

FluencyAR: Augmented Reality Language Immersion

Shanna Li Ching Hollingworth
University of Calgary
Calgary, Alberta, Canada
shanna.hollingwor1@ucalgary.ca

Wesley Willett
University of Calgary
Calgary, Alberta, Canada
wesley.willett@ucalgary.ca

ABSTRACT

FluencyAR is an augmented reality second language learning tool centered around the concepts of language immersion and self-talk. For many second language learners, advancing into upper levels of fluency can be difficult without sufficient opportunities to practice. Traditional solutions of tutoring or finding exchange partners are often inconvenient or limiting. FluencyAR provides situational conversation practice in highly self-directed practice sessions that imitate environments where the target language is dominant. We utilize augmented reality to allow users to practice their target language with immediate feedback at any time, and from any location. Using ChatGPT and the physical space of the user, we can produce unique and challenging conversation prompts relative to a user's surroundings, ensuring that sessions remain interesting.

ACM Reference Format:

Shanna Li Ching Hollingworth and Wesley Willett. 2023. FluencyAR: Augmented Reality Language Immersion. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct)*, October 29–November 1, 2023, San Francisco, CA, USA. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3586182.3616670>

1 INTRODUCTION

Most second language learners will reach a point where they struggle to increase in fluency. Existing apps and tools are more often designed for beginners because at upper levels improvement relies more on practical application. Physical language immersion is the ideal solution, but not always viable in practice. Therefore, finding a language exchange partner or tutor is the next best option, but this introduces problems as well such as scheduling issues, and communication issues if neither party is fluent in both languages. Another potential drawback is the anxiety[5] of making a mistake in front of another person, which may hinder the learner's willingness to speak on topics they are less familiar with.

Alternatively, within the language learning literature, it has been noted that thinking in the target language works well as a form of "rehearsal"[3] for conversation, allowing for practice anytime, anywhere, and without performance anxiety. It would naturally follow that speaking this train of thought out loud would be at least equally if not more effective. In this paper, we title this version of the technique as "self-talk", and define it as the act of talking out loud to oneself about any random topic in the target language.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
UIST '23 Adjunct, October 29–November 1, 2023, San Francisco, CA, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0096-5/23/10.
<https://doi.org/10.1145/3586182.3616670>

Language Immersion	FluencyAR	Self-Talk
✓ Practice language constantly	✓ Practice language in everyday settings	
✓ Forces learners to construct sentences on new topics using existing knowledge		
✓ Feedback via social cues	✓ Feedback via recognized speech	✗ Limited feedback
✓ Learn new words and phrases from real situations		✗ Limited new content
✗ Often requires travel to access	✓ Accessible AR tool	✓ Accessible daily practice
✗ Performance anxiety	✓ Reduced social pressure	

Figure 1: FluencyAR combines benefits of language immersion and self-talk.

However, this method does not allow for any feedback, and speakers may find it difficult to think of what to talk about. In response to this challenge, our FluencyAR prototype uses augmented reality (AR) technology alongside a simulation of object recognition and ChatGPT to provide unique and challenging self-talk conversation prompts rooted in reality. The result is an experience which emulates many of the positive aspects of language immersion, combined with the benefits of self-talk, as illustrated in Figure 1. We can also provide real-time and post session feedback for the user to review.

2 RELATED WORKS AND CONTRIBUTION

In a systematic review of the mixed reality language learning space published in 2021, Qiu et al.[8] found research had centered more on virtual reality(VR) than on AR. At the time of review, there were 87 articles addressing VR solutions in the space, and only 21 addressing AR solutions. In one of the studies addressing AR, Cheng et al.[2] found that location-based AR promotes an increase in user engagement and interest, which would result in a more focused learning effort. Thus, we know that not only is there much room for more work in the AR space, it has shown to be effective in engaging users.

With that said, there are many notable works using AR to teach languages. These works include Explorez[4] (an AR mobile app that turned University of Victoria campus into an immersive French environment), VocabEncouter[1] (a program that uses NLP to partially convert sentences on the web into the target language), and VocabuLARY[7] and WordSense[6], which both use object recognition to annotate real world objects with vocabulary and descriptors.

