A multisensory cognitive-AI framework integrating Phonetic-Aware Image Association and personalized pronunciation feedback for Real-Time Dyslexia Assistance

Yamuna Vasanth
Department of computer science and engineering
School of Engineering
Dr.A.Rajesh,
Professor, Department of computer science and engineering
School of Engineering

Abstract: Dyslexia is a common learning difficulty that affects reading, spelling, and pronunciation skills, making it challenging for many learners to keep up with traditional teaching methods. To address this issue, this paper presents a multisensory learning system supported by artificial intelligence that helps learners improve their reading and speaking abilities in an interactive way. The proposed approach combines phonetic-based word breakdown, image support for better understanding, real-time pronunciation feedback, and an adaptive learning mechanism that adjusts to each learner's performance. The system was evaluated by comparing it with a non-adaptive baseline model. The results show clear improvements across all key areas. Reading accuracy increased from 72% to 91%, while pronunciation accuracy improved from 68% to 89%. The error rate was reduced from 32% to 11%, and the average response time decreased from 4.8 seconds to 2.6 seconds. In addition, learner engagement improved significantly, with scores increasing from 65 to 88. These improvements show that combining visual, auditory, and adaptive learning strategies can make the learning process more effective and engaging. Overall, the proposed system provides a practical and scalable solution for supporting learners with dyslexia. It offers real-time assistance, personalized learning, and improved outcomes, making it suitable for use in both classroom environments and assistive learning applications.
Keywords: Dyslexia, Adaptive learning, phonetic processing, pronunciation correction

1. INTRODUCTION

Dyslexia is a common learning disorder that affects a child's ability to read, spell, and pronounce words correctly. It is mainly related to difficulties in phonological processing, which makes it hard for learners to connect letters with their corresponding sounds. Because of this, many children with dyslexia struggle in traditional classroom environments and require specialized learning support. In recent years, Artificial Intelligence (AI) has been widely used to support dyslexia detection and learning. Studies such as Y. Alkhurayyif and A. R. Sait (2024) show that machine learning and deep learning techniques can effectively identify dyslexia patterns from behavioral and linguistic data. Similarly, an interdisciplinary review by K. Dabaghi et al. (2026) highlights the growing role of AI in building intelligent educational tools for dyslexic learners. However, these works mainly focus on detection rather than learning assistance.

Recent research also explores personalized and neuroscience-based approaches. For example, R. Niu et al. (2025) emphasize the importance of combining AI with cognitive science to create adaptive learning systems. In addition, machine learning-based models have been used to provide personalized strategies for learners, as demonstrated by A. Zingoni et al. (2024). While these approaches improve personalization, they still lack real-time interaction and feedback mechanisms. Multisensory learning has also been shown to be effective for dyslexic students. Studies such as R. Eryilmaz and E. Balci (2025) demonstrate that combining visual and auditory inputs can improve reading performance. Similarly, phonics-based interventions have been proven to enhance literacy skills over time, as reported by V. R. Oliveira et al. (2026). Furthermore, phonological processing has been identified as a key factor in dyslexia, as shown in the machine learning study by N. Ding et al. (2025). However, most of these systems are either non-adaptive or lack AI-driven feedback. Earlier works such as O. L. Usman and R. C. Muniyandi (2020) explored deep learning models for detecting dyslexia using neuroimaging data. Although effective, such approaches are computationally complex and not suitable for real-time educational use.

From the above studies, it is clear that existing approaches focus on either detection, multisensory learning, or personalization, but very few systems integrate all these aspects into a single framework. In particular, there is a lack of systems that combine phonetic learning, visual reinforcement, real-time pronunciation correction, and adaptive intelligence in a unified manner. To address this gap, this paper proposes a multisensory cognitive-AI framework that integrates phonetic-aware learning, image-based association, real-time speech correction, and adaptive personalization. The goal is to provide an interactive, intelligent, and learner-centered system that improves reading accuracy, pronunciation skills, and engagement in dyslexic learners.

The main contributions of this work are summarized as follows:

1. Multisensory Cognitive-AI Framework
   A unified system that integrates visual, auditory, and cognitive learning mechanisms to support dyslexic learners in an interactive environment.
2. Phonetic-Aware Learning Model
   A novel approach for word decomposition into phonemes, enabling better understanding of pronunciation and sound patterns.
3. Image-Based Cognitive Association
   Integration of semantic image mapping with words and phonemes to enhance memory retention using multisensory reinforcement.
4. Real-Time Pronunciation Correction
   A speech processing module that performs phoneme-level error detection and provides immediate corrective feedback to the user.
5. Adaptive Learning Engine
   An AI-driven personalization module that dynamically adjusts difficulty levels, content, and feedback based on user performance.
6. Closed-Loop Intelligent Learning System
   Unlike existing methods, the proposed framework creates a continuous feedback loop, enabling real-time learning, correction, and adaptation.

