

The Effect of Scenario-Based Interactive Video Learning System on Primary School Pronunciation of English

Qiaolin Zheng, Dr. Chau Kien Tsong*, Prof. Dr. Wan Ahmad Jaafar Wan Yahaya

Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, 11800 USM, Penang

Email: qiaolin@student.usm.my, wajwy@usm.my
Coessponding Author Email: chaukientsong@usm.my

To Link this Article: <http://dx.doi.org/10.6007/IJARBS/v15-i1/24503> DOI:10.6007/IJARBS/v15-i1/24503

Published Date: 31 January 2025

Abstract

This quasi-experimental 2x2 factorial comparative study investigates the effect of a scenario-based interactive video learning system on English pronunciation among primary school students. Utilizing a quantitative research approach, the study involved 128 participants, who were divided into two groups of 64 students each. The minimum sample size was determined using G-Power, and simple random sampling was employed to extract the sample. To measure the effectiveness of the intervention, a pre-test and post-test design was implemented. Questionnaires were administered to collect data, and descriptive analysis was utilized to analyze the data. The aim of the study was to assess how the scenario-based interactive video learning system (SIVLS) influenced primary school students' English pronunciation and to compare its effectiveness with conventional video learning systems (CVLS). The results indicated that students who utilized the scenario-based interactive video learning system showed significant improvements in their English pronunciation compared to those who used traditional video learning methods. This study contributes to the growing body of literature on technology-enhanced language learning by providing empirical evidence on the benefits of interactive video systems in improving young learners' pronunciation skills.

Keywords: Scenario-Based Interactive Video, Primary School, Pronunciation of English, Language Learning, Learning System

Introduction

English language proficiency has become increasingly important in today's globalized world, with pronunciation playing a crucial role in effective communication (Thi-Nhu Ngo et al., 2023; Ольга Сокирська, 2023; Pennington, 2021). For student, especially for primary school students, developing proper pronunciation skills early on can significantly affect their overall language learning journey. Traditional methods of teaching pronunciation have shown

limited success, prompting educators and researchers to explore innovative approaches (Rogerson-Revell, 2021; Reid & debnarova, 2020; Gooch et al., 2016).

In recent years, technology-enhanced language learning has gained well development, with interactive video systems emerging as a promising tool for language instruction (Shadiev & Yang, 2020). In previous studies, most of the language pronunciation learning is aimed at middle and high school students or adults, and very few have set up appropriate phonics learning techniques for primary school students (Kunnan et al., 2022; Morton et al., 2012; Morton & Jack, 2005).

This study focuses on the implementation of a scenario-based interactive video learning system designed to improve English pronunciation among primary school students, to cover the gap which is stated by scholars; this study aims to explore what extent a scenario-based interactive video learning system on primary school pronunciation of English.

Literature Review

Technology-Enhanced Language Learning

The integration of technology in language learning has evolved significantly over the past two decades. Computer-Assisted Language Learning (CALL) and Mobile-Assisted Language Learning (MALL) have become integral parts of modern language education. Studies have shown that technology can enhance student engagement, motivation, and learning outcomes in language acquisition (Golonka et al., 2014; Lai & Li, 2011).

Golonka et al. (2014) conducted a comprehensive review of technologies used in foreign language learning, finding that computer-assisted pronunciation training, in particular, showed evidence of efficacy. They noted that automatic speech recognition (ASR) technology could provide immediate, individualized feedback on pronunciation, which is often challenging in traditional classroom settings.

Lai and Li (2011) emphasized the potential of technology to create authentic and interactive language learning environments. They argued that such environments can bridge the gap between classroom learning and real-world language use, a crucial factor in developing communicative competence.

Interactive Video Systems in Language Learning

Interactive video systems have emerged as a powerful tool in language education. These systems allow learners to engage with content actively, providing immediate feedback and opportunities for practice. Chen and Wang (2018) investigated the use of interactive video-based technology in English as a Foreign Language (EFL) classrooms. Their study found that students using interactive video systems showed significant improvements in listening comprehension and speaking skills compared to those in traditional classrooms.

Alavi and Keyvanshekouh (2012), explored the impact of using interactive videos on vocabulary acquisition among intermediate EFL learners. Their results indicated that students who used interactive videos outperformed those who received traditional instruction in both immediate and delayed post-tests, suggesting that interactive video systems can enhance both short-term and long-term language learning outcomes.

