Evaluating AI Applications for Enhancing Listening Comprehension among EFL students in English language in Real-World Scenarios

Wafa' A. Hazaymeh[1], Turky Alshaikhi[2], Mohammad Osman Abdul Wahab[3], & Mohamad A. Khasawneh[4]

[1] College of Education, Al-Ain University, United Arab Emirates

[2] College of Education and Arts, Institute of English Language Teaching, University of Tabuk, Saudi Arabia

[3] Department of English, Faculty of Languages and Translation, King Khalid University, Abha, KSA

[4] Special Education Department, King Khalid University, Saudi Arabia

Correspondence: Mohamad A. Khasawneh, Special Education Department, King Khalid University, Saudi Arabia. E-mail: drkhasawneh@kku.edu.sa

Abstract

This paper assesses the extent to which Artificial Intelligence technologies improve listening comprehension in English among foreign language learners in Saudi Arabia. In this study, both the quantitative assessment of the changes in the listening comprehension levels of the participants as well as qualitative feedback that focused on the use of AI applications by the learners were collected. It was established that listening comprehension improved, particularly in the contextualized area, and that learner engagement and motivation also improved. At the same time, difficulties connected with accent recognition, technical problems, and the lack of depth in the comments were also discussed. This is the main lesson from this study about AI, in applications, much as the tools are useful and provide great benefits, they should be used as a supplement rather than a substitute to traditional methods. Further works should be aimed at enhancing the AI in applications to various linguistic contexts and enhancing the feedback.

Keywords: AI tools, EFL (English as a Foreign language), listening comprehension, English language learning, Saudi Arabia, educational technology

1. Introduction

Comprehension of spoken materials is still a major problem area in learning English as a foreign language. As more and more importance of English language factor in Saudi Arabia, especially in the context of academic achievement and career opportunities, the mastery of spoken language in environments that imitate real-life situations should be considered an educational objective. Listening comprehension is far from simply sounding or words; it calls for a connection of information about the setting, speech, accents, rate, and idioms (Itzchakov et al., 2023). However, one look at it, those aspects are crucial and foreign language learners, with those in KSA inclusive, find it hard to get it. This challenge is most apparent in real-life situations of language use, which often are less formal, less bi-directional, and less controlled than the input students are exposed to in classroom learning (Leonardo, 1998). Therefore, students struggle to generalize what they learn in class about listening to the realistic and spontaneous communication that they are likely to encounter in daily life (Otte, 2006).

In recent research, the said issues have been explained, and technology has been deemed capable of solving all these problems, especially AI in language education. AI solutions like speech recognition systems or adaptive learning technologies have been tested for the general benefit of custom-made learning materials and for mimicking real-life listening scenarios (Chen et al., 2022). These tools can adapt to the range and level of the input language, inform the learner immediately about their result, and give multiple practices that sometimes classroom learning misses out on (Huang et al., 2024). Conversely, the use of AI tools in improving listening skills as a subprocess of listening remains one that has received little attention the most especially when it comes to endeavours in real-life situations. Although research has shown that, through an AI system, listening practice can be rather effective and efficient in rather controlled conditions (Itzchakov et al., 2023), what is less clear is how helpful AI can be in recreating the particular situations that are so important in the process of developing listening skills.

In Saudi Arabia, more and more, AI-based language learning technologies are introduced in an educational context, however, the applicational outcome of such tools on listening comprehension is underemphasized in the literature. Almaleki (2020) presented the first real discussion of how AI affects Saudi education and found that these tools can help attend to language advancement. However, they did not investigate listening comprehension, which is an area of concern among learners, especially those learning a second language. Besides, the majority of students in Saudi Arabia are culturally and linguistically diverse and experiencing challenges such as a lack of practice in English spoken by native speakers, the use of AI in language learning becomes even more important. Their use of AI tools that introduce

the learners to actual spoken language, including acoustic accents, rates, and vulgar expressions, can be effective in preparing the learners for actual real-life use. However, there is still very little work done about the efficacy of these tools in helping Saudi learners enhance their real-life listening comprehension.

The current literature on AI in language learning is, to some extent, limited – it mainly concerns the use of AI in language learning from a cognitive viewpoint, e.g. the ability of AI to offer immediate and accurate feedback (Huang et al., 2022). Although this feedback may help the learner improve language development within specific settings, it becomes necessary to consider whether these AI tools help learners develop listening skills in real-life scenarios. AI programs are quite specific about the language input they provide—usually clean audio alongside paralinguistic cues, which looks very different from (and often more structured than) real-life discourse. A set of specific difficulties may include impairment of learners' understanding of naturalistic speech patterns including reduced forms, fillers, hesitations, and other non-standard English expressions that may not be accurately mimicked by AI systems (Rahko, 2017). Consequently, what is required is AI tools that provide more than the rational and polished language typically associated with language learning software and provide learners with the real-life usage of the language they will be exposed to when outside the classroom.

However, as with any intelligent and adaptable system, the issue now is to identify how new models of instruction can develop usable paths given the dynamic and multifaceted aspects of real-life listening situations that learners encounter. Abdellatif et al. (2024) explained that effective AI for listening comprehension should not only mimic linguistic variations but also be contextually diverse, that is, situational context, noise, and interaction. Such factors may not be incorporated in present AI instruments in most linguistic AI applications mainly emphasize the linguistic correctness of a translation as opposed to relativity. If these elements are omitted, there is an emergence of a situation whereby learners attain the fluency to understand clear and scripted speech in a second language while lacking the ability to cope with the dynamics of everyday communications in a second language.

However, language learning with the help of AI presents pedagogical and technological issues of great importance. On the capacity side, such tools, if used for personalizing learning, entail a high level of technology pedagogy, including internet connectivity, innovative software, as well as preparation of teachers to use teaching and learning resources. Despite some progress in Saudi Arabia regarding digitization in education, a digital divide between the urban and rural areas challenges the delivery of AI-based learning (Alotaibi & Alshehri, 2023). Furthermore, using artificial intelligence in Saudi schools and universities, many educators are still not ready for the use of artificial intelligence in teaching and the use of its devices. Therefore, though AI has the potential to contribute greatly to the education system, integrating it into language classrooms has to be done in ways that will provide benefits to learners as well as teachers while optimizing AI-specific characteristics.

The present study aims to fill these research gaps by assessing the facilitative role of AI tools that can enhance Listening Comprehension in authentic contexts among Saudi learners of foreign languages. Thus, the present research is intended to offer empirical data on the effectiveness of the AI-based platforms imitating genuine uses of NL in diverse contexts for facilitating the development of learners' comprehension and processing of spoken language. The knowledge generated from this research will also be beneficial not only to the theory on the application of AI in the acquisition of languages but also the practical application by educators and policymakers in Saudi Arabia on the use of AI in language programmes.

1.1 Problem of the Study

English as a foreign language learners also need to understand spoken language and this aspect is widely considered as one of the most difficult (Almaleki, 2020). Foreign language learners in Saudi Arabia, especially those learning English have limited listening comprehension in real-life situations. Studies with technologically enhanced action-based listening show that in classroom-based listening activities, students are provided with a limited amount of listening material which is slow and mostly eased, unlike real-life listening. Therefore, learners face challenges when trying to apply what they have learned in real-life situations such as conversations, the media or in classes, and other formal contexts where speech is fast and includes cultural references and multiple accents and idioms. There is one major difficulty that comes from the absence of real-life listening situations – learners' inability to constructively comprehend and participate in communicative events. Since AI-based tools contain the potential to improve listening skills by giving practice in comprehending actual native language input, the effect of these technologies on listening comprehension in Saudi learners was still unexplored. To this end, the present study is designed to fill this gap to assess the use, relevance, feasibility, and impact of AI applications in improving listening comprehension in real-life situations of Saudi Arabians learning foreign languages.

