Speech Recognition and Pronunciation Apps and EFL Learners: A Study of Efficacy in Developing Speaking Skills

Ali Hussein Salim Babekir[1]

Abstract

Technology has entered EFL classrooms in a big and irreversible manner. The latest addition to the foreign language classroom are the speech recognition and pronunciation apps which can be honed as valuable language learning tools. The current study examines the efficacy of the free trial pack of FluentU, an automatic speech recognition app in enhancing the speaking skills of 60 sophomore Saudi male learners at Qassim University. Using a pre-post intervention approach, the study exposed the participants to two weeks of FluentU use and compared their perceptions about it in the pre and post stages. Teachers' feedback about any changes in learner behavior were also sought. Result showed positive perceptions of the learners towards the efficacy of the app in motivating them to acquire English speaking skills, while the teacher reported enhanced confidence and willingness in the learners to speak in English. However, no perceptible enhancement in grammar, content, syntax, and vocabulary was reported. The study concludes with recommending the use of Automatic Speech Recognition (ASR) apps in the speaking classes to ensure greater learner motivation and engagement and the advantage of immediate feedback.

Keywords: *EFL learners, FluentU, pronunciation Apps, speech recognition software.*

Introduction

An enriched and positive learning environment is a prerequisite for effective learning to take place. Tileston (2004) noted that the classroom environment has to be conducive to learning and motivation to learn is a key factor. Motivation clusters around three things: Autonomy, mastery, and purpose (Pink, 2009). Each one of these components is the subject matter of deep exploration but the scope of the current study concerns itself the most with autonomy. Autonomy refers to self-direction which can be linked to tasks wherein learners are free to independently demonstrate their learning, via opportunities that are creative (Lan, 2022). Moreover, learners' progress over time rather than the conventional time-bound 'completion of syllabus' will nurture autonomous learning principles that employ techniques that ensure optimum learning. Technology available in mobiles phones has given learners autonomy like never before, not only because it allows them to learn anywhere and anytime (Al-Ahdal & Shariq, 2019), but also, because motivation to learn is not inhibited by peer censure and other detrimental factors in this mode (Ali & Bin-Hady, 2019). On the contrary, motivation is spurred by opportunities to learn with their favorite playmate (Alahdal & Al Ahdal, 2019). From the early days of online dictionaries that also offered audio pronunciation, language learning apps also known as automatic speech recognition (ASR) apps have readily addressed the needs of

[1] Department of English and Translation, College of Arts and Sciences in Methnab, Qassim University, Saudi Arabia,
a.babekir@qu.edu.sa

an increasingly cosmopolitan global workplace where language is no longer seen as a barrier in communication.

From non-English teachers communicating with an international learner group in their classes to EFL learners interacting with their native English-speaking teachers, language apps have actually shrunken linguistic boundaries (Islam & Stapa, 2021). What makes them even more attractive is the artificial intelligence (AI) powered ability to cater to a large number of foreign language users with equally diverse accents and pronunciations, enabling them to steer their language use with ease into the standardized forms of English. This is in direct contrast to the conventional EFL classroom where the development of speaking is seen as a form of skill-learning to facilitate which certain items to be 'learnt' are isolated and presented separately to the learners (Al-Ahdal, 2020). One example of such 'items' is vocabulary which is to be learnt not in-context as one learns the mother tongue, but out of context as a foreign language. Thereafter the learners are expected to practice them and eventually, master them, again, in isolation as opportunities for contextual use are missing or very few.

ASR apps have today come a long way with the learning process gamified to engage the language learners in a unique fun way as they acquire and improve their language (MacWhinney, 2017). Duolingo is one of the very popular language apps with wide users base given the numerous languages that it deals with and also because it caters to different levels of language learners (Nushi & Eqbali, 2017). Using a combination of tasks anchored to the four language skills (listening, speaking, reading, writing), it targets the global language proficiency achievement of its users. An immersive and highly interactive language learning environment is ensured by Rosetta Stone, which works on a range of visual and audio clues to help users develop their language skills. Duolingo uses the natural language acquisition approach and also has a variety of languages in its database (Shortt et al., 2023). Development of practical conversation skills are the target of Babel which caters to 14 different languages with short lessons centred around everyday situations.

