- ORIGINAL ARTICLE -

Inclusive Learning through Immersive Virtual Reality and Semantic Embodied Conversational Agent: A case study in children with autism

Aprendizaje Inclusivo mediante Realidad Virtual Inmersiva y Agente Conversacional Semántico Encarnado: Un caso de estudio en niños con autismo

Yoselie Alvarado[1,2], Roberto Guerrero[1] and Francisco Serón[3]

[1] *Laboratorio de Computación Gráfica, Universidad Nacional de San Luis, Argentina*
[2] *Instituto de Física Aplicada, CONICET, Universidad Nacional de San Luis, Argentina*
{ymalvarado, rag}@unsl.edu.ar
[3] *Instituto de Investigación de Ingeniería de Aragón, Grupo GIGA, Universidad de Zaragoza, España*
seron@unizar.es

Abstract

Sustainable Development Goal 4 from the United Nations 2030 Agenda, focus on ensuring inclusive and equitable quality education and promoting lifelong learning opportunities for all. Inclusive education is a key component of sustainable development goal 4, and assistive technology is a critical factor in achieving it. In this context, this paper introduces the development of an immersive virtual reality system with conversational skills for inclusive learning processes. The idea of this work is to provide an assistive technology to motivate, educate, and train students with disabilities in a more inclusive way. In order to improve the user experience, the system was developed to work in a CAVE-like immersive environment using Natural Language Processing through a Semantic Embodied Conversational Agent. This work highlights that assistive technology can promote educational, psychological, and social benefits for students with disabilities. The use of assistive technology can enable academic engagement and social participation and be transformative from a psychological perspective. A case study was conducted on children with autism, which showed encouraging results of the system as an assistive technology for teaching and learning.

Keywords: Assistive Technologies, Inclusive Education, Immersive Virtual Reality, Natural Language Processing, Autism.

Resumen

El Objetivo de Desarrollo Sostenible 4 de la Agenda 2030 de las Naciones Unidas se centra en garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje permanente para todos. La educación inclusiva es un componente clave del Objetivo de Desarrollo Sostenible 4, y la tecnología de apoyo es un factor crítico para lograrlo. Este trabajo presenta el desarrollo de un sistema de realidad virtual inmersiva con habilidades conversacionales para procesos de aprendizaje inclusivo. La idea de este trabajo es proporcionar una tecnología de apoyo para motivar, educar y formar a estudiantes discapacitados de una forma más inclusiva. Para mejorar la experiencia del usuario, el sistema fue desarrollado para trabajar en un entorno inmersivo tipo CAVE utilizando Procesamiento del Lenguaje Natural con un Agente Conversacional Semántico Encarnado. Este trabajo pone de relieve que la tecnología de apoyo puede promover beneficios educativos, psicológicos y sociales para los estudiantes con discapacidad. El uso de tecnología de apoyo puede permitir el compromiso académico y la participación social y ser transformador desde una perspectiva psicológica. Se realizó un caso de estudio en niños con autismo, que mostró resultados alentadores del sistema como tecnología de asistencia para la enseñanza y el aprendizaje.

Palabras claves: Tecnologías de apoyo, Educación Inclusiva, Realidad Virtual Inmersiva, Procesamiento del Lenguaje Natural, Autismo.

1. Introduction

Nowadays, computers have become one of the most essential objects in people's daily lives. These devices are not only used as a hobby or entertainment tool, but are also very important for academic or work tasks. In addition, the constant evolution of hardware devices provides more and more computing capacity and, therefore, allows the execution of more complex applications.

In recent years, Virtual Reality (VR) technologies have begun to be used as an assessment and intervention tool for multiple purposes. Virtual environments provide users with safe access to interactive, true-to-life situations that would otherwise be inaccessible to them due to motor, cognitive and psychological limitations. Among these new emerging technologies, a Cave Automatic Virtual Environments (CAVE) stand out, where in particular the user is immersed in computer-generated environments that simulate reality through the use of real-time interactive devices that send and receive information (peripherals such as glasses, helmets, gloves, suits, among others). Simultaneously, the “telepresence” or illusion of “being there” is controlled by sensors that capture the user's actions and adjust accordingly what is displayed on the screen in real time. Currently, they have achieved great significance for hospital projects [1], health benefits [2], scale cognition and learning [3] and learning in undergraduate courses [4].

At the same time, increasing dependency of humans on computer-assisted systems has led to researchers focusing on more effective communication technologies that can mimic human interactions as well as understand natural languages and human emotions. In this context, virtual reality helps a lot to deceive the user, but it needs several means to make user interaction easier and increase credibility.

