the participants to actively read aloud the materials provided to them (it may be noted that in the conventional classroom, the teacher corrects the pronunciation and prosodic deviations, and has the entire group repeat the corrected versions).

The next stage was one of self-assessment as each participant procured a copy of the machine transcribed text and compared this with the original material provided to them. they were then asked to mark the words that the software identified incorrectly (because the participant pronounced it wrong) and record them in a self-maintained pronunciation sheet. In the following session they were asked to go back to Speechnotes, but this time, verify the correct pronunciation(s) and prosodic features from the Text to Speech function (Task 2), practice the corrections and repeat task one i.e., Dictate, check the output again for deviations, and record on the pronunciation sheet against the original mistake. At the end of the study period, these sheets were shared with the researcher who summarized the output of all the 24 participants and identified the leading problem areas as recorded in Table 2 below in the form of frequencies. All 26 words and phrases used in the tasks were reported to have been incorrectly identified though with different frequencies.

Table 2: Task 2

No.	Word/ phrase	Frequency of incorrect identification in Dictate function in Task 1 (EG)	Nature of incorrect identification	Frequency of incorrect identification in Dictate function in Task 2 (EG)
1	tasks	4	Consonant cluster broken	0
2	splash	7	Consonant cluster broken	2
3	finish	18	confusion between /ɪ/ and /e/	6
4	phone	21	Diphthong /əʊ/	4
5	day	20	Diphthong /eɪ/	4
6	pay	23	Diphthong /eɪ/	5

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

7	cupboard	12	Consonant cluster broken, substitution of /p/ with /b/	3
8	property	16	Soft /b/	3
9	happy	21	Soft /b/	4
10	favorite	13	Substitution of /v/ with /f/	4
11	flavor	15	Substitution of /v/ with /f/	3
12	kitchen	18	Substitution of /ʧ/ with /ʃ/	5
13	valley	14	Substitution of /v/ with /f/	3
14	situation	16	Substitution of /ʧ/ with /ʃ/	4
15	ringing	19	Words ending in the continuous tense, median and word final /ŋ/	8
16	I want fish and chips	23	Problems with stress timed patterns where (want, fish, chips) are to be produced as stressed syllables	7
17	properly	15	Soft /b/	3
18	strike	4	Consonant cluster broken	0
19	blowing	12	Consonant cluster broken, word final /ŋ/	8
20	always	19	Diphthong /eɪ/	3
21	We want to go out tonight	21	Problems with stress timed patterns	6
22	telephone	19	Diphthong /əʊ/	4
23	Ali is a very good student	19	Problems with stress timed patterns	8
24	break	17	Diphthong /eɪ/	3
25	grab	5	Consonant cluster broken	0
26	global	16	Diphthong /əʊ/	5

Using the above data, the P value and statistical significance computations indicated:

1. A two-tailed P value less than 0.0001: This value shows that the difference in incorrect identification, and hence learners' pronunciation and general speaking ability with respect to English prosody before and after the intervention to be extremely statistically significant. In other words, the intervention is highly likely to have positively impacted these skills.

2. Mean: In Task 1 the group mean of incorrect identification of content was 15.65 against 4.04 in Task 2 in the post intervention stage. This indicates that group performance improved drastically in the post stage.

3. The difference in averages of incorrect identification frequency between Task 1 and Task 2 is 11.62. This implies that, on average, there were 11.62 more incorrect identifications in Task 1 compared to Task 2, or that the frequency of errors in the pre stage is much higher than the same in the post stage.

4. **The confidence interval of this difference is 95%, from 9.87 to 13.36** which indicates that one can be 95% certain that the true difference in the frequency of incorrect identification between Task 1 and Task 2 lies somewhere between 9.87 and 13.36.

In summary, the results show a statistically significant difference between the two tasks in terms of incorrect identification, with Task 1 showing consistently higher errors than Task 2. The confidence interval further confirms the reliability of this difference estimate.

