

A Virtual Reality Framework for Immersive Dyslexia Intervention

Anushka Kamble

Sakshi Shilimkar

Sy Msc (CS)

Sy Msc (CS)

E-mail:kambleanushka7778@gmail.com

E-mail:shilimkarsakshi59@gmail.com

Abstract—Visual Education Resources (VISER) represents a Virtual Reality (VR) ecosystem engineered to assist learners with dyslexia by providing a highly immersive, participatory, and frictionless educational environment. This research highlights the technical deployment of the architecture, focusing on the synergy between Natural Language Processing (NLP) and WebVR to facilitate Automatic Text Visualization (ATV). Furthermore, the study explores the integration of customizable 3D humanoid entities within the virtual space.

Keywords— *VISER; Dyslexia; VR; WebVR; NLP; ATV*

I.INTRODUCTION

Evolution through Virtual Reality for Dyslexic Learners (VISER).Dyslexia is characterized as a neurodevelopmental learning difference that primarily hinders an individual's ability to process phonological components of language, resulting in difficulties with reading fluency, orthography, and verbal comprehension. Crucially, these challenges are independent of cognitive intelligence; however, a lack of timely intervention often precipitates a decline in academic confidence and overall school performance. In the context of the Indian education system, data from the Dyslexia Association of India and NIMHANS indicate a prevalence rate of 5% to 15% among school-aged children. With India's vast youth demographic, this equates to roughly 35 to 40 million students requiring specialized support—a figure that presents a logistical crisis in classrooms often burdened by high student-to-teacher ratios.

Conventional pedagogical interventions—ranging from intensive one-on-one phonics tutoring to tactile multisensory tools—face significant barriers to scalability. These methods are frequently cost-prohibitive and struggle to captivate the attention of modern, digitally-native students. To bridge this gap, we introduce VISER (Visual Education Resources). Built upon a robust Model-Template-View (MTV) architecture, VISER is a WebVR platform designed to be a scalable, low-cost solution for Indian schools. By leveraging a centralized backend and an expansive 3D asset library, it allows for a high degree of teacher-student interaction within a simulated space.

The VISER ecosystem is composed of four distinct technical modules:

- 1] The Foundational Course Module: This component focuses on early cognitive milestones such as shape recognition, colour theory, and numeracy. By utilizing 360-degree visuals, it transforms abstract numbers into concrete objects; for instance, selecting the digit "2" populates the virtual field with two discrete apples, reinforcing the concept of quantity through spatial memory.
- 2] The Immersive Quiz Module: To reduce "test anxiety," this module replaces traditional paper-based exams with interactive assessments. Students interact with 3D objects to demonstrate knowledge, receiving immediate feedback and automated performance analytics.

3] Dynamic Object Generation (DOG): Powered by an NLP engine, this module parses complex sentences and automatically constructs corresponding 3D environments. This helps students visualize spatial relationships and "see" the narrative of a sentence, which is often a major struggle for those with dyslexia.

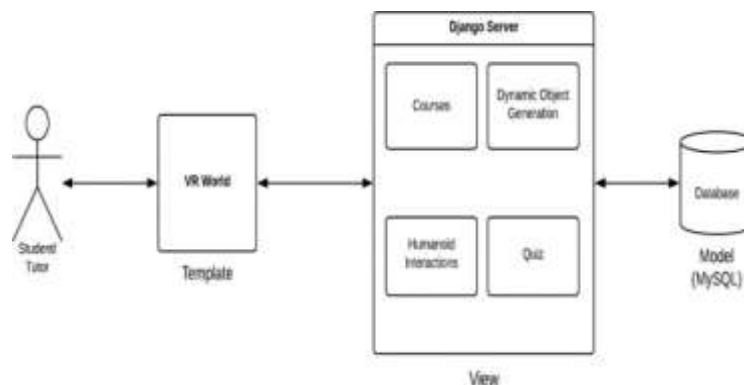
4] Humanoid Interaction & Animation: This module introduces programmable avatars that respond to natural language commands. By visualizing actions like "dancing" or "running," it aids in the acquisition of verbs and enhances the student's linguistic-to-visual mapping.