1.2 Research Questions

1. How effective are AI applications in improving listening comprehension for English language learners in Saudi Arabia?
2. To what extent do AI applications expose learners to authentic, real-world listening scenarios that contribute to the development of listening skills?
3. How do English language learners perceive the use of AI tools for listening comprehension compared to traditional learning methods?

1.3 Significance of the Study

The findings of this research provide substantial theoretical and practical advancements in language learning specifically with the Ministry of Saudi Arabia in mind. The information gathered in this study is beneficial for educators and policymakers wishing to

investigate and apply innovative technology tools to enhance listening comprehension in FL learners. In this paper, a research gap on how AI-based applications may facilitate the development of practical listening skills to address real-world listening contexts is addressed through the assessment of AI-based applications intended to mimic real-world listening scenarios. The findings of the study would be beneficial to educators to help them integrate the utilisation of AI tools in a learning environment that supports the development of language comprehensiveness while making the learners more prepared to deal with the natural use and practice of the language. Nevertheless, this research is especially timely for Saudi Arabia as the use of technology within education constitutes a strategy. Acknowledging that the application of AI in language learning can be beneficial, this study contributes to the national drive to enhance the teaching methodologies used in schools and to offer students increasingly fascinating and ultimately comprehensible tuition.

Furthermore, it provides the existing literature a fill in the understanding of using AI within language learning. To the best of the author's knowledge, the current study is one of the few where the potential of AI has been examined to enhance listening comprehension in real-world scenarios. This paper further extends knowledge on how the skill of listening can be enhanced given the emerging AI technologies by providing more concrete insight and ideas. In addition, the present research underlines the need for contextualised language learning, a kind of learning that focuses on the practical use of information as opposed to learning that happens in a vacuum. In light of this, the study demonstrates working applications of AI in the micro-level learning processes which the more traditional socio-cultural approach has posited as fundamental to future models of language teaching.

1.4 Terms of the Study

The period of this study was six months during the late 2024 and early part of 2025. Altogether, 150 foreign language learners from many universities in Saudi Arabia were involved in the present study. The study was structured around two main phases: an assessment phase, in which participants' listening comprehension levels were then determined before an increasing phase in which learners will interact with AI-based listening comprehension resources. These were intended to replicate the normal world listening situations so that learners were presented with various degrees of difficulty depending on their learning progress. The post-test section was administered three months later to afford the participants time to practice with the AI tools to superior their listening comprehension.

Throughout the study, the students were split between an experimental group which engaged in the use of AI applications in the class and a control group, which learned through conventional means such as lectures. The AI tools used in the study included speech recognition capabilities, natural language dialogues that were completed with the clients, and audio-visual clips that demonstrated real-life uses of language. The data was collected by pre-and post-tests for listening comprehension as well as by the satisfaction questionnaires focusing on participants' attitudes towards the use of AI in the language learning process. Qualitative data drawn from interviews with the participants were also used to enhance understanding of the experiences of the participants in using the AI applications.

1.5 Limitations of the Study

However, despite these successful research findings, the following limitations emerged in this study that need to be addressed. One of the many limitations was the shorter time for the intervention phase. While using three months guaranteed the experimental group enough time to interact with the AI tools, it might not have been enough to document the benefits of the AI-based learning approach towards the improvement of the participant's listening comprehension in the long run. Thus, the results of future studies could be more accurate if longer exposure to AI applications had been chosen to determine how stable the improvement in listening comprehension is.

A fourth limitation was the size and relevance of the sample as well as the lack of variation in the participants. The sample included 150 foreign language learners within universities The participants were all selected from urban universities which may limit the generalization of the results for the sampling of Saudi Arabia universities. Education from the learner's perspective will be different and learners in the rural area may not have had interaction with some of the technology tools in the lesson. Further, the use of tech as a medium and language acquisition may vary by previous experience of the subject with technology or learning through language respectively. The results can be generalized barely because the participants of the survey were mainly selected from urban areas; hence, a broader representation of learners from rural areas could have painted the full picture of the influence of AI in diverse learning environments.

Furthermore, the AI tools enabling this study were developed to mimic real-life conditions and yet they could not capture certain uncertainty and variation that are often found in conversations. It is for this reason that real-world listening entails things like noise, interference, and issues to do with multi-talker and non-standard English which the AI applications might take some time to mimic. That is why the results of this study may not always be indicative of the difficulties learners experience in genuinely changing, realistic settings. More studies of this nature could be carried out to investigate designing new AI tools that can effectively mimic realistic human interaction.

Finally, while this research focused on the learners' attitudes towards the AI applications, the responses obtained through the use of questionnaires and interviews were likely to contain distortions due to self-reporting. Their perceptions of AI tools could have also been aligned with their experience with technology or their willingness to embrace technological learning instruments. Future research could include more quantitative aspects of motivation, as well as engagement of learners to complement the qualitative data analysis.

2. Literature Review and Previous Studies

The place of aural skills in the learning context is an acknowledged fact that remains one of the more difficult areas for foreign language learners. This is especially the case in some countries such as Saudi Arabia; despite English being taught as the second language, listening

comprehension skills are comparatively lower than other sections such as reading and writing abilities. This raises difficulty due to many factors including the nature of natural spoken language and accent variations, the use of informal language and quick delivery of messages (Stuart-Smith et al., 2013). This section discusses various works done on the topic of listening comprehension, difficulties learners encounter, and how AI tools are being examined in an attempt to solve the difficulties already mentioned.

It was also earlier deemed by many theorists and researchers as one of the most complex facets of a language learner's acquisition process. In contrast to reading and writing in which both are best practised through meaning-form connections with a good measure of control being exercised on the actual input provided, listening requires learners to comprehend spoken language in real-time with limited opportunities for review or rectification (Vandergrift, 2011). This becomes quite a challenge, especially for foreign language learners who do not have any close interaction with people who use the foreign language. In Saudi Arabia, although English is taught as a foreign language and used in the classroom, it is not used in real-life contexts, the learners fail to map their listening learnt in the classroom to real-life listening contexts. Studies have indicated that listening tasks in classrooms are often minimal in interference with naturalistic features such as different accents, colloquialisms, and informal tone of speech.

Language learners themselves, for example, are likely to be struggling with some biochemical and mechanical problems, for instance, misunderstanding of specific terms and expressions, failure to grasp important points because of fast flow of speech, and inability to process language in interaction with others (Sapir, 1921). Therefore, the existing practices of practising listening comprehension, which mainly involve listening to slow and clear speech in rather limited circumstances, can hardly equip learners with better skills for listening in natural communicative situations (Vandergrift, 2017). Therefore, the objective areas establish the necessity of the learning tools which provide the true-to-life imitation of the listening sessions.

The application of information technology especially the use of Artificial Intelligence in language acquisition has expanded significantly over the last decade. The use of AI-based applications presents several advantages such as flexibility of learning, feedback sections and dialogues that cannot be obtained by conventional language learning approaches (Strambi & Bouvet, 2003). AI has been adopted in Learning and Instruction because of its capabilities to adapt to learners' needs and offer individualized lessons, and instant feedback. Many AI applications have been deployed for language learning such as speech recognition systems, feedback systems, and AI tutors.

The study confirms that AI is of significant benefit about certain aspects of language learning, including vocabulary and grammar, but finds that the application of the technique in succeeding at the listening comprehension level has been less well-researched. More recent research indicates that with the use of leading artificial intelligence-based platforms, learners can be exposed to real speech and actual communication situations that are much closer to real-life interactions and thereby allow the learner to listen in less controlled conditions (Richards, 2008). These tools which employ a form of artificial intelligence to learn the learners' level of proficiency with the knowledge improvement of audible comprehension may help provide better approaches than class training (Gilakjani & Ahmadi, 2011). Nevertheless, although the use of AI has its potential benefits and advantages, it is far from perfect, primarily when it comes to recreating actual-life interaction.

