Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025 ASEE
Paper ID #47725

BOARD # 83: WIP: AR AniMotion: Augmented Reality Application for Enhancing Speech Skills in Children with Speech Difficulties

Mr. MALEK EL KOUZI, Queen's University

Malek EL Kouzi is currently pursuing a Ph.D. at Queen's University and holds a Master's in Digital Media from Carleton University. With over 15 years of experience in educational leadership, he serves as a Professor at Algonquin College and a Teaching Assistant at Queen's University. Malek is profoundly dedicated to the integration of Augmented Reality in science, especially at elementary levels. Through his publications, he has crafted multiple augmented reality tools to enhance student comprehension in lessons. His recent research focuses on the collaboration within augmented reality educational applications and its impact on student skills. Personal Website: https://malekelkouzi.com/ Google Scholar : https://scholar.google.ca/citations?user=9y

Haley Clark, Queen's University
Richard Reeve, Queen's University

AR AniMotion: Augmented Reality Application for Enhancing Speech Skills in Children with Speech Difficulties- A Work in Progress

Abstract. Roughly 3% of children experience speech delays or difficulties globally [1]. A “speech delay” is defined as a child's speech or language development significantly communicative milestones based on phases of language development, language acquisition, and reading comprehension skills [2,3]. Although speech delays are commonly rooted in neuronormative presumptions [4], speech delays can have long-term impacts on children [5, 6]. Some children experience a lack of well-being, lack social skills, and struggle to develop meaningful relationships [5]. With the large importance of supporting students in interactive and engaging manners and the rise of expansive technology, AR An- iMotion leverages advanced AR and artificial intelligence (AI) technologies to create an engaging, immersive learning environment for children. Augmented reality, or AR, is a technology app that relates digital information to the real world through images, videos, and or 3D models to make an environment more visible and easier to manipulate. Therefore, this application supports closing the gap in engaging technology to support positive long-term impacts for children with speech difficulties. This application was rooted in the literature highlighting how current speech therapy methods often involve repetitive exercises that may disengage young children over time. This work can potentially revolutionize speech therapy delivery by integrating cutting-edge AR technology with sound pedagogical principles.

Keywords: Augmented Reality, Speech Delays, Interactive Learning, Educational Technology, Speech to Text.

1 Introduction

Speech delay is a situation where a child does not grow according to typical developmental markers in speech and language development but significantly slows down their communication. This can effectively influence the development of socialization skills [7]. A child may have difficulties forming relationships and fitting in, presenting emotional upheavals, and this can also impact their academic performance, making it highly crucial for early intervention [4]. Customizing learning approaches enhances engagement, and the rich content makes practice more motivating for children, supporting their communication skill development [8].

Augmented Reality (AR) creates immersive environments that merge real-world interactions with digital feedback. By offering immediate and personalized responses, AR tools can transform repetitive exercises into enjoyable activities, keeping children engaged in their learning process [9]. Other adopters of AR in the education sector have attested to benefits such as enhancing learners' attentiveness, improving their performance, and facilitating interactivity and learning experiences [9, 10, 11].

Currently, speech therapy applications are available, and the focus is on the auditory and visual feedback sections [12,19]. Still, none of them uses the actual AR functions to make speech therapy fun for the kids and practical at the same time. In developing AR AniMotion, the design was influenced by a comprehensive literature review, which underscored the importance of interactive learning tools for children with speech difficulties. Studies indicate that when children are provided with real-time feedback, it significantly improves their learning outcomes and motivates them to engage more actively in their education. With its ability to combine the real world with digital enhancements, Augmented Reality has emerged as a powerful educational tool, particularly in creating immersive learning experiences. AR AniMotion builds on these findings by providing auditory feedback and a strong visual stimulus, helping children connect the words they produce and the actions they see.

In the initial testing phase, the research team thoroughly reviewed the application. We conducted internal testing, evaluating its functionality, user experience, and educational effectiveness. Feedback was used to refine the user interface and improve the responsiveness of the speech-to-text engine, ensuring a seamless interaction between the child's speech and the application's output. The application is ready to be tested in real-

In world classrooms or therapy settings, approval for ethics is pending. With the speech-to-text technique incorporated into AR, possibilities to make timely responses in a format that will be engaging and, at the same, engaging children more often and with more passion in speech therapy sessions. This paper will seek to fill this gap by developing an AR application tailored to support speech therapy to build on the benefits already proven in using AR applications in other fields.

2 Literature Review

Past research suggests that integrating practical educational tools rounds the tool's effectiveness, ability to stimulate student engagement, and enjoyment a student gains from the activity. With the introduction of technology in the classroom, many educators and researchers alike warrant the benefits of digital technologies in the school [9, 12, 13]. As traditional classroom instruction and tools fall short for some students, introducing technology is vital to quality education [12]. This means that with the use of technology in teaching and learning processes, students' engagement, motivation, and the amount of information they can absorb are boosted immensely [10, 14, 12]. Furthermore, students with more visual and auditory participation create multimedia tools that use better learning modes and enable them to understand complex areas easily [15]. Specifically, learning technologies in special education for students with LD predetermine how students with learning disabilities pay attention and promote improved academic learning to make teaching-learning more inclusive [16].