Ultimately, VISER aims to foster an inclusive digital classroom that aligns with the goals of the Rehabilitation Council of India, ensuring that high-quality, specialized education is available to students in both urban centers and rural districts alike.

II. PORPOSED SYSTEM

In the following sections, we have described the overall system architecture along with algorithms used for text visualization in virtual world using natural language processing. Apart from this, we have also shown the implementation of system along with examples and results.

A. SYSTEM ARCHITECTURE



The VISER framework is built upon a decentralized, open-access architecture following the **Model-Template-View (MTV)** paradigm. This structural choice ensures a clear separation of concerns, allowing for a seamless experience whether the user is engaging as an educator or a learner.

- **The Template Layer:** Acts as the immersive interface, encompassing the entire 3D Virtual Reality environment where user interaction occurs.
- **The View Layer:** Functions as the computational engine of the platform. It manages the logic, handles data flow between components, and executes backend processing tasks.
- **The Model Layer:** Serves as the centralized repository, housing the comprehensive database of 3D assets, user profiles, and educational materials required for the system's operation.

Modules in VISER

The VISER system has four main modules designed to help dyslexic students learn better by using visual and hands-on methods.

1. Foundational Course Module

The Course Module is engineered to promote early-stage cognitive growth through specialized instructional programs. It addresses fundamental academic building blocks such as geometric shapes, chromatic recognition, numerical quantities, and spatial orientation. For instance, within the **Quantity Sub-module**, numeric symbols are rendered as tangible 3D entities. If a learner interacts with the number "2," the system dynamically generates two discrete 3D apples within the 360-degree environment. This spatialized pedagogy leverages visual memory and reduces the cognitive load associated with traditional text-heavy instruction, making it particularly effective for India's diverse, multilingual student population.

2. Immersive Assessment (Quiz) Module

This module transitions away from high-stakes testing toward an exploratory evaluation model. The Quiz Module utilizes immersive 3D scenarios to gauge student comprehension through direct interaction with virtual objects. Key features include instantaneous performance feedback and the generation of automated progress analytics. By moving beyond the conventional paper-and-pencil examinations prevalent in Indian schools, this VR-assisted method significantly lowers learner anxiety while fostering active participation and experiential knowledge retention.

3. Dynamic Object Generation (DOG) Module

The DOG Module serves as the bridge between linguistic input and visual representation by utilizing advanced **Natural Language Processing (NLP)**. Its primary goal is to help dyslexic learners deconstruct and visualize complex syntax. The engine parses raw text to identify specific objects, their unique attributes (size/color), and their relative spatial coordinates to construct a live 3D scene.

Example: A sentence such as *"A miniature brown sphere is situated on the right, while a massive elephant stands in the distance on the left,"* is instantly translated into a visual landscape. This helps the student map prepositional phrases and adjectives to physical space, strengthening sentence comprehension and imaginative faculties.

4. Humanoid Interaction & Animation Module

This component features programmable 3D avatars capable of performing a wide range of kinesthetic actions, such as running, jumping, or dancing, in response to natural language prompts. For example, if a student inputs "Maya is dancing," the system identifies the subject and the specific verb, triggering the corresponding animation in real-time. This interactive capability allows students to customize their virtual guides, assisting in the acquisition of action-based vocabulary and enhancing linguistic fluency through visual reinforcement.

III. Importance of VR for Indian Dyslexic Students

The Indian educational infrastructure faces unique systemic hurdles, including overcrowded classrooms, a deficit in specialized pedagogical training for educators, and a sharp divide in resource accessibility between metropolitan and rural sectors. Furthermore, the multilingual nature of the student body adds a layer of complexity for those struggling with phonological processing.

