Check for updates

Spelland: Situated Language Learning with a Mixed-Reality Spelling Game through Everyday Objects

Chia, Hsu
National Taiwan University of Science and Technology
d11010002@gapps.ntust.edu.tw

Yu, Chen
National Taiwan University
r11922026@ntu.edu.tw

Yu-Jung, Liu
National Taiwan University of Science and Technology
m11110302@mail.ntust.edu.tw

Yu-Cheng, Chang
National Taiwan University
r11922021@ntu.edu.tw

Min-Jui, Lee
National Taiwan University
r11944018@ntu.edu.tw

ABSTRACT

This work explores the use of mixed-reality (MR) technology to enable situated language learning using *everyday objects* in the environment around the *learners*. The learning method is based on Presentation, Practice, and Production (PPP), which cultivates the habit of independent learning through repetition, practice, and demonstration. In our game design, the *learners* first interact with real-world objects via MR, and the objects' spelling and their pronunciation will appear (Presentation), the *learners* repeat the pronunciation (Practice) to collect the letters of this objects, and finally the *learner* use the collected letters to spell out the *target words* (Production), which then transform into interactive 3D objects. We designed the learning experience and content tools using gestural UI, voice input, and object-to-word engine. Children in the preliminary user study found the game to be immersive, helpful in learning the spelling of the *everyday objects* and the *target words*, and additionally showed increased interests in learning about other nearby objects after playing the game.

CCS CONCEPTS

• **Human-centered computing** → Interaction design; Interaction design; Empirical studies in interaction design.

KEYWORDS

Mixed Reality, Tangible Interaction, Learning Experience, Language Learning, Situated Learning

ACM Reference Format:
Chia, Hsu, Yu, Chen, Yu-Jung, Liu, Yu-Cheng, Chang, and Min-Jui, Lee. 2023. Spelland: Situated Language Learning with a Mixed-Reality Spelling Game through Everyday Objects. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3544549.3583830

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3583830

1 INTRODUCTION

Learning a second language(L2) from an early age is generally understood as a valuable foundation for young *learners*' long-term language proficiency development. In general, spelling is a fundamental skill for learning a new language. Language educators also consider that spelling is essential for young children, and learning to spell enhances children's reading and writing. However, spelling can be a difficult and frustrating task for kids. Designing a good game could make learning to spell fun and engaging. We started by interviewing with L2 language teaching experts and found that situated language learning with *everyday objects* can effectively stimulate kids' curiosity and motivate them to learn new words relentlessly. Situated learning refers to the type of learning that takes place in a specific context and within a particular social and physical environment. Prior work [2] showed the potential of mixed reality in language learning. This work extends the situated learning concept and combines gestural and voice input into the game to immerse young *learners* in an input-rich, natural and meaningful context. Furthermore, a *facilitator* plays an important role in guiding young children to the learning activity. We built an additional lesson planning tool for parents and teachers. Finally, we conducted user testing with 1st to 3rd-grade Chinese children and found that *Spelland* is capable of encouraging kids to learn spelling and explore new words from surrounding objects in real life.

2 RELATED WORK

It has been shown that people are more motivated to learn when they find the content with respect to the situation is valuable or interesting [13]. Previous works have explored applying situated language learning in different aspects. MicroMandarin [1] and VocaBura [5] utilized the user's geolocation information to provide related content and a considerable amount of AR/MR systems that support language learning in real-world contexts, focus on providing labels of new words corresponding to real-world objects [2] [3] [4] [6] [8]. Due to its ability to contribute interactively to context-specific, just-in-time information, AR/MR is the preferred technology when it comes to situated learning. It could thus produce more immersive and engaging learning environments. Draxler et al. [3] developed a handheld AR app for learning vocabulary and case grammar by dynamically creating quizzes based on real-world objects in the learner's surroundings. Santos et al. [8] presented a handheld-AR system that displays text, images, animation, and sound on real-world objects. AR head-mounted devices (specifically, Microsoft HoloLens) are used in some applications to provide more immersive experiences. ARbis Pictus [6] labeled real-world objects with the corresponding vocabulary. AR's approach was more effective and enjoyable than a traditional flashcard-based technique, and participants could remember words better. VocabulARy [4] presented a system augmenting real-world objects with both first and second-language (language to learn) words. It visualized homophonic keywords to enhance memory retention. Generally, these studies showed AR might lead to better retention of words and improve student attention and satisfaction. Learning with AR/MR is also widely explored for children's education. MOW [7] presented a desktop-AR game for elementary students to learn vocabulary in different languages and found that children using AR had superior learning progress than those who used only traditional methods. TeachAR [9] presented a system combining desktop AR and speech recognition with teaching young children English colors, shapes, and spatial relationships and found that it improves their knowledge gain and enjoyment. AR-Poetry [14] introduced a handheld AR app to show interactive 3D models related to the texts in the book. WordSense [2] introduced a platform with object detection functionality and embedded rich augmentation on real-world objects, including words, definitions, example sentences, and related audiovisual content. It defines a "Serendipitous" (situated, incidental) Learning concept with MR technology. We addressed this concept in this work and made an engaging game with gestural and voice input for an immersive playful learning experience for kids.

