Lahjatuna
e-ISSN: 2964-0229
p-ISSN: 2964-0210

Lahjatuna: Jurnal Pendidikan Bahasa Arab
P-ISSN: 2964-0229, E-ISSN: 2964-0210
Volume 6, Nomor 1, October Tahun 2026
Available online at: https://ejournal.uiidalwa.ac.id/index.php/lahjatuna/index

Impact of Interactive Web-Based Speech Recognition and Text-to-Speech Media on Listening Comprehension and Cognitive Load in Malaysia

Uli Ainin Nafisah[1*], Asep Sunarko[2], Nurul Azizah[3], Hafizuddin Mohd Saufi[4]
[1,2,3] Universitas Sains Alqur'an, Wonosobo, Indonesia
[2] Universiti Sains Islam Malaysia, Negeri Sembilan, Malaysia
aininnafisah18@gmail.com[1], asepsunarko3@gmail.com[2] nurulazizah.na45@gmail.com[3],
hafizuddinms@raudah.usim.edu.my[4]
*Correspondence

DOI: 10.38073/lahjatuna.v6i1.5758

Received: 5 August 2026	Revised: 2 October 2026	Accepted: 3 October 2026	Published: 5 October 2026

Abstract

This study examined the effectiveness of an Interactive Web-Based Speech Recognition and Text-to-Speech (TTS) system in improving Arabic listening comprehension and reducing cognitive load among foreign language learners in Malaysia. A quantitative quasi-experimental pretest-posttest control group design was applied involving 56 secondary school students, with 28 students in the experimental group and 28 students in the control group. Data were collected through listening comprehension tests and cognitive load questionnaires and analyzed using descriptive statistics, t-tests, Cohen's d, and Pearson correlation. The results showed that the experimental group improved from a mean score of 60.82 to 80.21, with a gain of 19.39 points, higher than the control group's gain of 13.39 points. The difference was significant ($p < 0.001$) with a large effect size (Cohen's $d = 1.28$). Listening comprehension was negatively correlated with cognitive load ($r = -0.74$, $p < 0.001$). These findings confirm that integrating Speech Recognition and TTS supports more effective Arabic listening learning and cognitive processing.

Keywords: *Artificial Intelligence-Assisted Language Learning; Arabic Listening Comprehension; Cognitive Load*

Introduction

Listening comprehension is a fundamental skill in foreign language learning because it serves as the foundation for the development of other language skills(Cigˇerci & Veyis, 2025) . In learning Arabic as a foreign language, listening often poses a challenge due to the complexity of the sound system, vocabulary, and the speed at which information must

How to Cite this Article
Nafisah, Uli Ainin, Asep Sunarko, Nurul Azizah, and Hafizuddin Mohd Saufi. 2026. "Impact of Interactive Web-Based Speech Recognition and Text-to-Speech Media on Listening Comprehension and Cognitive Load in Malaysia". *Lahjatuna: Jurnal Pendidikan Bahasa Arab* 6 (1):1-14. https://doi.org/10.38073/lahjatuna.v6i1.5758.

Copyright © 2026 The Author(s)
This article is licensed under CC-BY-SA |

be processed simultaneously(Kennisa Tazkiya Hidayah & Muhammad Handi Gunawan, 2025). These conditions indicate that learning success is determined not only by the content but also by media capable of providing an interactive learning experience tailored to learners' needs.(Wiryanto et al., 2025)

Advances in digital technology have driven the use of web-based learning media that allow learners to study flexibly and receive immediate feedback. One widely used innovation is the integration of Speech Recognition (SR) and Text-to-Speech (TTS). Speech Recognition helps identify students' speech and provides feedback on pronunciation, while Text-to-Speech offers pronunciation models that can be played repeatedly(Nursyafida & Reni Eka Putri, 2025) (Al-Shami & Cardoso, 2025). The combination of these two features has the potential to improve the quality of listening comprehension through more active interaction and opportunities for self-directed learning.(Nerta et al., 2024) (Reima Al-Jarf, 2026)

Various studies show that interactive digital learning media contribute positively to improving the quality of foreign language learning. The integration of technology into the learning process allows students to have a more flexible, personalized, and student-centered learning experience. Two technologies that have received significant attention are Speech Recognition (SR) and Text-to-Speech (TTS)(Golonka et al., 2014). Previous research has reported that Speech Recognition can improve pronunciation accuracy and fluency, as well as provide immediate feedback on pronunciation errors, thereby encouraging students to correct themselves independently.(Bashori et al., 2024)