In this landscape, VR-centric solutions like **VISER** offer a standardized, highly engaging alternative to conventional methods. By deploying digital classrooms that require less reliance on expensive private interventions, VISER aligns with the inclusive education mandates set forth by the **Rehabilitation Council of India (RCI)**. It democratizes access to specialized tools, ensuring that disability rights and quality education are upheld across diverse socio-economic backgrounds.

A. Performance Evaluation Framework

To rigorously assess the efficacy of the VISER platform, we utilize a multi-dimensional evaluation strategy. The system's success is quantified through:

- **NLP Precision:** Accuracy in parsing and visualizing user queries.
- **Comprehension Gains:** Measurable improvements in reading and interpretative skills.
- **Comparative Analytics:** A "Pre-vs-Post" assessment of student performance metrics.
- **Immersive Analytics:** Tracking user dwell-time and active engagement levels within the 360-degree environment.
- **Knowledge Retention:** Longitudinal testing to determine the durability of learned concepts.

B. Expected Results

The implementation of VISER is projected to catalyze significant improvements in the academic lives of dyslexic learners. Primary outcomes include enhanced **spatial-visual reasoning**, a marked reduction in **academic anxiety**, and heightened **information recall speeds**. By transforming abstract text into a tangible experience, the system fosters a deeper, more intuitive connection with the curriculum.

C. ALGORITHM

This section delineates the procedural logic governing the **Dynamic Object Generation (DOG)** and **Humanoid Animation** modules.

1. The DOG Logic (Text-to-Scene Synthesis)

The DOG algorithm follows a serialized pipeline to convert natural language into a structured 3D layout:

1. **Linguistic Pre-processing:** The system first sanitizes the input query by filtering out non-essential "stop words" and tokenizing the string into discrete lexical units.
2. **Syntactic Tagging:** A **Part-of-Speech (POS)** tagger is employed to categorize each token.
3. **Object Identification:** The algorithm specifically isolates nouns, which serve as the primary entities to be rendered in the virtual space.

2. Data Representation via Tuples

Once the linguistic analysis is complete, the data is structured into a mathematical tuple for the rendering engine.

- **Single-Object Format:** For isolated entities, the data follows the structure:

$$$$$ = \{ \{ \text{Noun, size, color, X, Y} \} \} $$$$$

- **Multi-Object Mapping:** For complex scenes involving various entities, a dictionary-based mapping is used:

$$M = \{ \text{'Noun_1': [size, color, X, Y], 'Noun_2': [size, color, X, Y]} \}$$

Parameter Definitions:

- **Noun:** The identifier for the 3D model asset.
- **Size:** A five-point scale (XS to XL).
- **Colour:** The specific texture/material attribute (e.g., Red, Hex code).
- **X-Axis:** Spatial orientation (Left/Right/Verticality).
- **Y-Axis/Distance:** Depth perception (Proximity/Distance).

3. Scene Rendering Phase

Following tuple generation, the system initiates a database query to retrieve the corresponding **.obj (geometry)** and **.mtl (material)** files. These assets, along with their calculated spatial coordinates, are passed to the **A-Frame template** for real-time 3D synthesis.