3 INTERFACE INNOVATION

Compared to existing AR/MR language learning applications, the game targets early second-language *learners* aged 6 to 8 and uses mixed reality to provide ample opportunities to practice spelling and learn new vocabulary. We adopted the situated learning strategy [15] into the game to enable *everyday objects* as collectible words. To provide a better gaming experience, we used gestural and voice input throughout the interaction process. Therefore, *learners* can naturally point/manipulate the physical/virtual objects and practice pronunciation and spelling of the corresponding word in a fun way. The bubble effect, letter-collection task, and word-to-object trick made the kids unweariedly play with *everyday objects*.

In addition, we allow the *facilitator* to set up a learning context for their children, place *everyday objects* to learn around them, and follow the Silent way to guide the *learner*. To sum up, we introduce *Spelland*, a MR situated language learning game that enables the *facilitator* to design an interactive environment for the *learner* to increase spelling skills.

4 GAME DESIGN

Our game aims to familiarize children with the context of words and spark a desire to deconstruct and spell them out. The learning process is divided into two tasks. First, *learners* have to look up the surrounding objects and collect the letters by speaking out the objects' names. In this task, *learners* can practice the pronunciation and spelling of *everyday objects*. Second, *learners* can use the collected letters to spell the *target word*. Once the *target word* has been correctly spelled, the *target word* will transform into an *interactive 3D object* in mixed reality. Through this task, *learners* should remember the correct order of the *target word* and try to explore the demanded letters from *everyday objects*.

The whole learning activity utilized the Presentation, Practice, and Production (PPP) [10] teaching method. The purpose of these three stages is to cultivate the habit of independent learning through repetition, practice, and demonstration [11]. In our game design, the *learner* will interact with real-world objects in MR, and the object's word along with its pronunciation will appear (Presentation), the feedback will guide the *learner* to repeat the word (Practice), correct pronunciation will obtain the letters of this word, then the *learner* can use the obtained letters to spell out the *target words*, correct spelling will transform the *target words* into *interactive 3D objects* for completing the task (Production). Through the process, children can understand the word form and meaning and improve their learning effectiveness. Figure 1 shows the initial storyboard of the game. In VR mode, *learners* can overcome spatial constraints and learn new words by exploring beyond their daily experiences. This includes foreign festive customs, fairy tales, and ballads. In our prototype, *learners* enter a magical forest in VR mode where they can explore the environment and change the weather. They can also obtain treasures by spelling out target words, which enhances their overall gaming experience.

During the game process, the *facilitator* triggers tasks through appropriate interventions to help *learners* overcome difficulties in understanding the tasks or using the target language. In our game design, a *facilitator* is someone who helps to make learning easier by guiding *learners* through the process of acquiring knowledge or skills. On the other hand, the *learner* is actively engaged in acquiring new knowledge or skills [16]. Facilitators play a critical role in the learning process by creating a supportive and engaging environment that encourages *learners* to participate actively in the learning experience. This allows *learners* to feel comfortable asking questions and making mistakes.

5 PAPER PROTOTYPE

In order to test our concept, we built a paper prototype and conducted a pilot study with two kids, one in preschool and one in 1st grade of elementary school, via a Wizard of Oz approach. The test's goal is to simulate the letter-collecting and spelling tasks and see if this process is appropriate for children's English vocabulary learning.