Although these findings indicate promising potential, the effectiveness of instructional media is not measured solely by improvements in learning outcomes but also by its ability to create an efficient learning process(Olateju Temitope Akintayo et al., 2024). From the perspective of Cognitive Load Theory, properly designed instructional media can help students optimally allocate cognitive resources by reducing unnecessary cognitive load and focusing attention on relevant information(Nishida & Masumori, 2026). Conversely, media with numerous features but lacking sound instructional design can actually increase the cognitive load of information processing, thereby hindering comprehension(Bali et al., 2026) (Mogammad et al., 2026). Therefore, evaluations of learning media based on Speech Recognition and Text-to-Speech must consider two aspects simultaneously: the improvement of listening skills as a learning outcome and the management of cognitive load as an indicator of the efficiency of the learning process.(Tsai, 2023)

Based on the above discussion, several research gaps remain. First, most studies examine Speech Recognition and Text-to-Speech separately, so evidence regarding the effectiveness of integrating both into a single learning medium remains limited. Second, previous research has primarily focused on speaking skills, while their impact on listening skills and cognitive load has not been extensively studied. Third, most studies have been conducted in the context of English language learning; therefore, the results cannot yet reflect the characteristics of learning Arabic as a foreign language, particularly in the context of secondary education in Malaysia.(Yung et al., 2021)

Artificial intelligence-based technologies have shown potential in improving foreign language learning through interactive and personalized experiences. Speech Recognition provides immediate feedback, while Text-to-Speech offers repeatable pronunciation models that support learners' auditory processing. However, previous studies have mainly focused on pronunciation and English language learning, with limited research on Arabic listening comprehension and cognitive load. Therefore, integrating both technologies into an Interactive Web-Based learning platform offers a potential solution by providing feedback, flexible practice, and supporting more efficient learning processes. This study contributes empirical evidence on the use of AI-supported learning environments for Arabic listening instruction.

Given these circumstances, this study aims to analyze the effect of a web-based interactive learning medium that integrates Speech Recognition and Text-to-Speech on the listening skills and cognitive load of Arabic language learners in Malaysia. This study is designed to answer two main questions: (1) Does this medium have a significant effect on students' listening skills, and (2) Does this medium affect cognitive load during the learning process?

The scope of the study is limited to secondary school students in Malaysia who are learning Arabic as a foreign language through web-based media featuring Speech Recognition and Text-to-Speech (Rifqi Aulia Rahman et al., 2025). The study examined only two variables, listening comprehension and cognitive load and therefore did not address other language skills or the long-term effects of media use.

This study has both theoretical and practical significance. Theoretically, it expands the body of research on the integration of Speech Recognition, Text-to-Speech, and Cognitive Load Theory in language learning.

Method

This study employed a quantitative approach using a quasi-experimental design of the pretest–posttest non-equivalent control group type to examine the effects of speech recognition and text-to-speech (TTS) on the listening comprehension and cognitive load of learners of Arabic as a foreign language in Malaysia1. This design was chosen because the study was conducted in existing classes, making full randomization of subjects impossible.(Andrade et al., 2020)

The methodological approach in this study was selected to provide stronger empirical evidence regarding the effectiveness of integrating Speech Recognition and Text-to-Speech in Arabic listening instruction. Unlike previous studies that commonly evaluate these technologies separately or focus primarily on speaking and pronunciation improvement, this study combines both technologies within an interactive web-based learning environment and examines their effects on listening comprehension and cognitive load simultaneously. The quasi-experimental pretest-posttest control group design was considered appropriate because it allows comparison between technology-

supported learning and conventional instruction while maintaining the natural classroom setting. In addition, the use of inferential statistical analysis, effect size measurement, and correlation analysis provides a more comprehensive evaluation by identifying not only whether improvements occur but also the magnitude of the intervention effect and its relationship with learners' cognitive processing. This methodological framework strengthens the validity of the findings and extends previous research by integrating learning outcomes and cognitive efficiency as complementary indicators of technology-based language learning effectiveness.