Natural User Interfaces (NUI) are a type of human-computer interaction that attempt to give the user machine communication skills through intuitive verbal and non-verbal actions. In this way, the systems offer an invisible form of communication with the user and reduce the cognitive load required to perform the interaction [5].

For this purpose, including some form of natural user interface is an attractive solution to this interaction issue. Embodied Conversational Agents (ECA) are a new kind of Human-Computer Interface that is embodied and has conversational skills [6]. ECA promise to increase the quality of communication between humans and computers, as they are designed to communicate and interact in a human-like manner.

In order to hold a conversation, they must exploit a part gamut of natural language research, from speech recognition and keywords detection to synthetic voice generation and speech synthesis.

Natural Language Processing (NLP) is one of the effective technologies that can be integrated with advanced technologies (such as machine learning, artificial intelligence, and deep learning) to improve the process of understanding and processing the natural language. It has become embedded in our daily lives: automatic machine translation is ubiquitous on the web and in social media; text classification keeps our email inboxes from collapsing under a deluge of spam; search engines have moves beyond string matching and network analysis to a high degree of linguistic sophistication; dialog systems provide an increasingly common and effective way to get and share information [7]. Over the last two years, the field of Natural Language Processing has witnessed the emergence of several transfer learning methods and architectures which significantly improved upon the state-of-the-art on a wide range of NLP tasks [8]. For example, to study methods to mitigate gender bias in NLP [9] or their applications in commerce, e-governance, healthcare, education, among others [10].

Education is the most essential ingredient in the development and empowerment of individuals, and inclusion in education irrespective of the varied socio-economic differences and the differences in “abilities” and “disabilities”, undoubtedly makes this foundation much stronger. Nowadays, because of the new educational politics of inclusion (both at higher, initial and middle level) there is an urgent need for technologies that facilitate such inclusion. All this is driven by the ONU 2030 Agenda, especially in the Sustainable Development Goal (SDG) 4 and SDG 10. These new inclusion policies require rethinking the tools to support traditional teaching methods; one possibility is to make use of technologies that facilitate the teacher's work.

In this context, the ability to change the virtual environment relatively easily, to grade task difficulty and to adapt it according to the students' capabilities are important advantages of VR, since these features are essential to learning.

The present work is an achievement from the joint effort between Laboratorio de Computación Gráfica (LCG) [11] and Grupo de Informática Gráfica Avanzada (GIGA) [12]. It combines immersive virtual reality with conversational skills as an assistive technology for inclusive education by using a CAVE-like environment, ECA and NLP. The ultimate goal is to motive, educate and train in an inclusive way.

In particular, the LCG has so far focused on the development of immersive environments that support multimodal and especially gestural interaction. The result of LCG efforts has been IVI CAVE (Immersive Virtual Innovation CAVE), a low-cost and powerful engine to manage real-time interaction within an immersive environment [13, 14]. The GIGA research has focused so far on developing interactive virtual agents that support multimodal and emotional interaction. The results of that effort have been Maxine, a powerful engine to manage real-time interaction with virtual characters [15] and VOX System, an embodied conversational agent exploiting linked data [16].

The paper is organized as follows: section 2 describes already existing examples of assistive technologies, section 3 details the system technical aspects, section 4 describes previous experiences, section 5 lists the possible disability that could be helped by this assistive technology focusing on learning, section 6 presents a case study. Finally, the conclusions section outlines the contributions of the work.

2. Related Works

The inclusive education for learners with disabilities is often not possible without their access to fit-for-purpose Assistive Technology (AT) as the barriers to their education are often environmental.

This technology consists of a range of devices and services that work to support students, augmenting their existing abilities, compensating for, or bypassing difficulties they may experience. Some AT has been specifically developed for functional use, while others, particularly emerging technology, can be adapted for, or used, in an assistive capacity. Where the AT promotes social interaction, curriculum access and the ability to express understanding, there is the potential for heightened inclusion in the classroom [17].

Some examples of AT are: the development of a robot with a gripper that enables children with motor and speech impairments to manipulate educational materials [18], a role-playing game method to teach social skills and to improve the ability of children with autism to recognize emotions correctly [19], a mobile-health based assistive technology for the healthcare of dementia and Alzheimer's disease [20], an educational robotics used in activities with students who have diverse special needs in the context of inclusive education [21], a study about E-learning in higher inclusive education [22], a remote teaching approach for equal and inclusive education in rural areas [23], a social virtual reality robot for music education to children with autism [24], a mobile application for blind students [25], a system of transformable tactile tools on VR for blind children [26], a serious game for job training of persons with developmental disabilities [27], a robot-assisted occupational therapy for children with down syndrome [28], an assistance application of people with cognitive disabilities in tasks for their social inclusion [29], and an assistive technology for students with reading and writing disabilities [30], among others.

