Outcomes of VR, AR and MR Technologies in K-12 Language Education: A Review

Yu Shen¹, Dongqing Zhou¹, and Yang Wang²*

¹ College of Foreign Languages, Wenzhou University, Wenzhou, China; Email: Shenyu0722@163.com (Y.S.), 2226045570@qq.com (D.Z.)
² College of Mechanical and Electrical Engineering, Wenzhou University, Wenzhou, China
*Correspondence: wyfellow@163.com (Y.W.)

Abstract—Since the 21st century, Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR) technologies have been increasingly used in the field of language education to promote student motivation, engagement, and so on. This paper reviews the results of 180 reviews of empirical studies on the use of VR, AR, and MR to enhance K-12 teaching or learning from 2012 to 2022 and summarizes the trends. Based on Bloom’s classification theory of instructional goals, namely Cognitive Goals, Affective Goals, and Behavioral Goals, the outcomes of 53 papers are summarized and analyzed from these three dimensions to find the function of VR, AR, and MR in K-12 Language Education, which can improve literacy, creative thinking, communication, collaboration, confidence, participation, and enjoyment in the digital era. And it also promotes self-learning, enables multisensory learning, and reduces cognitive load. Despite the technological complexity and problems associated with the use of AR in education, VR, AR, and MR applications can successfully improve language learning. Compared with previous research approaches, this study first constructs a VR, AR, and MR-supported learning model based on their characteristics respectively and the relevance of these technologies.

Keywords—virtual reality, augmented reality, mixed reality, K-12 language education

I. INTRODUCTION

Consolidating language knowledge often requires a lot of boring and repetitive exercises. Even with the help of pictures and videos, it is difficult for children and adolescents to arouse their interest and stay focused in traditional classrooms. For example, mastering phonetic symbols is an important task in elementary school English curricula, but without vivid animations and guidance, students may fail to gain overall understanding simply by watching the teacher’s mouth. Moreover, creating authentic language environments is often limited by a variety of factors, such as lack of teaching materials, safety concerns, or geographical distance hardness.

To address these challenges, researchers have resorted to computer technology, which contributes to students’ learning, acquisition of language skills, and development of language application skills. Generation Z were born in the digital age where technology permeates almost every aspect of their lives. They required more personalized, mass data, and enjoyable education [1]. For example, instead of watching videos of earthquakes to understand the process of disaster, students can use virtual reality technology to immerse themselves in a real earthquake environment from visual, auditory, and tactile sensory to clearly see the whole process and the immense power of earthquakes to better carry out language learning on the topic of earthquake.

In recent years, researchers have seen progress in computing technologies, Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR) are the favorable ones. To be specific, VR immerses the user in an artificial world [2], while AR allows virtual objects to be placed in the actual world [3]. MR describes real-world settings that allow people to interact with virtual experiences [4].

As VR, AR and MR technologies developing, they are increasingly being applied to other areas, including language education. These technologies contribute by creating immersive settings for first or second language learners, giving them a competitive advantage over traditional multimedia. Supported by a variety of technologies such as computer graphics, affective computing and sensor technologies, that can greatly enhance the user experience [5]. What’s more, with the multimedia, high-quality interaction and self-openness offered by VR, AR and MR technologies are much more powerful than traditional media. With these characteristics, they can successfully enhance language acquisition while overcoming the limits of traditional media. Despite the many potential uses of these technologies in education, some educators have questioned how to overcome the barriers to K-12 technology integration when schools are hard to integrate [6]. In addition, the use of them is too complex to operate correctly.

