

Adaptive Learning Platform for English Language Development in Autistic Children in Sri Lanka

ADAPTIVE ENGLISH PRONOUNCIATION SUPPORT MODULE FOR AUTISTIC CHILDREN

Project Proposal Report

Samaranayake N G (IT22266378)

R26-IT-006

B.Sc. (Hons) Degree in Information Technology

Department of Information Technology,

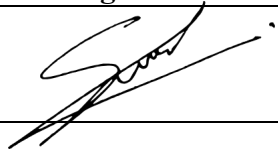
Sri Lanka Institute of Information

Technology, Sri Lanka

March 2026

Declaration

I declare that this is our own work, and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or Institute of higher learning, and to the best of our knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text. Also, I hereby grant to the Sri Lanka Institute of Information Technology the nonexclusive right to reproduce and distribute my dissertation, in whole or in part, in print, electronic, or other medium. I retain the right to use this content in whole or in part in future works (such as articles or books).

Name	Student ID	Signature
Samaranayake N G	IT22266378	

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Name	Date	Signature
Ms. Manori Gamage	20.03.2026	

Abstract

Children with Autism Spectrum Disorder (ASD) often experience difficulties in communication and language learning due to differences in cognitive processing, attention, and sensory sensitivity. In Sri Lanka, this challenge is further intensified because English is commonly taught as a second language, requiring additional support in pronunciation, vocabulary, and comprehension. Traditional teaching methods and learning materials are generally designed for typically developing students and may not adequately address the individualized learning needs of autistic children. This research aims to design and develop an adaptive pronunciation support component for Grade 3 autistic children to improve foundational English pronunciation learning. The proposed component provides multimodal learning support through images, audio pronunciation, visual pronunciation guidance, repeated practice opportunities, and activity-based reinforcement. The system also records student speech, performs pronunciation similarity evaluation against a reference pronunciation, and adaptively selects the next word based on the student's previous performance. In addition, detailed performance records are presented through a teacher dashboard to support progress monitoring and personalized instruction. The study follows a design and development research approach. The expected outcome is a teacher-supported adaptive learning component that can enhance pronunciation practice, provide structured feedback, and improve teacher awareness of student progress in English language learning.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my research supervisor, Ms. Manori Gamage, for her continuous guidance, valuable feedback, and support throughout the development of this research proposal. Her insights and academic expertise have been instrumental in shaping the direction of this research. I would also like to extend my sincere appreciation to my co-supervisor, Mrs. Suriya Kumari, for her guidance, encouragement, and constructive suggestions that contributed significantly to improving this research work. I would also like to thank the Sri Lanka Institute of Information Technology (SLIIT) for providing the academic environment and resources necessary for conducting this research. Special appreciation is extended to my research group members for their collaboration and valuable discussions that helped refine the overall system concept. Finally, I would like to express my heartfelt gratitude to my family and friends for their encouragement, patience and continuous support throughout my academic journey.

Table of Contents

Declaration	2
Abstract	3
ACKNOWLEDGEMENT	4
Table of Contents	5
LIST OF FIGURES	6
LIST OF TABLES	6
LIST OF ABBREVIATIONS.....	7
1. INTRODUCTION	8
1.1 Background and Literature Survey	8
1.2 Research Gap	9
1.3 Research Problem	10
2. OBJECTIVES	11
2.1 Main Objective.....	11
2.2 Specific Objectives	11
3. METHODOLOGY	12
3.1 Tools and Technologies	13
3.2 System Architecture.....	14
3.3 Component Overview Diagram.	15
4. Work Breakdown Structure	16
5. Gantt Chart.....	17
6. PROJECT REQUIREMENTS.....	18
6.1 Stakeholder identification	18
6.2 Functional Requirements	18
Student-side functions.....	18
Teacher-side function.....	19
6.3 Non-Functional Requirements	19
7. Existing applications and their limitations.....	20
Main focus	20
Key strengths	20
Limitations	20
How the proposed component defers.....	20
8. COMMERCIALIZATION PLAN.....	21

8.1 Value proposition.....	21
8.2 Revenue model.....	22
8.3 Cost estimation.....	22
8.4 Target market and market potential	22
8.5 Competitive advantage.....	23
8.6 Intellectual property considerations	23
9. BUDGET AND BUDGET JUSTIFICATION	24
9. REFERENCES	25
10. APPENDICES	26
10.1 Risk Analysis	26
10.2 Ethical Considerations and Consent Documentation.....	27