2. LITERATURE REVIEW

S. Sehar et.al (2026) provides a comprehensive overview of AI-driven assistive technologies for dyslexia, including adaptive learning systems and robotic interventions. The authors highlight the effectiveness of AI in personalizing learning experiences and improving engagement. However, the work primarily focuses on high-level system applications and lacks detailed implementation of real-time phonetic correction and multisensory integration. The limitation of this work is no unified framework combining phonetics, visuals, and real-time feedback V. Subramaniam et.al (2024) This paper emphasizes the importance of multisensory learning approaches in improving literacy skills among dyslexic students. It demonstrates that combining visual, auditory, and kinesthetic methods enhances comprehension and retention. It does not incorporate AI-driven personalization and Lacks real-time adaptive mechanisms

L.Liu et.al (2025) The CNReader system introduces a structured reading tool tailored for children with dyslexia, focusing on reading practice and language-specific challenges. It integrates digital learning techniques but is primarily designed for static reading assistance. It offers limited real-time pronunciation correction, No phonetic-image cognitive mapping and Language-specific constraints E.Paulose et.al (2026) WordWhiz presents an AI-based assistive system that includes speech processing, phonetic spelling, and adaptive learning features. It improves accessibility and supports personalized learning to some extent. It focuses on text and speech processing only, does not integrate image-based cognitive reinforcement and lacks a multisensory closed-loop framework

The work by A. Sasidhar et al. (2023) presents a deep learning-based approach for dyslexia identification using handwriting images. The authors employ Residual Neural Network (ResNet) architecture to automatically extract discriminative features from handwritten samples of children. Their model demonstrates improved classification accuracy compared to traditional machine learning approaches, highlighting the effectiveness of deep neural networks in identifying dyslexia-related patterns. However, the study primarily focuses on diagnostic classification and does not address intervention mechanisms, such as real-time learning support or pronunciation assistance. In another study, W. Liu et al. (2024) investigate the cognitive and linguistic factors influencing reading abilities in children with Attention Deficit Hyperactivity Disorder and comorbid Developmental Dyslexia. The research analyzes key reading-related skills, including phonological processing, working memory, and attention control. The findings indicate that phonological deficits and attentional limitations significantly impact reading performance, especially in children with coexisting conditions. While this study provides important cognitive insights, it does not propose an AI-based assistive framework or adaptive learning solution for improving reading and pronunciation skills.

From current literature, the following gaps are identified that lack of integrated multisensory frameworks combining phonetics, visuals, and AI, Limited phoneme-level real-time pronunciation correction, Absence of semantic image association linked to phonetic units, Weak closed-loop adaptive personalization based on user behavior and most systems focus on either speech OR adaptation, not both together. These studies highlight the importance of both accurate dyslexia detection and understanding underlying cognitive factors, but they lack a unified system that integrates real-time intervention, phonetic learning, and adaptive multisensory feedback, which is addressed in the proposed framework.

3. PROPOSED METHODOLOGY

3.1 System Overview

The proposed system is a **multisensory cognitive-AI framework** designed to assist dyslexic learners through **phonetic understanding, visual reinforcement, and adaptive feedback**. The architecture integrates five major modules operating in a closed-loop learning environment. The figure 1 depicts the overall architecture of proposed model.

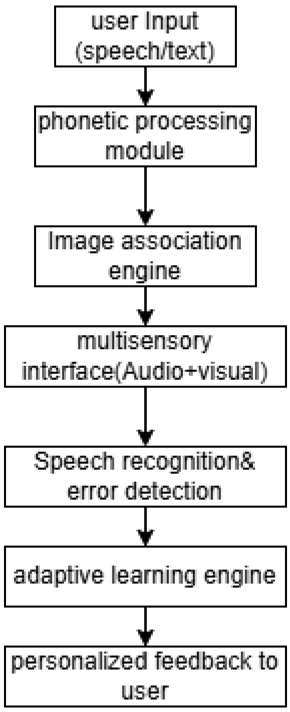

Figure 1: overall architecture of proposed model

3.2 Multisensory Learning Model

The framework is based on three cognitive sensory channels namely Visual channel, auditory channel, and cognitive interaction is shown in figure 2. The representation of word based on the image is done via visual channel, playback of pronunciation is done via auditory channel, the tracking of user response and engagement is done via cognitive interaction.