This study noticed two flaws in prior analyses after assessing the outcomes of VR, AR, and MR technologies in language acquisition from 2012 to 2022. For starters, the majority of the papers examined only looked at studies that were published before 2019, Solak et al. [7] looked at articles from 1995 to 2015, Lin et al. [8] from...
2004 to 2013, and Khoshnevisan et al. [9] from 2007 to 2017, Parmaxi et al. [10] conducted a systematic analysis of studies published from 2014 to 2019, identifying the main benefits of augmented reality in language learning. In contrast, studies published in the last three years in this field have not been reviewed. In addition, most of the reviewed studies investigated the benefits of VR, AR and MR technologies for language education, and many specific aspects of VR, AR and MR have not been thoroughly discussed, such as how can augmented reality and virtual reality can be integrated into language learning and how they specifically promote language learning. Thus, evaluating how virtual reality, augmented reality, and mixed reality are utilized in language acquisition, the major findings, why VR, AR, and MR are useful in improving language learning, and what the future implications are for the study are overwhelmingly well-deserved. In summary, the previous reviews on technologies in language education only focus on the advantages of them, rather than providing a comprehensive and in-depth discussion of how they can facilitate language learning. Therefore, this paper discusses the definitions of language education, AR, VR, and MR at the beginning, and then from Bloom’s classification dimensions of teaching goals, that is Cognitive Goals, Affective Goals, and Behavioral Goals to classify the learning outcomes of the selected 53 papers. Not only the status and significance of research and development are illustrated, but also valuable research directions are finally proposed.

II. REVIEW METHODS

A. Inclusion and Exclusion Criteria

It is overwhelmingly significant to review based on high-quality publications. Hence, in this study, a preliminary search was performed in April 2022, with the application of Bloom’s logic (virtual reality or VR or augmented reality or AR or mixed reality or MR in subject terms) AND (“language education” or “language teaching” or “language learning” in abstract) AND (“primary school” or “elementary school” or “primary education” or “high school” or “k-12” in abstract). To ensure both quality and accuracy, only peer-reviewed journal papers with full text available have been included. This paper establishes the following inclusion and exclusion criteria (Table I), and reviews each paper to determine whether it is eligible for analysis.

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students used VR/AR/MR</td>
<td>Not using VR/AR devices to learn</td>
</tr>
<tr>
<td>The participants were primary or secondary school or high school students</td>
<td>For preschool children, special education, college students, teachers and other adult learners</td>
</tr>
<tr>
<td>Learning of language</td>
<td>Non-language subjects</td>
</tr>
<tr>
<td>Written in English</td>
<td>Written in other languages</td>
</tr>
</tbody>
</table>

B. Definition of Language Learning

Influenced by behavioral theories, early research on language teaching and learning mainly focused on language and teaching methods. However, when cognitive theories flooded, researchers discovered that language learning is a process in which learners develop their own systems in social practice. Compared to the traditional learning, the real use of a language is both complicated and hard to implement.

The requirements of sociality and autonomy in the language learning process often struggle to meet. Today, the use of technology has become a major focus of research and offers new solutions. Technology should be a driving force in creating a technological environment for language learning. In the 21st century, the role of technology in language learning has changed dramatically, from a learning tool to a created environment where learners could interact and corporate with others [11]. In recent years, the boosting VR, AR, and MR technologies have brought hope to language learners.

C. Definition of VR

VR is a 3D virtual world through which users get visual simulations that allow them to feel in an environment [12]. Virtual reality appeared in the 1960s, but it did not draw much attention before 2000. In recent years, virtual reality has advanced and grown significantly. With a three-dimensional space where they can experience their own learning, its competitiveness identified some benefits of virtual reality in language learning, including visual support, increased interest, and authentic learning environment [13].

D. Definition of AR

In language education, such as English, Chinese, French, Korean, Turkish, etc. Augmented reality is a three-dimensional (3D) technology that displays digital data in the real environment [14]. This environment is built on augmented reality technology, which allows users to engage organically with virtual subjects. It has a favorable influence on learning by improving performance, increasing motivation and involvement, and encouraging learners to collaborate.

E. Definition of MR

Extending along the virtual unity between AR and VR is MR. Mixed Reality is a combination of them that provides the ability to physically interact with virtual objects in the real world [15]. By allowing the complete body to be combined with both real and virtual aspects, MR differs from AR, which simply superimposes digital items onto the actual environment [16]. MR is more capable and expensive to process than AR and VR because of its capacity to interact with digital overlays in the real environment. Continuous technological advancements will allow for more consideration of its capabilities in K-12 educational settings.
III. STATISTICAL RESULTS FROM SELECTED STUDIES

The learning results of 53 papers were divided into three dimensions using Bloom’s classification theory of instructional goals. As shown in Fig. 1, 36 papers set cognitive goals, of which 7 reckon mixed effects; 30 papers established affective goals, of which 6 reckon mixed effects. And 5 papers aiming to improve behavior are all positive.

