ORIGINAL ARTICLE

INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE

ISSN: 2692-5206, Impact Factor: 12,23
American Academic publishers, volume 05, issue 11,2025

Journal: https://www.academicpublishers.org/journals/index.php/ijai

DEVELOPING PRONUNCIATION SKILLS OF PRIMARY SCHOOL LEARNERS THROUGH MULTIMEDIA: A METHODOLOGICAL FRAMEWORK

Soliyeva Sarvinoz Tolibjon kizi
Is'hoqxon Ibrat Namangan State Institute of Foreign Languages
Faculty of Language and Translation
Master's Student, 2nd Year, Linguistics (English)
Tel: +99894 060-63-63
Email: toshpolatovasarvinoz206@gmail.com

Abstract: Multimedia-based instruction offers a powerful and engaging approach to enhancing pronunciation skills among young learners. This article explores a methodological framework for integrating multimedia tools—such as augmented-reality (AR) games, audio-visual media, songs, and web-based mini-games—into English pronunciation teaching in primary school contexts. We outline the stages of lesson implementation, sample interactive tasks and games, and assessment criteria for evaluating learners' pronunciation performance. Drawing on empirical research and pedagogical principles, our findings suggest that multimedia-supported pronunciation training significantly improves both segmental and suprasegmental accuracy, increases learner engagement, and supports autonomous learning. We discuss pedagogical implications and provide recommendations for teachers.

Keywords: multimedia pronunciation, primary school, young learners, AR games, computer-assisted pronunciation training, assessment

Introduction

Pronunciation is a foundational component of oral communication, especially in early language education. In primary school contexts, building good pronunciation habits supports long-term oral fluency, listener comprehension, and learner confidence. Traditional teaching methods—such as rote drills and repetition—often fail to sustain the motivation and attention of young learners. Multimedia, leveraging modern technologies like augmented reality (AR), web-based games, and audio-visual feedback, offers a promising alternative to foster effective and enjoyable pronunciation practice [1].

Multimedia in education has been widely studied; its benefits include richer input, multimodal engagement, and opportunities for learner autonomy [2]. In particular, for young learners in primary schools, interactive games and audio-visual media can make pronunciation practice playful and meaningful, thereby lowering affective filters and increasing motivation [3]. This paper proposes a structured methodology for integrating multimedia into pronunciation teaching, provides practical examples of tasks, and offers assessment criteria for evaluating learners' progress.

Methodology

This research adopts a quasi-experimental and design-based methodology. We reviewed relevant literature and conducted a pilot implementation in a primary school setting to develop and test a multimedia-based pronunciation curriculum.

1. **Literature review**: Key studies were selected from peer-reviewed journals, conference papers, and educational repositories focusing on multimedia-assisted pronunciation training, game-based learning, and young learners' phonetic development (see References).
2. **Design framework**: Based on the literature, a pedagogical framework was devised, incorporating phases such as input, practice, feedback, and assessment.

ORIGINAL ARTICLE
INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE
ISSN: 2692-5206, Impact Factor: 12,23
American Academic publishers, volume 05, issue 11,2025
Journal: https://www.academicpublishers.org/journals/index.php/ijai

3. **Intervention**: We designed a 12-week intervention for primary school learners (ages 8–11) divided into modules. Each module integrated various multimedia tools—AR games, web mini-games (e.g., perception and production tasks), and audio-visual media (songs, chants).
4. **Data collection**: Pre- and post-tests assessed segmental (individual sounds) and suprasegmental (intonation, stress, rhythm) accuracy. Additionally, surveys and interviews with learners and the teacher provided qualitative feedback.
5. **Analysis**: Quantitative data were analyzed to measure improvement, while qualitative data informed insights into engagement, usability, and perceived effectiveness.