Figure 2 provides the context for the paper prototype. We use paper-cut glasses as an MR headset and hand-made letter cards for the spelling task. In the experiment, the wizard asked the participant to point at one of the surrounding *everyday objects* (e.g., a puppet duck). Then the wizard hid the letter cards of the selected object (e.g., 'd', 'u', 'c', 'k') in the palm without the kid's notice. Once the participant spoke out the object's name, the wizard shook it and dropped the letter cards underneath it. The same process was repeated for the other objects. Afterward, the participant was guided to use letter cards to spell out a word in the task list (e.g., 'car'). In this step, the participants were asked to let their eyes closed for ten seconds so the wizard could have enough time to take out the letter cards and put a physical toy (e.g., a car) instead. The study showed that the children were amazed by the tricks, such as letter-dropping and word-to-object. The children showed high interest in other

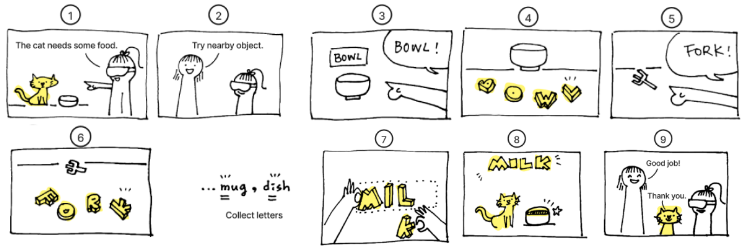

Figure 1: Storyboard. (Yellow for virtual objects) The learner asked to feed something to a hungry virtual cat. The *facilitator* guides the *learner* to search for surrounding objects to collect letters by practicing the pronunciation of *source words* (*everyday objects*' names). The *learner* should consider a *target word* that can solve the task (i.e., Milk that can be feed to the cat) and try to use collected letters to spell the *target word*. Once the *target word* has been correctly spelled, it will turn into an animated object as a reward.

Figure 2: Paper Prototype: (1) Environment setup, (2) Letters dropped from *everyday object*, (3) Spelling *target word*, (4) *Target word* turn into an *interactive 3D object*.

objects in the room and could spell the *target word* well in the game. The result encouraged us to move on to the next stage.

6 IMPLEMENTATION

This section describes the design and technical details of the interactive process, including the *learner* and *facilitator* side. The prototype was built by Meta Quest Pro using the Unity game engine.

6.1 Interactive design

6.1.1 Learner Side. The primary activity the *learner* will engage in while playing the game is letter interaction, including collecting letters from already-existing interactive objects and spelling words.

1. Obtaining letters from *everyday objects*

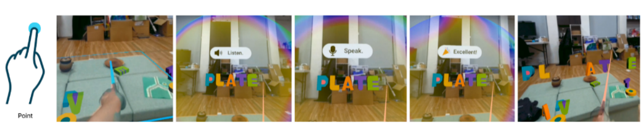

Figure 3: An example of obtaining letters from a "plate".

To obtain letters, the *learner* needs to Point to the *everyday object* they want to interact with and read out the *source word* of it. The complete process is as follows:

Pointing: When the *learner* extends the index finger (Point), a beam of light is emitted, and if the beam meets an *everyday object*, it changes color.

Listening: A bubble wrapping *Alphabet Bricks* spelling the *source word* will appear above the *everyday object* facing the *learner* once the beam has been on it for one second. The bubble will then play the pronunciation of the *source word*, with a panel in the bubble showing "Listen".

Speaking: After the pronunciation example, the panel changes to "Speak", and this is the time for the *learner* to speak out the *source word*.

Reward: If the *learner* pronounces the *source word* correctly, the panel changes to "Excellent," and the bubble will advance towards the *learner*'s direction before popping, or it may pop right away if it is close enough. The *learner* can interact with the *Alphabet Bricks* inside the bubble as they fall to the ground. When a *source word* is mispronounced, the process returns to **Listening**, giving the *learner* another chance.

2. *Alphabet Bricks*

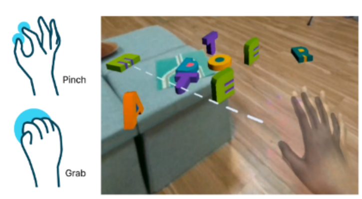

Figure 4: Gestures Pinch and Grab and an example of grabbing in a distance

We designed *Alphabet Bricks*, to realize the interaction with letters. *Alphabet Bricks* are affected by gravity and can be grabbed and released by the *learner*. The *learner* can conveniently grab *Alphabet Bricks* either up close or from a distance with a dotted line indicating which is being chosen. More specifically, grab interaction can be triggered by the gestures Pinch and Grab.

3. Spelling a word with *Alphabet Bricks*

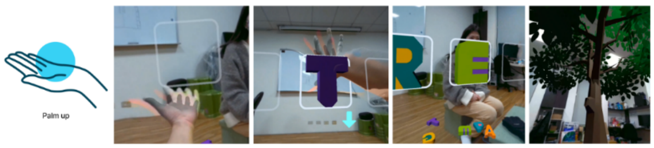

Figure 5: An example of spelling "tree".