LIST OF FIGURES

Figure 1: System overview	14
Figure 2: component overview	15
Figure 3: WBS	16
Figure 4: Gantt Chart	17

LIST OF TABLES

Table 1: Existing systems and limitations	21
Table 2: Cost estimation	22
Table 3: Budget and Justification	24
Table 4: Risk Analysis	27

LIST OF ABBREVIATIONS

Description	Abbreviations
autism spectrum disorder	ASD
Computer-Assisted Pronunciation Training	CAPT
Mel-frequency cepstral coefficients	MFCC
Dynamic time warping	DTW
English as a second language	ESL

1. INTRODUCTION

Pronunciation plays an important role in the development of foundational English language skills, especially for children who are learning English as a second language. For children with autism spectrum disorder (ASD), learning pronunciation can be more challenging due to differences in communication, attention, sensory sensitivity, and learning pace. In Sri Lanka, this challenge becomes more significant because many autistic children are required to learn English within educational settings that are largely designed for typically developing students. As a result, they may need more structured, repetitive, and personalized support to develop basic pronunciation skills effectively.

This research focuses on the design and development of an adaptive pronunciation support component for Grade 3 autistic children. The component is intended to assist students in learning English pronunciation through a multimodal approach that includes images, audio pronunciation, flashcards, and visual pronunciation guidance. It also incorporates activity-based reinforcement to help students practice and recall learned words in an engaging and supportive manner.

A key feature of the proposed component is its ability to record student pronunciation and compare it with a correct reference pronunciation to generate a similarity-based score. While detailed performance results are made available only to teachers through a dashboard, students receive simple and supportive feedback such as visual indicators and encouraging animations. In addition, the component includes an adaptive learning mechanism that selects the next word based on the student's previous performance, thereby enabling a more personalized learning experience.

Overall, the proposed component aims to support foundational English pronunciation learning in a way that is structured, adaptive, and teacher-assisted, while also helping teachers monitor student progress more effectively.

1.1 Background and Literature Survey

English pronunciation plays a foundational role in early language learning, supporting intelligibility, listening comprehension, and classroom participation. In digital learning contexts, Computer-Assisted Pronunciation Training has emerged as a key approach, allowing learners to hear model speech, practice, and receive automated feedback.

These challenges are amplified for children with autism spectrum disorder, a developmental condition affecting communication, social interaction, and learning. Conventional classroom instruction-which often relies on verbal explanation and fast-paced, uniform delivery-tends to be less effective for autistic learners. Research consistently highlights the value of structured, predictable, and visually supported environments for this population, where pictures, cues, and guided prompts reduce cognitive load and help connect sounds with meaning.

Technology-assisted learning has gained traction as a promising avenue for autistic learners. Studies indicate that educational technologies can support language and vocabulary development in children with ASD, and a 2024 systematic review specifically noted growing interest in school-based language learning technologies for this group, while calling for more targeted interventions. Mainstream tools like ELSA Speak, Duolingo, and Google Read Along already demonstrate the

technical feasibility of speech-based English support-yet these platforms are designed for general learners and lack autism-sensitive scaffolding.

Key gaps remain in the existing literature. Most commercial tools do not support activity-embedded pronunciation capture, where speech is collected naturally during tasks rather than through explicit recording steps. Equally, teacher-centered monitoring features-such as tracking category-wise errors or identifying repeated pronunciation difficulties for individualized remediation-are largely absent. Adaptive learning in current CAPT systems also tends to rely on single performance scores rather than richer patterns like sound-category confusion or response hesitation, which are more informative for special education contexts.

From a technical standpoint, practical pronunciation support can be implemented through speech feature extraction and similarity analysis, avoiding the need for computationally intensive model training-making it a feasible approach for design-oriented academic development targeting Grade 3 learners with ASD.