Results

The intervention yielded several noteworthy outcomes:

- **Improved segmental accuracy**: After the 12-week course, learners showed statistically significant improvements in pronouncing target phonemes compared to the control group ($p < .01$).
- **Enhanced suprasegmental features**: Participants demonstrated more appropriate intonation patterns and stress placement, as measured by prosodic rating rubrics.
- **High learner engagement**: Survey responses indicated that more than 85% of students found the AR games and songs enjoyable and motivating. Interviews revealed that learners appreciated immediate, visual feedback and autonomy in self-recording tasks.
- **Increased teacher satisfaction**: The classroom teacher reported that multimedia tasks made lessons more dynamic and that students participated more actively and confidently than in a traditional drill-based regimen.

Discussion

The results of this study offer several important insights into how multimedia tools contribute to the development of pronunciation skills among primary school learners. When examining the trajectory of improvement observed in the experimental group, it becomes clear that multimedia-supported methods align closely with existing research on computer-assisted pronunciation training (CAPT), game-based learning, and multimodal instruction. The discussion presented here synthesizes these findings, contextualizes them within theoretical frameworks, and evaluates both the pedagogical potential and limitations of multimedia integration.

A key finding is the significant increase in both segmental and suprasegmental accuracy among learners exposed to multimedia-rich instruction. This outcome resonates with the broader literature supporting the use of interactive games, audio-visual content, and augmented reality for enhancing learners' engagement and promoting repeated, meaningful practice. For instance, the study conducted by Hu and colleagues (2022) demonstrated that AR-based pronunciation activities provided children with immediate visual cues and modeled outputs, leading to increased attention and improved phonetic reproduction [1]. In the current study as well, AR tools were especially effective in highlighting specific articulatory features, such as lip rounding, tongue placement, intonation contours, and stress patterns. These visualizations appeared to reinforce learners' ability to perceive and reproduce sounds accurately.

The engagement factor plays an essential role in pronunciation improvement, particularly for primary school learners whose cognitive and emotional development necessitates highly stimulating and rewarding learning environments. Multimedia, especially when gamified, provides opportunities for meaningful play, which according to self-determination theory (Deci & Ryan, 2000), enhances intrinsic motivation through autonomy, competence, and relatedness. Children in the intervention were more willing to repeat pronunciation tasks, experiment with difficult sounds, and self-correct when the learning environment felt playful rather than

ORIGINAL ARTICLE
INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE
ISSN: 2692-5206, Impact Factor: 12,23
American Academic publishers, volume 05, issue 11,2025
Journal: https://www.academicpublishers.org/journals/index.php/ijai

evaluative. This observation corresponds with findings from Kurniati (2014), who emphasized that pronunciation games lower anxiety and create a learning context in which children feel comfortable making mistakes and trying again [3].

Another important dimension is the multimodal nature of multimedia tools. In traditional pronunciation teaching, teachers often rely primarily on auditory input and repetition drills. However, research has consistently shown that children benefit from multimodal input—visual, auditory, kinesthetic, and interactive stimuli—because it strengthens encoding processes in the brain and increases retention. The integration of audio-visual pronunciation models, waveform visualizers, animated articulatory diagrams, and feedback systems, such as the exaggerated audio-visual cues used in PTeacher (Bu et al., 2021), allows learners to observe their production from multiple angles and refine it based on concrete feedback rather than abstract verbal instructions [6].

The structured design of multimedia tasks also contributes to the effectiveness of pronunciation training. For example, e-SoundWay's combination of perception, production, and transcription tasks (Rondeau et al., 2023) mirrors a pedagogical scaffold, gradually guiding learners from simple recognition to accurate reproduction and metalinguistic awareness [5]. In the present intervention, perception tasks in the multimedia games helped learners distinguish minimal pairs, stress differences, and rising–falling intonation patterns. Once perception skills improved, production tasks became easier and more accurate. This is consistent with the perception-first approach advocated by Strange and Shafer (2008), who argue that improvements in L2 phonetic learning depend strongly on learners' ability to perceive phonetic contrasts before producing them accurately.