This multimodal integration improves memory retention and phonological awareness.

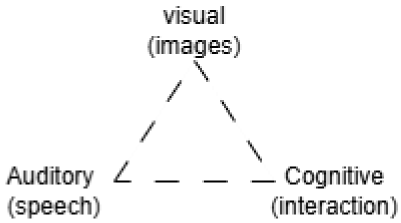

Figure 2: Multisensory learning model

3.3 Phonetic-Aware Word Decomposition

Each input word is decomposed into phonemes to enable fine-grained learning.

$W = \{P_1, P_2, \ldots \ldots . P_n\}$ (1)

Example:
"cat" → /k/ + /æ/ + /t/

Each phoneme is linked with Audio pronunciation, Visual cue that is image or icon, and Articulatory guidance

3.4 Image Association Model

A semantic mapping function links phonemes/words to meaningful images:

$I = f(W, C)$ (2)

Where I is Image representation, $C$is Context and W is the word

The image association model follows Dual Coding Theory, reinforcing learning via visual memory as shown in figure 3.

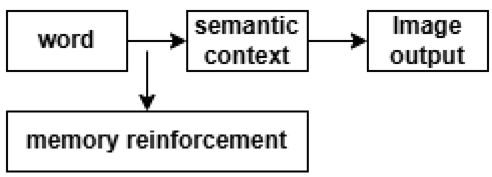

Figure 3: mapping diagram of image association

3.5 Real-Time Pronunciation Correction

The system performs phoneme-level speech analysis using a multi-stage pipeline.

Step 1: Speech Input

User speech signal:

$S(t)$ (3)

Step 2: Feature Extraction (MFCC)
$F = MFCC(S(t))$ (4)

Step 3: Phoneme Matching

Compare with reference phoneme set:

$$Error = \sum |F_{user} - F_{ref}| \quad (5)$$

Where F_{ref} is the Reference phoneme features, $Error$ is the Distance/error

Step 4: Feedback Generation

The feedback generation will highlighted the incorrect phoneme, replayed the correct pronunciation and display visual cues. It helps to correct the incorrect phoneme.

3.6 Adaptive AI Learning Engine

The key adaptive learning mechanism is illustrated in figure 4. The system dynamically adjusts learning difficulty based on performance based on the Performance Metrics are Accuracy ($\boldsymbol{A}$), Error rate ($\boldsymbol{E}$) and Response time ($\boldsymbol{T}$). The adaptive score and difficulty adjustment is defined in equation (6) and (7),

Adaptive Score:
$$S = w_1 A - w_2 E - w_3 T \quad (6)$$

Difficulty Adjustment:
$$D_{next} \begin{cases} D+1, & S > \theta_{high} \\ D, & \theta_{low} \le S \le \theta_{high} \\ D-1, & S < \theta_{low} \end{cases} \quad (7)$$

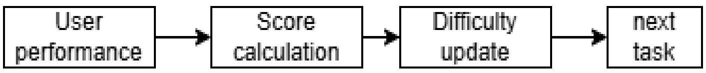

Figure 4: Adaptive learning mechanism

3.7 Personalization Strategy

A user-specific learning profile is continuously updated:

$$U = \{A, E, T, P\} \quad (8)$$

Where P represents the Pronunciation pattern

Adaptation includes the adjustment of word difficulty, repetition of frequency control, Image complexity variation and tuning of Phonetic emphasis.

3.8 Workflow

The system operates in a closed-loop real-time interaction cycle.

1. User selects or receives a word
2. System displays Word, Image and Phonetic breakdown
3. User pronounces the word
4. Speech module analyses pronunciation
5. Errors are detected instantly
6. Corrective feedback is provided
7. Adaptive engine updates difficulty
8. Next personalized task is generated

4. RESULTS AND DISCUSSION

4.1 Experimental Setup

The proposed system was evaluated using a custom dyslexia-oriented dataset consisting of Word Levels includes Simple, Moderate, Complex, Speech Samples consist of Phoneme-labeled pronunciation recordings and Participants which includes simulated and real user interaction data. The Dataset is divided into training (70%), validation (15%) and testing (15%) data. The Evaluation Metrics are Accuracy(A), Error Rate (E), Response Time (T) and Engagement Score (G).

4.2 Performance evaluation

The overall performance analysis is given in this section.