Some other apps have more than one use, such as Memrise, which also boosts the memory in addition to language. The technique is to use spaced repetition and mnemonic devices to create long term memory of language use (Seibert Hanson & Brown, 2020). The feature that allows its users to tailor their own courses makes this app unique. Greater audio-based learning is the USP of Pimsleur which has comprehensive language lessons specially designed to enhance users' pronunciation and listening skills in many languages. A valuable native speaker experience in language learning is ensured by Tandem which allows users to chat, engage in video calls, or text messages with native users of the language which gives them authentic language exposure in addition to being a great platform for learning of the target cultures and real-lives of the native speakers (Rahimi & Fathi, 2022). Busuu offers courses in 12 languages and follows the CEFR (Common European Framework of Reference for Languages) standards with interactive exercises, vocabulary practice, and speaking exercises with native speakers. This app, thus, nurtures authentic language use opportunities (Abreu Avellaneda, 2022). HelloTalk is also a similar app as it connects users with native speakers through a medium of their choice such as audio, video, or text and also aids them with the unique translation and correction features. Vocabulary acquisition through visual and interactive exercises is the mainstay of Drops which helps learners pick new words and phrases related to certain themes. Lastly, FluentU uses the very interesting film and news articles content with interactive subtitles and language comprehension quizzes to help the users. This study explores the impact of free FluentU version of developing EFL students at Qassim University. Most of the ASR apps offer free as well as paid options. This study used the 14 days free trial option of the FluentU app to learn English. The research questions that the study aimed to answer are:

Research question

1.   What is the perception of Saudi EFL learners about the efficacy of FluentU in motivating them to learn speaking skills?

2.   How do the Saudi EFL teachers evaluate the role of an ASR app in learners' speaking output?

Literature review

The outcomes of multiple research studies on how well ASR supports ESL/EFL speaking have been conflicting. For instance, Gorjian et al. (2013) and Neri et al. (2008) reached entirely distinct outcomes when comparing the effects of an ASR programme with a conventional technique (such as manual and teacher-fronted training) on ESL/EFL student pronunciation. While Gorjian et al. (2013) found that after 10 training sessions, those who had been learning pronunciation using the conventional method significantly underperformed, those who had been using the ASR program (such as Praat software), Neri et al. (2008) found that there was little variation between the two groups' language acquisition effectiveness after just four weeks of instruction. Along with the inconsistent findings seen throughout investigations, the impact of ASR is different in different studies. In addition to the conflicting results observed between investigations, the impact of ASR varied even within investigations.

Evers and Chen (2022) looked into how an ASR transcription software (i.e. Speech notes) performed on a variety of pronunciation characteristics (such as accented Ness, understanding, and spontaneous speech). The accented-ness between the two sets of participants did not significantly differ, but understanding and impulsive speech did differ significantly, according to the researchers. In a related study, Evers and Chen (2021) discovered differences in the afterward magnitude of effects of the experimental groups, spoken and visual style learners honing their vocalizations with Speech notes, in their performance on reading and spontaneous speaking tasks. The various pronunciation measurement techniques, ASR programs, and treatment lengths applied in investigations.

The conflicting findings seen in the primary research appear to be a result of the different ASR programmes, speech measurement techniques, and treatment lengths utilised in the investigations. An explanation of the causes of the contradictory results and the overall effect size may be provided by a meta-analysis (Lipsey & Wilson, 2001).This meta-analytic study investigated the general impact of automatic speech recognition (ASR) on ESL/EFL student speaking competency. Data from 15 studies with 38 effect sizes identified between 2008 and 2021 were subjected to a meta-analysis. The meta-analysis's findings revealed that the overall magnitude of ASR's effect is moderate (g = 0.69).

Students of English as a foreign language (EFL) frequently struggle with speaking. Speaking as frequently as you can with native speakers of the target speech is one technique to advance your foreign language speaking abilities (Al-Tamimi et al., 2020). However, this is not always simple in areas with a small population of native speakers of the target language, such as Saudi Arabia. Another significant obstacle that EFL students must overcome is language anxiety. Students often find it difficult to speak up in class because they are frightened of speaking incorrectly (Al Nakhalah, 2016). Students' anxiety may increase if teachers correct their mistakes all the time. The challenge is how to give learners the right assistance with little involvement from instructors is a major difficulty in EFL learning. Chang (2012) details the design of a study on the application of text-to-speech (TTS) synthesis in learning English pronunciation with the goals of reducing language anxiety in EFL students and enabling them to acquire pronunciation with little teacher assistance. When performing exercises that concentrate on the supra-segmental level, the study used TTS synthesis as a pronunciation model. While segmental

level coverage was not intended for the study, TTS synthesis enabled learners to acquire unknown word pronunciation by listening to the synthesised speech.