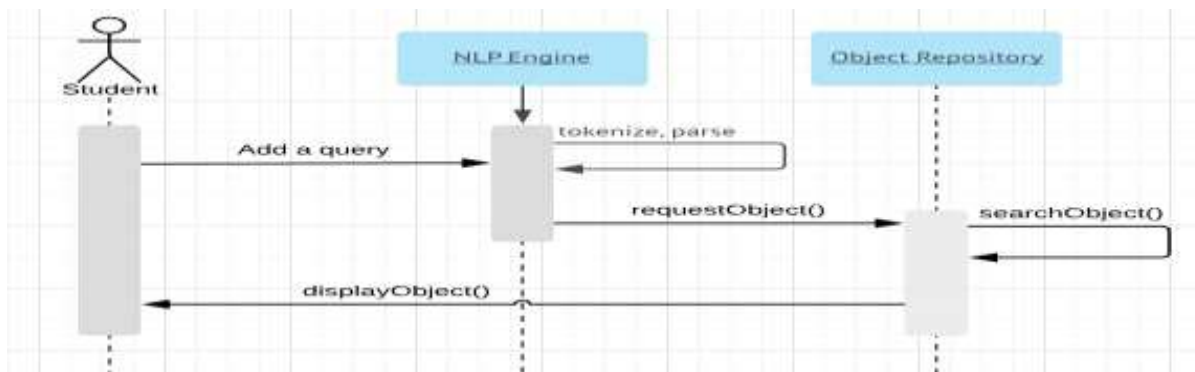


Figure 2: Process flow for DOG

Ex1. Show me a green apple on the left side of the brown table. Here, green is connected with the apple, brown is connected with the table, and left is connected with the apple. So, the tuple will be like this: {'apple': [None, green, left, None], 'table': [None, brown, None, None]}

Ex2. Show me an apple and a table. If the spatial distance and direction are not mentioned in the query, the objects will be shown from left to right as mentioned. Here, the apple will be on the left side of the table. The tuple formed will be: {'apple': [None, None, None, None], 'table': [None, None, None, None]}

Ex3. Show me a laptop which is up and far from me.

Here, the tuple formed will look like:
['laptop', None, None, up, far]

Ex4.

Show me an apple, car and table.
Here, the tuple formed will look like:
{'apple': [None, None, None, None], 'car': [None, None, None, None], 'table': [None, None, None, None]}

D.HUMANOID ACTIONS

This model operates through a systematic pipeline designed to interpret and visualize human movements within the 3D space. The process is governed by the following steps:

1. Linguistic Processing & Intent Extraction

Upon receiving a natural language query, the system initiates a multi-stage NLP refinement:

- **Normalization:** Non-informative "stop words" are filtered out to isolate the core intent.
- **Morphological Analysis (Stemming):** Lexical variations (e.g., "running") are reduced to their root forms (e.g., "run") to match the animation database.
- **Semantic Tagging: A Unigram Part-of-Speech (POS) Tagger** identifies specific entities (avatars) and their associated actions (verbs).

2. Conflict Resolution & Identity Mapping

The system prioritizes identity based on the specificity of the query:

- **Implicit Queries:** If no name is mentioned, the system defaults to the user's pre-selected avatar.
- **Explicit Queries:** If a specific name (e.g., "Mohan") is detected, the system overrides the default preference to render that specific character.
- **Multi-Agent Coordination:** For complex inputs involving multiple subjects (e.g., "*Akshay is kicking while Mohan is dancing*"), the system dynamically calculates the horizontal distribution of characters (Left, Center, Right) to avoid visual overlap.

3. Asset Execution

The backend architecture organizes character data in a hierarchical folder structure. Each avatar's directory contains specialized **.fbx** files. When a valid action-subject pair is identified, the system fetches the corresponding animation file (e.g., *mohan\exercise.fbx*) and applies the unique scale and coordinate parameters stored in the MySQL database.

The execution of humanoid animations follows a sophisticated logic to ensure visual accuracy and user-centric customization. The system interprets queries based on the following prioritized scenarios:

- **Scenario A: Implicit Subject Mapping** If the user requests an action without specifying a subject (e.g., "*Show me how to dance*"), the system automatically maps the action to the user's pre-selected "preferred avatar" (e.g., Akshay), resulting in the rendered output: "*Akshay is dancing.*"
- **Scenario B: Explicit Subject Override** When a query explicitly mentions a character name (e.g., "*Mohan is exercising*"), the NLP engine overrides the default user preference to render the specified character.
- **Scenario C: Multi-Agent Coordination** For complex syntax involving multiple actors (e.g., "*Akshay is kicking and Mohan is dancing. Conversely, I am punching*"), the system calculates spatial distribution. It

renders the primary subject on the left, the secondary on the center-stage, and the user’s own avatar (reverting to preferred settings) on the right to visualize the entire sequence simultaneously.