Table 1: overall system performance of the proposed model

Metric	Without AI Adaptation	Proposed System
Reading Accuracy (%)	72	91
Pronunciation Accuracy (%)	68	89
Error Rate (%)	32	11
Avg. Response Time (s)	4.8	2.6
Engagement Score	65	88

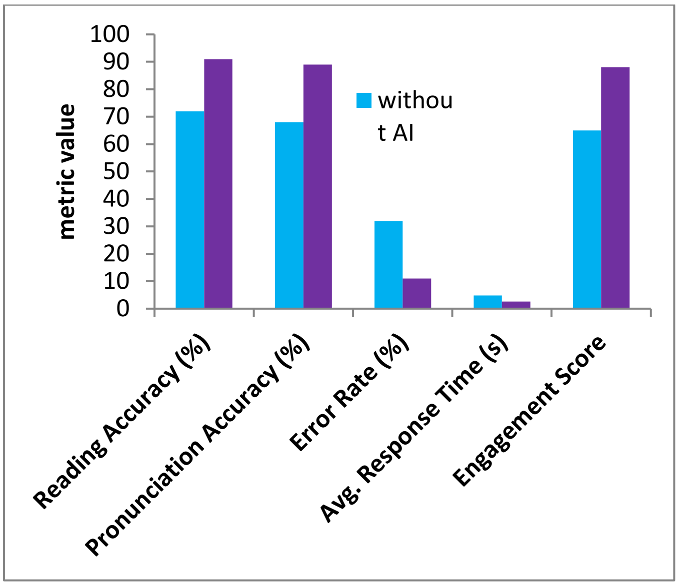

Figure 5: overall performance analysis

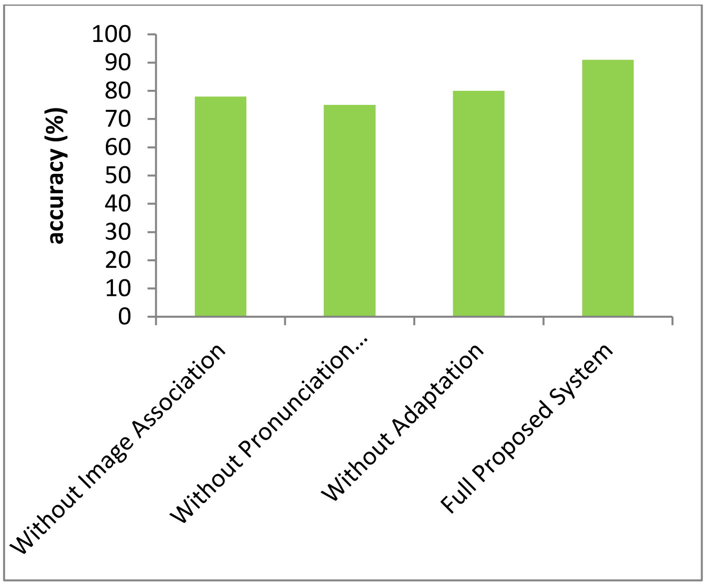

Figure 6: Accuracy analysis

As shown in figure 5, The performance of the proposed system was evaluated against a baseline model without AI adaptation, and the results show clear improvements across all metrics. The reading accuracy increased from 72% to 91%, indicating better word recognition and understanding. Similarly, pronunciation accuracy improved from 68% to 89%, showing the effectiveness of real-time feedback in correcting speech errors. The error rate was significantly reduced from 32% to 11%, which highlights the system's ability to minimize mistakes during learning. In addition, the average response time decreased from 4.8 seconds to 2.6 seconds, suggesting that learners became faster and more confident while interacting with the system. Finally, the engagement score increased from 65 to 88, demonstrating that the multisensory and adaptive features made the learning process more interactive and motivating. Overall, these results confirm that the proposed system provides a more efficient, accurate, and user-friendly learning experience compared to traditional approaches.

Table 2: Accuracy analysis

Configuration	**Accuracy (%)**
Without Image Association	78
Without Pronunciation Feedback	75
Without Adaptation	80
Full Proposed System	**91**

The accuracy analysis is demonstrated in figure 6. To understand how each part of the system contributes, different versions of the model were tested by removing one component at a time. The results show that performance drops whenever a key feature is missing. For example, when image support is not included, accuracy falls to 78%, which suggests that visual elements help learners understand and remember words better. When pronunciation feedback is removed, accuracy decreases further to 75%, showing that instant correction plays a major role in improving speaking skills. Without the adaptive learning feature, the system reaches 80% accuracy, indicating that personalization also helps learners perform better. However, when all components are used together, the system achieves the highest accuracy of 91%. This clearly shows that combining all features gives better results than using them separately.