An investigation was conducted by Umezawa et al. in 2023 on the applicability of ASR in English speech practice. They also focussed on analyzing ASR's impact on EFL learners learning. The usefulness of the strategy was assessed by investigators who created and implemented it as a way to monitor the daily learning outcomes of self-study pronunciation practice. A native language speaker precisely evaluated the spelling and grammar on the initial day of independent research and the final day (seven days afterward) to see whether it was appropriately articulated. Additionally, the authors employed the speech recognition feature each day of the seven-day experiment to see if the pronunciation was picked up. The study came to the conclusion that the iOS speech recognition feature is superior than others and recognizes pronunciation even when it is mispronounced, making it unfit for self-learning pronunciation. It should be mentioned that the outcomes of this study will probably rely on the version of the speech recognition function used. So outcomes may change in future investigations if a different version of the speech recognition function is applied.

Sun (2023) used an explanatory sequential design to investigate the effects of using ASR with peer correction on the development of second language (L2) spoken language and language proficiency among EFL students. The goal was to determine whether this method, as opposed to conventional teacher-led feedback and teaching, may be an effective tool for improving L2 pronunciation and speaking abilities. The statistical examination of the data revealed that the EG fared better than the CG in all tests of L2 speech pronunciation, namely understanding and accentedness. Additionally, compared to the CG, the EG showed considerably improved global speaking ability. The interviews' qualitative analysis found that the vast majority of participants in EG found it effective in improving their L2 grammar and statements, it also enhanced their speaking skills. The findings of this investigation imply that using ASR equipment in conjunction with peer reinforcement can be a powerful strategy for improving L2 articulation and speaking abilities among English as a Foreign Language students. The EG's improved scores on speech and speaking tests when compared to the CG show the promise of utilizing ASR equipment in language learning settings. Furthermore, the positive comments from the EG participants highlight the benefits of utilizing ASR technology as a supplementary tool in language acquisition environments.

For Arab learners, one difficulty is the change in sound from their mother tongue (Arabic) to the target language (English). Sidgi and Shaari (2017) determined whether employing ASR EyeSpeak software for automatic speech recognition may help Iraqi English learners pronounce words more clearly in this study at the Department of English at Al-Turath University College in Baghdad, Iraq, with a pretest-posttest design over the course of one month. According to the findings, utilizing EyeSpeak software significantly improved students' speaking of English as evidenced by the test results they received after employing the program.

Despite being covered in many EFL instruction programmes, pronunciation teaching is frequently lacking (Bin-Hady & Hazaea, 2021). Almusharraf's (2022) research on the developing subject of reasoning aimed to address the knowledge gap and examine Saudi EFL students' self-assurance, behaviours that are and practise in mastering pronunciation. In order to do this, the study enrolled 336 Saudi EFL students focusing in the English language at a Saudi Arabian university as a convenient sample. Findings revealed that the research participants' pronunciation trust was stronger than neutral and that they had a very favourable opinion of pronunciation that sounded like a native. Intriguingly, this study found no statistically significant distinction in pronunciation confidence between individuals who had completed a phonetic(s) course and those who had not. This study therefore exhorts teachers to be mindful of their students' pronunciation needs, provide

pertinent information, and provide additional opportunity for practise of diverse techniques.

The development of ASR software has made it possible for educators to identify pronunciation issues in learners using ASR-based pronunciation assessment. While there are more opportunities for pronunciation practise with ASR-based pronunciation training. Xiao and Park (2021) examined the efficiency of ASR technology in identifying spoken English problems as well as the mindsets of educators and learners towards using ASR technology for both evaluation and learning of pronunciation. Five Chinese EFL students took part in telling students to read examinations, including tests with human and ASR scoring. To assess the degree of overlap, the two tests' diagnoses of various spelling mistakes were evaluated. The results showed that there were segmental level discrepancies between subjective rating and automated rating. Additionally, it was discovered that the ASR technique satisfied the various pronunciation learning objectives of learners. The study's implications will offer information about how to analyze and learn English pronunciation using ASR technology.