Apart from these findings, the data also shows that there are certain areas of mis-pronunciation or speaking that are more prominent or problematic than others in the case of Saudi EFL learners. The English language content that the study tested were consonant clusters, word median and final /ŋ/, Diphthong /əʊ/,

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

Diphthong /eɪ/, problems with stress timed patterns, Problems with stress timed patterns, substitution mistakes. The least problematic of the tested items were consonant clusters: Arabic has rare blends as compared to English which has these on the higher side, leading Arab speakers of English to break the clusters with vowel insertions. However, in this study we found that this was the least of the problems which may be attributed to the fact that with many years of exposure to the language, learners overcome this lacuna which (to teachers of English) may appear to be quite jarring and hence, inviting constant correctional feedback. Substitution of Substitution of /v/ with /f/ followed consonant cluster mistakes with the frequency ranging between 13-15 followed by another substitution mistake of replacing /tʃ/ with /ʃ/ as the former does not exist in Arabic leading to substitution with the familiar latter sound. Similar reverts were found in the use of the soft /b/ for /p/ as the latter is missing in Arabic phonology. In fact, where the consonant cluster comprised this sound, there was a digression from the behavior pattern exhibited in other blends as the participants inserted the /b/ for /p/ (in 'cupboard') to articulate the token as /kʌb.bɜːd/. High frequency in incorrect identification were seen in the diphthongs (/eɪ/, /əʊ/) ranging between 16-23. This can be explained by the fact that Arabic has lesser number of these (8) compared to English which has as many as 20 vowel and diphthong sounds. The /ŋ/ sound also saw high casualty as its application is rather restricted in Arabic (to before /k/). Lastly, problems with stress timed patterns were reported to be the highest as Arabic has a different prosodic repertoire than English since reducing the typically weak vowels and emphasizing the content words is quite difficult for the native Arabic speaker to master as their mother tongue is characterized by an energetic, staccato delivery that is typical of Arabic.

However, in the post intervention stage, a great deal of these pronunciation and prosody issues were tackled (statistically significant values obtained) due to the intervention as was disclosed by the participants in the semi-structured interviews. These interactions were also focused around the research questions in addition to eliciting information on whether and why pronunciation and prosody are learning targets. All 24 participants averred that they aspired to obtain native-like speaking ability as far as the pronunciation and prosody were concerned though their reasons for this varied. A large section (n=18) conceded that the conventional pronunciation training was not effective at all as they were wary of speaking in the class for fear of peers judging them and the teacher too could not devote much time to the activity. The third finding from the interviews was that using ASR for pronunciation and prosody training was very effective as the participants could practice even on their mobile phones in the small windows of time that they had such as while traveling, found it so much fun, and were not afraid of peer judgement. Overall, the intervention was reported to be extremely useful by 17 of the participants, easy to use by 21 of the participants, and preferable to conventional methods by all 24 participants.

Finally, the IELTS scores for the speaking component of the group were compared to identify changes in performance. While the pre intervention speaking band score of the group was between 4.0-5.0 (B1 Intermediate), in the post intervention stage it was between 5.5 -7.0, placing them in the B2 Upper Intermediate to C1 Advanced group.

Conclusion

The results of this study indicate that the development of automatic speech recognition (ASR) technology has made it possible to diagnose learners' pronunciation issues using ASR-based pronunciation assessment. Xiao and Park (2021) investigated the options for pronunciation practice with ASR-based pronunciation training and positively concluded ASR technology's efficacy in identifying English pronunciation problems and the attitudes of educators and students toward its use as a tool for learning and pronunciation evaluation. Additionally, it was discovered that students' diverse spoken English understanding requirements were met with the aided ASR technology. Secondly, participants in this study reported a preference for the ASR based pronunciation tasks rather than the conventional method as being more engaging, much like Yang (2024) which examined how AI affects learning motivation in EFL (English as a Foreign Language) classes and concluded that there existed a strong relationship between student motivation and AI use. Other prior research too focuses on how AI-assisted language learning affects EFL learners' self-regulated learning, English learning outcomes, and L2 motivation as discussed in the review of literature. It has been demonstrated that AI-mediated training improves learning outcomes, self-regulated

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

learning, and motivation for learning English. Furthermore, interview results point to learners' favorable opinions of AI in language learning.