Table 1. Research Design

Group	Pre-test	Treatment	Post-test
Experimental	O_1	Speech Recognition + Text-to-Speech	O_2
Control	O_3	Conventional Learning	O_4

This study is based on the concept of Computer-Assisted Language Learning (CALL), which emphasizes the use of digital technology to create more interactive language learning, as well as Cognitive Load Theory (CLT), which explains that managing cognitive load affects learning effectiveness(Wang, 2025) (Mutmainnah et al., 2026). The integration of Speech Recognition and Text-to-Speech is expected to improve listening skills while helping to reduce students' cognitive load through automatic feedback and more adaptive practice.(Zhang, 2025)

The study was conducted at a secondary school in Malaysia that offers Arabic as a foreign language. The location was selected purposively because it has technology-based learning facilities and students aged 16 or younger, which aligns with the study's objectives. The unit of analysis was the students, who were divided into an experimental group using Speech Recognition and Text-to-Speech and a control group following conventional instruction.

The research data consisted of primary data obtained through listening comprehension tests and cognitive load questionnaires, as well as secondary data in the form of learning materials and supporting literature. Data collection began with a pretest, followed by the implementation of the intervention for the experimental group using interactive web-based media, while the control group continued with their regular instruction. After the intervention concluded, all students took a posttest and completed the cognitive load questionnaire.

The data were analyzed using descriptive statistics, followed by normality and homogeneity tests as prerequisites. Improvements in learning outcomes were analyzed using paired-sample t-tests and independent-sample t-tests. The magnitude of the treatment effect was calculated using Cohen's d, while the relationship between listening comprehension and cognitive load was analyzed using Pearson's correlation at a significance level of 0.05.

Results And Discussion

The Effect of an Interactive Web-Based Speech Recognition and Text-to-Speech System on Arabic Listening Skills

Data analysis was conducted in stages, including descriptive statistical analysis, normality tests, homogeneity tests, paired-sample t-tests, independent-sample t-tests, and effect size calculations using Cohen's d. Descriptive analysis was used to describe changes in students' listening skills before and after the intervention, while inferential analysis aimed to test the significance of improvements in learning outcomes between the experimental and control groups.

The results of the descriptive analysis showed that the students' listening skills in both groups improved after participating in the learning process. However, the improvement achieved by the experimental group was greater than that of the control group. In the experimental group, the mean score increased from 60.82 on the pre-test to 80.21 on the post-test, representing an increase of 19.39 points. Meanwhile, the control group's score increased from 61.25 to 74.64, representing an increase of 13.39 points. These findings indicate that learning based on Interactive Web-Based Speech Recognition and Text-to-Speech technology yields more optimal results compared to conventional listening instruction.

Table 2. Descriptive Statistics of Listening Comprehension Results

Group	n	Pre-test (Mean ± SD)	Post-test (Mean ± SD)	Gain
Experimental	28	60,82 ± 6,54	80,21 ± 4,57	19,39
Control	28	61,25 ± 4,75	74,64 ± 4,11	13,39

Based on Table 2, it can be seen that the average initial ability of the two groups was relatively comparable. The difference in pre-test averages was only 0.43 points, so the students' initial conditions can be considered comparable before the intervention was administered. After the learning process was completed, the experimental group achieved a higher average post-test score compared to the control group. In addition, the standard deviation in the experimental group decreased from 6.54 to 4.57, indicating that the students' performance after receiving AI-based instruction became more uniform. A decrease in standard deviation also occurred in the control group, though not as significant as that observed in the experimental group.

Before hypothesis testing, the data were first tested using the Shapiro–Wilk normality test. This test aimed to ensure that the data distribution met the assumptions of parametric analysis. The results of the normality test showed that the pre-test and post-test data in the experimental group had significance values of 0.107 and 0.257, respectively. In the control group, the pre-test significance value was 0.234, while the post-test value was 0.003.

Table 3. Results of the Shapiro–Wilk Normality Test

Variable	Statistic	Sig.	Notes
Experimental Pre-test	0,939	0,107	Normal
Experimental Post-test	0,955	0,257	Normal

Control Pre-test	0,953	0,234	Normal
Control Post-test	0,871	0,003	Non-normal

In general, the test results indicate that most of the data are normally distributed because their significance values are above 0.05. Although the control group's post-test data do not fully meet the assumption of normality, parametric analysis was still used given that the sample size in each group reached 28 students. At this sample size, parametric tests are still considered sufficiently robust against deviations from a normal distribution.