Cámara-Arenas et al. (2023) set out to find the difference between automatic pronunciation and ASR for L2, EFL learners. The purpose of the investigation was to determine whether pronouncing words through ASR is more effective than traditional instruction and to determine which teaching method—individual work, pair work, or group work—is most effective for teaching pronunciation through ASR. Four groups— three experimental and one control—were created by chance. The ASR approach was employed to instruct the three groups participating in the experiment. They practiced pronunciation with the Tell Me More Performance English programme. The same teacher used routine instruction to teach the control group. The findings demonstrated that there were statistically significant differences between the mean scores of the experimental (ASR) and control (regular instruction) groups in favour of the latter. Additionally, the results showed that there was a statistically significant difference in favour of the individual work approach in the mean scores between pair-work, group-work, and individual work. The results also showed that there was no significant relationship between the experimental group students' performance on the pronunciation post-test and the level (word, phrase, and real-life dialogue). The dialogue level, however, was where pupils did better than the word and sentence levels. The study made several recommendations for teachers in this area.

Methods

Research design

This study adopted a pre-post-survey quantitative approach to collect perceptional data in three phases. The study took place in the first semester of the academic year 2022-2023. The study explored the Saudi EFL learners' perceptions of the efficacy of an ASR app, FluentU, in motivating them to acquire speaking skills in English. The researcher himself has had a few years exposure to the EFL situation at a Saudi university, and during his tenure he has observed that lack of motivation to learn, learning lethargy, and the perception that English speaking is a massive challenge for them, has been a major obstacle in the learning process.

Participants

The study selected a convenience sample of 60 male EFL learners enrolled in two sophomore classes at Qassim University, Saudi Arabia. In addition to this, 2 teachers. The students are enrolled in the second year at the English department. They were asked for their contest to participate in the study. The researcher told them the purpose of the study and assured them they could withdraw at any time they wished. The age declaration on

the FluentU app was made in the bracket between 18-24 years as the participants' median age was 21.2 years.

Procedures and data collection

Phase 1

Two EFL teachers were also interviewed before the intervention to answer the second research question. This was the first phase. Before the planned intervention, it was verified that all the learners had access to the internet to seamlessly use the ASR app in the study. After the ethical issues were spelt out, the researcher obtained formal consent of the participants and counselled them on the purpose of the study. This began the second phase of the study, the first part of which was to collect data on students' perceptions of motivation to learn English speaking skills. For this purpose, the English version Gardner's (2004) attitude/ motivation test battery (AMTB) was used in a slightly abridged form (only 10 of the original 12 items included here). The questionnaire determines users' feelings about their motivation in speaking English and other similar constructs. The same instrument was again used after the intervention to gauge if there were any changes in participants' motivation to learn English speaking skills. These results are summarized in Table 1. In addition the teachers were also interviewed again to see if there were any changes in the level of participation of the learners in speaking classes and how the ASR app might have impacted the same.

Phase 2

The second phase of the study was the actual engagement with the ASR app for two weeks in the first semester of academic year 2023. This was a phase of autonomous interaction with the app and the participants were free to take the lessons anywhere and anytime and also counselled not to miss any session in their own good. As the researcher is also a faculty member at the same department where the learners are enrolled, he formed a WhatsApp group to provide support to the learners during the intervention period as it was an out-of-class activity. He encouraged them to post regular updates on their experience with the app. However, the data gathered in this were not for use in the study.

The FluentU app is highly interactive. It takes the user through a few non-challenging steps before they begin the lessons. The first amongst these is the assurance of a customized experience, onto selection of the user's current proficiency level. The researcher had informed the participants that their official proficiency level is 'lower intermediate' and they should declare the same on the app. The daily learning goal was set at 30 minutes. From here on, the participants were on their own though they were still engaging with their regular EFL class in this period.

Results

RQ1: What is the perception of Saudi EFL learners about the efficacy of FluentU in motivating them to learn speaking skills?

This study adopted a pre-post intervention survey to collect data. Accordingly, Gardner's (2004) attitude/ motivation test battery (AMTB) was deemed as the right instrument in gauging learners' pre and post intervention perceptions of motivation to acquire English speaking skills.