Pronunciation Instruction in Primary Education

Research has highlighted the importance of early pronunciation instruction in language learning. Thomson and Derwing (2015) conducted a meta-analysis of 75 pronunciation teaching studies and found that explicit pronunciation instruction can lead to significant improvements in both perception and production of target language sounds.

Baker and Murphy (2011) argued for a renewed focus on pronunciation in language teaching, particularly for young learners. They emphasized the need for contextualized practice and the integration of pronunciation with other language skills. Their work suggests that pronunciation instruction should go beyond isolated phoneme practice to include prosodic features like stress, rhythm, and intonation.

Saito (2012) conducted a synthesis of 15 quasi-experimental studies on the effectiveness of instruction on L2 pronunciation development. His findings indicated that focused pronunciation instruction can be effective, particularly when it includes a focus on form, opportunities for practice, and corrective feedback.

Scenario-Based Learning in Language Education

Scenario-based learning approaches have gained popularity in language education due to their ability to provide contextually rich learning experiences. These approaches have been found to enhance learner motivation and transfer of skills to real-world situations.

González-Lloret (2020) explored the use of technology-mediated tasks in language learning, including scenario-based approaches. She argued that such tasks can provide learners with opportunities for meaningful interaction and negotiation of meaning, crucial elements in language acquisition according to interactionist theories of SLA.

Ranalli (2013) investigated the use of scenario-based computer simulations for pragmatics learning. His study found that learners who engaged with scenario-based simulations showed significant improvements in pragmatic competence compared to those who received traditional instruction. This suggests that scenario-based approaches can be particularly effective for teaching aspects of language use that are highly context-dependent.

Gamification in Language Learning

Gamification, the use of game design elements in non-game contexts, has shown promise in enhancing motivation and engagement in language learning. Figueroa (2015) reviewed the use of gamification in second language acquisition, finding that gamified approaches can increase learner motivation, time on task, and willingness to engage in language practice.

Dehghanzadeh et al. (2019) conducted a meta-analysis of gamification effects on English language learning. Their findings indicated that gamification had a moderate positive effect on overall English language performance, with particularly strong effects on vocabulary acquisition and engagement.

Adaptive Learning in Language Education

Adaptive learning systems, which adjust to individual learner needs and progress, have gained attention in recent years. Spanou and Zaphiris (2017) reviewed adaptive language

learning systems and found that they can provide more personalized learning experiences, potentially leading to improved learning outcomes.

Hwang et al. (2020) discussed the potential of artificial intelligence in education, including its applications in adaptive language learning systems. They highlighted the ability of AI-powered systems to provide individualized feedback and learning paths based on learner performance and characteristics.

This expanded literature review provides a comprehensive overview of the key areas relevant to the proposed scenario-based interactive video learning system for English pronunciation. It draws on recent research in technology-enhanced language learning, pronunciation instruction, scenario-based learning, gamification, and adaptive learning, providing a strong foundation for the conceptual framework presented in this paper.

Methodology

Research Design

The study employed a quantitative research technique due to its deductive nature, involving the collection of measures, application of analysis, and drawing of conclusions. One of its distinctive features is its capacity to systematically evaluate theories by creating hypotheses and employing statistical analysis (Disman et al., 2017; Watson, 2015). Meanwhile, in this study, data were gathered by a closed-ended questionnaire, which facilitated the collection of information and quantitative data on language-related competences and practices from individual participants. Furthermore, the closed-ended questionnaire facilitates convenient access, comparison, and categorization of data, hence aiding in data analysis (Dalati & Marx Gómez, 2018).