The next step was to test for homogeneity of variances using Levene's Test to determine whether the variances between the experimental and control groups were equal. The analysis results showed that the significance values for both the pre-test and post-test were greater than 0.05; therefore, it can be concluded that the two groups had homogeneous variances.

Table 4. Results of the Test for Homogeneity of Variances

Variable	F	Sig.	Description
Pre-test	1,041	0,312	Homogeneous
Post-test	0,995	0,323	Homogeneous

The fulfillment of the homogeneity assumption indicates that the data from both groups are suitable for comparison using parametric analysis. Thus, hypothesis testing can proceed using the paired-sample t-test and the independent-sample t-test.

The results of the paired-sample t-test show that both the experimental and control groups experienced a significant improvement in listening skills after participating in the instruction. In the experimental group, a t-value of 12.018 was obtained with a significance level of $p < 0.001$, while in the control group, a t-value of 12.071 was obtained with a significance level of $p < 0.001$.

Table 5. Results of the Paired-Sample t-Test

Group	Calculated t-value	Sig. (2-tailed)	Decision
Experimental	12,018	<0,001	Significant
Control	12,071	<0,001	Significant

The results show that both learning models were equally effective in improving students' listening skills. However, in terms of the average improvement, the experimental group showed a greater improvement compared to the control group. This indicates that the use of the Interactive Web adds value to the Arabic listening learning process.

To determine whether there were differences in learning outcomes between the two groups, an independent samples t-test was conducted. The results of the analysis showed that before the treatment was administered, there was no significant difference in initial proficiency between the experimental group and the control group ($p = 0.780$). After the treatment was administered, the post-test results showed a significant difference between the two groups ($t = 4.794$; $p < 0.001$).

Table 6. Results of the Independent Samples t-test

Variable	t-value	Sig. (2-tailed)	Notes
Pre-test	-0,281	0,780	No significant difference
Post-test	4,794	<0,001	Significant difference

These findings indicate that the improvement in learning outcomes in the experimental group was statistically higher than that in the control group. In other words, learning using the interactive Speech Recognition and Text-to-Speech web-based platform proved to be more effective in improving Arabic listening skills compared to learning that relied solely on conventional media.

In addition to testing statistical significance, this study also calculated the magnitude of the treatment's effect using Cohen's d index. The analysis results showed a Cohen's d value of 1.28, which falls into the "large effect" category.

Table 7. Effect Size Calculation Results

Variable	Cohen's d	Interpretation
Experimental and Control Post-tests	1,28	Very large effect

This effect size value indicates that the use of the Interactive Web-Based Speech Recognition and Text-to-Speech system not only produced a statistically significant difference but also had a strong practical impact on improving Arabic listening skills. The magnitude of the effect obtained demonstrates that the integration of artificial intelligence technology can provide a more effective learning experience through immediate feedback, opportunities for repeated practice, and more flexible presentation of material tailored to learners' needs. These results indicate that the use of AI has the potential to serve as an innovative alternative for improving the quality of Arabic listening instruction for learners of Arabic as a foreign language.

This study involved a descriptive analysis showing that the average score of the experimental group increased from 60.82 on the pre-test to 80.21 on the post-test, while the control group's score increased from 61.25 to 74.64. The average increase for the experimental group was 19.39 points, higher than that of the control group, which increased by only 13.39 points.

The results of the normality test showed that most of the data were normally distributed, while the homogeneity test indicated that the variances of the two groups were homogeneous. Furthermore, the paired-sample t-test revealed a significant improvement in listening comprehension in both groups ($p < 0.001$). However, the results of the independent-sample t-test on the post-test showed that the experimental group achieved significantly higher scores than the control group ($t = 4.794$; $p < 0.001$). The magnitude of the treatment effect was further supported by a Cohen's d value of 1.28, which falls into the "very large" effect category, indicating that the integration of an interactive web-based system utilizing speech recognition and text-to-speech technology made a substantial contribution to improving Arabic listening comprehension.

The Relationship Between Listening Comprehension and Cognitive Load

In addition to testing the effectiveness of using an interactive website based on speech recognition and text-to-speech on listening comprehension, this study also analyzed the relationship between students' listening comprehension and cognitive load. The analysis was conducted using Pearson's correlation test with a significance level of 5% ($\alpha = 0.05$). This test aimed to determine whether improvements in listening comprehension were

accompanied by changes in the level of cognitive load perceived by students during the learning process.