Table 1. Saudi EFL students' perception in the pre and post rounds

		1=Weak, unfavorable, very little	2	3	4= Strong, favorable, very high
My motivation to learn English in order to communicate with English speaking people is	Pre	41	5	7	6
	Post	7	7	19	27
My interest in foreign languages is	Pre	38	15	3	4
	Post	2	5	36	17
My desire to learn English is	Pre	43	8	9	0
	Post	0	13	34	13
My attitude toward learning English is	Pre	52	2	4	2
	Post	9	6	32	13
My attitude toward my English teacher is	Pre	39	11	8	2
	Post	8	15	20	17
My motivation to learn English for practical purposes is	Pre	26	10	13	11
	Post	2	2	25	31
I worry about speaking English outside of class is	Pre	53	2	1	4
	Post	11	5	23	21
My attitude toward my English course is	Pre	50	8	0	2
	Post	3	1	28	28
My motivation to learn English is	Pre	31	17	5	7
	Post	6	11	19	24
I worry about speaking in my English class is	Pre	55	4	0	1
	Post	13	12	22	13

Table 1 indicates negative perceptions of the EFL learners towards motivation to acquire English speaking skills. In the pre-intervention stage, bulk of responses are concentrated towards low motivation and hence, efficacy scores. For example for questionnaire item, My attitude toward learning English is…the pre and post intervention responses are 52 and 9 respectively indicating the poor perception of 52 participants for the item in the pre stage as against only 9 in the post stage. Similarly, 43 participants exhibited negative perceptions for My desire to learn English is…in the pre intervention stage as against 0 in the post intervention stage. This trend is unanimous for all the items of the survey with clear support for the ASR app-based speaking skills acquisition. In other words, to answer the first research question, the perceptions of the Saudi EFL learners about the efficacy of FluentU ASR app to improve motivation to learn speaking skills is positive as the highly tipped responses in the post intervention data indicate.

RQ2: How do the Saudi EFL teachers evaluate the role of an ASR app in learners' speaking output?

The teachers were asked to speak their opinion about the speaking participation of the learners in the study, and whether the ASR app FluentU based intervention had proven useful to them. The following themes were identified in their responses:

1. Individual feedback is difficult to give in the EFL classes due to the poor speaking skills of the learners and also because the teacher-learner ratio does not favor such an endeavor. With this aspect tackled by the ASR based approach, learners were less inhibited about making mistakes and more natural in their language use.

2. The speaking class atmosphere was earlier one rife with tension and fear as most of the learners were concerned of what their classmates will say if they spoke out with their limited proficiency. After the intervention, however, learners exhibited more

confidence especially as many of them came armed with the hedge expressions ('well', 'um' etc.) which are part of the authentic native speaker's repertoire. These hedges enabled them to process information in the time gap and also elicited appreciation from their peers for such native like usage.

3.   Learners were more confident in speaking, though the intervention had not actually tackled their grammatical, syntactical, or vocabulary deficiencies, it had motivated them enough to look beyond what they did not know and focus on what they did know. In addition, unlike the pre intervention stage when learners were always keen not to catch the teacher's attention, in the post intervention stage, they were rather keen to show off their newly acquired confidence to their peers.

Discussion

Results indicated that Saudi EFL learners had negative perception towards speaking before the intervention. The majority of them reported weak motivation at the beginning of the course. Their perceptions however, improved due to the intervention they received. This finding confirmed the importance of ASR in developing students' speaking skills. This findings agreed with Zou et al. (2023) who looked into the possibility that the various automatic feedback offered by AI speech evaluation algorithms could help English as a foreign language (EFL) students improve their speaking skills. That was a study with forty Chinese EFL students and both quantitative and qualitative data were gathered. The results showed that most participants believed that the feedback from the AI speaking assessment programme helped them improve their speaking skills. A similar sentiment has been voiced by the EFL teachers in the current study and they also opined that this is definitely a big pro in favor of the ASR app based approach. Additionally, the results in Zou et al. (2023) revealed a considerable improvement in the mean speaking skill ratings between the pre- and post-tests of the participants. Zou et al. (2023) also examined speaking anxiety, self-perceived pronunciation competence, and speaking ability in foreign language education institutions. Ur Also confirmed that (1996) which noted that higher motivation leads to better speaking ability in foreign language learning situations.