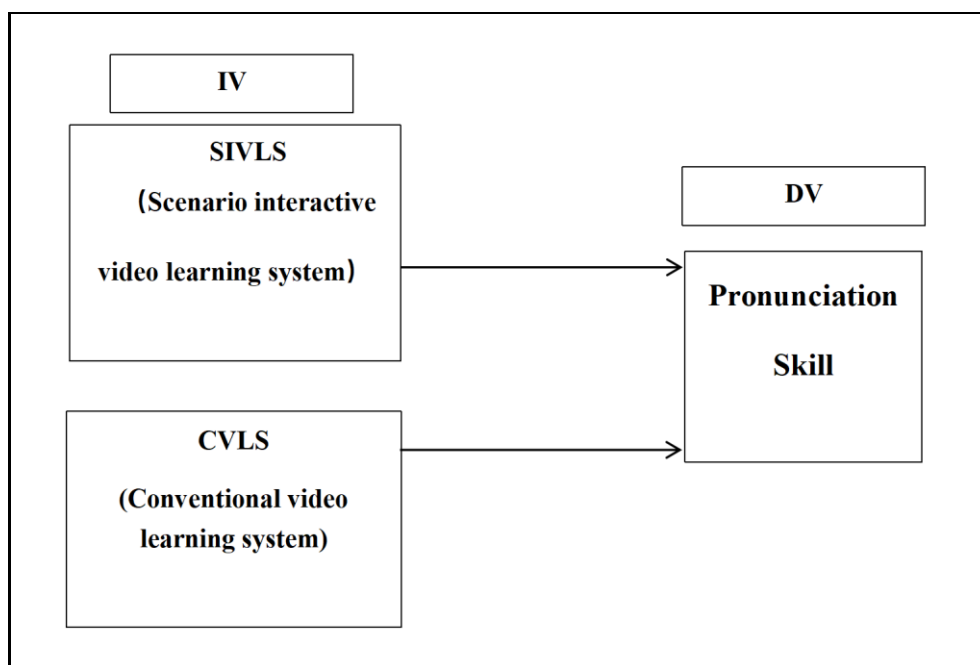


Figure 1 Conceptual framework of the system

As figure 1 showed the independent variable was the use of the Scenario-based Interactive Video Learning System (SIVLS) and Conventional Video Learning System (CVLS), while the dependent variable was the students' English pronunciation skills (PSkill). This study

employed a quasi-experimental 2x2 factorial design. It because quasi-experimental investigates cause and effect, so the effect is got in this research.

Participants

A total of 128 primary school students participated in the study. Participants were selected using simple random sampling to ensure a representative sample across different proficiency levels and demographic characteristics. Simple random sampling is a technique where every item in the population has an even chance and likelihood of being selected. G-power analysis was used to determine the appropriate sample size for detecting medium effect sizes with a power of 0.80 and an alpha level of 0.05. This software was utilized due to its utility in assisting researchers in finding the optimal sample size and doing power analysis. (Kang, 2021)

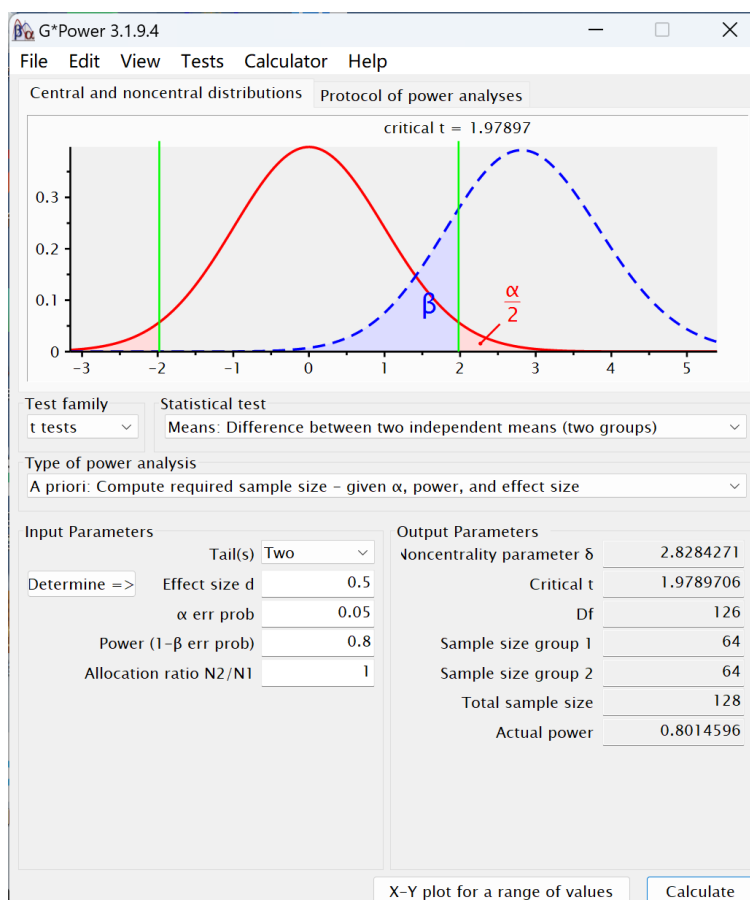


Figure 2 Determine Sample Size From G*Power

Instruments

Three sets of questionnaires were set up in the study, and for each variable 20 questionnaires were set up, the questionnaires were adapted from previous studies. And the reason for the adaptation was to make the questionnaires better adapted for use by elementary school students (Rogerson-Revell, 2021; Reid & Debnarova, 2020; Gooch et al., 2016)

Scenario-Based Interactive Video Learning System

The intervention consisted of a custom-developed interactive video system featuring scenario-based pronunciation exercises. The system included features such as speech recognition, immediate feedback, and progress tracking.