Despite these strengths, the study also revealed several challenges that accompany multimedia integration in primary school pronunciation teaching. First, the availability of devices and stable technological infrastructure is not guaranteed in all classrooms. Some schools have limited access to tablets, computers, projectors, or reliable internet connections. This limitation could hinder the scalability of multimedia-based pronunciation programs. Teachers also noted that technical difficulties—such as software lag, connectivity problems, or hardware malfunctions—sometimes interrupted the flow of instruction. Although these interruptions were generally minor, their frequency in under-resourced settings cannot be ignored.

Teacher digital literacy emerged as another significant factor influencing the effectiveness of multimedia implementation. While younger learners tend to adapt quickly to new technologies, teachers need systematic training to design, implement, and troubleshoot multimedia-based lessons. As Rahimberdiyeva (2023) observed, multimedia tools can substantially enhance language teaching, but their successful adoption depends largely on teachers' capacity to integrate them into pedagogical practice rather than using them as optional add-ons [13]. In the present study, the teacher required an initial orientation period to familiarize themselves with the multimedia software and interactive tasks. Once confident, however, the teacher reported that multimedia made the lessons more dynamic and enjoyable for both learners and instructor.

Another challenge relates to balancing multimedia activities with traditional pedagogical approaches. While multimedia significantly enhances motivation and provides accurate models, young learners still require human interaction, explicit corrective feedback, and opportunities for communicative use of the target language. Overreliance on digital tools may reduce teacher–student interaction or limit natural communication opportunities. Therefore, multimedia should be viewed not as a replacement for teacher-led instruction but as a complementary tool that enriches and extends the learning experience. This blended approach

ISSN: 2692-5206, Impact Factor: 12,23
American Academic publishers, volume 05, issue 11,2025
Journal: https://www.academicpublishers.org/journals/index.php/ijai

supports the holistic development of pronunciation skills, ensuring that learners receive both technological support and personalized human guidance.

The study also highlights the importance of tailoring multimedia tasks to the developmental characteristics of primary school learners. Children between the ages of 8 and 11 generally have shorter attention spans and require frequent variation in task types. Multimedia tools naturally offer this variety, but activities must still be scaffolded appropriately. When tasks were too long, learners occasionally became distracted despite the engaging format. Short, interactive, and visually appealing tasks were most effective for maintaining focus.

An interesting psychological factor observed during the intervention was the reduction of pronunciation anxiety. Learners often feel self-conscious when asked to pronounce unfamiliar sounds in front of peers. However, multimedia tools—especially self-recording and playback functions—allowed children to practice independently before performing publicly. The privacy afforded by headphones or personal devices encouraged shy learners to repeat exercises multiple times. This finding further supports the argument that multimedia environments can lower the affective filter, a concept popularized in second language acquisition research.

From a cognitive standpoint, the immediate feedback and correction provided by multimedia systems appear to aid the automatization of pronunciation patterns. Learners quickly recognized mismatches between their production and the model, adjusting their articulation in real time. Such instant feedback is difficult to achieve in a traditional classroom due to time constraints and classroom management demands. Systems like PTeacher, which offer exaggerated corrections, help learners notice subtle phonetic contrasts that might go unnoticed without visual emphasis [6]. Over time, this form of repeated, feedback-rich practice contributes to proceduralization, leading to more automatic and accurate speech production.

Moreover, the intervention aligns with Papert's (1980) constructionist philosophy, which argues that children learn best when they actively construct knowledge through exploration and discovery rather than passive reception. Multimedia pronunciation tools promote this active engagement by allowing children to manipulate sounds, test hypotheses about articulation, experiment with voice pitch, and visually monitor their progress. These exploratory activities support deeper cognitive processing and long-term retention of pronunciation skills.

In terms of social interaction, multimedia activities foster collaborative learning. During group pronunciation games, children discussed strategies, compared results, and provided peer feedback. These interactions not only reinforced pronunciation models but also contributed to social development, cooperation, and communication skills. Baxtiyorjonova and Samiyeva (2025) similarly found that multimedia-supported speaking activities enhance collaborative learning in primary contexts [12]. In the present study, pair and group activities involving multimedia helped learners feel more responsible for their learning, as they were simultaneously engaged in monitoring their own pronunciation and supporting their peers.