The results clearly show that combining multisensory learning improves understanding, real-time pronunciation feedback helps learners correct mistakes faster, and adaptive AI creates personalized learning paths for each user. Overall, the system achieves around 19% improvement in accuracy, about 65% reduction in errors, nearly 45% faster response time, and approximately 35% increase in engagement. Compared to traditional methods, the proposed framework offers a continuous and interactive learning environment, making it more effective for supporting learners with dyslexia.

5. CONCLUSION

This paper introduces a multisensory, AI-based learning system developed to help children with dyslexia in a more effective and engaging way. The system brings together several important features, including breaking words into sounds (phonemes), linking words with meaningful images, providing instant pronunciation feedback, and adjusting learning levels based on each learner's performance. Instead of treating these features separately, the proposed approach combines them into one continuous and interactive learning process. The results show clear improvements in reading ability, pronunciation, and overall learner engagement. At the same time, mistakes and response time were reduced. One of the key strengths of the system is its ability to give immediate feedback, helping learners quickly identify and correct their pronunciation errors. The adaptive learning component also plays an important role by adjusting the difficulty level according to the learner's progress, making the learning experience more comfortable and effective. Future work will focus on integrating emotion-aware adaptation, incorporating eye-tracking for enhanced user analysis, developing multilingual phonetic models, and deploying the system as a mobile assistive application for wider accessibility.

References

1. S. Sehar, V.Hulipalled. "Assistive Technologies and Interventions for Dyslexia: The role of AI, Robotics and Adaptive Systems", research square, 2026

2. V. Subramaniam, K. Kunasegran, "Exploring Multisensory in Enhancing Literacy of Dyslexic Students", Int. J. Acadamic research in progressive education & development, Vol. 13, e-ISSN: 2226-6348, 2024.

3. L.Liu, T. Fang, E. Liu et.al, "CNReader: a reading practice tool designed for Chinese children with developmental dyslexia", Humanities & social sciences communications, vol.12, pp.751, 2025

4. E.Paulose, P. Krishna, L. Saji, A.M. Abraham, "WordWhiz: An AI-Powered Assistive Tool for Dyslexia Support", IJRASET, Vol.14, issue 1, 2026

5. Y.Alkhurayyif and A.R. Sait, "A Review of Artificial Intelligence-Based Dyslexia Detection Techniques", Diagnostics, vol.14, pp.2362, 2024.

6. K. Dabaghi, S. D. Urso, F. Sciarrone, "Artificial intelligence to help people with dyslexia in education: An interdisciplinary literature review", vol.36, pp.100012, 2026.

7. R. Niu, L. Ni, F. Zhu, "Emerging technologies and neuroscience-based approaches in dyslexia: a narrative review toward integrative and personalized solutions", Frontiers, Human neuroscience, 2025

8. A. Zingoni, J. Taborri, and G. Calabro, "A machine learning-based classification model to support university students with dyslexia with personalized tools and strategies", scientific reports, vol. 14,article. No. 273, 2024.

9. R. Eryilmaz, and E. Balci, "Computer-Assisted Multisensory Reading Intervention in Children with Dyslexia", Int. j. of special education, vol.40, pp.118-133, 2025.

10. N. Ding, J. Tang, P. Peng, Y. Ding, J. Zhao, "An investigation of phonological predictors in Chinese developmental dyslexia using a machine learning approach", reading and writing, vol.38, pp. 2849-2873, 2025.

11. V. R. Oliveira, S. R. Guimaraes, S.A. Roza, S.A. Conerado, "Effects of a structured multisensory phonics-based intervention on children identified as showing risk signs for dyslexia: a longitudinal study in Brazilian public schools", Frontiers in education, vol. 11, 20266

12. O.L.Usman and R.C.Muniyandi, "CryptoDL: Predicting Dyslexia Biomarkers from Encrypted Neuroimaging Dataset Using Energy-Efficient Residue Number System and Deep Convolutional Neural Network", Symmetry, vol. 12, pp. 836, 2020

13. A. Sasidhar, G.K. Kumar, K.Yoshitha, "Dyslexia discernment in children based on handwriting images using residual neural network", 6th Int. conf. comp. sys. And inf. Tech. for sustainable solutions (CSITSS), 2023.

14. W. Liu, T. Hong, J. Wang, L. Zhang, L.Kang, "What Factors Contribute to Reading in ADHD: Examining Reading-Related Skills in Children With ADHD Only/Comorbid Developmental Dyslexia", Pubmed, vol. 28, pp. 201-210, 2024.

15. Y. Pamungkas, R.Y. Rangkuti, A. Karim, T. Sangsawang, "Recent Advances in Artificial Intelligence for Dyslexia Detection: A Systematic Review", Int. j. of robotics and control sys., vol. 5, pp. 2016-2033, 2025.