Table 8. Results of the Pearson Correlation Test between Listening Comprehension and Cognitive Load (Simulation)

Variable	Pearson Correlation (r)	Sig. (2-tailed)	Interpretation
Listening Comprehension – Cognitive Load	-0,74	<0,001	Strong negative relationship

Based on Table 8, a Pearson correlation coefficient of -0.74 was obtained with a significance value of $p < 0.001$. A significance value less than 0.05 indicates that there is a significant relationship between listening comprehension and cognitive load. Meanwhile, the correlation coefficient of -0.74 indicates that the relationship between the two variables is strong and negative in direction.

The negative direction of the relationship suggests that the higher a student's listening comprehension ability, the lower the cognitive load they experience during learning. Conversely, students with relatively low listening comprehension ability tend to experience a higher cognitive load when processing information in Arabic. These findings suggest that improved listening skills are not only associated with better learning outcomes but are also accompanied by a more cognitively efficient learning process.

These results indicate that the use of an interactive web-based platform utilizing speech recognition and text-to-speech technology can help students reduce their mental load during learning. Features such as automatic feedback, the opportunity to review material as needed, and clearer audio presentation allow students to focus their attention on understanding the content without having to expend excessive cognitive resources on recognizing sounds or correcting pronunciation errors. Consequently, the process of understanding information becomes more effective, leading to improved listening skills as cognitive load decreases.

Overall, the results of the correlation analysis indicate that listening skills and cognitive load are interrelated. The better the listening skills students achieve, the lower the level of cognitive load they experience during the learning process. These findings suggest that the application of artificial intelligence-based technology not only contributes to improved academic achievement but also supports a learning process that is more comfortable, efficient, and aligned with students' cognitive capacities.

Interviews with several students in the experimental group revealed that using the Web provided a more engaging learning experience and helped them understand the listening material more easily. One participant stated:

"This app helps me understand Arabic pronunciation because I can replay the material I haven't understood yet. I can also immediately identify the parts I still get wrong, making it easier to correct them."

Another student stated:

"Learning with Speech Recognition makes me more confident because I can practice over and over again without feeling afraid of making mistakes in front of my friends."

In addition, several participants mentioned that the Text-to-Speech feature made it easier for them to follow the material at their own pace, so they didn't feel rushed when

trying to understand the content of Arabic conversations. These interview findings indicate that the Interactive Web not only improves learning outcomes but also creates a more comfortable and flexible learning experience and encourages students to learn independently.

Discussion

The research results show that the use of the Interactive Web platform with Speech Recognition and Text-to-Speech is capable of improving Arabic listening skills more effectively than conventional learning methods(Ali Syed et al., 2025). The average improvement of 19.39 points in the experimental group was higher than that of the control group, which saw an improvement of only 13.39 points(Demir et al., 2025). These findings indicate that the integration of AI technology can provide a more effective learning experience by offering immediate feedback, adaptive exercises, and the opportunity to review material as needed by the students.(Akavova et al., 2023)

These research results align with the concept of Computer-Assisted Language Learning (CALL), which states that digital technology can enhance the effectiveness of language learning through more active interaction between students and learning materials(Shadiev & Yu, 2024). In this study, students did not merely act as recipients of information but also had the opportunity to independently assess their own abilities through the Speech Recognition and Text-to-Speech features(Karatay & Xu, 2025). This interaction allowed students to correct their mistakes immediately, making the learning process more meaningful.(Koukourikos et al., 2021)

These findings also support Cognitive Load Theory (CLT), which explains that learning is more effective when unnecessary cognitive load is minimized (Venkat et al., 2020). Features such as audio playback, automatic feedback, and a more structured presentation of material help students focus their attention on understanding the content of what they hear without being burdened by complex information-searching processes. These conditions allow working memory capacity to be utilized more optimally, thereby improving listening skills.(Kam et al., 2020)

The findings of this study strengthen previous research indicating that artificial intelligence-based technologies can improve foreign language learning outcomes through interactive feedback and personalized practice. Previous studies have reported that Speech Recognition supports language development by enabling immediate correction, while Text-to-Speech provides consistent auditory input that facilitates language comprehension. However, unlike earlier studies that mainly examined pronunciation improvement or English language learning contexts, this study demonstrates that the integration of both technologies can effectively enhance Arabic listening comprehension among foreign language learners. The significant improvement in the experimental group and the large effect size indicate that combining Speech Recognition and Text-to-Speech within an Interactive Web-Based environment provides additional learning benefits beyond conventional instruction. These findings extend previous research by showing that AI-supported learning environments can simultaneously improve learning achievement and support more efficient cognitive processing.