Listening comprehension using AI involves attempting to mimic real-life situations to expose the learner to speech that is used in actual daily life. Such tools usually include capabilities such as speech recognition capability as well as the natural language processing ability to enable the learners to be able to understand spoken language, recognize the words as well as decipher the meaning of words from context. For instance, some AI platforms apply the application of conversational writing with actual native speakers; others apply the application of machine studying algorithms that adapt the degree of listening tasks based on the learner's progress (Hassani et al., 2016).

A specific example relates to applications in language learning where there exist listening assignments facilitated by AI whose pace, content difficulty, and speaking tone depend on the learner's level of comprehension. It is for such reasons that such platforms can mimic conversations in as many actual conditions as there are, from simple conversation to business conversation thus exposing the learners to various listening practices. This simulation of such casual conversation is particularly important for engaging the learners to develop speech for real-life contexts in which there is background noise, interruption or informal language use. According to Schmidt & Strasser (2022) learners who practice with listening aids that are facilitated by AI claim to have gained in speed and fluency as well as in their capacity to comprehend rapidly spoken, informal, and real-time language.

However, the following problems are drawn out; Increased Student Traffic Although these benefits are apparent, difficulties persist. For instance, although AI can generate various listening scenarios, these applications are inadequate to predict the complexities of actual speech (Kamm et al., 1985). Crosstalk, interruptions and cultural factors are typical of real-life listening that may not be well simulated by most AI applications. Thus in their quest to use tools and structures to enhance comprehension, particularly as instantiated by current AI technologies, learners may still be left exposed to the reality of real-life human conversation. Secondly, most AI interfaces are designed to use sets of pre-recorded or scripted language, which hardly mimics real-life unscripted speech. Such limitations serve as the rationale for studying the degree to which AI-based tools can mimic the different ways that natural language is used.

There are several papers which examined the application of AI in language learning focusing on such tasks as vocabulary, grammar and writing. Nonetheless, there are relatively few studies committed to listening comprehension, which is still a challenging and research-deficit domain. A social experiment which was conducted in Saudi Arabia by Zhai & Wibowo (2023), explored the applicability of AI in language acquisition in general but with disregard to listening comprehensibility in particular. Unlike other elements of language proficiency that their studies indicated might be boosted by AI applications, the authors' works offered limited insights into how AI might

enhance listening.

Some works initiated up to date include detailed analysis of the use of AI in listening comprehension. For instance, AbuSahyon et al. (2023) established that the implementation of AI enhancement boosted the listening comprehension mastery of foreign language students, especially when learning formal language. However, these tools were unable to mimic the nature of authentic and diverse language that learners encounter in day-to-day life. In the same context, Yu et al. (2023) stressed the possibility of using AI in the way it provides an imitation of the listening conditions that refer to real-life communication; however, he noted that the difference between AI-based controlled tasks and real-life communication is vast. Li et al. (2024) also underlined that AI-based speech Recognition should take into account certain aspects such as noise or other competing voices in the environment to prepare the learner for real life.

The lack of research in this area is the logical premise for a study that delves into determining the impact of AI tools toward enhancing the practice of LCM in other realistic settings including Saudi Arabia, where English is taught, but hardly used in daily parlance. The present research intends to fill this void by assessing the utility of the AI tools in improving the section of listening for the learners of the foreign languages in Saudi Arabia by assessing the genuine issue of global communication which is encountered by the students.

3. Methods

The purpose of this study was to determine the role of AI applications in improving listening skills in relevant contexts among Saudi students learning a foreign language. Secondary data was analyzed to answer the research question because of its multifaceted nature a comprehensive methodological approach was used. To this end, the study adopted both quantitative and qualitative data collection methods to ensure the gathering of holistic data about the utility of AI tools in language learning.

3.1 Participants

Participants of the study were 150 foreign language learners from different Saudi Universities such as the King Saud University, the King Abdul Aziz University, and the Imam Muhammad bin Saud University. These participants were purposively recruited from a large university and each of them was recruited based on gender, academic major and language proficiency level to achieve proportionality through the use of st retreated random sampling technique. All subjects involved in the survey course were learners in English language courses and their age was restricted between 18 and 25. Level of proficiency was also measured in which about 50% of the students fell under the beginner level, 30% intermediate level and only about 20% under the advanced level using the Common European Framework of Reference (CEFR). The sample diversity used here enabled a consideration of how the performance of the AI tools may differ depending on one's L2 proficiency. The ethical approval for the study was sought from the University Research and Ethical Committee of the universities under which the study was carried out and all the respondents participated voluntarily.

3.2 Research Design

A quasi-experimental, pre-test/post-test design was used for this study, with participants divided into two groups: a selected sample which is also known as an experimental group and the second, which is called a control group. In the experiment, the students in the experimental group utilized Applications in their understanding of the comprehension listen, in contrast, the students in the control group employed only conventional classroom techniques. This design was chosen because it facilitated the assessment of learning gains of the two groups to a specific extent, about some external variables such as prior knowledge and experience with technology.

The comprehension quizzes used in the experimental group were selected from AI tools that approximate real-life listening situations. The use of these tools in interaction incorporated features such as speech recognition and dialogues as well as feedback mechanisms. They were meant to familiarize learners with both regional dialects and diverse accents and spontaneous speech and colloquialisms as opposed to the structural or normal read, providing a divergent mode of practising comprehension than that achievable in a classroom setting. The control group also remained in their traditional listening comprehension practice that includes dialogue, textual exercises, and audio without the elements of AI found in the tools.

3.3 Data Collection

Data collection for this study was conducted in three phases: an assessment test before the implementation of any changes (pre-test), the implementation of the changes (intervention phase), and an assessment test after the changes (post-test). Before the intervention, all participants took a pre-test aimed at evaluating the initial listening comprehension level of the participants. This test comprised audio clips on increasing difficulty that featured various real-life scenarios; social discourse, academic discourse and various forms of media. The listening tasks were succeeded by comprehension questions that checked whether the participants could extract data, scrutinize hints and grasp cultural implications. This was aimed at assessing the participants' listening skills before the study so that any improvements that might be noticed were the results of the study alone.

The intervention in the study took three months for outpatients and one month for inpatients within the hospitals. In this phase, the experimental group used the AI-based listening comprehension tools twenty minutes a day three days a week. For the AI applications to be interactive, listening tasks in the form of dialogues, academic lectures and news broadcasts were included to resemble natural communication. The control group, on the other hand, participated in typical class listening tasks for the same amount of time using suits such as CDs or online audio that were recorded and could not provide the flexible parameters of the AI tools. The topics selected for both groups' contents were as well matched in terms of the material covered as well as the level of the language used.

After the intervention, both groups completed the post-test, which consisted of 36 questions that were similar to the pre-test but used different audio files to prevent the subjects from recalling the content of the audio material used at the time of the pre-test. The post-test was also designed to determine whether there was a change in listening comprehension ability in both groups. Besides the post-test, the participants of the experimental group also filled in the questionnaire and did the semi-structured interview on the perceived usefulness of AI tools, motivation and learning experience.

3.4 Instruments

Based on the chosen instruments, this study controlled aspects of listening comprehension and made sure that aspects of validity and reliability were met. Habits of listening comprehension test both at the pre-and post-test data was also used in this study was developed in this study. The test was divided into three sections: After reducing by half, the pretest and posttest assessed four categories: (a) factual comprehension, which involved factual recall and understanding of specific content; (b) inferential comprehension, which examined the participants' ability to grasp the inferences made in spoken language; (c) cultural comprehension, which covered the learners' knowledge and understanding of cultural expressions and references in spoken language. It was administered in a pilot form to a few students with a view of ascertaining the clarity validity and reliability of the test. The internal consistency reliability value was found to be 0.88 by Cronbach's alpha method.

The particular applications used in the experimental group were selected for language learners and imputed such tools as speech recognition, learning paths, and contexts to mimic language usage. Such applications presented activities like dialogue, quiz questions, and speech exercises, which allowed learners to articulate answers verbally, in a manner that imitates normal communication. The tools offered feedback from the perspective of accuracy, fluency and comprehension skills thus enabling the learners to boost their listening ability as they learn. The AI applications also changed in response to the implementation of learning objectives, which defined the level of difficulty after each learner's performance.