It is evident that AT has been widely reported in the literature, and a significant percentage of them are focused on improving the education of people with some type of disability. Consequently, AT constitutes a huge opportunity to contribute new technologies to the already existing ones. From our research, we have not found studies applying the joint interaction of VR, NLP and ECA as AT, even less using the Spanish language to interact, this being the purpose of our study.

3. System Technical Aspects

This work is part of a master thesis in computer science. It consisted of the implementation of a system that included Natural Language Processing with a Semantic Embodied Conversational Agent (SECA) within an Immersive CAVE-like environment [31]. Due to the unique qualities of each component, they are also significantly different.

This section describes the technical details of the proposed system. Accordingly, the following subsections detail both components.

3.1. Immersive Virtual Reality Environment

The immersive virtual reality environment is composed of a CAVE-like Environment and a Semantic Embodied Conversational Agent (SECA) supported by a platform called Multi-VRMedia that connects users through different virtual reality technologies.

Fig. 1. User while interacting with the SECA within the CAVE-like environment, which displays the scenario of a virtual library.

3.1.1. CAVE-like Environment

The CAVE-like environment should incorporate devices that facilitate interaction with the user through multiple means of communication. Therefore, the implementation of the module requires a physical installation that allows, among other aspects, the projection of images and an adequate interaction.

In this proposal, the physical installation is a room with three walls, each of which includes an overhead projector that will allow to visualize the virtual world in a coherent way, offering immersion in the virtual environment to be designed. The complexity of the modeled virtual worlds is a key factor in the performance of this kind of immersive systems, so the implementation of a CAVE-like environment must use high-capacity computers equipped with high-end graphics processors. The hardware architecture of the CAVE-like environment currently corresponds to a centralized approach (in the past it was distributed), in which only one computer performs the complete processing of the three walls of the environment. For this purpose, it divides the image assigning the corresponding image portion to each screen.

Based on the above, the main hardware devices required in the physical implementation of the immersive environment are:

Processing devices. A powerful computer which includes a seventh-generation processor and a video board with the capability of generating high-definition images. Because the image to be generated must be split for multi-viewing, the video board used has multiple video outputs.

Input Devices. One of the main tasks performed by the CAVE-like environment is the interaction with the user. To this end, the environment includes traditional input devices such as keyboard and mouse as well as allowing communication through more natural interaction devices such as wireless microphone.

Output Devices. In addition to the input, the interaction with the user involves a constant exchange of information for which the use of output devices such as audio or display devices is necessary. The display devices include the use of 3 screens that allow rear projection, 3 high resolution projectors with ultra-short distance projection capabilities, and a support structure that will allow rigidifying and calibrating both projectors and screens.

3.1.2. A computing Platform for Immersive Visualization

The CAVE-like subsystem was developed to work on a computing platform for immersive collaborative 3D virtual world visualization (See Fig 1). It allows for the use of geographically distributed VR media, called a multi-VRmedia. Remote players can navigate and interact through a 3D scenery in a multi-VRmedia. During navigation, players can exchange information in order to cooperatively solve the observed problems as in a multiplayer game way [13].

A computing platform consists of a hardware architecture and a software framework, where the combination allows application software to run. Typical platforms include a computer's architecture, an operating system, programming languages and related user interface (run-time system libraries or graphical user interface). Fig. 2 shows an overview of the platform. A system to visualize scenarios in a multi-virtual reality media environment has been defined. Such a system will provide the necessary structure for attributes definition, rendering and collaborative multi-visualizations, as well as the necessary interactive resources.

The distributed multiple-display virtual reality hardware components used in this work are a Desktop Driving Set, Head Mounted Displays (HMDs), Data Gloves, Motion Trackers and a CAVE-like virtual environment.

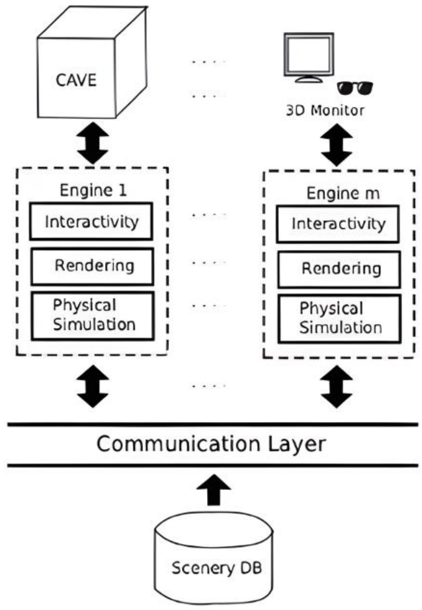

Fig. 2. Computing Platform Multi-VRMedia.