1.2 Research Gap

Existing research has demonstrated technology-assisted interventions show promise for autistic learners and CAPT systems are well established in ESL contexts, these two bodies of work have developed largely in isolation-and their intersection remains a clear gap in the literature. Research on digital tools for children with ASD focuses predominantly on vocabulary, communication, and literacy, with pronunciation receiving little attention. CAPT research, while technically mature, is built almost entirely on neurotypical and adult learner assumptions, using drill-based methods that do not align with how autistic children learn. The few pronunciation interventions designed for autistic learners remain narrow in scope, and their authors themselves acknowledge that integrated, automated systems are still underdeveloped-explicitly calling for real-time speech assessment and adaptive feedback as future directions. Individually, the building blocks exist: visual scaffolding is well supported for ASD, lightweight MFCC and DTW-based speech scoring is technically feasible, and learning analytics dashboards are recognized as effective for teacher decision-making. Yet none of these have been brought together into a system where pronunciation practice is embedded naturally within reinforcement activities, progression adapts to each learner's error patterns, and teachers receive meaningful category-wise insights for follow-up. This gap is especially pronounced in Sri Lanka, where Grade 3 marks the formal start of ESL instruction, assistive technology use in autism classrooms remains basic and inconsistent, and local speech technology research is concentrated on Sinhala-language intervention rather than English pronunciation support. An integrated, autism-sensitive ESL pronunciation system designed around this specific context is therefore both absent from the literature and genuinely needed in practice.

1.3 Research Problem

Children with Autism Spectrum Disorder (ASD) often experience difficulties in developing language, communication, and comprehension skills due to differences in cognitive processing, attention, and sensory sensitivity. In Sri Lanka, this challenge is further compounded because English is typically taught as a second language, which requires additional effort in vocabulary acquisition, pronunciation, reading, and understanding basic concepts. Traditional classroom teaching methods and existing learning materials are largely designed for typically developing students and do not adequately accommodate the unique learning needs of autistic children. Furthermore, special education teachers often manage several students with varying abilities, making it difficult to provide personalized instruction tailored to each child's learning pace and preferences. As a result, many autistic children struggle to effectively learn foundational English language concepts and may find it difficult to participate in mainstream educational assessments and government examinations to which they are being taught for.

2. OBJECTIVES

2.1 Main Objective

The main objective of this research is to design and develop an adaptive pronunciation support component to assist Grade 3 autistic children in learning foundational English language skills through multimodal practice, pronunciation assessment, and personalized learning progression, while enabling teachers to monitor student performance and progress effectively.

2.2 Specific Objectives

- **To collect and organize** Grade 3 English learning content relevant to autistic students, including alphabets, word categories, reference pronunciations, images, and activity materials.
- **To design and develop** a multimodal learning interface that provides pronunciation support through flashcards, images, audio playback, and visual pronunciation guidance such as lip movement cues, in order to suit the learning needs of autistic children.
- **To implement** a speech acquisition and preprocessing module to capture and prepare student pronunciation attempts for analysis.
- **To develop** a pronunciation similarity assessment mechanism to compare student speech with expected pronunciations and generate word-level performance scores.
- **To design** activity-based reinforcement tasks that assess student understanding and pronunciation of learned words through structured learning interactions.
- **To implement** an adaptive learning mechanism that selects the next learning word based on the student's pronunciation performance and relevant linguistic complexity factors.
- **To develop** a teacher-centered dashboard for monitoring individual student progress, category-wise performance, pronunciation records, and learning trends.
- **To evaluate** the proposed component in terms of functionality, usability, and effectiveness in supporting foundational English pronunciation learning among Grade 3 autistic students.

3. METHODOLOGY

This study adopts a **design and development research approach** to design, develop, and evaluate an adaptive pronunciation support component for **Grade 3 children with ASD**. The component is intended to support foundational English pronunciation learning through **multimodal instructional support, activity-based speech capture, phoneme-aware pronunciation analysis, adaptive content sequencing, and teacher-centered progress monitoring**. The module is developed as a web-based interface using **React Native** and the learning process begins with the selection of either alphabet pronunciation or word pronunciation. The teacher first selects the relevant word category based on the Grade 3 English syllabus, such as animals, classroom objects, fruits, or daily-use vocabulary. The teacher then selects the initial target word for the learner. Once the target word is selected, the system presents a set of multimodal learning supports to help the learner become familiar with the word before evaluation takes place. These supports may include **flashcards or picture-based prompts (real and animated), audio pronunciation, phoneme-level highlighting, and visual pronunciation guidance** such as lip movement cues or lip-sync-based assistance. The learner is allowed to interact with these supports multiple times before moving to the reinforcement stage. This phase is designed to reduce learner anxiety, improve word familiarity, and create a more supportive and autism-sensitive environment for pronunciation learning.