A. Cognitive Goals

Thirty-six papers focused on achieving cognitive goals. Thirty-two papers concluded that VR, AR, and MR technologies can be effective in improving course performance, enhancing understanding of abstract concepts and linguistic phenomena, promoting students’ knowledge, improving students’ language skills, and achieving an overall improvement in literary quality; or assess the effectiveness of students’ mastery of the listening, speaking, reading, and writing process, while connecting prior knowledge with the new one.

Fig. 2 shows different skills in achieving cognitive goals. Among the many goals of language learning, phonics or speaking is the first goal of VR language learning, with 18 empirical studies on speaking improvement. These strengths of VR/AR shown here demonstrate that VR/AR-assisted language learning places a high premium on how to use the language in real life. In addition, vocabulary (12), listening (8), and writing (6) were mentioned several times. In contrast, less attention was paid to grammar and reading, with less than 3.

1) Knowledge

Since listening, reading, and writing require a certain base of words, accumulating words by memorizing them is the only way to learn a language well. We analyze the first level of knowledge—cognitive goals—by taking the role of VR, AR and MR technologies on word memorization as an example. Researchers discovered that AR or VR technology is beneficial in increasing pupils’ word recognition, according to Ibrahim et al. [17], Solak and Cakir [18], and Vedadi and Somaiyeh [19]. They attributed this to the fact that AR and VR present word knowledge in a real language environment, which facilitates students’ understanding of word meanings and consolidates morphosyntactic associations. In the study conducted by Ibrahim et al. [20], students learned vocabulary using Microsoft’s AR head-mounted display, HoloLens. Learners walked around the room with the AR device and saw the real stuff labeled with meaning annotations. The findings show that employing virtual labels helps students acquire vocabulary more effectively because AR technology allows pupils to retain words more thoroughly by presenting objects in front of them.

However, the main goal of AR, VR, and MR-assisted devices in these 14 papers is basic vocabulary learning, which suggests that they are still mainly used among basic language skills without further application. Therefore, apart from creating environments with imaginative and innovative features, VR/AR technologies should demonstrate how they can reproduce the real social context of the target language to provide more visual, imaginative learning activities and content.

2) Comprehension

Alakärppä’s study is one of the limited studies that used the AR application Aurasma to practice speaking and listening. Students were asked to collect the correct objects in a cluttered room [21]. Listening to an oral description of a certain stuff, they were asked to organize it in their own language to retell their group members. By describing the object in the target language, their apprehension will be improved, but not profoundly and thoroughly.

In addition, Kurilovas [22] investigated spoken language learning with VR support and found that students who used VR tools performed better in terms of fluency, pronunciation and content. Using Google Cardboard and Google Expedition, participants were asked to play the part of tour guides to introduce famous Chinese sites to the audience. Since VR technology helps create opportunities to speak by presenting verbal and visual information, it can improve students’ speaking skills indeed.

3) Application

Unlike the traditional Computer-Supported Collaborative Learning (CSCL) environment, in the
virtual environment, learners can interact face-to-face in real time through roles and corporate with teams.

For students to practice communication skills, Lee et al. [23] designed a virtual gaming environment. Students accomplished the tasks in the game Ghosts Out by interacting with their friends. The game was discovered to induce a lot of speaking, and students would utilize their previously taught information to converse with other participants on their own. Also, the use of AR and VR technology to support writing offers learners with visual content that allows them to have a deeper understanding of writing content to facilitate applied practice.

4) Analysis
The use of VR/AR ensures a stable organized communicative activity and continuous interaction. It also does so by creating a platform for communicating, collaborating, recording, and analyzing students’ performance.