To spell words, the *learner* needs to open the Spelling Board and attach *Alphabet Bricks* to it; once the *target word* is spelled, the corresponding *interactive 3D object* appears. The *learner* can open the Spelling Board by making the gesture Palm Up with their left hand, and it will appear above their palm and face them, remaining in that position until they make the gesture again to move it. The Spelling Board initially appears as a translucent box; after an Alphabet Brick is set, the box turns opaque white, and translucent boxes are then added on the left and right sides (to allow for the placement of further *Alphabet Bricks*). The *learner* can also simply grab the *Alphabet Bricks* and pull them off of the Spelling Board. The box where the removed Alphabet Brick was inserted will change back to translucent, and the boxes with no *Alphabet Bricks* on either side will be removed.

6.1.2 Facilitator Side. In addition to providing guidance during the game, the *facilitator*'s other crucial task is to prepare the gaming environment for the *learner* to play in before it begins. The *facilitator* should first login to the official website of our game and choose one or more words from a list of *target words* that we have established for this game. Following the *facilitator*'s selection, the system will offer some *source words* that can be utilized to reform the *target words*, which are some items that are easier to reach in everyday life, or *everyday objects*. The *facilitator* can add or remove *source words*, and the algorithm will determine whether the new *source words* can still produce the *target words*. When logging in from the game side, the information will be acquired from the game. The *facilitator* should then locate the real-world items in the *source words* and place them into the gaming environment. The *facilitator* may enter the game and select Setup Mode to set the position and rotation of *target words*. If not, *target words* will apply the default setting.

The information above pertains to MR Mode; if the *facilitator* wants to include VR Mode in the game, he or she must either choose the *target words* (such as Gate) that allow for VR Mode transition in MR Mode or decide to start the game in VR Mode. Last but not least, choose the environment theme to utilize in VR Mode (forest, volcano, underwater, human body, etc.). The chosen theme will determine the *target words* and *source words* in VR Mode. Due to the use of Wizard of Oz in the prototype we use for user testing, we must manually align the everyday objects with the virtual content in game (see 6.2.1). We streamlined the procedure by placing all functions in the game's setup mode and fixing the *target words* and source words used for testing.

6.2 Technical details

6.2.1 Object-to-word engine. Interacting with real-world objects is a crucial aspect of MR mode. Our objective is to automatically identify items in the surroundings and collect data on words, spatial positions, shapes, etc. We are unable to directly retrieve Passthrough camera data (real-world environment images) from Meta Quest Pro because of Meta's existing regulations. We first attempted to implement a workaround by online-casting the game screen to a webpage. Next, in order to perform object detection, the game screen is sent to an OpenCV application. The data is subsequently transmitted online to our game on Meta Quest Pro. The outcome, however, wasn't generally acceptable due to issues with internet latency, a lack of depth information for identifying objects' spatial placements, and an uneven transition from the processed image to MR perspective. Due to the inadequate results, we simulate the experience by employing a Wizard of Oz technique in our prototype. Before the game starts, we manually set up the interactive objects by positioning invisible 3D models in the appropriate locations. We make this a practical way to play by offering scalable 3D models and allowing the user (*facilitator*) to specify the word of the object by selecting from a menu or by speaking the word directly.

6.2.2 Text-to-Speech and Speech-to-Text. As mentioned in 6.1.2, an example pronunciation is needed for each *everyday object*, and accurately pronouncing the word is a required for *learners* to obtain *Alphabet Bricks*. In practice, we use the Oculus Voice SDK powered by Wit.ai to implement the functionalities, for its high-quality voice experience with low latency, cross-platform support and compatibility with Oculus app development in Unity. Text-to-Speech is employed since it's more scalable compared to simply using pre-recorded audio files. Speech recognition processes the player's voice into text to validate the correctness of the *learner*'s pronunciation.

6.2.3 Automatic Speech Recognition (ASR) and language learning. Kumar et al. [12] designed two mobile games for children to learn English words with understanding. Their studies showed that productive training involving vocalizing words is significantly superior to receptive vocabulary training. Yaniafari et al. [17] collated recent research showing that ASR is beneficial in assisting pronunciation learning. In our prototype, we emphasize the importance of correctly recognized pronunciation as the learning objective, but the *learner* may need help identifying pronunciation problems with this. To enhance our system for more accurate pronunciation acquisition, we plan to take phonemes into account and add orthographic and phonological hints in future development.