During the post-intervention assessment of the experimental group, a questionnaire was completed by the participants. The questionnaire gauged participants' perceptions of the AI tools with a focus on the perceived usefulness, interaction and motivation. The questionnaire consisted of closed-ended Likert scale items from 1-strongly disagree to 5- strongly agree and open-ended items to collect qualitative responses. The validity of the questionnaire was ascertained from a pilot study and had a Cronbach's alpha of 0.85. About the participants: Five sets of semi-structured interviews were completed with 20 participants from the experimental group. The interviews were focused on the outcomes and participants' perceptions of the AI tools' effectiveness, the difficulties with its usage and their motivation. The interviews were audiotaped and transcribed for thematic analysis.

3.5 Data Analysis

Quantitative and Qualitative data were used in this study since a combination of both enabled me to get the most complete picture of how the AI tools impacted listening comprehension.

The quantitative data analysis consisted of mean and standard deviations for both the experimental and control groups' pre-test and post-test results. To establish any differences between the two groups, with specific regard to the listening comprehension scores before the experiment and after the experiment, a paired-sample t-test was employed. Then, to compare the performance of the experimental and control groups in the post-test phase, an independent-sample t-test was conducted. Such an analysis made it easy to compare how effective the AI tools were as compared to the traditional methods. Moreover, an analysis of variance of multiple regression was used to examine the influence of participants' proficiency level on the increase in listening comprehension while assuming age and gender.

Both the interviews and questionnaires administered were analyzed through thematic analysis. The purpose was to examine the common tendencies of participants' attitudes toward AI-based tools, their learning processes, and their attitudes to AI. The respondents' answers were thus analyzed, and the themes were developed to offer information on the experience of learners and their reasons to learn. Coders were also assisted by the use of NVivo in coding and handling the thematic analysis.

4. Results

To ensure that the results of the main statistical tests conformed with the parametric and variance tests, some prescreening tests were conducted on the results obtained from both the experimental and control groups.

To compare the variances of the experimental and control groups' pre-test Listening Comprehension scores, Levene's test was performed. The findings of the test are shown below:

Table 1. Homogeneity of Variance Test (Levene's Test)

Group	Mean Pre-test Score	Standard Deviation	Levene's Test for Equality of Variances	p-value
Experimental	35.5	5.2	0.983	0.375
Control	36.1	5.0		

Levene's test gave an F-value of 0.375 which is higher than the table value of 2.60 at 0.05 significance level which therefore shows that the variance in the experimental and control group pre-test scores are equal. Hence, the assumption of equality of variances is valid, and thus we can go for more parametric tests like t-tests without having to worry about unequal variances.

Next, the normality of the pre-test listening comprehension scores for both groups was assessed using the Shapiro-Wilk test. The results are summarized in the table below.

Table 2. Normality Test (Shapiro-Wilk Test)

Group	Pre-test Score Normality (Shapiro-Wilk)	p-value
Experimental	0.983	0.572
Control	0.967	0.422

In the first experimental group, the p values of the pre-test were 0.572, while in the second experimental group, the p values of the pre-test were 0.422, which was greater than the specified level of 0.05. This means that the pre-test scores of both experimental and control groups are normally distributed and as such, the use of parametric tests is possible.

Once the homogeneity and normality assumptions were confirmed, the pre-test and post-test scores of both groups were analyzed using paired-sample t-tests to assess the within-group changes, and independent-sample t-tests to compare the differences between the two groups. The results are presented in the following tables.

Table 3. Pre-test and Post-test Comparison within the Experimental Group (Paired-Sample t-test)

Group	Pre-test Score	SD	Post-test Score	SD	t-value	p-value
Experimental	35.5	5.2	45.2	3.8	-10.42	0.000

The paired-sample t-test revealed a significant improvement in the experimental group's listening comprehension scores from the pre-test (M = 35.5, SD = 5.2) to the post-test (M = 45.2, SD = 4.8), with a t-value of -10.42 and a p-value of 0.000. This suggests that the AI-driven intervention had a significant positive effect on the listening comprehension skills of participants in the experimental group.

Table 4. Pre-test and Post-test Comparison within the Control Group (Paired-Sample t-test)

Group	Pre-test Score	SD	Post-test Score	SD	t-value	p-value
Control	36.1	5.0	39.5	4.9	-4.28	0.000

The control group also demonstrated a statistically significant improvement in their listening comprehension scores from the pre-test (M = 36.1, SD = 5.0) to the post-test (M = 39.5, SD = 4.9), with a t-value of -4.28 and a p-value of 0.000. However, the magnitude of the improvement in the control group is smaller compared to the experimental group, suggesting that traditional methods were less effective than AI-based methods in enhancing listening comprehension.

Table 5. Between-Group Comparison at Post-test (Independent-Sample t-test)

Group	Post-test Score (M ±SD)	SD	t-value	p-value
Experimental	45.2	4.8	4.87	0.000
Control	39.5	4.9		

The rise of the post-test indicated that the experimental group that solved the problems performed significantly better in the achievement test (M = 45.2, SD = 4.8) than the control group (M = 39.5, SD = 4.9 with t-value 4.87 pn value 000. This result is evidence that the extent of enhancement in the students' level of listening comprehension was many folds higher when the developed AI-supported tools were used in contrast to the more conventional approach.

A multiple regression analysis was also carried out to analyse the result to see if the learning improvement in listening comprehension was affected by or about the learners' initial proficiency level. The findings are provided below.

Table 6. Regression Analysis: Influence of Proficiency Level on Improvement

Predictor Variable	B	SE B	β	p-value
Pre-test Score	0.75	0.15	0.32	0.000
Proficiency Level (Beginner)	2.32	1.05	0.14	0.03
Proficiency Level (Intermediate)	1.98	1.02	0.12	0.04

In the regression equation, the pre-test was found to be a significant predictor of post-test improvement with B = 0.75 indicating that learners with higher pre-test percentage had higher improvement; p<0.05. Further, the proficiency level of the learner also emerged as a significant predictor with both beginner (B = 2.32, p-value = 0.03) and intermediate learners (B = 1.98, p-value = 0.04) experiencing a relatively better improvement as compared to the advanced learners. This means AI tools could be especially helpful for new and mid-level learners because presumably, AI helps to compensate for basic listening comprehension deficiencies.

To know the impact of the intervention on learners' motivation, the motivation scores of the pre-test and post-tests of both the experimental and control groups were compared. Motivation scores were obtained on a Likert scale and they were conducted pre and post-intervention.

Table 7. Motivation Levels (Pre-test vs. Post-test)

Group	Pre-test Motivation Score	SD	Post-test Motivation Score	SD	t-value	p-value
Experimental	3.4	0.8	4.5	0.6	-8.21	0.000
Control	3.5	0.7	3.9	0.7	-3.50	0.001

The experimental group received a test of motivation of 3.4 (0.8) pre-test and 4.5 (0.6) as a result of the post-test, t = -8.21, p < 0.05. This shows that the AI tools highly boosted enthusiasm among learners. On the other hand, the control group recorded a slightly higher increase in motivation from 3.5 (SD = 0.7) to 3.9 (SD =.7) with t=-3.50, p=.001. Therefore, although the control group also had an increase, it was relatively lower as compared to the other groups – underlining the prospect of AI for enhancing learner's interest.

The table below provides the results of the questionnaires filled out by the participants on the rated effectiveness of the AI tools. Our responses were on a Likert scale with options that included 1 – Strongly Disagree; 2 – Disagree; 3 –Neutral; 4 – Agree; 5- Strongly Agree.