VR/AR stands out because it provides a free and active learning environment that transcends time and space. It is conducive to enhancing learners’ reasoning skills, critical thinking skills, and developing creative ways [24]. This leads to clearer organization of the material, and detailed elucidation of some theories. For example, Lorenzo et al. used VR technology to help students compare and discriminate the pronunciation of different vowels, and guide them to spell words on their own [25]. Comparing their own scores with their peers at the same second increased their self-efficacy. In the bottom-up model, where information is processed with attention to detail in the language cognition process, the input information is gradually processed from lower to higher levels, with individual vowels learned first, words and sentences later, enhancing the learners’ own inductive, reasoning, and problem-solving skills.

5) Synthesis
Synthesis is the process of comprehensively processing elements based on analysis to solve problems in an integrated and creative way. It is a high-level requirement because it emphasizes identity and originality. VR, AR, and MR are suitable for contextual learning because they create an environment in which reality and the virtual overlap. In such an environment, task-based language learning, there are targeted activities, rich social scenarios, and immediate feedback and guidance from the teacher. With these advantages, learners can enjoyably complete complex work through engagement and collaboration with others.

6) Evaluation
This is the highest level of educational goals in the cognitive domain, requiring a rational and profound judgment of the essential value of things.

AR technology provides cultural learning by allowing learners to interact with cultural-related content, while virtual reality technology places learners in cultural situations to enhance their understanding of culture. Chang et al used VR to promote cultural understanding [26]. Students used Google Cardboard to watch 360-degree videos that immersed them in cultural contexts, which allows learners to closely explore myths of the target culture, thereby promoting cultural understanding and integrating internal and external data and information to make an inference objectively.

B. Affective Goals
Fig. 3 shows secondary dimensions of affective goals in 30 papers mentioned. The affective domain teaching agenda was proposed in 1964 and divided into five levels, 30 papers on affective goals can be further classified to motivation (12), attitude (8), participation (5), satisfaction (4), interests (4), self-efficacy (3), cognitive load (3), technology acceptance (3), expectation of success (2), etc.

![Figure 3. Secondary dimensions of affective goals in 30 papers mentioned.](image)

1) Reactions
There is a consensus that AR, VR, and MR enhance learners’ enjoyment, stimulate their curiosity, and increase motivation. When role-playing games in language teaching, the findings showed that VR and MR-based language learning workshops supported knowledge co-construction. Following the cue prompts, they can complete role tasks voluntarily with high motivation to learn, which could facilitate higher-order learning.

In addition, several studies on AR learning have focused on the impact of AR on motivation and attitudes. Han and Jo [27] found that AR improved learner satisfaction; Lu [28] revealed increased confidence; Ibanez [29] mentioned increased enjoyment; Cheng [30] claimed that AR increased student participation; Zhang [31] claimed that AR increased student interest in learning. Although numerous researches have been conducted on students’ motivation and attitudes because of AR-assisted learning, only a few have looked at motivation and attitudes in all aspects.

2) Organizational value system
A study by Di Serio showed that by using investigation equipment, some groups failed to discover the source of the haunting [32]. The data showed that the probability of task failure was approximately 8%. Since the game poses a challenge to learners, conflicts and contradictions will arise in communication and cooperation due to differences in values, if eliminated, it can lead to the establishment of an internally consistent value system.

The use of these three technologies can stimulate learners’ motivation, attitude and interest in learning, which in turn can lead to more active participation in
learning, resulting in better satisfaction, expectations, good mood, more healthy values. Most studies conclude that the use of them has higher technology acceptance than non-VR, AR and MR technologies and lower cognitive load than or the same cognitive load as the counterparts.

This is since the virtual worlds and objects generated by VR or AR give students with experiences that may substitute the actual world, as well as chances for inquiry-based learning, which allows students to have fun while learning. Chang [33] discovered, for example, that an AR-based flipped learning method not only improved students’ project achievement but also their group self-efficacy.

However, there are still some studies that have reached negative conclusions. The dimensions are technology acceptance, self-efficacy, and expectation. This because the use of technologies is too complex to operate correctly, or the information provided is insufficient, making learning difficult. Wu et al. [34] found that the experimental group that used AR had lower motivation to learn than the control group that did not use AR. The authors attributed the main reason for this to learners’ unfamiliarity with the materials and devices, which posed some learning challenges. Chen [35] discovered that the perceived utility of AR was related to age. The authors predicted that it is more difficult for the older pupils to understand or follow the teacher’s directions. Learners did not appreciate the desktop VR experience, according to Shin, since it did not provide intense immersion [36].