7 USER TESTING

We conducted a round of user testing, which included a semi-structured interview after the game. Based on the content of the semi-structured interview, we evaluated whether the interactive process of *Spelland* was able to generate interest, whether the children could operate it smoothly, and whether it could enhance children's interest and motivation in learning vocabulary for daily life.

7.1 Participants

In the user testing, we recruited three children and their parents to participate. The ages of the three child participants were 7, 8, and 9 years old (Grade 1, Grade 2, and Grade 3 in primary school). In terms of their English learning situation at school, the Grade 1 participant has currently learned 13 English letters out of 26, the Grade 2 participant has learned all 26 English letters, and the Grade 3 participant has learned all 26 English letters and basic vocabulary and sentences.

7.2 Procedure

In the experiment, the *facilitator* was played by the researcher, and the experiment was conducted in the laboratory to ensure there was enough space for the children to explore. During the experiment, the parents of the three children participants observed their children's use of the MR system to learn English vocabulary. First, because there was no object recognition function, the researcher would set up the *everyday objects* (mapping real-world objects) in *facilitator* mode according to the *source word*. In this experiment, we set up two *target words* in MR mode: 1. *Target word* [tree], *source words* [teapot], [paper]. 2. *Target word* [gate], *source words* [glove], [plate]. Then, the researcher would first explain the operation method to the child. After the explanation, the researcher will give the device to the child, and guide the child to explore and find the *source words* to spell the *target word* in MR, by pointing to the *everyday objects* to trigger word bubbles. According to the guidance of the interaction mechanism, after judging that the child has pronounced the word correctly, the child will get letters. When the child successfully completes the *target words* [tree] and [gate], they will enter VR mode. In VR mode, there are two *target words* set: [rain], [chest], *source words* [chair], [stone] [sign]. After the child successfully spells out these two *target words*, the experiment is completed. After the experiment, the researcher will separately invite the parents and the child to participate in a 3-minute semi-structured interview.

8 RESULT AND DISCUSSION

Whether the interactive process is able to generate interest. The participants felt very interested in the game. P1 described what was interesting: "Spelling is fun because you can find the letters, the process of finding the letters is fun, the letters falling down is also fun, the bubble will run out is also very surprising, and the appearance of objects in the spelling is amazing."

Whether the children could operate it smoothly. In terms of content, the participants were able to operate the gestures and speech recognition in the correct way. P1 found it the easiest to use: "Feels very easy to operate", P2 and P3 spent more time picking up *Alphabet Bricks*, but it was easy to put the *Alphabet Bricks* into the spelling box. The situation of not finding letters usually occurs due to distance problems or being obscured. In terms of hardware, P2 felt that the MR headset was heavy and a little tight, while the other participants did not report being uncomfortable at all.

Whether it could enhance children's interest and motivation in learning vocabulary for daily life. After using it, the participants wanted to learn more about the vocabulary of other *everyday objects*. Especially with the use of spelling, because the interactive way of the *Alphabet Bricks* made it interesting for them and made them feel more confident in speaking English.

We also conducted semi-structured interviews with parents, who did not interfere with the experiment but rather observed the children's learning behaviors and provided feedback. In general, parents were happy to see their children playing the game and learning new words. Regarding the future of teaching with MR/VR in education, all parents agreed that if teachers have time control and guide the students, they can use MR/VR to teach. As for the technical part, one mother expressed concern that the special design of the letters might affect the children's cognition and thought, suggesting that the letters should be displayed together with their upper and lower case forms and that they should not fall to the ground but float in the air for better recognition. But we found that all children could recognize the letters well and felt natural to search for the letters on the floor. This requires further investigation with various designs.

9 CONCLUSION

Our game allows children to develop an interest in understanding new vocabulary and motivates them to learn through the process of collecting and reorganizing words. Initial interview results show positive feedback, with children generally finding the interactive process interesting and quickly discovering letters and spelling words correctly. In terms of interactive design, children were particularly impressed by the special effects of words, such as the appearance of *Alphabet Bricks*, the bursting of bubbles, and the falling of *Alphabet Bricks*. In terms of gesture recognition, we observed that when children pinch letters, they will retract their arms, causing their hands to move out of the headset range and affecting recognition accuracy. In the future, the sensing technique should be improved or clear guidance can be provided on the scene to remind children not to exceed the operational range. In terms of speech recognition, it effectively recognized children's pronunciation, showing expected results. Overall, We discussed teaching methods, abilities, and challenges of integrating gesture, object, and speech recognition in *Spelland*. We believe mixed reality could be a powerful learning tool in the future.