3.1.3. Semantic Embodied Conversational Agents (SECA)

In recent years major efforts have been made to promote the use of virtual assistants by several applications; ECA are computer-generated images with life-like facial features and body movements, capable of performing a collaborating role as instructors. Semantic Embodied Conversational Agents improve the quality of interaction through visual information. In particular, SECA further enhance the agent concept by using semantically structured and labeled information as a knowledge base, allowing them to reason and react to changes in the environment [16, 32]. As the real-time interaction is important for this work, the conversational aspect not the visual aspect was only considered. The "Reasoning" subsection of NLP details how this source of semantic information is constituted and Fig. 1 shows the SECA in the CAVE-like environment.

3.2. Natural Language Processing

Natural language processing is the set of methods for making human language accessible to computers. A language can be defined as a set of rules or set of symbols where symbols are combined and used for conveying information or broadcasting the information. Since all the users may not be well-versed in machine specific language, Natural Language Processing (NLP) caters to those users who do not have enough time to learn new languages or get perfection in it. It came into existence to ease the user's work and to satisfy the wish to communicate with the computer in natural language, and can be classified into two parts i.e. Natural Language Understanding (NLU) (with the due reasoning) and Natural Language Generation (GNL) which evolves the task to understand and generate the corresponding text [32].

In this work, an NLP subsystem was developed with the SECA. As the real-time interaction is important for this work, the conversational aspect was only considered. This system's modality involves simulating inherent human skills like listening and talking. The interaction between the SECA and the user is performed in real-time through the user's voice (by speech to text conversion) and a synthetic voice (by text to speech translation).

As a result, in order to establish the aspects of the NLP subsystem of SECA, the main involved components are the following.

3.2.1. Natural Language Understanding (listening)

While interacting, a human user talks to the system, which transforms the user's speech to a textual representation by using an Automatic Speech Recognition (ASR) module. The module takes as input the user's speech utterance that comes from a microphone and gives a resultant text from parsing the word string produced by speech recognition and forms an internal semantic representation based on keywords contained into a grammar. Each word said by the user is analyzed for the recognizer and compared through a phonetic transcriber. All the words picked up by the microphone are contrasted with words defined by rules in a grammar. According to the rules of grammar, possible sentences are analyzed and compared with the received sentence, then every word gets a confidence value. The ASR module takes this confidence value to determine if a word is accepted or rejected. If the word is accepted, then it is stored as a valid keyword. If the word is rejected, then the system must inform the user about the failure. From previous analysis with users, it is known that the rejection rate is approximately 50%. When all the words are accepted, it allows for a query as a text string consisting of keywords. Because the process performed by the ASR is completely computational and is performed in parallel, it is not possible for the user to know the time required to execute this operation and it is usually represented in nanoseconds.

3.2.2. Reasoning (thinking)

This module generates a response based on the input, the current state of the conversation and the dialog history. For this, it extracts the meaning of the utterance from keywords, manages the dialog flow and produces the appropriate actions for the target domain in on-line Dbpedia [33]. This approach gives complete control over system knowledge and expressions to the scriptwriter who creates the responses. When an interactor comes up to the SECA and asks it a question, the system driving the NLP analyzes the interactor's question and selects the appropriate response from the Dbpedia collection. If this process fails, then the system notifies the user about the fault. From previous analysis, it is known that the fault rate is approximately 10%.

3.2.3. Natural Language Generation (talking)

It is a Text-To-Speech (TTS) module carrying out the generation of the synthetic output voice from the text that comes as a response from the Reasoning Module.

This module is performed on three main situations:

If one word is rejected, then the SECA says "Disculpe, no he oído bien" which takes about 3 seconds, if the Dbpedia query fails, then the SECA says "Lo siento. En este momento no se pueden obtener los artículos solicitados" which takes about 7 seconds; and if the SECA found an answer from Dbpedia, the delay could be between 40 and 300 seconds according to its length.

4. Previous Experiences

The system presented here, as well as previous works, has been evaluated with a variety of users. In this section, we present two user experiences with different user groups.

Teachers, students, and librarians. Usability, user experience, and emotional response were evaluated using the CAVE-VOX system, with a library assistant as the application. All the quality variables analyzed yielded positive results, particularly those related to emotional aspects, as they showed high levels of positive stimulation in users [31].

Patients with Alzheimer's disease or dementia. The experiment consisted of testing the system, followed by questionnaires for patients and therapy professionals. The primary objective of the experience was to allow Alzheimer's patients to explore their memories. This exploration was considered to be a support for the intervention and as a complement to traditional treatments. Considering the point of view of a person with Alzheimer's disease or dementia, both the degree of satisfaction and the degree of emotional response analyzed were positive [34].