C. Behavioral Goals

Five papers focused on the goal of behavioral attainment, which shows that the use of VR, AR, and MR can improve students’ behaviors. Yoon and Wang [37] found the interaction time and teamwork time between AR users has significantly increased than non-AR users. This suggests that AR devices increase participation and affect cooperation between teams to some extent.

1) Operation

This refers to the student’s ability to act on cues, but not imitative observations, such as performing or practicing actions as instructed. This means that students should be able to perform independently. For example, after the teacher explains how to pronounce vowels, how to raise the tongue, etc. In addition, there is an article used generic data from multimodal learning analysis in Legault’ studies [38]. It uses Magnetic Resonance Imaging (MRI) to observe neural activity during word learning and AR headsets to record eye movements at the same time.

2) Habituation

Immersion and presence directly relate to the experience of participation and learning gains in virtual environments [39]. Research has shown that students are able to develop or consciously use language in the virtual environment, which can help them form habits of muscle memory more quickly.

IV. CONCLUSION

Research on the educational uses of VR, AR, and MR for language has skyrocketed in the past four years. Several studies have highlighted the potential of them to provide deeper language learning [40–45]. Wu et al. [46] studied the impact of virtual reality on learners, finding that most learners achieved language and emotional achievements using these three technologies. In real life, cognitive, affective, and behavioral aspects often occur simultaneously. For example, when students write (behavioral), they are also memorizing and reasoning (cognitive), while they develop emotional response to the task (affective). Therefore, teachers often need to set goals in all three areas simultaneously before and during instruction.

Meanwhile, rapid advances in technology have made it easier and less expensive to consume and invent the software and hardware needed for AR and VR. While MR access remains a challenge because the need for computing is still beyond the scope of most K-12 school environments. In addition, most studies proposed or used software or application design. However, most of them highlight the impact of that software or app, rather than research that focuses on consideration of learning opportunities in the digital environment [47].

The biggest challenge in the literature review is finding research that examines the use of these technologies for K-12 educational purposes. The lack of research in these areas is due to two main reasons. Firstly, the rapid proliferation of AR and VR technologies, accessible through the ubiquity of smartphones, which make it possible now; Secondly, the varying availability in technology between different regions.

V. FUTURE WORK

The current study focuses on using AR, VR and MR technologies to present knowledge to students. In this perspective, these technologies are primarily viewed as an alternative to traditional methods of distributing or displaying information. On the other hand, it’s crucial to consider schools’ capacity to consume as well as their desire to create and generate utilizing emerging technology. From students’ perspective, students must develop digital skills to handle and interpret this massive amount of data, rather than merely consume it, since data rises rapidly year after year. In the digital age, these technologies also raise new questions about how we represent, consume, and manipulate information in non-text representations, and further study into AR technology is needed in K-12 education.

Students now demand a degree of digital literacy that is not already required in the curriculum. Although VR, AR and MR technologies have been in development for more than two decades, the educational applications of these technologies remain superficial [48]. These reviews mainly focus on the research focus and advantages of them, but do not make a comprehensive and in-depth discussion on how VR, AR and MR technologies can promote language learning.
CONFLICT OF INTEREST
The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS
Yu Shen conducted the research; Dongqing Zhou and Yang Wang analyzed the data; Yu Shen wrote the paper; all authors had approved the final version.

ACKNOWLEDGMENT
The deepest and sincerest gratitude goes to the professors in the College of Foreign Languages at Wenzhou University. I would also gratefully acknowledge the support of my dearest friends and research group mates for their encouragement friendship and support since the beginning of my study. I am extremely grateful to my parents for their motivation, patience love and care during my education and life journey. Without their help, this paper would not have been accomplished. My final thanks go to all the people who have directly or indirectly supported me in completing this paper.

REFERENCES