Table 8. Perceived Effectiveness of AI Tools (Questionnaire Responses)

Item	Mean (M)	Standard Deviation (SD)	Percentage Agreement (%)
The AI tools helped me improve my listening skills	4.6	0.7	92%
The AI tools were engaging and interactive	4.4	0.8	88%
I feel more confident in real-world listening situations due to the AI tools	4.5	0.6	90%
The AI tools provided useful feedback	4.7	0.6	94%

The results of the experimental group's questionnaires also show that they have fairly high levels of satisfaction with the AI tools. The mean score of most items was greater than 4.4; the item that received the highest agreement was ''The AI tools provided useful feedback' '(94%) This indicates that the participants realised that the AI tool improved their listening skills and that the feedback was helpful to them. The tool kits yield relatively higher engagement (88%) and confidence building (90%) and these real-world statistics again endorse the efficacy of the tool kits.

To compare the results of the beginner, intermediate, and advanced users of the tools which incorporated AI, a post-hoc test was conducted. The pre-and post-test means by proficiency level are as follows:

Table 9. Breakdown of Improvement by Proficiency Level (Pre-test vs. Post-test)

Proficiency Level	Pre-test Score	SD	Post-test Score	SD	t-value	p-value
Beginner (n = 50)	28.7	4.5	43.2	4.2	-15.46	0.000
Intermediate (n = 45)	37.1	4.8	46.3	4.5	-9.24	0.000
Advanced (n = 55)	44.2	3.9	47.8	3.7	-4.75	0.000

The results therefore also reveal that the AI intervention caused a change in proficiency level across all the groups. Again, the largest improvement was in the beginner group and was from 28.7 (4.5) to 43.2 (4.2); t=-15.46 and p= 0.000. This points to the fact that AI tools had the biggest contribution to beginners most probably due to the simple fact that they bridged some gaps in listening skills. The intermediate group used in the present study also exhibited a significant improvement, and a t-value of -9.24 for the present measure and a p-value of 0.000 indicating statistical significance The advanced group demonstrated a similar, though slightly smaller, improvement: it increased from 44.2 (SD = 3.9) to 47.8 (SD = 3.7, t (14) = -4.75,

To answer the question of whether the frequency of using the AI tools influenced the results the following table illustrates the post-test scores according to frequency of usage.

Table 10. AI Tool Usage Frequency and Effectiveness

Usage Frequency	Post-test Score (M ±SD)	SD	t-value	p-value
1-2 times per week	42.3	5.1		
3-4 times per week	45.1	4.3	2.01	0.045
5+ times per week	47.6	4.0	5.47	0.000

The findings show that as the use of tools incorporating AI was higher, it was met with a more significant enhancement in listening comprehension. Among the participants, the post-test total scores mean were highest with 5 or more use of the AI tool per week (M = 47.6, SD = 4.0), t = 5.47, p = 0.000. The group that used the tools 3-4 times every week also a had higher improvement mean score of 45.1 with a standard deviation of 4.3 Ci = 2.01 and PI = 0.045; the group that used the tools 1-2 times every week had the least improvement. This has implications that imply that with increased usage of the developed AI tools the learners will boost the comprehension of the target language in the listening domain.

The following table indicates the correlation which exists between listening comprehension scores and motivation scores (post-test).

Table 11. Comparing Motivation and Listening Comprehension Improvement

Motivation Change (Post-test - Pre-test)	Post-test	Pre-test	Correlation Coefficient (r)	p-value
Low (1-2 points)	4.2	3.1	0.35	0.002
Moderate (3-4 points)	7.5	4.2	0.55	0.000
High (5+ points)	10.8	4.5	0.75	0.000

Analyzing the results of changes in motivation levels and listening comprehension, the difference is statistically significant. Learners who underwent higher upgrade in motivation (5+ points) yielded the best improvement in the post-test listening comprehension mean of 10.8 (SD=4.5) with rho 0.75 and p 0.000. Those who have moderately improved their motivation (3-4 points) demonstrated good improvement in their listening knowledge (M=7.5, SD=4.2) with a co-relation of 0.55 custom and the learners who have low improvement in motivation (1-2 points) realized relatively better improvement in the listening comprehension.

5. Discussion

5.1 The Influence of AI identification on Listening Comprehension as Perceived by the Students

Overall participants had positive perceptions regarding the use of the tools with regards to enhancing their listening comprehension. The value of the AI intervention was considered high as far as the building of actual listening skills is concerned.

> *"The AI tools helped me understand fast conversations, especially in the context of real-world situations like interviews or news reports. I used to find them too difficult, but the tools made them more accessible."* — Participant 7, Experimental Group.

Our Participant C's comment underscores the usefulness of using AI tools in emulating realistic listening situations. The comment is in concordance with recent studies conducted by Novawan et al. (2024) that pointed out the views that the current AI tools in the realm of language learning can accurately simulate real listening assessments. However, it also highlights the problem of teaching listening in a conventional model when the resources are beyond students' capabilities.

> *"The feedback provided by the AI was useful, but sometimes I felt it didn't explain the mistakes in detail. I think I would improve faster if I had more explanations."* — Participant 12, Experimental Group.

Finally, regarding the feedback given by the AI, the participants expressed certain discontent with the explanation depth we gave them. This is consistent with the observation by Gajos & Mamykina (2022) that the feedback from an AI model tends to be broad rather than specific and may, therefore, not help learners grasp their mistakes. Future work should investigate such feedback mechanisms to implement systems with more detailed feedback capabilities.

> *"In the beginning, I didn't think AI would help much, but after a few weeks, I realized it was making a big difference. I could understand things I couldn't before, like specific accents."* — Participant 3, Experimental Group.

This is an indication of one of the main issues that learners might have when implementing technology, as Pinho et al. (2021) pointed out, the learners may resist the implementation of new technologies. But gradually, real changes emerged as the participant extended time spent with AI tools, which can be explained by Dwivedi et al. (2021) stating that repeated use of technology-based treatments improves the results.

> *"I think the AI tool works well for basic tasks, but I still need the teacher's support for more advanced listening comprehension."* — Participant 15, Experimental Group.

This quote speaks of the shortcoming when it comes to the use of AI tools in listening – that is where the tasks involved are more comprehensive. According to Pedro et al. (2019), while using AI for the identified foundational level works well, such tools fall short of providing the depth necessary for higher-level students, thus the idea of combining the use of AI with traditional teaching.

5.2 Motivation and Participation with the Help of AI Instruments

Several authors claimed that AI tools enhanced learner participation and motivation due to the elements of interactivity and customisation they offer. Even here, the level of engagement depended on the usage frequency as well as the learner's level of experience.

> *"I enjoyed using the AI because it felt like I was playing a game rather than studying. The immediate feedback made me feel like I was getting better."* — Participant 9, Experimental Group.

This quote echoes the elemental concept of gamifying AI tools and aligns with Rivera & Garden's (2021) research that incorporates gamified learning experiences to enhance learner engagement and motivation. The immediacy of feedback is also viewed broadly about sustaining motivation, A picture of similar conviction is painted by Jung et al. (2019) who stated that when feedback is offered as readily to the learners, it offers a sense of accomplishment to the learners.

> *"At first, I thought the tasks were easy, but after a while, I saw that I was progressing. It pushed me to keep trying."* — Participant 2, Experimental Group.

From this comment, one may get an indication that there is some sort of flow in the AI-supported tools such that it can take a learner in as a beginner and graduate him as a master without undue stretching of his/her abilities. As observed by Fan et al. (2024), when an AI application prescribes to the learner's learning curve, the increase in motivation is a result of the sense of accomplishment obtained from the use of the application. But this may also pose the question of how effectively the tools can address the learners at different proficiency levels though validated by Shroff et al. (2019).

> *"I used to feel bored during listening lessons in class, but with the AI, I looked forward to the practice sessions. I was more excited to listen to new materials."* — Participant 5, Experimental Group.

This particular participant has a forward and backward transition of motivation owing to the use of AI. According to Diwan et al. (2023), the individualized learning nature of AI generally results in higher learner engagement in language learning. This means that AI tools can

effectively enhance the language acquisition process in a way that breaks the monotony that learners come across while in class.