These systems prioritize teaching vocabulary and providing slow transitions for students, which is useful for lower level students. However, improvement to upper levels require more opportunities to practice the words and grammar they have spent so long learning, and receive feedback on their attempts.

3 FLUENCYAR

With these challenges in mind, we have designed a language learning solution called FluencyAR to mimic language immersion and self talk while mitigating drawbacks. It is highly self directed and

is not intended to teach, only to guide and correct. As a result, our target user is the upper level second language student.

FluencyAR is implemented using the Unity Engine, prototyped on desktop but envisioned as headset-based AR, to best imitate immersion. We employ a Wizard of Oz approach for the object recognition, by hardcoding a list of objects in the area and randomly selecting one. Using the selected object, we then send a HTTP request to the ChatGPT API, and parse the returned prompt and keywords. For real time speech recognition and translation, we work with the Microsoft Azure API.

The goal of FluencyAR is to allow the user to practice the target language in a way that closely imitates the fast-paced thinking and nature of real life, while still being supported and corrected. In order to achieve the larger goal of this system, we have defined 5 sub-goals; *Encouraging Self-Talk Practice (G1)*, *Encouraging Quick Recall (G2)*, *Supporting Quick Translation (G3)*, *Instant Feedback (G4)*, and *Higher-Level Reflection (G5)*. These goals were designed as part of the initial prototyping stage, inspired by the natural progression of learning a second language and adapting to conversations.

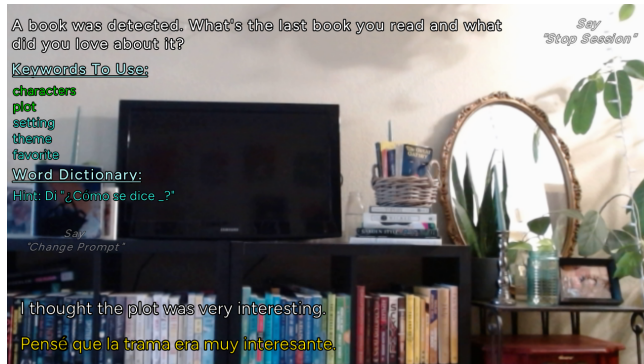
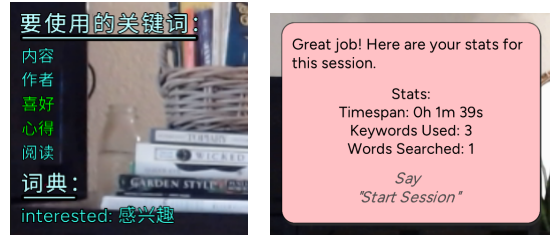


Figure 2: The FluencyAR system (Spanish speaker learning English)

To encourage self-talk practice (G1), we look at employing gamification tactics, and generating discussion prompts written in the target language based on the user's real life environment. As previously mentioned, a common barrier for self-talk is not knowing what to say, so the provision of prompts helps to encourage practice. Prompts are generated by ChatGPT as shown at the top of Figure 2, so we can ensure that they are always relevant and unique, keeping the practice sessions fresh. In case the user does not like the prompt, there is an option to easily regenerate.

Although this system is not designed to teach new vocabulary, the ability to quickly recall and use new words in a sentence is a crucial part of language fluency, and something that students will be required to do in real life conversations. As a result, encouraging quick recall is our second sub-goal (G2). As shown at the top of Figure 3(a), five vocab words related to the prompt are generated, which the user must include in their response. Successfully included words will turn green.

The next sub-goal (G3) is supporting quick translation. As shown at the bottom of Figure 3(a), a user can ask how to say a word in their native language and the translation will appear on screen. In an ideal real life conversation, one could ask their conversation partner



(a) ChatGPT Keywords/Word Search (English speaker learning Mandarin)

(b) End of session stats

Figure 3: FluencyAR System Closeups

how to say something, so adding this as an additional element of support does not undermine the immersive experience.