Recommendations

Based on the findings, this study recommends greater intervention and integration of AI tools like Speechnotes int eh Saudi EFL classrooms. Teachers must assess learner needs regularly (as they are dynamic) and be updated with the latest tech and AI options that can be used as educational aids. Action research while applying new pedagogies should be encouraged to accurately gauge the relationship between teaching and learning needs fulfilment. The AI platform is expanding like never before, and this has opened new vistas in educational technology as well. Institutions and teachers should be proactive in staying at par if not ahead of the learners in ensuring that education remains relevant and not redundant.

Limitations

The researchers engaged an all-male sample in this study. This can be a limitation in applying findings to other contexts. Only one AI based ASR tool has been tested here, future studies may experiment with other options, larger and varied learner groups to reach generalizable conclusions.

References

Aladini, A. (2023). Artificial intelligence applications' impact on improving EFL University Academic writing skills and their logical thinking. Educational Sciences, 31(4). https://doi.org/10.21608/ssj.2023.320166.

Al-Ahdal, A.A.M.H. (2020). Overcoming pronunciation hurdles in EFL settings: An evaluation of podcasts as a learning tool at Qassim University Saudi Arabia. Asian EFL Journal Research Articles, 27, 86-101

Alhalangy, A. G. I., & AbdAlgane, M. (2023). Exploring the impact of AI on the EFL context: A case study of Saudi universities. Journal of Intercultural Communication, 23(2), 41–49. https://doi.org/10.36923/jicc.v23i2.125

Chen, L. (2023). Special issue on automatic speech recognition. Applied Sciences, 13(9), 5389. https://doi.org/10.3390/app13095389

Couper, G. (2017). Teacher cognition of pronunciation teaching: Teachers' concerns and issues. TESOL Quarterly, 51, 820–843.

Crowther, D., Trofimovich, P., Isaacs, T., & Saito, K. (2015). Does a speaking task affect second language comprehensibility? The Modern Language Journal, 99, 80–95.

Evers, K., & Chen, S. (2021). Effects of automatic speech recognition software on pronunciation for adults with different learning styles. Journal of Educational Computing Research, 59(4), 669-685. https://doi.org/10.1177/0735633120972011

Gilakjani, A. P., & Ahmadi, M. R. (2011). Why is pronunciation so difficult to learn? English Language Teaching, 4(3), 74–83.

Jiang, M., Jong, M., Lau, W., Chai, C., Wu, N. (2023). Effects of automatic speech recognition technology on EFL learners' willingness to communicate and interactional features. Educational Technology & Society, 26, 37-52.

Isaac, J. C. (2019). Increasing students' levels of fluency by reinforcing sounds and syllabic recognition in their first language of Spanish. (Master thesis). Northwestern College

Lingjie, Y., & Xiaojuan, L. (2025). The effect of artificial intelligence tools on EFL learners' engagement, enjoyment, and motivation. Computers in Human Behavior, 162,108474. https://doi.org/10.1016/j.chb.2024.108474 .

Liu, J., Liu, X., & Yang, C. (2022). A study of college students' perceptions of utilizing automatic speech recognition technology to assist English oral proficiency. Front. Psychol. 13, 1049139. https://doi.org/10.3389/fpsyg.2022.1049139

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. Cogent Education, 10(2). https://doi.org/10.1080/2331186X.2023.2236469

Ngo, T., Chen, H., & Lai, K. (2023). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. ReCALL. 36. 1-18. 10.1017/S0958344023000113.

Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1210187.

Wenqi, X., & Moonyoung, P. (2021). Using automatic speech recognition to facilitate English pronunciation assessment and learning in an EFL context: pronunciation error diagnosis and pedagogical implications. International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 11(3),74–91. https://doi.org/10.4018/IJCALLT.2021070105

Xiao, W., &, & Park, M. (2021). Using automatic speech recognition to facilitate English pronunciation assessment and learning in an EFL context. International Journal of Computer-Assisted Language Learning and Teaching, (IJCALLT), 11(3), 74-91.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 979 – 987
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6384

Yang, T. (2024). Impact of artificial intelligence software on English Learning motivation and achievement. SHS Web of Conferences (Vol. 193, p. 02011). EDP Sciences.3. https://doi.org/10.1051/shsconf/202419302011