> *"Sometimes, I got frustrated because the system didn't recognize my answers or said I was wrong when I wasn't."* — Participant 10, Experimental Group.

Most of the reactions received for the AI tools were positive but there was some dissatisfaction with the performance of the Technology. This concurs with Chen et al. (2023) who argue that in time, the AI systems might fail to understand learner inputs thus resulting in frustration. This has brought out the need to ensure better accuracy in the responses that the AI provides to ensure a better learner experience.

5.3 Challenges in Using AI Tools

However, certain difficulties were identified when it comes to the application of AI tools: technology-related; and the lack of clear descriptions of the feedback provided.

> *"The AI kept freezing sometimes, especially when there were a lot of people using it at the same time. It was frustrating because I couldn't complete my tasks."* — Participant 8, Experimental Group.

This participant presents a typical concern in technology-supported interventions – technical problems. Thus, Wilson (1996) pointed out that technical aspects can make the process challenging as infrastructure problems can occur in a large group since it is not designed for that type of learner. These problems show why timely reliable technical support is desirable when implementing learning systems using AI techniques.

> *"Some tasks felt repetitive after a while. I think the system could provide more variety to keep me engaged."* — Participant 6, Experimental Group.

Then the participant emphasizes that it is important for learner engagement to have variation in the tasks that are assigned. Barreno et al. (2006) stated the importance of learning systems that are flexible and include various exercises to avoid doing repetitive work. AI tools should always be improving to present more progressive tasks to ensure learners are engaged and lost with what to do next.

> *"I think the AI doesn't fully understand the different accents we might hear in Saudi Arabia. It worked well with American English, but some local accents were difficult for it to process."* — Participant 11, Experimental Group.

This comment draws the curtain over a shortcoming of most AI tools particularly in handling different accents; in a multinational country like Saudi Arabia, this poses a major challenge. Mamun et al. (2024) stated that, although conventional AI programs perform rather well in deciphering ordinary tones, regional accents present emerging difficulties. While native language processing is fairly well recognized, this suggests that there is still the need to continue advancing AI tools to tackle additional varieties of language use.

> *"I would prefer more interaction with my teacher. Sometimes the AI can't explain things as well as a real person can."* — Participant 14, Experimental Group.

This participant echoes one of the most recurring complaints users have with such tools, namely, their inability to mimic real human conversation. According to Meijer et al. (2009), even though these solutions give quick feedback, they lack the personal approach a teacher takes with a student. This implies that AI tools should work as complementary to human trainers in contexts where learning languages is required.

The assessed qualitative data mostly confirm the positive effect of the AI tools for the improvement of listening comprehension skills and participants' attitudes toward the task. However, the following findings are revealed with the gathered data: various technical problems, less depth in feedback, and applicability of challenges in accommodating different accents. Such concerns complement the issues discussed in the literature including the fact that AI feedback might miss out or offer imprecise comments (Cavalcanti et al., 2022) and the limitation of technology in identifying regional variations.

5.4 Technology Aspects on Use of AI Tools/Listening Comprehension Enhancement

Something that the quantitative results depicted was that there was a significant boost in listening comprehension for the learners who employed AI assets. This research finding resonates with existing scholarly taps of AI as being able to deliver customised, differentiated learning interventions that can meet the specific needs of learners. Gilakjani & Ahmadi (2011) stressed the fact that AI-assisted tools particularly those that present the learner with real-life listening scenarios, have been proven to enhance listening comprehension because the learner listens to real natural spoken language. This study also found that learners felt that AI-enhanced materials enabled them to understand news reports and interviewers' rapid speech, which supports the outcome of Liu et al. with the accounts of how AI is effective in learning to cope with fast pace and context switch compared with the traditional classroom lessons.

However, the study also revealed a critical issue: AI tools for listening comprehension were further helpful for usual items and passages but they were less capable of performing in regional accents as are usual in KSA. Non-standard accents are also based on one's ability to speak and as observed by Auer (2004), systems trained largely based on formal speech may not be as efficient at handling non-standard dialects as realised in this study. Some of the complaints made by participants included issues to do with the inability of the AI to comprehend Saudi Arabic dialects as this reveals a major weakness in AI language comprehension. From my standpoint of critical thinking, the above limitation has a profound revelation; there is a significant weakness inherent in the use of Artificial Intelligence in embracing contextual linguistic differences. Although AI tools work in some restricted settings, their lack of ability to consider aspects of

the regional language, excluding the standard Arabic dialect restricts their applicability in Saudi Arabia which has very diverse dialects from the standard Arabic as depicted in the Saudi Arabian Education Ministry. A question arises, is the current state of AI language tools ready to facilitate learners in the multilingual multicultural learning environment?

Even accent recognition suggests the problem of how well AI models represent the societies in which they are developed and deployed. Fletcher also stated that these AI systems are built with prejudice to their designers and inventors, so it might be quite reasonable to assume that the virtual assistants do not understand regional accents because their creators did not add them into the models to focus on. To enhance the efficiency of these AI tools, voice recognition developers must comprise a wider data set in terms of accents and dialects. This should be a focus in the next sequel for further studies to enable designers to develop tools that will work best for different learners.

5.5 Learner Engagement and Motivation

Perhaps one of the most beneficial aspects of this study was that there was a noticeable upturn in the learners' interest in the subject and their learning motivation, especially among the group of learners who earlier were unsuccessful under the conventional approaches to language acquisition. In a similar line of reasoning, Yang et al. (2023) have observed that AI-based gamification improves the interest of learners and this paper's participants did not differ from this trend. Due to the ability to engage with the AI tools and the instant feedback they provide, learners were motivated, similar to what Namaziandost and Rezai (2024) have pointed out, whereby real-time feedback helps learners see their progress or improvement which can be very motivating to learners.

However, as this study acknowledges, though the publicly posted interactions elicited high levels of engagement, it is important to look at the depth of that engagement. According to Khaleghi et al. (2021), learning with AI tools involves mainly gamification and engaging students through providing feedback, but such strategies enhance general learning instead of GBL procedures. Some of the participants in this research also described the AI tools as entertaining but complained that such feedback gave little substance. They explained that while the feedback encouraged the learners, it was not always specific enough from which they could deduce the nucleus of their error. This relates to Chan et al. (2024) and they have to find fault with the applicatory AI-based feedback systems, as they are seldom universal and encompassed in the given contexts to provide the learners with feedback by which they can develop.

Moreover, as a researcher, I am sceptical regarding the current emphasis on interaction levels in AI tools, as it is a cognitive achievement at the cost of real learning. However, engagement should not be traded for greater utility where AI systems could be better serving learners by offering more meaningful, constructively sound feedback. To my mind, the difficulty is to combine the fun aspect with feedback possibilities that offer a complete and more long-term learning experience. Concerning the future of AI in language learning, direction should be made towards effective interactivity and meaningful learning instead of simple interaction.

5.6 Challenges in Using AI Tools

From the study, the following difficulties in the application of AI tools were noted despite a positive impression. To start with, difficulties concerning the technical side, such as freezing and lagging, were also reported by the participants. This is in sync with Reich (2020) who noted that technical disruptions were among the major challenges ceding AI implementation in education. Some participants said that the AI tools stop working during the usage of the tools and more frequently when many people use the tools at the same time. These technical issues raised frustrations and interrupted the learning process More technical support and scalability were needed when utilising technologies within learning environments, concerning the AI technologies that were applied in the current study.

However, some participants complained about the limited number of listening tasks given by the AI tools, which they termed monotonous. Zhou & Hou (2024) also observed another major drawback of AI tools in language education; its activities consist of repetitive operations that may be uninteresting to learners after some time. Some participants observed that even if some tasks appeared to be stimulating at first, numerous similar listening activities caused them to become less motivated. This implies that designers of AI tools should include more variety and the learning paths that engage the learners should vary and at the same time complicate the learners. For instance, AI systems could be designed to provide learners with different tasks some of which are less complex and of different contents than others to accommodate the learner's needs.