The final two sub-goals are both related to feedback, which is important for growth and traditionally a big missing piece when practicing alone. The first is instant feedback (G4). FluencyAR provides a real-time transcription of the speech in both the native and target languages, as shown at the bottom of Figure 2. By comparing what they meant to say to the native language translation, a user will have a better idea of whether they actually said what they meant to say. The higher-level reflection sub-goal (G5) refers to post-session feedback, as shown in Figure 3(b). The user can also review a locally saved dual language transcript of their session. This allows the user to review the session or even obtain feedback from a native speaker at a later time, and to see their improvement between sessions and over the long term.

Although a full system evaluation has not been conducted at this time, a handful of peers with second language learning experience were asked to test the system, and provide some commentary. The consensus was that FluencyAR is challenging yet engaging, and the ability to immediately search new words is very helpful. However, some users felt that they were too focused on answering the prompt to pay attention to the real time native and target language transcription. A couple of individuals shared that they would be highly inclined to incorporate FluencyAR into their everyday language study.

4 CONCLUSION

The combination of augmented reality and self-talk into a language learning system provides users with many opportunities to test themselves on topics they may not usually need to speak about, and challenges them to use new vocabulary in combination with their own language ability to form a coherent answer to the prompt. Future research directions include a full system evaluation, the incorporation of real-time object detection, as well as implementation on an augmented reality headset.

ACKNOWLEDGMENTS

We would like to thank Shivesh Jadon, Aaron Lengyel, and Azren Raju for their suggestions and contributions.

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [RGPIN-2021-02492] and the Canada Research Chairs Program.

REFERENCES

- [1] Riku Arakawa, Hiromu Yakura, and Sosuke Kobayashi. 2022. VocabEncounter: NMT-Powered Vocabulary Learning by Presenting Computer-Generated Usages of Foreign Words into Users' Daily Lives. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 6, 21 pages. <https://doi.org/10.1145/3491102.3501839>
- [2] Kun-Hung Cheng and Chin-Chung Tsai. 2012. Affordances of Augmented Reality in Science Learning: Suggestions for Future Research. *Journal of Science Education and Technology* 22 (Aug 2012), 449–462. <https://doi.org/10.1007/s10956-012-9405-9>
- [3] Maria C.M. deGuerrero. 1999. Inner Speech as Mental Rehearsal: The Case of Advanced L2 Learners. *Issues in Applied Linguistics* 10 (Jul 1999). <https://doi.org/10.5070/L4101005010>
- [4] Bernadette Perry. 2015. Gamifying French language learning: A case study examining a quest-based, augmented reality mobile learning-tool. *Procedia-Social and Behavioral Sciences* 174 (2015), 2308–2315.
- [5] Yoshiko Saito and Keiko K. Samimy. 1996. Foreign Language Anxiety and Language Performance: A Study of Learner Anxiety in Beginning, Intermediate, and Advanced-Level College Students of Japanese. *Foreign Language Annals* 29 (May 1996), 239–249. <https://doi.org/10.1111/j.1944-9720.1996.tb02330.x>
- [6] Christian David Vazquez, Afika Ayanda Nyati, Alexander Luh, Megan Fu, Takako Aikawa, and Pattie Maes. 2017. Serendipitous Language Learning in Mixed Reality. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI EA '17). Association for Computing Machinery, New York, NY, USA, 2172–2179. <https://doi.org/10.1145/3027063.3053098>
- [7] M. Weerasinghe, V. Biener, J. Grubert, A. Quigley, A. Toniolo, K. Pucihar, and M. Kljun. 2022. VocabulARy: Learning Vocabulary in AR Supported by Keyword Visualisations. *IEEE Transactions on Visualization & Computer Graphics* 28, 11 (nov 2022), 3748–3758. <https://doi.org/10.1109/TVCG.2022.3203116>
- [8] Xing yue Qiu, Chuang-Kai Chiu, Lu-Lu Zhao and Cai-Feng Sun, and Shu jie Chen. 2021. Trends in VR/AR technology-supporting language learning from 2008 to 2019: a research perspective. *Interactive Learning Environments* 0 (Jan 2021), 1–24. <https://doi.org/10.1080/10494820.2021.1874999>