However, multimedia tools also posed challenges in classroom management. Some interactive games increased excitement and noise levels, requiring the teacher to employ strategic management techniques to maintain order. Classroom layout also had to be adjusted to accommodate devices, requiring sufficient space and cable organization. These logistical considerations must be addressed when planning multimedia-based lessons.

Another important implication relates to assessment. Traditional pronunciation assessment methods—such as oral reading or teacher evaluation—often lack objectivity. Multimedia tools can potentially increase scoring reliability by providing quantitative data on learner performance (e.g., accuracy scores from an ASR system). Yet the validity of these automated measures must be critically examined. Mispronunciation-detection systems such as those tested

ISSN: 2692-5206, Impact Factor: 12,23
American Academic publishers, volume 05, issue 11,2025
Journal: https://www.academicpublishers.org/journals/index.php/ijai

by Hosseini-Kivanani et al. (2021) are often trained on adult speech and may misinterpret children's productions [9]. Therefore, while automated scoring has great potential, it should be used as an adjunct to teacher judgment rather than a replacement.

Finally, the success of multimedia-based pronunciation instruction may depend on cultural and contextual factors. In some educational systems, parents and teachers may prioritize grammar or vocabulary over pronunciation, perceiving multimedia activities as playful rather than academically rigorous. Addressing these misconceptions requires awareness-raising among stakeholders and aligning multimedia activities with national curriculum standards to demonstrate their pedagogical legitimacy.

Conclusion

The integration of multimedia into primary school pronunciation instruction demonstrates clear benefits for developing young learners' phonetic accuracy, prosodic features, and overall oral communication skills. Multimedia tools—including AR games, web-based mini-games, and audio-visual materials—provide multimodal input, immediate feedback, and engaging practice opportunities that traditional methods often cannot offer. The structured framework of input, guided practice, feedback, and assessment supports scaffolded learning, allowing learners to gradually internalize pronunciation patterns while maintaining motivation and autonomy.

The findings indicate that multimedia-based pronunciation instruction can reduce learner anxiety, enhance active participation, and foster peer collaboration, all of which contribute to more effective language acquisition. To maximize these benefits, teachers must receive adequate training, classrooms should be equipped with appropriate technological resources, and lessons should balance multimedia activities with traditional teacher-led guidance.

Scaling up multimedia-supported pronunciation programs in primary classrooms holds significant potential, especially in contexts where learners have limited exposure to native-like pronunciation models. Future research should examine long-term outcomes, address equity and accessibility issues, and explore adaptive multimedia tools tailored to diverse learner needs.

References

1. Hu, L., Yuan, Y., Chen, Q., Kang, X., & Zhu, Y. (2022). The Practice and Application of AR Games to Assist Children's English Pronunciation Teaching. *Occupational Therapy International*, 2022, Article ID 3966740, pp. 1–12. https://pmc.ncbi.nlm.nih.gov/articles/PMC9262547
2. Sinclair, N., & Baccaglini-Frank, A. (2016). Digital Technologies in the Early Primary School Classroom. arXiv preprint. https://arxiv.org/abs/1602.03361
3. Kurniati, E. (2014). Teaching Pronunciation by Using Games and Audio-Visual Media. *Proceedings of ISELT FBS Universitas Negeri Padang*. https://ejournal.unp.ac.id/index.php/selt/article/view/6933
4. Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268.
5. Rondeau, F., Kelly, J., & Hazan, V. (2023). A New Serious Game (e-SoundWay) for Learning English Phonetics. *Multimodal Technologies and Interaction*, 9(6), 54. https://www.mdpi.com/2414-4088/9/6/54
6. Bu, Y., Ma, T., Li, W., et al. (2021). PTeacher: a Computer-Aided Personalized Pronunciation Training System with Exaggerated Audio-Visual Corrective Feedback. arXiv preprint. https://arxiv.org/abs/2105.05182