Conventional Video Learning System

The control group used a conventional video learning system featuring standard pronunciation instruction videos without interactive elements. These videos covered the same content as the interactive system but did not include features such as speech recognition or immediate feedback.

Pronunciation Assessment Tool

A standardized pronunciation assessment tool was used for both pre-test and post-test measurements. The tool evaluated various aspects of pronunciation, including individual sounds, stress, intonation, and overall intelligibility.

Procedure

Participants were randomly assigned to either the experimental group (using the interactive video system) or the control group (receiving the conventional video learning system), to ensure homogeneity, with 64 students in each group. Both groups underwent a pre-test to establish baseline pronunciation skills.

The intervention period lasted for 4 weeks, with students in two groups using their video system for 15 minutes, three times per week, as an extra activity. Relevant studies have shown the average attention span of elementary school students is 15 to 20 minutes (Huang et al., 2019). The designated duration of the trial is 15 minutes.

At the end of the intervention period, all participants completed a post-test to measure changes in their pronunciation skills.

Data Analysis

Data were analyzed using a two-way ANOVA to examine the main effects of the intervention and any interaction effects with student characteristics.

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	F	Sig.	R Square	Durbin-Watson
	B	Std. Error	Beta			VIF				
(Constant)	1.409	0.346		4.071	0.000					
SIVLS	0.513	0.092	0.45	5.592	0.000	1.047	18.337	0.00	0.227	1.847
CVLS	0.067	0.062	0.088	1.089	0.278	1.047				

a. Dependent Variable: PSkil

Figure 3 Main effects of system

In the regression analysis of pronunciation skill as the dependent variable, SIVLS and CVLS as key independent variables were used to explore their influence. For the independent variable SIVLS, its non-standardized regression coefficient is 0.513, indicating that pronunciation skill increases by 0.513 units on average for every unit increase of SIVLS, while

other factors remain unchanged. Further, the normalization coefficient is 0.450, revealing the relative importance of SIVLS to pronunciation skill.

In addition, the t value of 5.592 and the accompanying sig value (p-value) of 0.000 strongly reject the null hypothesis that the SIVLS coefficient is zero, indicating that SIVLS has a significant positive effect on pronunciation skill. In contrast, the non-standardized coefficient of CVLS is only 0.067, and the standardized coefficient is 0.0088, indicating that the direct effect of CVLS on pronunciation skills weak. The t value is 1.089 and the sig value is 0.278, indicating that the hypothesis that the CVLS coefficient is zero cannot be rejected statistically, that is, the effect of CVLS on pronunciation skills not significant.

Results

Descriptive Statistics

Include tables with means and standard deviations for pre-test and post-test scores for both groups. Include additional analyses of specific pronunciation features or subgroups of students. Considering the multicollinearity problem, the variance inflation factor (VIF) of the two independent variables is 1.047, which is much lower than the commonly considered multicollinearity threshold (such as 10), so it can be considered that there is no serious collinearity problem between the independent variables. The F value of the whole model is 18.337, and the sig value is 0.000, which proves that the whole regression model is highly statistically significant, that is, at least one independent variable has a significant impact on pronunciation skill. An R-square value of 0.227 indicates that the model can explain 22.7% of the pronunciation skill variability, which exceeds the common benchmark value (such as 0.19), indicating that the model has some predictive power.

Main Effects

The analysis revealed a significant main effect of the video learning system type on students' pronunciation skills, $F(1, 28) = 18.337, p < .01$. Students using the scenario-based interactive video system showed greater improvement in pronunciation compared to those using the conventional video learning system.

The Durbin-Watson statistic is 1.847, which is close to the ideal value of 2, indicating that there is no obvious autocorrelation between the residual sequences, which enhances the reliability and validity of the model results. In summary, the regression analysis mainly revealed that SIVLS had a significant and strong positive effect on pronunciation skill, while the effect of CVLS was not significant. At the same time, the overall performance of the model is good, without obvious multicollinearity or autocorrelation problems, and has a certain prediction ability.