E. ARCHITECTURAL LAYERS

The VISER platform leverages the **Model-Template-View (MTV)** framework, primarily facilitated by the python web framework, A-Frame for VR rendering, and MySQL for relational data management.

1. **The Template (Front-End):** Developed using A-Frame, this layer provides the interactive WebVR interface. Its **Entity-Component-System (ECS)** architecture allows for efficient 3D scene building within a standard HTML environment.
2. **The View (Business Logic):** Powered by Python’s **NLTK (Natural Language Toolkit)**, this layer serves as the system’s brain. It processes linguistic inputs, extracts spatial relationships, and coordinates data flow between the user and the database.
3. **The Model (Database):** A MySQL database stores critical metadata for 3D entities, including scale factors, spatial coordinates, and rotation vectors. This structured approach ensures that asset retrieval is both fast and consistent.

IV. SYSTEM VALIDATION VIA SYNTACTIC MAPPING

Given the unique challenges of quantifying immersion within a Virtual Reality domain, the performance of the VISER platform is evaluated through the accuracy of its **NLP** Orchestration Engine. The primary metric for success is the high-fidelity translation of linguistic input into spatially accurate 3D environments.

To standardize this evaluation, we have established a set of Syntactic Validation Rules. These rules ensure that the system correctly interprets various sentence structures and renders the corresponding assets without logical errors. Following are the abbreviations used in formulating the syntax rules:

- NNP - Proper Noun
- VBG – Verb
- NN – Noun
- JJ – Adjective
- PRP – Pronoun

Sr. No.	Syntax Rule	Example
1.	NNP VBG	Mohan is dancing
2.	(NNP VBG) + (NNP VBG)	Mohan is dancing, and Maya is Running
3.	(NNP VBG) +	Mohan is dancing and running is

	(VBG PRP)	being performed by me
4.	(NNP VBG) + (VBG NNP)	Maya is running, and dancing is being performed by Mohan
5.	(NNP VBG)+ (NNP VBG)+ (PRP VBG)	Akshay is kicking, and Mohan is dancing. On the other hand, I am punching
6.	(PRP VBG) + (NNP VBG) + (NNP VBG) +(VBG NNP)	I am running, and Mohan is dancing. On the other hand, Akshay is punching and rolling is being performed by dadi

Table I: Syntax Rule for Humanoid Action module

The fidelity of the VISER system is verified through a structured set of syntax rules. These rules categorize various English sentence patterns based on the density of nouns, adjectives, and verbs. When processed by the NLP engine, these strings are mapped to specific 3D rendering protocols.

Table II: Syntax Rule for Dynamic Object Generation module

Sr. No.	Syntax Rule	Example
1.	NN JJ JJ	The apple is red in color, and it is big
2.	NN JJ	The apple is on the right/left Side
3.	NN	Show me a zebra/computer/spoon etc.
4.	JJ JJ NN JJ	Show me a big red apple which is on the left side
5.	JJ NN JJ	Show me a big apple which is close/closer to me
6.	(NN) + (NN)	Show me an apple and an elephant

V. RESULTS

The results of the various modules and their representation in the open world are shown here.



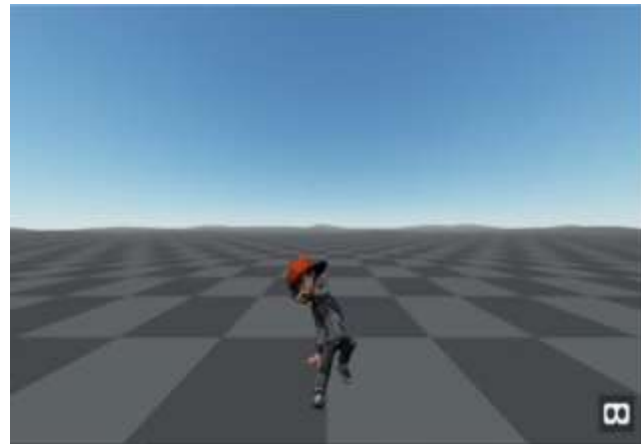
Shapes being taught in a 3D world
Quantity being taught in 3D world



