Kashihara, A. et al. (Eds.) (2024). Proceedings of the 32nd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Examining Augmented Reality's Influence on Pronunciation Training: Insights from PinyinGuo's Application and Comparative Avatar Testing

Daria SINYAGOVSKAYA
Linda Berry Stein College of Fine Arts & Humanities, Jacksonville University, USA
dsinyag@ju.edu

Abstract: This study introduces PinyinGuo, an immersive language learning application developed in Unity, leveraging LipSync plugin integration from RogoDigital for accurate articulation of facial sounds in AR avatar. MongoDB integration enhances user data processing and personalization. An online experiment evaluates the impact of AR technology by comparing 3D and AR avatars on mobile phones for speech training and language learning, with preliminary results indicating promising outcomes for AR on mobile platforms. Through rigorous usability testing, PinyinGuo demonstrates the transformative potential of AR in language education and speech training, highlighting the role of AR technology in reshaping speech training applications.

Keywords: Augmented Reality, pronunciation training, peer assessment, immersive education.

Figure 1. PinyinGuo AR version.

1. Introduction

Language learning is a complex process that requires extensive practice and effective feedback (Bitchener, 2008). Immersive technologies like Augmented Reality (AR) open new possibilities for enhancing speech training and pronunciation through interactive avatars and virtual practice environments. Studies by Zhu et al. (2022) and Wei et al. (2020) have shown that AR can significantly improve language learning by providing realistic practice scenarios that engage learners more deeply than traditional methods. Sykes et al. developed the AR game Mentira to teach Spanish, where spatial context played a vital role in the learning process by requiring users to complete a quest in the original town where the story took place (Holden & Sykes, 2011a). This paper introduces PinyinGuo, an immersive language learning application that leverages AR to enhance speech training by simulating real-world scenarios and fostering collaborative peer assessment.

The objectives of this study are to:

- Develop an AR-based immersive learning application for language education.
- Evaluate the comparative effectiveness of AR and non-AR (3D) avatars in speech training.

PinyinGuo aims to improve learner engagement through interactive environments with AR, offering an opportunity for self-paced and personalized practice. By providing immediate feedback via virtual avatars, responsive UI, and collaborative peer learning, the application seeks to address common challenges in pronunciation training.

2. Related Work

2.1 Augmented Reality in Language Learning

AR technology has been explored as a tool that can place virtual objects onto real-life surroundings. For example, Thorne et al. developed an AR game that positively influenced learners' interest in environmental education. Similarly, Sykes et al. developed the AR game Mentira to teach Spanish, where spatial context was crucial. AR overlays synthetic imagery onto the real world, situating it in the environment, and falls toward the reality side of the Reality-Virtuality continuum (Milgram et al., 1995). Popular AR applications like Pokémon GO (Paavilainen et al., 2017) and Snapchat lenses demonstrate the potential of AR in engaging users by altering their surroundings.

AR offers several benefits for language learning, including the ability to contextualize linguistic concepts in real situations and improve engagement through interaction. Studies have shown that AR-based learning, especially when combined with gamification, can increase motivation and self-efficacy among learners (Taskiran, 2019). While commercial mobile apps like Duolingo provide gamified learning content, they often lack immersive contexts or active learning opportunities, a gap AR can fill.

2.2 Mobile-Assisted Language Learning

Mobile-Assisted Language Learning (MALL) apps like Duolingo [1] and Busuu [2] have popularized gamified language education, emphasizing short and engaging learning sessions. While gamification is crucial in retaining user engagement, commercial apps often lack feedback mechanisms necessary for comprehensive pronunciation practice. Incorporating AR technology can address these gaps by providing immersive feedback and fostering collaborative learning.

Research on gamified learning suggests that game-like mechanics, such as leaderboards and rewards, can improve user engagement (Mekler, 2013). However, these mechanisms should be balanced with pedagogical goals to ensure learners receive meaningful feedback while having fun (Hamari et al., 2014).