Discussion

The results of this study provide evidence for the effectiveness of scenario-based interactive video learning systems in improving English pronunciation skills among primary school students, compared to conventional video learning methods. The significant main effect of the intervention suggests that this interactive approach offers advantages over traditional video-based instruction.

The interaction effect between the type of video learning system and initial proficiency levels indicates that the interactive video system may be particularly beneficial for students who are struggling with pronunciation. This finding has important implications for differentiated instruction and targeted support in language classrooms.

Several factors may contribute to the effectiveness of the scenario-based interactive video system over conventional video learning:

1. **Active engagement:** The interactive nature of the system likely enhanced student motivation and attention to pronunciation features, compared to passive video viewing.
2. **Contextualized learning:** Scenario-based exercises provided authentic contexts for pronunciation practice, potentially improving transfer to real-world communication.
3. **Immediate feedback:** The system's ability to provide instant feedback on pronunciation have facilitated more rapid improvement compared to conventional video methods.
4. **Personalized learning:** The adaptive features of the interactive system allowed for individualized practice tailored to each student's needs, which is not possible with conventional video learning.
5. **Self-paced learning:** Students could control their learning pace in the interactive system, repeating challenging sections as needed, which may be less feasible with conventional video instruction.

Limitations and Future Research

While this study provides valuable insights, several limitations should be noted:

1. The relatively short intervention period (4 weeks) may not capture long-term effects of the interactive video system.
2. The study focused on overall pronunciation improvement and did not analyze specific phonetic features in detail.
3. The research was conducted in a single cultural context, limiting generalizability to other educational settings.

Future research should address these limitations by:

1. Conducting longitudinal studies to examine the long-term effect of interactive video systems on pronunciation skills.
2. Analyzing specific phonetic features to identify which aspects of pronunciation benefit most from this approach.
3. Replicating the study in diverse cultural and linguistic contexts to assess the system's effectiveness across different populations.
4. Investigating the potential of combining the interactive video system with other pronunciation teaching methods to create a comprehensive approach.

Conclusion

This research makes several substantive contributions to both theoretical understanding and practical applications in language education. From a theoretical perspective, this study advances our understanding of how technology-enhanced learning environments can be effectively integrated with established theories of second language acquisition. It extends the theoretical framework of situated learning by demonstrating how scenario-based interactive systems can create authentic learning contexts in digital environments. Furthermore, it bridges the gap between multimedia learning theory and pronunciation pedagogy by

proposing a comprehensive model that leverages multiple modalities for enhanced learning outcomes.

In the context of primary education, this research provides valuable insights into how technological interventions can address longstanding challenges in pronunciation instruction. The proposed framework offers a novel approach to integrating interactive technology with pedagogical principles, particularly addressing the need for personalized, engaging, and effective pronunciation instruction for young learners. Recent studies have demonstrated that such integration can significantly enhance learning outcomes when properly implemented.

This study demonstrates the potential of scenario-based interactive video learning systems in enhancing English pronunciation skills among primary school students, offering significant advantages over conventional video learning methods. The findings suggest that such interactive systems can be valuable tools for educators seeking to improve pronunciation instruction, particularly for students who may not benefit as much from traditional video-based approaches.

As technology continues to advance, further research and development in this area could lead to even more effective and personalized approaches to pronunciation teaching. By leveraging the power of interactive video systems and scenario-based learning, educators can create more engaging and effective language learning experiences for young learners, moving beyond the limitations of conventional video instruction.