Teachers reported that ASR apps participated in developing Saudi EFL students' speaking skills. This finding matches with Sidgi and Shaari (2017) who found that utilizing EyeSpeak software significantly improved Iraqi EFL learners speaking of English as evidenced by the test results they received after employing the program.

Conclusions

The findings in the current study have established that ASR app based approach does motivate EFL learners to acquire speaking skills and boosts their confidence in learning. This finding resonates. Advances in artificial intelligence (AI) technology have enhanced language learners' usage of computerized voice assessment devices for speaking practice. EFL students. If learners can be motivated to learn, half the classroom challenge is overcome. The findings in this study showed that ASR based learning is highly enjoyed by the EFL learners and achieves motivational goals. Technology is used in teaching and learning through the employment of programmes like computer-assisted language learning (CALL). ASR is a promising new technology that aids in language learning. It's crucial to incorporate such technology when providing pronunciation instruction in a learning environment, especially to ensure that learners are pronouncing words correctly.

Recommendations

Though the speech recognition and pronunciation app in this study has proven to be useful in enhancing learners' speaking skills, there are certain challenges with their use, especially with a diverse language group as well as with very poor or very proficient users. Human exchanges are not unilateral as many features other than the lexis, such as, body language, tone, speed, intonation etc. play an equally significant part in the response mechanism. These are difficult to be captured by language apps, moreover, a small vocabulary size will have restricted usage domain and a very large vocabulary will affect the speed of the app. In short, human to human communication is multimodal, which is so far not available in speech recognition and pronunciation apps. Another dimension of this is that man-machine communication is not as natural as its human counterpart as colloquial and irregular use of words mark the latter. With low proficiency users as in the case of EFL learners in Saudi Arabia, this can lead to less than desirable attainment of learning objectives. In addition, morpheme, phonic, and word boundary ambiguities in speech of foreign language learners can mislead the AI powered apps and ultimately, lead to poor learning.

Limitations

This study only tested the efficacy of Automatic Speech Recognition and Pronunciation apps in motivating the Saudi EFL learners to acquire English speaking skills. However, as much as pronunciation, content and that too, correct content and presentation are also important learning objectives. Due to the limited scope of the study, the effects of ASR on these components could not be gauged in this study. Moreover, the participants were all males and whether the results arrived at here apply to other genders is a question for future studies to answer, it remains a limitation of this study.

References

Abreu Avellaneda, J. N. (2022). Enhancing English vocabulary acquisition of Colombian high school students through vocabulary learning techniques, mobile apps, and web 2.0 technologies. (Unpublished Master thesis). Universidad Nacional Abierta y a Distancia

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.

Al-Ahdal, A. A. M. H., & Shariq, M. (2019). MALL: Resorting to Mobiles in the EFL Classroom. The Journal of Social Sciences Research, 90-96.

Alahdal, A., & Al Ahdal, A. A. M. H. (2019). Effectiveness of collaborative learning as a strategy in the teaching of EFL. Opción: Revista de Ciencias Humanasy Sociales, 20, 1026-1043.

Ali, J.K.M., & Bin-Hady, W. R. A. (2019). A study of EFL students' attitudes, motivation and anxiety towards WhatsApp as a language learning tool. Arab World English Journal (AWEJ) Special Issue on CALL (5). 289-298 https://dx.doi.org/10.24093/awej/call5.19

Almusharraf, A. (2022). EFL learners' confidence, attitudes, and practice towards learning pronunciation. International Journal of Applied Linguistics, 32(1), 126-141.Almusharraf, Asma. (2021). EFL learners' confidence, attitudes, and practice towards learning pronunciation. International Journal of Applied Linguistics. 32. https://doi.org/10.1111/ijal.12408

Almusharraf, Asma. (2021). EFL learners' confidence, attitudes, and practice towards learning pronunciation. International Journal of Applied Linguistics. 32. 10.1111/ijal.12408.

Al-Tamimi, N. O., Abudllah, N., & Bin-Hady, W. R. A. (2020). Teaching speaking skill to EFL college students through task-based approach: problems and improvement. British Journal of English Linguistics, 8(2), 113-130.