3. Methodology

3.1 Interface Design

The app interface was refined based on usability testing feedback. Participants noted challenges with font size and button contrast, leading to adjustments that increased the size of pinyin tone marks, text, and buttons for better visibility and interaction. The app also

[1] https://www.duolingo.com/
[2] https://www.busuu.com/

includes a 360-degree AR avatar that demonstrates facial articulations, allowing learners to observe subtle mouth movements necessary for tonal languages. The practice screen integrates visual cues such as progress bars and motivational messages to encourage ongoing practice.

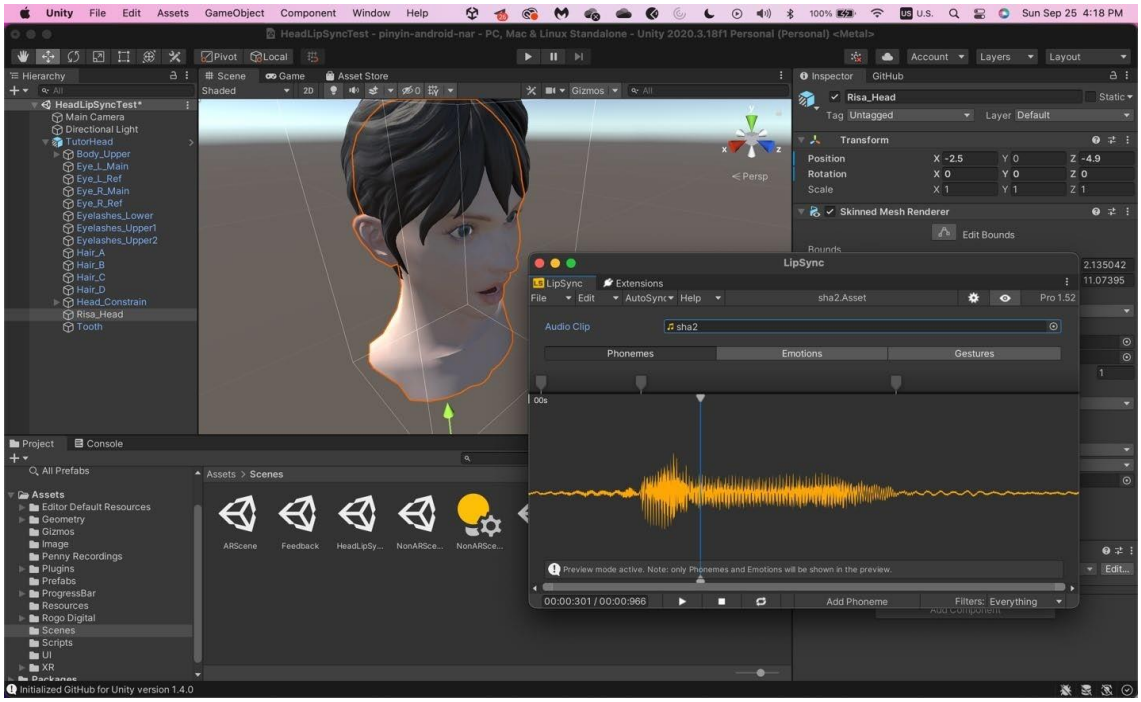

Figure 2. Rogo Digital Phoneme right width adjustment for the phoneme sha in second tone.

3.2 Content Design

The learning content is based on linguistic theory and second-language (L2) pedagogy, particularly the first lesson from Zadoenko and Huang's (Zadoenko & Huang, 2019) Chinese Language Beginner course. This content was chosen to cover the four tones of Standard Chinese within a short learning period (Li & Thompson, 2009).

To ensure clarity and functionality, each screen of the app was wireframed, emphasizing tone marks and minimizing distractions to help learners focus on pronunciation. The app design incorporated twenty syllables into the exercises, ensuring a balanced practice of all four tones. The practice screen was designed to provide learners with immediate feedback, supported by visual cues that reinforce the importance of tone accuracy in pronunciation.