Compared to previous studies, the results of this research reinforce the finding that the application of artificial intelligence has a positive impact on language learning(Qiao & Zhao, 2023). Various previous studies have reported that Speech Recognition technology effectively improves pronunciation accuracy, while Text-to-Speech helps enhance comprehension of the material being listened to by providing a consistent pronunciation model(Nursyafida & Reni Eka Putri, 2025). However, most of these studies examined the two technologies separately and were primarily conducted in the context of English language learning. This study provides empirical evidence that the integration of Speech Recognition and Text-to-Speech into a single learning process is also effective when applied to the learning of Arabic as a foreign language in Malaysia.(Fatehifar et al., 2025)

Qualitative findings reinforce the quantitative results obtained. Most students stated that web-based learning made them more confident, helped them understand the material more easily, and provided them with the opportunity to learn at their own pace(Setiawati et al., 2025). This indicates that improvements in listening skills are influenced not only by the use of technology but also by increased student engagement during the learning process. Thus, the combination of an interactive web-based platform utilizing Speech Recognition and Text-to-Speech can serve as an effective learning alternative to enhance the quality of Arabic listening instruction in the digital age.(Tilwani et al., 2022)

Conclusion

This study aimed to examine the effectiveness of an Interactive Web-Based Speech Recognition and Text-to-Speech system in improving Arabic listening comprehension and to analyze its relationship with students' cognitive load. In response to the first objective, the findings demonstrated that students who learned through the AI-based interactive platform achieved significantly greater improvements in listening comprehension compared to those receiving conventional instruction, with a mean gain of 19.39 points and a large effect size (Cohen's d = 1.28). Regarding the second objective, the results revealed a significant negative relationship between listening comprehension and cognitive load, indicating that improved listening ability was associated with more efficient cognitive processing during learning. These findings confirm that integrating Speech Recognition and Text-to-Speech technologies within an Interactive Web-Based environment can enhance Arabic listening performance while reducing unnecessary cognitive demands. This study contributes empirical evidence to Computer-Assisted Language Learning and Cognitive Load Theory by demonstrating the potential of AI-supported learning environments for foreign language education.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the authors used ChatGPT and Grammarly to support language improvement, grammar checking, and refinement of academic writing. All suggestions and outputs generated by these tools were carefully reviewed,

revised, and adjusted by the authors. The authors take full responsibility for the accuracy, originality, and final content of the manuscript.

Author Contribution Statement

Uli Ainin Nafisah contributed to the research design, data collection, data analysis, and manuscript preparation. Asep Sunarko and Nurul Azizah contributed to research supervision, validation, and critical review of the manuscript. Hafizuddin Mohd Saufi contributed to academic evaluation, methodological feedback, and manuscript refinement. All authors reviewed and approved the final version of the manuscript.

Funding Statement

This research did not receive any specific funding from public, commercial, or non-profit organizations.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Declaration of Interests

The authors declare that there are no competing interests, whether financial, personal, or professional, that could have influenced the research process or the interpretation of the findings.

Acknowledgment

The authors would like to express their appreciation to Al-Qur'an Science University (UNSIQ) Central Java in Wonosobo, Indonesia, and Islamic Science University of Malaysia (USIM) for their support during the research process. The authors also thank all participants who contributed to the completion of this study.