Ambra Neri, Ornella Mich, Matteo Gerosa & Diego Giuliani (2008) The effectiveness of computer assisted pronunciation training for foreign language learning by children, Computer Assisted Language Learning, 21:5, 393-408, DOI: 10.1080/09588220802447651

Bin-Hady, W. R. A., & Hazaea, A. N. (2021). EFL students' achievement and attitudes towards flipped pronunciation class: correlational study. PSU Research Review, 6(3), 175-189. https://doi.org/10.1108/PRR-09-2020-0029

Cámara-Arenas, E., Tejedor-García, C., Tomas-Vázquez, C. J., & Escudero-Mancebo, D. (2023). Automatic pronunciation assessment vs. automatic speech recognition: A study of conflicting conditions for L2-English. Language Learning & Technology, 27(1), 1–19. https://hdl.handle.net/10125/73512

Chang, A.H.Y. & Kwan, Alvin. (2012). EFL students' perception of the use of text-to-speech synthesis in pronunciation learning. Proceedings of the European Conference on e-Government, ECEG. 524-527.

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

Evers, K. & Chen, S. (2022) Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults. Computer Assisted Language Learning, 35(8), 1869–1889. https://doi.org/10.1080/09588221.2020. 1839504

Gardner, R. C. (2004). Attitude/motivation test battery: International AMTB research project. Canada: The University of Western Ontario.

Gorjian, B., Hayati, A. & Pourkhoni, P. (2013) Using Praat software in teaching prosodic features to EFL learners. Procedia - Social and Behavioral Sciences, 84: 34–40. https://doi.org/10.1016/j.sbspro.2013.06.505

Gorjian, B., Hayati, A. & Pourkhoni, P. (2013) Using Praat software in teaching prosodic features to EFL learners. Procedia - Social and Behavioral Sciences, 84: 34–40. https://doi.org/10.1016/j.sbspro.2013.06.505

Islam, M. S., & Stapa, M. B. (2021). Students' low proficiency in spoken English in private universities in Bangladesh: reasons and remedies. Language Testing in Asia, 11, 1-31.

Lan, Y. (2022). The role of teachers' grit and motivation in self-directed professional development. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.922693

Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. SAGE Publications, Inc.

MacWhinney, B. (2017). A shared platform for studying second language acquisition. Language Learning, 67(S1), 254-275.

Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. Computer Assisted Language Learning, 21(5), 393-408. https://doi/Ambra

Nushi, M., & Eqbali, M. H. (2017). Duolingo: A Mobile Application to Assist Second Language Learning. Teaching English with Technology, 17(1), 89-98.

Pink, D. (2009). The puzzle of motivation. TED Global.

Rahimi, M., & Fathi, J. (2022). Employing e-tandem language learning method to enhance speaking skills and willingness to communicate: the case of EFL learners. Computer Assisted Language Learning, 1-37. https://doi.org/10.1080/09588221.2022.2064512

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2023). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. Computer Assisted Language Learning, 36(3), 517-554.

Sidgi, L. F. S., & Shaari, A. J. (2017). The Effect of Automatic Speech Recognition Eyespeak Software on Iraqi Students' English Pronunciation: A Pilot Study. Advances in Language and Literary Studies, 8(2), 48-54Sidig, Lina & Shaari, Ahmad. (2017). The Effect of Automatic Speech Recognition EyeSpeak Software on Iraqi Students' English Pronunciation: A Pilot

Study. Advances in Language and Literary Studies. 8. 48. http://doi.org/10.7575/aiac.alls. v.8n.2p.48

Sun, W. Sun, Weina. (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. http://doi.org/10.3389/fpsyg.2023.1210187

Tileston D. W. (2004). What every teacher should know about classroom management and discipline. Thousand Oaks, CA: Corwin Press.

Umezawa, K., Nakazawa, M., Nakano, M., & Hirasawa, S. (2023). Utilizing automatic speech recognition for English pronunciation practice and analyzing its impact. IEEE

Ur, P. (1996). A course in language teaching: Practice and theory.

Cambridge: Cambridge University Press.

Xiao, W., & Park, M. (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. http://doi.org/10.4018/IJCALLT.2021070105

Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An Investigation Into Artificial Intelligence Speech Evaluation Programs With Automatic Feedback for Developing EFL Learners' Speaking Skills. SAGE Open, 13(3). http://doi.org/21582440231193818.