Concerning the disadvantages of intelligent technologies in learning, participants' complaints about the absence of individual feedback is another essential constraint. According to VanLehn (2006), even though the reactions occur in real-time, AI systems rarely explain why a learner's answer is incorrect. This was evident in the current study where several learners wanted the teacher to explain why a particular answer was wrong. From my critical point of view, this limitation lays a fact that although AI can bring efficiency and scalability it does lack the depth created by human teachers. The primary implications for future iterations of such AI systems are to incorporate more and/or better feedback to guarantee that the system will be able to accommodate the learner as he or she may require it most.

5.7 Teachers in AI-Based Language Learning

Another revelation made by this research was that, whereas the AI application held several benefits for the learners, most of the participants wanted the support of the teachers for other difficult listening tasks. This contrasts with Al-thresher (2024) who reasoned that common AI language tools are useful as initial language development instruments, but cannot provide the kind of teaching and coaching human instructors offer especially for advancing language assignments. Some of the things that learners said included the following; Several learners complained that they need more specific instructions from the instructors whenever they face challenges with the complex listening materials. This implies that, although AI can help to teach basic skills and provide feedback for them instantly, it cannot

mimic teachers' instructional knowledge in higher-order thinking tasks.

The blending of AI tools with conventional approaches to learning languages could therefore be the most effective learning model. Kim (2024) supported AI as an important supplement to human teachers and tutors instead of competing with them. Therefore, the most suitable models for the constant improvement of language learning results will be the use of AI that will be integrated with human instructors and personal trainers who will provide students with support and understanding of the context of situations. AI can provide learners with even more opportunities to improve their practice and provide feedback instantaneously, while teachers can provide learners with more in-depth information or help learners develop their valuable problem-solving skills.

In addition, Kim et al. (2022) pointed out that there is an important role of the teacher in orchestrating the use of AI tools and facilitating learning about the technology. For instance, teachers could assist learners in the following ways: explaining the AI-given feedback to learners, presenting learners with more challenging listening exercises to complete, and enabling learners to review their learning progress. Such an approach would help in guaranteeing that learners get to coordinate with both the artificial intelligence and human tutor in the teaching process and also in developing a better understanding of the whole teaching strategy in learning language.

6. Conclusion

Nevertheless, the research results collectively emphasise the year-on-year prospects of AI instruments in promoting English language acquisition in Saudi Arabia. Engagement and motivation which have been exhibited by the learners indicate that AI tools can complement traditional language education by helping the learners in real-life listening comprehension tasks. Although it is argued that AI can be usefully integrated into classrooms, this paper has also shown, based on the literature (Benhadj et al., 2019), that such use must be sensitive to the structure of classroom learning and the   LL of the region.

As a result, this paper offers a critique of the use of AI tools and how they can transform listening comprehension among FLLs in Saudi Arabia. Preventing falls consists of a large component of listening so it may effectively enhance listening skills in learners, especially for real-life situations as well as increase learner engagement. Nevertheless, there are some obstacles; there is a technical problem, accent recognition, and the lack of individualized feedback. Such difficulties imply that the AI tools are not sufficiently developed to handle all the language learning difficulties in a setting like Saudi Arabia where students are from diverse linguistic backgrounds.

Prospects for future investigations should be directed at fine-tuning the AI devices to address various accents or to supply more specific feedback in addition to enhancing dependability in an extensive application of the systems. On the same note, AI tools should be made to enhance traditional teaching methods but not to replace them, rather to create a synthesis of the two teaching methods that are the best in the market. The findings of this study suggest that future research programs in the application of AI to educational contexts should adopt a multifaceted perspective considering the utility and the potential of artificial intelligence, while recognising its imperfections, for prioritising the conditions of learning a language.

7. Recommendations

The findings of this study have been useful in identifying the possibility of using AI tools to improve the listeners' understanding of a foreign language among the students of Saudi Arabia. This study showed that utilising AI applications enhances listening comprehension with the likelihood of using the result in real life, for instance, while watching news broadcasts and interviews. The learners noted improvement in motivation and participation, thus gained from the use of the AI applications because they use interactivity with quick feedback. However, the study also discussed important limitations such as difficulties in recognizing the accent, technical problems, and problems with feedback, which are rather shallow and should be worked through in further developments of AI. This research, therefore, fits into the current debate about AI in education, especially in language learning environments. It emphasizes the need to create AI systems that will better work with the different linguistic environment characteristics of Saudi Arabia and to develop these systems to give more targeted and detailed feedback. Also, the results indicate that, while AI tools can enhance traditional classroom practices, AI cannot perform all teaching-learning activities that human teachers do, perhaps the complicated ones. Based on this analysis, the combination of an AI approach with traditional human instruction seems to be the most promising in the future for language learning.

Authors' contributions

Wafa' A. Hazaymeh and Turky Alshaikhi carried out the experiment. Mohammad Osman Abdul Wahab wrote the manuscript with support from Mohamad A. Khasawneh, Turky Alshaikhi and Wafa' A. Hazaymeh. Turky Alshaikhi helped supervise the project.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The author obtained permission from the Deanship of Scientific Research at King Khalid University for funding this work through Large Research Groups under grant number (RGP.2 / 68 /64).

Provenance and peer review

Not applicable

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

Abdellatif, M. S., Alshehri, M. A., Alshehri, H. A., Hafez, W. E., Gafar, M. G., & Lamouchi, A. (2024). I am all ears: listening exams with AI and its traces on foreign language learners' mindsets, self-competence, resilience, and listening improvement. *Language Testing in Asia*, *14*(1), 54. https://doi.org/10.1186/s40468-024-00329-6

AbuSahyon, A. S. A. E., Alzyoud, A., Alshorman, O., & Al-Absi, B. (2023). AI-driven technology and Chatbots as tools for enhancing English language learning in the context of second language acquisition: a review study. *International Journal of Membrane Science and Technology*, *10*(1), 1209-1223.

Al-khresheh, M. H. (2024). Bridging technology and pedagogy from a global lens: Teachers' perspectives on integrating ChatGPT in English language teaching. *Computers and Education: Artificial Intelligence*, *6*, 100218. https://doi.org/10.1016/j.caeai.2024.100218

Almaleki, W. S. A. (2020). *Saudi international students' perceptions of the utility of artificial intelligence and intelligent personal assistant tools in EFL learning* (Doctoral dissertation, Concordia University Chicago).

Alotaibi, N. S., & Alshehri, A. H. (2023). Prospers and obstacles in using artificial intelligence in Saudi Arabia higher education institutions—The potential of AI-based learning outcomes. *Sustainability*, *15*(13), 10723. https://doi.org/10.3390/su151310723

Auer, P. (2004). Non-standard evidence in syntactic typology-Methodological remarks on the use of dialect data vs spoken language data. *Trends in Linguistics Studies and Monographs*, *153*, 69-92. https://doi.org/10.1515/9783110197327.69

Barreno, M., Nelson, B., Sears, R., Joseph, A. D., & Tygar, J. D. (2006, March). Can machine learning be secure?. In *Proceedings of the 2006 ACM Symposium on Information, computer and communications security* (pp. 16-25). https://doi.org/10.1145/1128817.1128824

Benhadj, Y., El Messaoudi, M., & Nfissi, A. (2019). Artificial intelligence in education: Integrating serious gaming into the language class classdojo technology for classroom behavioral management. *IAES International Journal of Artificial Intelligence*, *8*(4), 382.

Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y. S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *2*, 100027. https://doi.org/10.1016/j.caeai.2021.100027

Chan, S., Lo, N., & Wong, A. (2024). Generative AI and Essay Writing: Impacts of Automated Feedback on Revision Performance and Engagement. *rEFLections*, *31*(3), 1249-1284. https://doi.org/10.61508/refl.v31i3.277514

Chen, C. C., Hung, P., Eğrioğlu, E., & Hsiao, K. L. (Eds.). (2022). *Deep Learning in Adaptive Learning: Educational Behavior and Strategy*. Frontiers Media SA.

Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers*, *25*(1), 161-182. https://doi.org/10.1007/s10796-022-10291-4

Diwan, C., Srinivasa, S., Suri, G., Agarwal, S., & Ram, P. (2023). AI-based learning content generation and learning pathway augmentation to increase learner engagement. *Computers and Education: Artificial Intelligence*, *4*, 100110. https://doi.org/10.1016/j.caeai.2022.100110

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... Williams, M. D. (2021). Artificial Intelligence (AI):

Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International journal of information management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., ... Gašević, D. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13544

Gajos, K. Z., & Mamykina, L. (2022, March). Do people engage cognitively with AI? Impact of AI assistance on incidental learning. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (pp. 794-806). https://doi.org/10.1145/3490099.3511138

Gilakjani, A. P., & Ahmadi, M. R. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement.

Hassani, K., Nahvi, A., & Ahmadi, A. (2016). Design and implementation of an intelligent virtual environment for improving speaking and listening skills. *Interactive Learning Environments*, *24*(1), 252-271. https://doi.org/10.1080/10494820.2013.846265

Huang, F., Wang, Y., & Zhang, H. (2024). Modelling Generative AI Acceptance, Perceived Teachers' Enthusiasm and Self-Efficacy to English as a Foreign Language Learners' Well-Being in the Digital Era. *European Journal of Education*, *59*(4), e12770. https://doi.org/10.1111/ejed.12770

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.

Itzchakov, G., Weinstein, N., & Cheshin, A. (2023). Learning to listen: Downstream effects of listening training on employees' relatedness, burnout, and turnover intentions. *Human Resource Management*, *62*(4), 565-580. https://doi.org/10.1186/s40468-024-00329-6

Jung, E., Kim, D., Yoon, M., Park, S., & Oakley, B. (2019). The influence of instructional design on learner control, sense of achievement, and perceived effectiveness in a supersize MOOC course. *Computers & Education*, *128*, 377-388. https://doi.org/10.1016/j.compedu.2018.10.001

Kamm, C. A., Dirks, D. D., & Bell, T. S. (1985). Speech recognition and the articulation index for normal and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *77*(1), 281-288. https://doi.org/10.1121/1.392269

Khaleghi, A., Aghaei, Z., & Mahdavi, M. A. (2021). A gamification framework for cognitive assessment and cognitive training: qualitative study. *JMIR serious games*, *9*(2), e21900. https://doi.org/10.2196/21900

Kim, J. (2024). Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies*, *29*(7), 8693-8724. https://doi.org/10.1007/s10639-023-12109-5

Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, *27*(5), 6069-6104. https://doi.org/10.1007/s10639-021-10831-6

Leonardo, Z. M. (1998). *Ideology, discourse, and school reform*. University of California, Los Angeles.

Li, Y., Hashim, A. S., Lin, Y., Nohuddin, P. N., Venkatachalam, K., & Ahmadian, A. (2024). AI-based visual speech recognition towards realistic avatars and lip-reading applications in the metaverse. *Applied Soft Computing*, *164*, 111906. https://doi.org/10.1016/j.asoc.2024.111906

Mamun, K. A., Nabid, R. A., Pranto, S. I., Lamim, S. M., Rahman, M. M., Mahammed, N., ... Khan, R. R. (2024). Smart reception: An artificial intelligence driven bangla language based receptionist system employing speech, speaker, and face recognition for automating reception services. *Engineering Applications of Artificial Intelligence*, *136*, 108923. https://doi.org/10.1016/j.engappai.2024.108923

Meijer, P. C., Korthagen, F. A., & Vasalos, A. (2009). Supporting presence in teacher education: The connection between the personal and professional aspects of teaching. *Teaching and teacher education*, *25*(2), 297-308. https://doi.org/10.1016/j.tate.2008.09.013

Namaziandost, E., & Rezai, A. (2024). Interplay of academic emotion regulation, academic mindfulness, L2 learning experience, academic motivation, and learner autonomy in intelligent computer-assisted language learning: A study of EFL learners. *System*, *125*, 103419. https://doi.org/10.1016/j.system.2024.103419

Novawan, A., Walker, S. A., & Ikeda, O. (2024). The new face of technology-enhanced language learning (TELL) with artificial intelligence (AI): Teacher perspectives, practices, and challenges. *Journal of English in Academic and Professional Communication*, *10*(1), 1-18. https://doi.org/10.25047/jeapco.v10i1.4565

Otte, J. L. (2006). *Real language for real people: A descriptive and exploratory case study of the outcomes of aural authentic texts on the listening comprehension of adult English-as-a-second language students enrolled in an advanced ESL listening course* (Doctoral dissertation, Loyola University Chicago).

Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable development*.

Pinho, C., Franco, M., & Mendes, L. (2021). Application of innovation diffusion theory to the E-learning process: higher education context. *Education and Information Technologies*, *26*(1), 421-440. https://doi.org/10.1007/s10639-020-10269-2

Rahko, L. (2017). *Transfer of Colloquial Elements in Translation from Finnish Speech into English Subtitles. A Case Study of the Television Series Pasila.*

Reich, J. (2020). *Failure to disrupt: Why technology alone can't transform education*. Harvard University Press.

Richards, J. C. (2008). *Teaching listening and speaking* (Vol. 35). Cambridge, England: Cambridge university press.

Rivera, E. S., & Garden, C. L. P. (2021). Gamification for student engagement: a framework. *Journal of further and higher education*, *45*(7), 999-1012. https://doi.org/10.1080/0309877X.2021.1875201

Sapir, E. (1921). An introduction to the study of speech. *Language*, *1*, 15.

Schmidt, T., & Strasser, T. (2022). Artificial intelligence in foreign language learning and teaching: a CALL for intelligent practice. *Anglistik: International Journal of English Studies*, *33*(1), 165-184.

Shroff, R. H., Ting, F. S. T., & Lam, W. H. (2019). Development and validation of an instrument to measure students' perceptions of technology-enabled active learning. *Australasian Journal of Educational Technology*, *35*(4). https://doi.org/10.14742/ajet.4472

Strambi, A., & Bouvet, E. (2003). *Flexibility and interaction at a distance: A mixed-mode environment for language learning.*

Stuart-Smith, J., Pryce, G., Timmins, C., & Gunter, B. (2013). Television can also be a factor in language change: Evidence from an urban dialect. *Language*, 501-536.

Vandergrift, L. (2011). Second language listening: Presage, process, product, and pedagogy. In *Handbook of research in second language teaching and learning* (pp. 455-471). Routledge.

VanLehn, K. (2006). The behavior of tutoring systems. *International journal of artificial intelligence in education*, *16*(3), 227-265.

Wilson, B. G. (1996). *Constructivist learning environments: Case studies in instructional design*. Educational Technology.

Yang, Q. F., Lian, L. W., & Zhao, J. H. (2023). Developing a gamified artificial intelligence educational robot to promote learning effectiveness and behavior in laboratory safety courses for undergraduate students. *International journal of educational technology in higher education*, *20*(1), 18. https://doi.org/10.1186/s41239-023-00391-9

Yu, C., Wu, L., Li, J., & Li, S. (2023). English Listening Teaching Mode under Artificial Intelligence Speech Synthesis Technology. *ACM Transactions on Asian and Low-Resource Language Information Processing*. https://doi.org/10.1145/3615866

Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, *4*, 100134. https://doi.org/10.1016/j.caeai.2023.100134

Zhou, C., & Hou, F. (2024). Can AI empower L2 education? Exploring its influence on the behavioural, cognitive and emotional engagement of EFL teachers and language learners. *European Journal of Education*, *59*(4), e12750. https://doi.org/10.1111/ejed.12750