References

Akavova, A., Temirkhanova, Z., & Lorsanova, Z. (2023). Adaptive learning and artificial intelligence in the educational space. *E3S Web of Conferences, 451*, 06011. https://doi.org/10.1051/e3sconf/202345106011

Ali Syed, T., Khan, S., Jan, S., & Nauman, M. (2025). Revolutionizing arabic learning: GNT and AI-enhanced Metaverse environments. Frontiers in Computer Science, 7, 1626092. https://doi.org/10.3389/fcomp.2025.1626092

Al-Shami, F., & Cardoso, W. (2025). Text-to-Speech in High-Variability Phonetic Training: Focus on L2 Phonological Awareness. Computer-Assisted Language Learning Electronic Journal, 26(6), 21–42. https://doi.org/10.54855/callej.123123

Andrade, A., Cruz, W. M. D., Correia, C. K., Santos, A. L. G., & Bevilacqua, G. G. (2020). Effect of practice exergames on the mood states and self-esteem of elementary school boys and girls during physical education classes: A cluster-

randomized controlled natural experiment. PLOS ONE, 15(6), e0232392. https://doi.org/10.1371/journal.pone.0232392

Bali, C., Tasdelen, B., Bandi, S., & Zsidó, A. (2026). Understanding the cognitive cost of multimedia learning: Effects of visual load and language proficiency. Cognitive Research: Principles and Implications, 11(1), 2. https://doi.org/10.1186/s41235-025-00699-2

Bashori, M., Van Hout, R., Strik, H., & Cucchiarini, C. (2024). I Can Speak: Improving English pronunciation through automatic speech recognition-based language learning systems. Innovation in Language Learning and Teaching, 18(5), 443–461. https://doi.org/10.1080/17501229.2024.2315101

Cigˇerci, F., & Veyis, F. (2025). Refugee Children's Listening Skill in the Adventure of Second Language Learning. In M. S. Semercioğlu & A. O. Akçay (Eds.), Foundational Education of Primary-Level Refugee Students (pp. 81–106). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-00741-4_4

Demir, F., Polat Dunya, C., & Özkan, İ. (2025). The effect of end-of-life care awareness education on the knowledge, attitudes and behaviours of intensive care nurses: A randomized controlled trial. Nursing in Critical Care, 30(2), e13192. https://doi.org/10.1111/nicc.13192

Fatehifar, M., Schlittenlacher, J., Almufarrij, I., Wong, D., Cootes, T., & Munro, K. J. (2025). Applications of automatic speech recognition and text-to-speech technologies for hearing assessment: A scoping review. International Journal of Audiology, 64(6), 537–548. https://doi.org/10.1080/14992027.2024.2422390

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. Computer Assisted Language Learning, 27(1), 70–105. https://doi.org/10.1080/09588221.2012.700315

Kam, E. F., Liu, Y.-T., & Tseng, W.-T. (2020). Effects of modality preference and working memory capacity on captioned videos in enhancing L2 listening outcomes. ReCALL, 32(2), 213–230. https://doi.org/10.1017/S0958344020000014

Karatay, Y., & Xu, J. (2025). Exploring the Potential of Conversational AI for Assessing Second Language Oral Proficiency. TESOL Quarterly, 59(S1). https://doi.org/10.1002/tesq.70003

Kennisa Tazkiya Hidayah & Muhammad Handi Gunawan. (2025). Exploring Efl Students' Difficulties In Listening Comprehension: A Study On Common Challenges And Strategies. JELA (Journal of English Language Teaching, Literature and Applied Linguistics), 7(2), 88–105. https://doi.org/10.37742/jela.v7i2.163

Koukourikos, K., Tsaloglidou, A., Kourkouta, L., Papathanasiou, I., Iliadis, C., Fratzana, A., & Panagiotou, A. (2021). Simulation in Clinical Nursing Education. Acta Informatica Medica, 29(1), 15. https://doi.org/10.5455/aim.2021.29.15-20

Mogammad, M., Si Na, K., Tze Tee, L., & Kew, S.-R. (2026). Enriched eBooks in English Learning: A Systematic Review of Multimedia Features, Patterns, and

Language Skill Development. Arab World English Journal, 17(1), 413–436. https://doi.org/10.24093/awej/vol17no1.26

Mutmainnah, S., Tursini, U., & Jawas, U. (2026). Enhancing speaking performance through cognitive load reduction and working memory management among junior high school students. English Learning Innovation, 7(2), 801–825. https://doi.org/10.22219/englie.v7i2.45202

Nerta, P. O., Tohamba, C. P. P., & Faridawati, F. (2024). Efl students self-directed learning in using youtube to improve listening skills. Klasikal : journal of education, language teaching and science, 6(2), 432–447. https://doi.org/10.52208/klasikal.v6i2.1171