Research suggests using interactive tools and technology can help speech therapists engage with children [17]. Current technologies in this area include diverse software applications used in speech therapy to rectify or develop speech while incorporating auditory and visual displays to facilitate pronunciation and rhythmic drilling in speech [8, 13]. In addition, Grossinho suggests that having built-in feedback and progress monitoring in speech technology can promote engagement and provide instant feedback for children to progress through the application [17]. For example, some applications use a cartoon character or a helping figure to facilitate speech practice [18]. The use of technology in speech and language pathology is developing, with roughly 33% of studies in this field focused on technological support between 2015-2019.

In some instances, the speech and language pathologist is restricted by barriers such as budgetary constraints and accessibility to technology [19]. Most tools applied in speech and language pathology are adapted to work on conventional two-dimensional interfaces, and users are not actively involved in such applications with the help of their senses. None of these technologies exploit the possibility of immersing users in the AR environment. Despite the extensively recognized benefits of AR in education, its application in speech therapy remains limited. Yet, many available tools have not taken this approach [17, 12]. Traditional speech therapy approaches can be boring and monotonous, eventually making children lose interest and leading to slow progress [5]. More research into this aspect is required to be done to come up with valuable results on the effectiveness of augmented reality in speech therapy. Again, developing AR applications tailored to speech therapies could fill this gap by engaging and inactivating the features of AR in support of early speech development of children experiencing speech delays. In the future, work should focus on obtaining ethical approvals, conducting clinical trials, and refining these applications based on trial feedback to address the specific needs of children and therapists [9].

3 Methodology

3.1 Design and Development

This application was built using Unity, a widely used and highly customizable game development environment with integrated support for Augmented Reality. To capture speech in real-time, a speech-to-text asset was bought from the online marketplace known as the Unity Asset Store. This asset was incorporated into the application to interpret the words spoken by the child. When the child speaks a word, the asset translates the audio into the text, corresponding to the asset's audio input. The 3D models of animals were also purchased from the Unity Asset Store section for the game development. AR contents had to be designed so that the employed AR detection system could easily identify their markers. These markers, when seen through the camera of the device, bring in the desired 3d animal model in the augmented reality environment. The application then asks the child to say a specific word

Displayed on the application screen. If the child correctly speaks the word, the speech-to-text power generates the data and animates the 3D animal model to affirm the correct answer and encourage the child. If the child doesn’t say the word correctly, the spoken word will be displayed, and the 3D animal will not move. Then, the app will display the word again, giving the child another chance to say the word correctly.

3.2 Application Functionality and Progress

A functional prototype of the AR Animotion application has been developed, showcasing the core features: AR marker detection, 3D model display, speech-to-text recognition, and animated feedback. The application engages children with speech delays by offering an interactive and motivating environment to improve their speech skills. Below are examples of how the application operates:

The application prompts the child to say a specific word (e.g."Eat"). Once the word is correctly spoken, the speech-to-text technology recognizes it and triggers the corresponding animation of the 3D animal model (Figure1). For this design, different animals are displayed in 3D animated models based on the child's speech. For instance, a crow model appears to eat when the child is prompted to say "eat", and the word is recognized correctly. The prototype includes a range of animals: crow, Orca, Dolphin, Duck, Goat, pigeon, cow, shark, Rhino, and Lion. Each performs different actions like "Run, Walk, Jump, fly, stop, fall, walk, swim, stop, eat, play". This variety helps keep the sessions engaging for children.

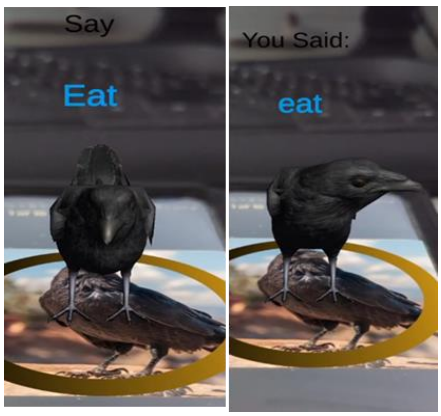

Figure 1Speech Prompt and Recognition

The application provides immediate visual feedback through animations when the child correctly pronounces the word. This instant reward system helps maintain the child's motivation and encourages more practice. These examples demonstrate the prototype's capability to recognize speech and provide engaging visual feedback (Figure2).

Figure 2. Feedback Mechanism

The application, developed using **Unity**, integrates a **voice-to-text API** to create an interactive learning experience. It works by displaying a **3D animal model** on a detected **AR marker**, as shown in **Figure 3**, along with a word that the player needs to say. The user presses the **blue button** to activate the voice input, and the system captures and converts the spoken word into text. If the recognized word matches the displayed word, the **3D animal performs a predefined movement** related to that word. This provides **real-time visual feedback** to reinforce correct pronunciation. Additionally, as shown in **Figure 3**, the application displays a confirmation message, **"You Said:"**, followed by the detected word to ensure the player sees what was recognized. This interactive approach **engages children in a multisensory learning process**, enhancing their speech skills through **visual, auditory, and kinesthetic interactions** in an immersive AR environment.