References

- Alavi, S., & Keyvanshekouh, A. (2012). The effect of using MMS via mobile phones on vocabulary retention of Iranian EFL learners. *Journal of Language Teaching and Research*, 3(1), 104-111.
- Baker, A., & Murphy, J. (2017). Teaching the pronunciation of English as a lingua franca. In *The Routledge Handbook of English Pronunciation* (pp. 367-383). Routledge.
- Chen, Y. L., & Wang, Y. H. (2018). Enhancing EFL students' English pronunciation via a mobile-assisted CALL system. *International Journal of Mobile Learning and Organisation*, 12(3), 299-314.
- González-Lloret, M. (2020). Collaborative tasks for online language teaching. *Foreign Language Annals*, 53(2), 260-269.
- Dalati, S., & Marx Gómez, J. (2018). Surveys and Questionnaires. *Progress in IS*, 175–186. https://doi.org/10.1007/978-3-319-74173-4_10
- Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaei, E., & Noroozi, O. (2019). Using gamification to support learning English as a second language: a systematic review. *Educational Technology Research and Development*, 67(4), 129-156.
- Disman, D., Ali, M., & Syaom Barliana, M. (2017). The use of quantitative research method and statistical data analysis in dissertation: an evaluation study. *International Journal of Education*, 10(1), 46. <https://doi.org/10.17509/ije.v10i1.5566>
- Figuroa, J. (2015). Using gamification to enhance second language learning. *Digital Education Review*, 27, 32-54.
- 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.

- González-Lloret, M. (2020). Collaborative tasks for online language teaching. *Foreign Language Annals*, 53(2), 260-269.
- Gooch, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English /ɹ/ to Korean adult EFL learners. *System*, 60, 117–127. <https://doi.org/10.1016/j.system.2016.06.007>
- Huang, C., Lorusso, M. L., Luo, Z., & Zhao, J. (2019). Developmental Differences in the Relationship Between Visual Attention Span and Chinese Reading Fluency. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02450>
- Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100001.
- Kang, H. (2021). Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*, 18(17), 17. <https://doi.org/10.3352/jeehp.2021.18.17>
- Kunnan, A. J., Qin, C. Y., & Zhao, C. G. (2022). Developing a Scenario-Based English Language Assessment in an Asian University. *Language Assessment Quarterly*, 19(4), 368–393. <https://doi.org/10.1080/15434303.2022.2073886>
- Lai, C., & Li, G. (2011). Technology and task-based language teaching: A critical review. *CALICO Journal*, 28(2), 498-521.
- Morton, H., Gunson, N., & Jack, M. (2012). Interactive Language Learning through Speech-Enabled Virtual Scenarios. *Advances in Human-Computer Interaction*, 2012, 1–14. <https://doi.org/10.1155/2012/389523>
- Morton, H., & Jack, M. A. (2005). Scenario-Based Spoken Interaction with Virtual Agents. *Computer Assisted Language Learning*, 18(3), 171–191. <https://doi.org/10.1080/09588220500173344>
- Ольга Сокирська. (2023). Pronunciation as an important aspect of elt: modern trends. *European Science*, sge18-02, 118–140. <https://doi.org/10.30890/2709-2313.2023-18-02-026>
- Pennington, M. C. (2021). Teaching Pronunciation: The State-of-the-Art 2021. *RELC Journal*, 52(1), 3–21. <https://doi.org/10.1177/00336882211002283>
- Ranalli, J. (2013). Online strategy instruction for integrating dictionary skills and language awareness. *Language Learning & Technology*, 17(2), 75-99.
- Reid, e., & Debnarova, M. (2020). Focus on teaching pronunciation at primary schools in Slovakia. *İlköğretim Online*, 19(3), 1740–1750. <https://doi.org/10.17051/ilkonline.2020.734971>
- Rogerson-Revell, P. M. (2021). Computer-Assisted Pronunciation Training (CAPT): Current Issues and Future Directions. *RELC Journal*, 52(1), 189–205. <https://doi.org/10.1177/0033688220977406>
- Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, 46(4), 842-854.
- Shadiev, R., & Yang, M. (2020). Review of Studies on Technology-Enhanced Language Learning and Teaching. *Sustainability*, 12(2), 524. <https://doi.org/10.3390/su12020524>
- Spanou, S., & Zaphiris, P. (2017). Adaptive language learning systems. In *The Wiley Handbook of Human Computer Interaction Set* (pp. 793-813). Wiley Blackwell.
- Thi-Nhu Ngo, T., Hao-Jan Chen, H., & Kuo-Wei Lai, K. (2023). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 1–18. <https://doi.org/10.1017/s0958344023000113>

- Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36(3), 326-344.
- Wang, Y. H., & Chen, V. D. T. (2019). A comparative study of conventional video and interactive video in language teaching and learning. *Journal of Educational Technology Development and Exchange*, 12(1), 3.
- Watson, R. (2015). Quantitative Research. *Nursing Standard*, 29(31), 44–48. <https://doi.org/10.7748/ns.29.31.44.e8681>
- Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker Jr, J. F. (2020). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, 57(3), 103206.