Colours being taught in 3D world
Quiz being taken on the taught



Way of entering textual input in the VR world



Humanoid action-Mohan is dancing

VI. CONCLUSION AND FUTURE SCOPE

This research presents the **VISER framework** as a transformative approach to inclusive education for dyslexic learners in India. By leveraging the tuple-based mapping methodology, the system successfully bridges the gap between static textual information and dynamic 3D visualization. The study demonstrates that an immersive, WebVR-based environment provides a critical alternative to traditional educational tools, allowing students to engage with a curriculum that prioritizes spatial memory and interactive exploration over decoding. With dyslexia impacting an estimated **5% to 15%** of the Indian student population, the need for scalable, low-cost digital interventions is paramount. VISER addresses this by offering a platform that is not only technologically robust but also culturally and logistically aligned with the needs of the Indian education

system. By utilizing **Natural Language Processing (NLP)** to automate scene generation, the system minimizes the cognitive load on students, fostering an environment where learning is both intuitive and impactful.

Future Research Directions

The evolution of the VISER project will focus on the following key areas:

- **Multilingual Integration:** Expanding the NLP engine to interpret and visualize regional Indian languages, specifically **Hindi, Marathi, and Tamil**, to cater to diverse linguistic backgrounds.
- **Adaptive Learning Algorithms:** Implementing **Artificial Intelligence (AI)** to monitor student performance and dynamically adjust the complexity of 3D scenes and assessment modules.
- **Institutional Deployment:** Conducting longitudinal field studies within government-run schools to empirically validate the system's long-term effect on reading comprehension and retention.
- **Ecosystem Synergy:** Integrating VISER with national digital literacy initiatives to ensure broad-scale accessibility for neurodivergent learners across all socio-economic sectors.

VII. ACKNOWLEDGMENT

We extend our sincere and heartfelt thanks to our esteemed Mentor **Prof. Swati Joshi Ma'am** for providing us with the right guidance and advice at the crucial junctures and for showing us the right way. We also take this opportunity to express a deep sense of gratitude to our college **PVG's College of Science and Commerce**.

VIII. References

- 1]Dyslexia Association of India
- 2]National Institute of Mental Health and Neurosciences (NIMHANS)
- 3]Priyanka Jain, R. P. Bhavsar, B. V. Pawar and Hemant Darbari" VRML for automatic generation of 3D Scene" International Journal of Computer Application Volume.2 (2018).
- 4]Vinit Sathe, Piyush Gupta, Karan Kaushik, Suvarna Bhat and Sachin Deshpande, "Virtual Reality Website (VR WEB)", in International Conference on Electronics, Communication and Aerospace Technology (2017).
- 5]Dr. R.S. Kamath and Dr. R.K. Kamat, "Development of an Intelligent Virtual Environment for Augmenting Natural Language Processing in Virtual Reality Systems" in International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Vol.2 (2013).
- 6] Priyanka Jain, Hemant Darbari and Virendra kumar C. Bhavsar, "Text Visualization as an Aid to Language Learning Disability", in Proceedings of the National Conference on e-Learning and e-Learning Technologies, (ELETECH 2013), Hyderabad, India, pp. 88-93, Jun. 28-29, 2013
- 7] NIH Research (2022). "Use of Virtual Reality in Children with Dyslexia." *PMC*.
- 8] Innovative Use of Machine Learning-Aided VR and NLP in Dyslexia Treatment, in *Dyslexia Diagnosis and Treatment*, IntechOpen, 2024. [Online]. Available: www.intechopen.com/chapters/1190685.



9] Maresca, G., et al. (2024). "Effectiveness of the Use of Virtual Reality Rehabilitation in Children with Dyslexia: Follow-Up after One Year." *Brain Sciences*.