5. Potential Applications

Students with disabilities face academic, psychological and social challenges within the education environment at every level. The concept of inclusive education has brought with itself the much needed share of equality in approach for the education of the “disabled” by giving them a leveled field to rightly exhibit their differential abilities, proving themselves capable enough to learn and perform together, at par with their non-disabled peers. And with this shift in approach, there also emerges the need and challenge to tailor the teaching strategies or the means of instructional delivery in the inclusive classrooms, to address the diverse learning needs of all learners in an equitable manner.

Not all disabilities are visible, so if assistive technology is an enabler for learning, then the processes by which children with disabilities are identified as users of assistive technology must take place as early as possible in the lifecycle.

Sustainable Development Goal 4 is focused on ensuring inclusive and equitable quality education and promoting lifelong learning opportunities for all. Inclusive education is a key component of SDG 4 and aims to ensure that all learners, regardless of their backgrounds, abilities, or disabilities, have access to quality education that meets their diverse needs and enables them to reach their full potential. Inclusive education is essential for achieving SDG 4 because it ensures that all learners have equal access to education and are able to develop the skills and knowledge they need to succeed in life. By promoting inclusive education, we can help to break down barriers that prevent individuals with disabilities from accessing education and enable them to become active participants in their communities and society as a whole.

Virtual Reality and Natural Language Processing technologies have the potential to help people with a wide range of disabilities, including:

Physical disabilities. VR can be used to create immersive environments that allow individuals with physical disabilities to experience activities and environments that may be difficult or impossible for them to access in the physical world. For example, individuals with limited mobility can use VR to explore virtual worlds or participate in virtual sports and games.

Visual impairments. NLP can be used to develop text-to-speech and speech-to-text technologies that enable individuals with visual impairments to interact with digital devices and access digital content more easily. VR can also be used to create immersive audio environments that provide spatial cues and help individuals with visual impairments navigate and interact with digital environments.

Hearing impairments. NLP can be used to develop speech recognition technologies that enable individuals with hearing impairments to communicate more easily. VR can also be used to create immersive visual environments that provide spatial cues and help individuals with hearing impairments navigate and interact with digital environments.

Cognitive impairments. VR can be used to create interactive training and rehabilitation programs that help individuals with cognitive impairments develop cognitive and social skills. NLP can also be used to develop conversational agents that provide personalized support and guidance to individuals with cognitive impairments.

6. Case study

The case study is a system application case that attempts to motivate children with autism to learn the letters of the alphabet. In order to achieve greater motivation, the *Lore* alphabet and numbers [35] were used as visual support. The test approach was individualized using as an example the name of each child who participated; in those cases of children who knew their name, the first and last name was used for the test.

Lore alphabets can change according to the topic, but the main idea is that letters have facial features, such as eyes, teeth, among other characteristics. Fig. 3 shows an example of a *Lore* alphabet.

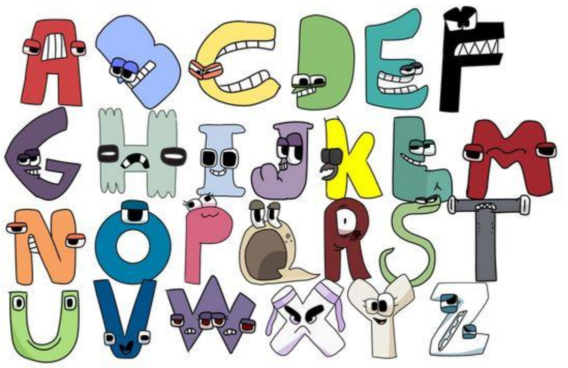

Fig. 3. *Lore* alphabet.

6.1. Children with autism

Children with Autism Spectrum Disorder (ASD) have unique characteristics and needs that require specialized approaches and support. Some key considerations when working with children with autism are: individualized approach, predictable environment, visual supports, use clear and concise language when communicating, recognize and address sensory sensitivities, among others.

According to the literature, when working with children with autism it is important to take into account cultural factors that may influence the child's experiences, perceptions, and preferences when designing interventions and supports.

6.2. Participants

The sample analyzed here is composed of 10 children between 3 and 7 years of age with varying degrees of autism. Due to the characteristics of the disorder, in order to reduce sensory sensitivity, it was necessary to make a previous preparation that consisted of asking the parents to show the children photographs of the laboratory, of the researcher in charge of the tests and of the SECA. It should be clarified that in order to gather the sample, a parent of a child with autism who had a network of contacts with other parents in the same situation was contacted.

6.3. Methodology

The application case consisted of showing the participants all the letters of the Lore alphabet and highlighting their name while the SECA spelled it letter by letter. In order to simplify the system for the children, the microphone was not used during testing. However, in order to provide a mechanism for the child to communicate with the system, a modified keyboard with Lore letters was provided. In this way the child was able to press a letter for the SECA to pronounce while enlarging the screen.