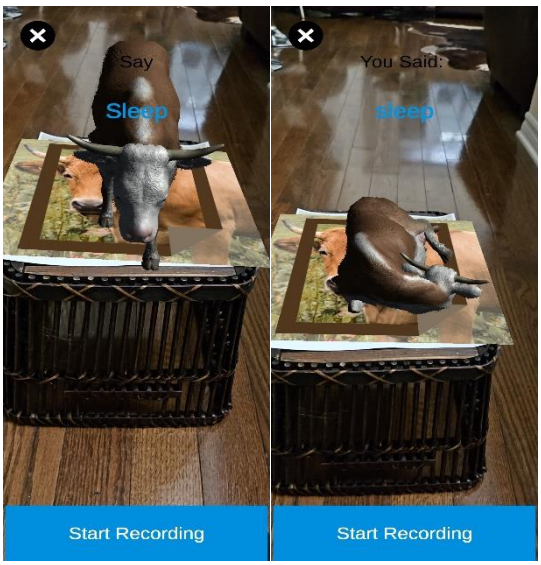

Figure 3:Application Functionality

4 Future Work

The application is expected to significantly impact children with speech delays by providing a motivating and enjoyable way to practice speaking. Future work includes obtaining ethical approval to conduct trials with children to evaluate the application's effectiveness in a real-world setting. Then, structured trials to collect data on the impact of the application on children's speech development. After collecting the data, we plan to refine the application to improve it. By leveraging the capabilities of AR and AI, this application aims to provide a novel, engaging, and effective tool for speech therapy, supporting the early development of speech skills in children with speech delay.

5 Conclusion

This work in progress highlights the potential of using AR and AI to create innovative tools for speech therapy. By making speech practice interactive and fun, the application aims to motivate children with speech difficulties to improve their speaking skills and support speech and language pathologists in providing interactive and supportive materials. Further development and testing will be crucial in realizing this potential and ensuring the application meets the highest efficacy, impact, and safety standards.

References

1. The Kids Health Organization. *Delayed Speech or Language Development*. Kids Health and Parents. https://kidshealth.org/en/parents/not-talk.html (2022).

2. Mayo Clinic Staff. Child development chart: Preschool milestones. https://www.mayoclinic.org/health/child-development/MY00136 (2011).
3. Piaget, J. The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer, *Measurement and Piaget.* McGraw-Hill (1971).
4. Cherney, J. L. Ableism and Intersectionality: A Rhetorical Analysis. In *The Palgrave Handbook of Disability and Communication* (pp. 99-111). Cham: Springer International Publishing (2023).

5. Lyons, R., & Roulstone, S. Well-being and resilience in children with speech and language disorders. *Journal of Speech, Language, and Hearing Research*, *61*(2), 324-344 (2018).

6. Sunderajan, T., & Kanhere, S. V. Speech and language delay in children: Prevalence and risk factors. *Journal of family medicine and primary care*, *8*(5), 1642-1646 (2019).

7. Bishop, D. V., & Edmundson, A. Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of*

Speech and Hearing Disorders, 52(2), 156-173 (1987).

8. Delmonte, R. Exploring speech technologies for language learning. Speech Lang. Technol, 71 (2011).

9. Bacca, J., Baldiris, S., Fabregat, R., Graf, S., & Kinshuk.. Augmented reality trends in education: A systematic review of research and applications. *Educational Technology & Society*, 17(4), 133-149 (2014)

10. Billinghurst, M., & Duenser, A. Augmented reality in the classroom.Computer, 45(7), 56-63 (2012).

11. Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, 62, 41-49 (2013).
12. Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. Understanding the role of digital technologies in education: A review. Sustainable Operations and Computers, 3, 275-285 (2022).
13. Schuler, B. Applicable applications: Treatment and technology with practical, efficient and affordable solutions. Auditory Processing Disorders: Assessment, Management and Treatment, 411-416 (2019).
14. Chen, C. J., Lee, I. J., & Lin, L. Y. Augmented reality-based self facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. Research in Developmental Disabilities, 36, 396-403 (2015).
15. Mayer, R. E. Multimedia learning (2nd ed.). Cambridge University Press (2009).

16. Wang, F., Kinzie, M. B., McGuire, P., & Pan, E. Applying technology to inquiry-based learning in early childhood education.
17. Grossinho, A., Cavaco, S., & Magalhães, J. An interactive toolset for speech therapy. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology* (pp. 1-4) (2014).
18. Rupal, A., Shilpa, G., Priti, S., & Sonal, P. Speech therapy software and application: A review. *International Journal of Advanced Research in Computer Science*, 8(3), 446-452 (2017).
19. Kuschmann, A., Nayar, R., Lowit, A., & Dunlop, M. The use of technology in the management of children with phonological delay and adults with acquired dysarthria: A UK survey of current speech-language pathology practice. International Journal of Speech-Language Pathology, 23(2), 145-154 (2021).