ACKNOWLEDGMENTS

This study was supported by the National Science and Technology Council, Taiwan, under grant NSTC-111-2221-E-011-133-. We thank Dr. Neng-Hao(Jones) Yu and Mike Chen for their invaluable guidance and support throughout this project.

REFERENCES

[1] Darren Edge, Elly Searle, Kevin Chiu, Jing Zhao, and James A. Landay. 2011. MicroMandarin: mobile language learning in context. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). Association for Computing Machinery, New York, NY, USA, 3169–3178. https://doi.org/10.1145/1978942.1979413

[2] 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 (CHI EA '17). Association for Computing Machinery, New York, NY, USA, 2172–2179. https://doi.org/10.1145/3027063.3053098

[3] Fiona Draxler, Audrey Labrie, Albrecht Schmidt, and Lewis L. Chuang. 2020. Augmented Reality to Enable Users in Learning Case Grammar from Their Real-World Interactions. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376537

[4] M. Weerasinghe *et al.*, "VocabulARy: Learning Vocabulary in AR Supported by Keyword Visualisations," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 11, pp. 3748-3758, Nov. 2022, doi: 10.1109/TVCG.2022.3203116

[5] Ari Hautasaari, Takeo Hamada, Kuntaro Ishiyama, and Shogo Fukushima. 2020. VocaBura: A Method for Supporting Second Language Vocabulary Learning While Walking. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 3, 4, Article 135 (December 2019), 23 pages. https://doi.org/10.1145/3369824

[6] A. Ibrahim, B. Huynh, J. Downey, T. Höllerer, D. Chun and J. O'donovan, "ARbis Pictus: A Study of Vocabulary Learning with Augmented Reality," in IEEE Transactions on Visualization and Computer Graphics, vol. 24, no. 11, pp. 2867-2874, Nov. 2018, doi: 10.1109/TVCG.2018.2868568

[7] J. Barreira, M. Bessa, L. C. Pereira, T. Adão, E. Peres and L. Magalhães, "MOW: Augmented Reality game to learn words in different languages: Case study: Learning English names of animals in elementary school," 7th Iberian Conference on Information Systems and Technologies (CISTI 2012), Madrid, Spain, 2012, pp. 1-6.

[8] Santos, M.E.C., Lübke, A.i., Taketomi, T. *et al.* Augmented reality as multimedia: the case for situated vocabulary learning. RPTEL 11, 4 (2016). https://doi.org/10.1186/s41039-016-0028-2

[9] C. S. C. Dalim, T. Piumsomboon, A. Dey, M. Billinghurst and S. Sunar, "TeachAR: An Interactive Augmented Reality Tool for Teaching Basic English to Non-native Children," 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Merida, Mexico, 2016, pp. 344-345, doi: 10.1109/ISMAR-Adjunct.2016.0113.

[10] Nunan, David. Research Methods in Language Learning. 1st ed. Cambridge University Press, 1992.

[11] Thornbury, Scott. How to Teach Speaking. 1st ed. Pearson Education ESL, 2005.

[12] Anuj Kumar, Pooja Reddy, Anuj Tewari, Rajat Agrawal, and Matthew Kam. 2012. Improving literacy in developing countries using speech recognition-supported games on mobile devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12). Association for Computing Machinery, New York, NY, USA, 1149–1158. https://doi.org/10.1145/2207676.2208564

[13] Pintrich, P. R. (2003). A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. Journal of Educational Psychology, 95(4), 667–686. https://doi.org/10.1037/0022-0663.95.4.667

[14] Nan Chao, Shengge Yang, Yuxian Qin, Zeming Song, Zhaofan Su, and Xiaomei Nie. 2022. AR-Poetry: Enhancing Children's Motivation in Learning Classical Chinese Poetry via Interactive Augmented Reality. In The Ninth International Symposium of Chinese CHI (Chinese CHI 2021). Association for Computing Machinery, New York, NY, USA, 162–166. https://doi.org/10.1145/3490355.3490518

[15] Lave, Jean, and Etienne Wenger. Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, 1991.

[16] Willis J. A framework for task-based learning. Longman; 1996.

[17] Putri Yaniafari, R., Olivia, V. and ., S. 2022. The Potential of ASR for Improving English Pronunciation: A Review. KnE Social Sciences. 7, 7 (Mar. 2022), 281–289. DOI:https://doi.org/10.18502/kss.v7i7.10670