Fig. 4. A child with autism interacts with the SECA within the CAVE-like environment during a test of the system. The CAVE-like displays the child's name on the front screen and the Lore alphabet on all its screens.

The test consisted of performing 3 repetitions of approximately 2 minutes per participant, the repetition method was chosen to decrease the times of each test, this due to the premature age of the participants. Fig. 4 shows a participant during one of the tests.

6.4. Evaluation

The evaluation consisted of two stages: an on-site test that measured the child's emotional response to the use of the system, and a post-test conducted by the parents at home, which attempted to measure the degree of learning.

6.4.1. *Lore* Emotional Scale

The evaluation in situ was based on the Self-Assessment Manikin (SAM) technique [36], an emotional scale, on which changes were made to the pictograms to correspond to Lore numbers. Fig. 5 shows the one implemented for this case study.

Fig. 5. *Lore* Scale.

The main features of the implemented scale include the alteration of sizes in the numbers as the scale increases and the incorporation of facial expressions ranging from sad to happy.

The results showed that the minimum value chosen was 3 and the maximum value was 5. Table 1 shows the average obtained from the five-point scale and its standard deviation. When observing the average obtained, the results exceed half of the scale, so it is possible to affirm that in general the emotional response is positive. However, the minimum value obtained corresponds to half of the scale, so there are aspects to improve in order to obtain better results. Among them, we could consider familiarizing the child with the system, especially with the SECA; the level of interactivity (minimum) of the participants and the appeal of the chosen topic (not attractive) to all the children who participated.

Table 1 Emotional Scale.

Average	*Standard Deviation*
3.73	0.70

6.4.2. *Learning* Assessment with *Lore* Alphabet

Subsequently, the parents of the participating children were asked to conduct an assessment at home. This test consisted of the children trying to spell the letters of their name, while the parents counted the correct answers.

For this purpose, each parent was given a sheet of paper with the child's name written in Lore letters. Fig. 6 shows an example of a sheet used for this test, which has a box at the bottom to mark the hits.

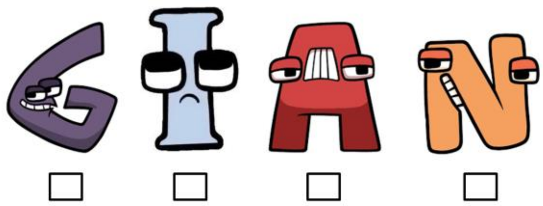

Fig. 6. Sheet with *Lore* Name.

The results for the number of hits showed a minimum percentage of 51% and a maximum of 97%. It showed that no child got all the letters right and that some of them only got half of them right. Table 2 shows the average of the percentages of correct answers and their standard deviation, considering that the scale is now a percentage scale. Taking into account the average, it is observed that the value corresponds to 74.6%, indicating that the general results show an average learning to pass. Additionally, the maximum of 97% indicates for this case a very favorable percentage of success, this value was obtained in a 7 year old child who is already in school, therefore it can be assumed that the child had some level of knowledge of the alphabet. On the other hand, the minimum value of 51% was obtained from a 4 year old child who is in kindergarten, so he has no previous formal knowledge of the alphabet. This last case allows us to obtain interesting data on the degree of real learning (without prior knowledge) provided by the system, indicating that it is close to 50%, i.e. in the worst case, the child learns half of it. This result may be a consequence of the maximum time of 6 minutes per single encounter (3 repetitions of approximately 2 minutes), which should be reinforced with more encounter time or more encounters in order to improve this result.

Table 2 Name Test.

Average	*Standard Deviation*
74.6%	15.00

7. Conclusions

This paper describes the development of an immersive virtual reality system with conversational skills. The AT was built into a CAVE-like environment allowing a high degree of realism and immersion. Additionally the use of natural interfaces as natural language processing increases the presence of the users. The aim is to create an assistive technology that can be used to teach and train learners with disabilities in a more inclusive way.

This work highlights that AT can promote educational, psychological and social benefits for Students with Disabilities (SWD). Assistive technology use can enable academic engagement and social participation and be transformative from a psycho-logical perspective.

The effectiveness of a system as an assistive technology for inclusive learning processes will depend on the specific application case being considered. However, some general conclusions can be drawn.

Our system can help to:

- support inclusive learning processes by providing personalized support and accommodations to learners with diverse needs and abilities.
- promote inclusive learning depending on how well they are designed, implemented, and integrated into the overall learning environment. It's important to ensure that assistive technologies are accessible, user-friendly, and well-supported by teachers and other education professionals.
- promote learner independence and self-determination by providing learners with more control over their learning experiences and enabling them to work at their own pace and in their own preferred ways.
- encourage collaboration and socialization among learners with diverse needs and abilities by providing opportunities for communication,

collaboration, and peer support.