3.4 Experimental Design

The study employed a comparative online experiment where participants were randomly assigned to either the AR or non-AR intervention, with both groups following a similar practice

regimen (Figure 4).

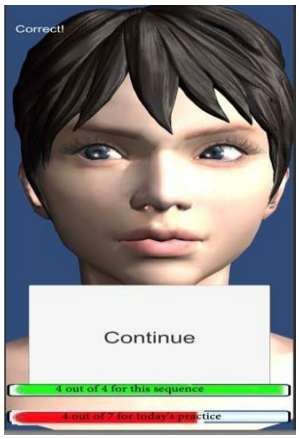

Figure 4. Non-AR Version

Participants were recruited via social media and email listservs, with inclusion criteria requiring adults aged 18-45. The study focused on participants without prior Chinese language experience, and no pre-test was conducted. The intended sample size was 40 participants, but only 13 enrolled, and just two completed the full seven-day study. This low completion rate reflects the challenge of maintaining engagement over a continuous week, particularly with a $40 Amazon gift card incentive, which proved insufficient.

3.5 Usability Testing

Usability testing for PinyinGuo was conducted in two rounds, involving iterative design and prototype refinement. Initially, two participants underwent think-aloud testing to identify navigation issues and provide feedback on learning content. This led to adjustments such as increasing the size of pinyin tone marks, text, and buttons for better visibility and easier interaction. During the intervention, additional usability testing was conducted by analyzing deployment data to assess what aspects of the application were effective and what needed improvement (Mirza-Babaei et al., 2020).

4. Findings

4.1 Usability Testing Results

Usability testing revealed that participants found PinyinGuo interactive and useful. Adjustments made to the user interface improved accessibility, and participants reported that the practice perception questionnaire showed increasing motivation and confidence. The LipSync integration effectively articulated mouth movements, enhancing learners' ability to understand phoneme articulation. Visual cues like eyebrow movements and mouth positioning helped participants understand subtle tonal differences.

4.2 Experimental Results

The seven-day online experiment indicated that participants were motivated to continue practice, finding it easier over time and gaining linguistic confidence (Figure 5). Retention rates were 15% higher among AR users, who also had fewer tonal errors compared to the non-AR group. The Intrinsic Motivation Inventory (IMI) results highlighted that participants

were more motivated in AR environments, with the visual and contextual feedback from AR avatars aiding in internalizing pronunciation rules.

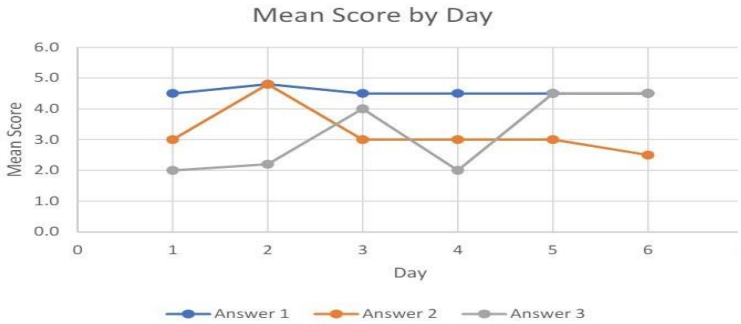

Figure 5. Feedback Results, Answer 1: I Want to Continue Practicing; Answer 2: Today's Practice Was Very Difficult For Me; Answer 3: I Had an Excellent Performance.

4.3 Study Limitations

The study was limited by a small participant pool and recruitment challenges. Technical issues with iOS UDID retrieval affected participant engagement, and the $40 Amazon gift card incentive was insufficient to retain participants over seven consecutive days. The cohort size varied due to attrition, making it difficult to generalize some findings. However, qualitative feedback provided valuable insights into refining the application's design and assessment mechanisms.