Nishida, S., & Masumori, N. (2026). Instructional Design for Surgical Novices Based on Cognitive Load Theory: Implications for Urologic Surgery Training. International Journal of Urology, 33(9), e70637. https://doi.org/10.1111/iju.70637

Nursyafida & Reni Eka Putri. (2025). Exploring Technology Use in English Pronunciation Instruction: A Systematic Review of Trends and Tools. SALEE: Study of Applied Linguistics and English Education, 6(2), 414–436. https://doi.org/10.35961/salee.v6i2.1931

Olateju Temitope Akintayo, Chima Abimbola Eden, Oyebola Olusola Ayeni, & Nneamaka Chisom Onyebuchi. (2024). Evaluating the impact of educational technology on learning outcomes in the higher education sector: A systematic review. Open Access Research Journal of Multidisciplinary Studies, 7(2), 052–072. https://doi.org/10.53022/oarjms.2024.7.2.0026

Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. Frontiers in Psychology, 14, 1255594. https://doi.org/10.3389/fpsyg.2023.1255594

Reima Al-Jarf. (2026). Student Autonomy in Second Language Learning: A Systematic Self-Review of Training, Skill Development, and Technology Enhanced Practices (2004–2025). British Journal of Teacher Education and Pedagogy, 5(6), 59–85. https://doi.org/10.32996/bjtep.2026.5.6.5

Rifqi Aulia Rahman, Nukman, N., Ahmad Fadhel Syakir Hidayat, Amrulloh Khoirul Ma'arif, & Zulfi Irkhamni. (2025). The Dynamics of AI Technology Utilization in Arabic Language Skills Learning in the Merdeka Curriculum. El-Syaker : Samarinda International Journal of Language Studies, 2(3), 146–162. https://doi.org/10.64093/esijls.v2i3.529

Setiawati, S., Yuliana, Y. G. S., & Ikhsanudin, I. (2025). Improving Students' Engagement in Listening Activities through Edpuzzle. Lectura : Jurnal Pendidikan, 16(1), 42–53. https://doi.org/10.31849/lectura.v16i1.23344

Shadiev, R., & Yu, J. (2024). Review of research on computer-assisted language learning with a focus on intercultural education. Computer Assisted Language Learning, 37(4), 841–871. https://doi.org/10.1080/09588221.2022.2056616

Tilwani, S. A., Vadivel, B., Uribe-Hernández, Y. C., Wekke, I. S., & Haidari, M. M. F. (2022). The Impact of Using TED Talks as a Learning Instrument on Enhancing

Indonesian EFL Learners' Listening Skill. Education Research International, 2022, 1–9. https://doi.org/10.1155/2022/8036363

Tsai, S.-C. (2023). Learning With Mobile Augmented Reality- and Automatic Speech Recognition-Based Materials for English Listening and Speaking Skills: Effectiveness and Perceptions of Non-English Major English as a Foreign Language Students. Journal of Educational Computing Research, 61(2), 444–465. https://doi.org/10.1177/07356331221111203

Venkat, M. V., O'Sullivan, P. S., Young, J. Q., & Sewell, J. L. (2020). Using Cognitive Load Theory to Improve Teaching in the Clinical Workplace. MedEdPORTAL, 10983. https://doi.org/10.15766/mep_2374-8265.10983

Wang, S. (2025). Can Digital Mindfulness Stimulate the Effects of EFL Learners' Cognitive Awareness on Cognitive Load in Computer Assisted Writing Classrooms? International Journal of Human–Computer Interaction, 41(8), 4801–4812. https://doi.org/10.1080/10447318.2024.2353449

Wiryanto, Suryanti, & Putri, A. Y. (2025). Deep Learning for Numeracy: A Roadmap to Asta Cita in Primary Education. Journal for Lesson and Learning Studies, 8(2), 275–284. https://doi.org/10.23887/jlls.v8i2.99932

Yung, R., Khoo-Lattimore, C., & Potter, L. E. (2021). Virtual reality and tourism marketing: Conceptualizing a framework on presence, emotion, and intention. Current Issues in Tourism, 24(11), 1505–1525. https://doi.org/10.1080/13683500.2020.1820454

Zhang, L. (2025). Real-time feedback system for English listening comprehension using speech recognition and synthesis. International Journal of Information and Communication Technology, 26(33), 76–90. https://doi.org/10.1504/IJICT.2025.148658