• facilitate equity and inclusion in education by providing learners with disabilities with the tools and resources they need to participate fully in the learning process and achieve their full potential.

In order to evaluate the possibilities of the system, a case study on children with autism of early school age was analyzed. A previous preparation of each child was very important in order to reduce sensory sensitivity during the in situ evaluation.

The results for the emotional test showed a degree of liking higher than the midpoint of the scale, indicating that the participants' liking falls between neutral and liking. When measuring the learning achieved in the evaluation, it was noted that the minimum is 51% of correct answers, i.e. those who had fewer correct answers managed to get half of the letters right, and the average percentage of correct answers is 74%. It is notable, in general, that the evaluation yields positive results, showing that the system is capable of providing a framework for learning experiences. However, in this case study, it was possible to note that the experience needs to be improved in order to allow for learning over time, which does not only correspond to a single encounter.

As future work, new experiences will continue to be tested with children with autism, and it is hoped to conduct customized case studies for other disabilities. The potential of this technology framework for other areas, such as healthcare, rehabilitation and entertainment, should also be explored.

Competing interests

The authors have declared that no competing interests exist.

Authors' contribution

YA implemented the system, conducted the experiments, analyzed the results and wrote the manuscript; RG and FS conceived the idea and revised the manuscript. All authors read and approved the final manuscript.

References

[1] L. Maftei and C. Harty, "Surprise: challenging design perceptions in immersive virtual reality environments? The case of designing a hospital project using a CAVE (Cave Automatic Virtual Environment)," *Archnet-IJAR: International Journal of Architectural Research*, vol. 15, no. 3, pp. 887-904, 2021.

[2] X. Chen, Z. Chen, Y. Li, T. He, J. Hou, S. Liu, and Y. He, "ImmerTai: Immersive motion learning in VR environments," *Journal of Visual Communication and Image Representation*, vol. 58, pp. 416-427, 2019.

[3] L. Wu, B. Sekelsky, M. Peterson, T. Gampp, C. Delgado, and K. B. Chen, "Immersive virtual environment for scale cognition and learning: Expert-based evaluation for balancing usability versus cognitive theories," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 66, no. 1, pp. 1972–1976, 2022.

[4] T. T. de Back, A. M. Tinga, and M. M. Louwerse, "CAVE-based immersive learning in undergraduate courses: examining the effect of group size and time of application," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, p. 56, Oct. 2021.

[5] G. C. Guerino and N. M. C. Valentim, "Usability and user experience evaluation of natural user interfaces: a systematic mapping study," *IET Software*, vol. 14, no. 5, pp. 451–467, 2020.

[6] J. Cassell, "Embodied Conversational Agents: Representation and Intelligence in User Interfaces," *AI Magazine*, vol. 22, pp. 67–84, 2001.

[7] J. Eisenstein, Introduction to natural language processing. MIT press, 2019.

[8] S. Ruder, M. E. Peters, S. Swayamdipta, and T. Wolf, "Transfer Learning in Natural Language Processing," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials*, Minneapolis, Minnesota: Association for Computational Linguistics, pp. 15–18, 2019.

[9] T. Sun et al., "Mitigating Gender Bias in Natural Language Processing: Literature Review," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy: Association for Computational Linguistics, pp. 1630–1640, 2019.

[10] M. Bahja, 'Natural Language Processing Applications in Business', E-Business - Higher Education and Intelligence Applications. IntechOpen, May 19, 2021.

[11] "Laboratorio de Computación Gráfica." Available at: `http://www.lcg.unsl.edu.ar`. Accessed on 2023/07/15.

[12] "Grupo de Informática Gráfica Avanzada." Available at: `http://giga.cps.unizar.es`. Accessed on 2023/07/15.

[13] Y. Alvarado, N. Moyano, D. Quiroga, J. Fernández, and R. Guerrero, "Augmented Virtual Realities for Social Developments," in Experiences between Europe and Latin America, chapter A Virtual Reality Computing Platform for Real Time 3D Visualization, Universidad de Belgrano, pp. 214–231, 2014.

[14] N. Jofré Pasinetti, G. Rodríguez, Y. Alvarado, J. Fernandez, and R. A. Guerrero, "Towards defining a CAVE like system performance evaluation," in

Proceedings of the XXVI Congreso Argentino de Ciencias de la Computación (CACIC), pp. 218-228, 2020.

[15] S. Baldassarri, E. Cerezo, and F. J. Seron, "Maxine: A platform for embodied animated agents," *Computers & Graphics*, vol. 32, no. 4, pp. 430–437, 2008.