5. DISCUSSION

6.1 AR vs. Non-AR Learning Outcomes

Preliminary results suggest that AR avatars have the potential to enhance pronunciation practice, increasing motivation and reducing errors in phoneme pronunciation compared to non-AR interventions. The immersive environment provides visual cues and interactivity that make learning more engaging, helping learners build confidence in real-world pronunciation.

6.2 Implications for Future Research

PinyinGuo's study provides a foundation for future research into the role of AR in immersive language learning. Expanding the participant pool would allow for more detailed statistical analysis of AR's effects on learning outcomes. Future research should also explore AI-driven corrective feedback and further enhance the learning experience. Additionally, exploring task-based pedagogy and AR's contextual learning advantages would help refine the app's methodology.

7. CONCLUSION

PinyinGuo exemplifies the transformative potential of AR technology in language education. By comparing AR and non-AR avatars, the study demonstrates how immersive virtual environments can reshape pronunciation training by improving learner engagement, motivation, and retention.

Future iterations of PinyinGuo could incorporate additional gamified elements, social media integration, and AI-enhanced corrective feedback to enrich the learning experience. With further research and refinement, AR could become a pivotal tool in enhancing language education and developing effective, inclusive pronunciation training applications.

References

Bitchener, J. (2008). Evidence in support of written corrective feedback. Journal of Second Language Writing, 17(2), 102-118. https://doi.org/10.1016/j.jslw.2007.11.004

Cheng, K. H. (2013). Affordances of augmented reality in science learning: Suggestions for future research. Journal of Science Education and Technology, 22(2), 449-462. https://doi.org/10.1007/s10956-012-9405-9

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

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. 2014 47th Hawaii International Conference on System Sciences, 3025–3034. https://doi.org/10.1109/HICSS.2014.377

Holden, C. L., & Sykes, J. M. (2011). Mentira: Prototyping AR for language education. International Journal of Computer-Supported Collaborative Learning, 6(4), 413-427. https://doi.org/10.1007/s11412-011-9125-1

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. In Proceedings of Telemanipulator and Telepresence Technologies (Vol. 2351, pp. 282-292). International Society for Optics and Photonics. https://doi.org/10.1117/12.197321

Mirza-Babaei, P., Stahlke, S., Wallner, G., & Nova, A. (2020). A Postmortem on Playtesting: Exploring the Impact of Playtesting on the Critical Reception of Video Games. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–12. https://doi.org/10.1145/3313831.3376831

Paavilainen, J., Korhonen, H., Alha, K., Stenros, J., Koskinen, E., & Mayra, F. (2017). The Pokémon GO experience: A location-based augmented reality mobile game goes mainstream. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 2493-2498). https://doi.org/10.1145/3025453.3025871

Peng, X., Xu, Y., Wang, W., & Wang, J. (2018). The impact of visual pronunciation aids on second language learners' pronunciation performance. Journal of Computer-Assisted Learning, 34(1), 28-39. https://doi.org/10.1111/jcal.12211

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2016). Augmented reality learning experiences: Survey of prototype design and evaluation. IEEE Transactions on Learning Technologies, 9(1), 3-15. https://doi.org/10.1109/TLT.2015.2445761

Sykes, J. M., & Holden, C. L. (2011). Mentira: Prototyping AR for language education. International Journal of Computer-Supported Collaborative Learning, 6(4), 413-427. https://doi.org/10.1007/s11412-011-9125-1

Taskiran, A. (2019). The effect of gamified augmented reality applications on motivation and self-efficacy of EFL students. Computers & Education, 139, 1-10. https://doi.org/10.1016/j.compedu.2019.05.011

Zhu, J., Zhang, X., & Li, J. (2022). Using AR filters in L2 pronunciation training: Practice, perfection, and willingness to share. Computer Assisted Language Learning, 35(4), 1-30. https://doi.org/10.1080/09588221.2022.2096202

Zadoenko, T. P., & Huang, S. (2019). Beginner Course of Chinese Language (5th ed., Vol. 1). Vostochnaya Literatura.