[16] F. J. Serón and C. Bobed, "VOX system: a semantic embodied conversational agent exploiting linked data," *Multimedia Tools and Applications*, vol. 75, no. 1, pp. 381–404, 2016.

[17] D. Chambers, "Assistive Technology Supporting Inclusive Education: Existing and Emerging Trends," in Assistive Technology to Support Inclusive Education, D. Chambers, Ed., in International Perspectives on Inclusive Education, vol. 14. Emerald Publishing Limited, pp. 1–16, 2020.

[18] P. Encarnação et al., "Using assistive robots to promote inclusive education.," *Disability and Rehabilitation: Assistive Technology*, vol. 12, no. 4, pp. 352–372, 2017.

[19] W.-T. Tsai, I.-J. Lee, and C.-H. Chen, "Inclusion of third-person perspective in CAVE-like immersive 3D virtual reality role-playing games for social reciprocity training of children with an autism spectrum disorder," *Universal Access in the Information Society*, vol. 20, pp. 375–389, 2021.

[20] K. Yousaf et al., "A comprehensive study of mobile-health based assistive technology for the healthcare of dementia and Alzheimer's disease (AD)," *Health Care Management Science*, vol. 23, no. 2, pp. 287–309, 2020.

[21] L. Daniela and M. D. Lytras, "Educational Robotics for Inclusive Education," *Technology, Knowledge and Learning*, vol. 24, no. 2, pp. 219–225, 2019.

[22] B. Meskhi, S. Ponomareva, and E. Ugnich, "E-learning in higher inclusive education: needs, opportunities and limitations," *International Journal of Educational Management*, vol. 33, no. 3, pp. 424–437, 2019.

[23] S. Stenman and F. Pettersson, "Remote teaching for equal and inclusive education in rural areas? An analysis of teachers' perspectives on remote teaching," *The International Journal of Information and Learning Technology*, vol. 37, no. 3, pp. 87–98, 2020.

[24] M. Shahab et al., "Utilizing social virtual reality robot (V2R) for music education to children with high-functioning autism," *Education and Information Technologies*, vol. 27, no. 1, pp. 819–843, 2022.

[25] A. Boza-Chua, K. Gabriel-Gonzales, and L. Andrade-Arenas, "Inclusive Education: Implementation of a Mobile Application for Blind Students," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, 2021.

[26] S. Yuko, W. Hiroki, and A. Naoki, "System Development of Transformable Tactile Tools on VR for Blind Children," in *Proceedings of the Fifteenth International Conference on Management Science and Engineering Management*, Eds., Cham: Springer International Publishing, pp. 368–378, 2021.

[27] J. Kwon and Y. Lee, "Serious games for the job training of persons with developmental disabilities," *Computers & Education*, vol. 95, pp. 328–339, 2016.

[28] V. Krasniqi, N. Ackovska, and K. Zdravkova, "Emerging Role of Robot-Assisted Occupational Therapy for Children with Down Syndrome," in *Proceedings of the UBT International Conference*, 2017.

[29] L. Cuascota, L. Guevara, R. Cueva, F. Tapia, and G. Guerrero, "Assistance application of people with cognitive disabilities in tasks for their social inclusion," in Proceedings of the 2019 14th Iberian Conference on Information Systems and Technologies (CISTI), IEEE, pp. 1–7, 2019.

[30] I. Svensson et al., "Effects of assistive technology for students with reading and writing disabilities," *Disability and Rehabilitation: Assistive Technology*, vol. 16, no. 2, pp. 196–208, 2021.

[31] Y. Alvarado, Agentes Semánticos Conversacionales y Entornos CAVE. Computer Science Master Thesis, Universidad Nacional de San Luis, 2022.

[32] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges," Multimedia Tools and Applications, vol. 82, no. 3, pp. 3713–3744, 2023.

[33] "DBpedia Association." Available at: `http://www.dbpedia.org`. Accessed on 2023/07/15.

[34] F. Seron et al., "A semantic memory bank assisted by an embodied conversational agent for mobile devices," *Engineering and Applied Sciences*, vol. 6, no 1, pp. 1-17, 2021.

[35] "Alphabet Lore" Available at: `http://www.imdb.com/title/tt21923830/`. Accessed on 2023/07/15.

[36] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no 1, pp. 49-59, 1994.

Citation: Y. Alvarado, R. Guerrero and F. Serón. *Inclusive Learning through Immersive Virtual Reality and Semantic Embodied Conversational Agent: A case study in children with autism.* Journal of Computer Science & Technology, vol. 23, no. 2, pp. 107-116, 2023.
DOI: 10.24215/16666038.23.e09
Received: April 17, 2023 **Accepted:** August 15, 2023.