Following the initial support phase, the learner is guided through reinforcement activities based on the selected word. These activities are designed as structured learning interactions that promote engagement and natural speech production. Examples include verbally responding to a picture prompt, repeating a model pronunciation after listening to it, or responding to a visual cue without directly viewing the written word. The learner's speech is captured naturally during these pronunciation-related reinforcement activities using the microphone.

The captured speech is then preprocessed and analyzed. Audio preprocessing includes steps such as silence trimming, normalization, and preparation of the audio signal into a comparable format for analysis. Relevant acoustic features are then extracted, with **MFCC** used as the primary feature extraction for pronunciation comparison. In addition to MFCC-based representations, timing-related information such as hesitation before speech, total response duration, and number of attempts may also be recorded as behavioral indicators. Each target word is linked to its expected phoneme sequence, and the learner's speech features are aligned with the reference pronunciation using a similarity-based comparison method called **DTW** in order to estimate phoneme-level pronunciation similarity and generate a word-level pronunciation score.

The **adaptive decision-making mechanism** in this component goes beyond a simple correct-or-incorrect evaluation model. After each pronunciation-related activity, the system considers repeated attempt patterns, hesitation behavior, learner performance history, and recurring phoneme-related difficulty patterns. If the learner performs at or above the expected level, the system proceeds with the next suitable word in the planned progression. However, if the learner performs below the expected level, the system selects the next learning item adaptively by choosing a pronunciation-related support word or reinforcement pathway rather than repeating the same item or skipping ahead. This adaptive selection is based on factors such as **phoneme similarity, phonics structure, vowel patterns, letter count, word difficulty, and previously observed learner weaknesses**. In this way, the component supports performance-based next-word selection.

A key methodological improvement in this study is the introduction of **phoneme-aware adaptive remediation mechanisms**. If a learner repeatedly struggles with words that share a similar sound pattern, the system identifies that recurring difficulty and routes the learner into a reinforcement or remediation pathway using simpler or related words. For example, if the learner experiences difficulty with multiple words containing a similar initial sound, the next selected word may target that same sound again in an easier, shorter, or more familiar context. This allows the adaptive mechanism to respond not only to the learner's final score, but also to likely phoneme-level sources of pronunciation difficulty. As a result, the component supports more meaningful, individualized, and responsive progression for ASD learners.

The backend and data management layer will be implemented using **Node.js with Express.js** or **Python FastAPI** to expose services for learner management, word retrieval, audio submission, pronunciation scoring, adaptive word selection, and teacher dashboard data access. Learner records, target words, phoneme mappings, pronunciation results, attempt history, and progress summaries, and student audio playbacks will be stored in a relational database such as **PostgreSQL**. The backend will act as the communication layer between the mobile learning application, the pronunciation analysis engine, and the teacher monitoring dashboard.

All detailed performance data are sent to a teacher dashboard and will be developed as a web-based monitoring interface using **React Native** and connected to the backend API for learner performance access. Where teachers can monitor learner performance in a comprehensive and actionable manner. This enables teachers to identify learner-specific needs and make more targeted instructional decisions.

Through this methodology, the study aims to design, implement, and assess a feasible, autism-sensitive pronunciation support component that integrates structured reinforcement, embedded speech analysis, adaptive progression, and teacher-centered intervention support.

3.1 Tools and Technologies

- **Backend:** FastAPI (Python) audio processing, scoring logic, adaptive decision-making, and dashboard data handling.
- **Database:** PostgreSQL storing structured data
- **API Testing:** Postman for testing and validating backend APIs and routing responses.
- **Mobile Application:** React Native with Expo for cross-platform mobile application for learner interaction
- **Audio Processing:** librosa, NumPy and SciPy for audio loading, preprocessing, and feature extraction
- **Authentication:** JWT - manage secure login and access control
- **Charts/Visualization:** Recharts - displaying progress data visually
- **ORM:** SQLAlchemy - provides a cleaner way to interact with PostgreSQL from FastAPI

3.2 System Architecture

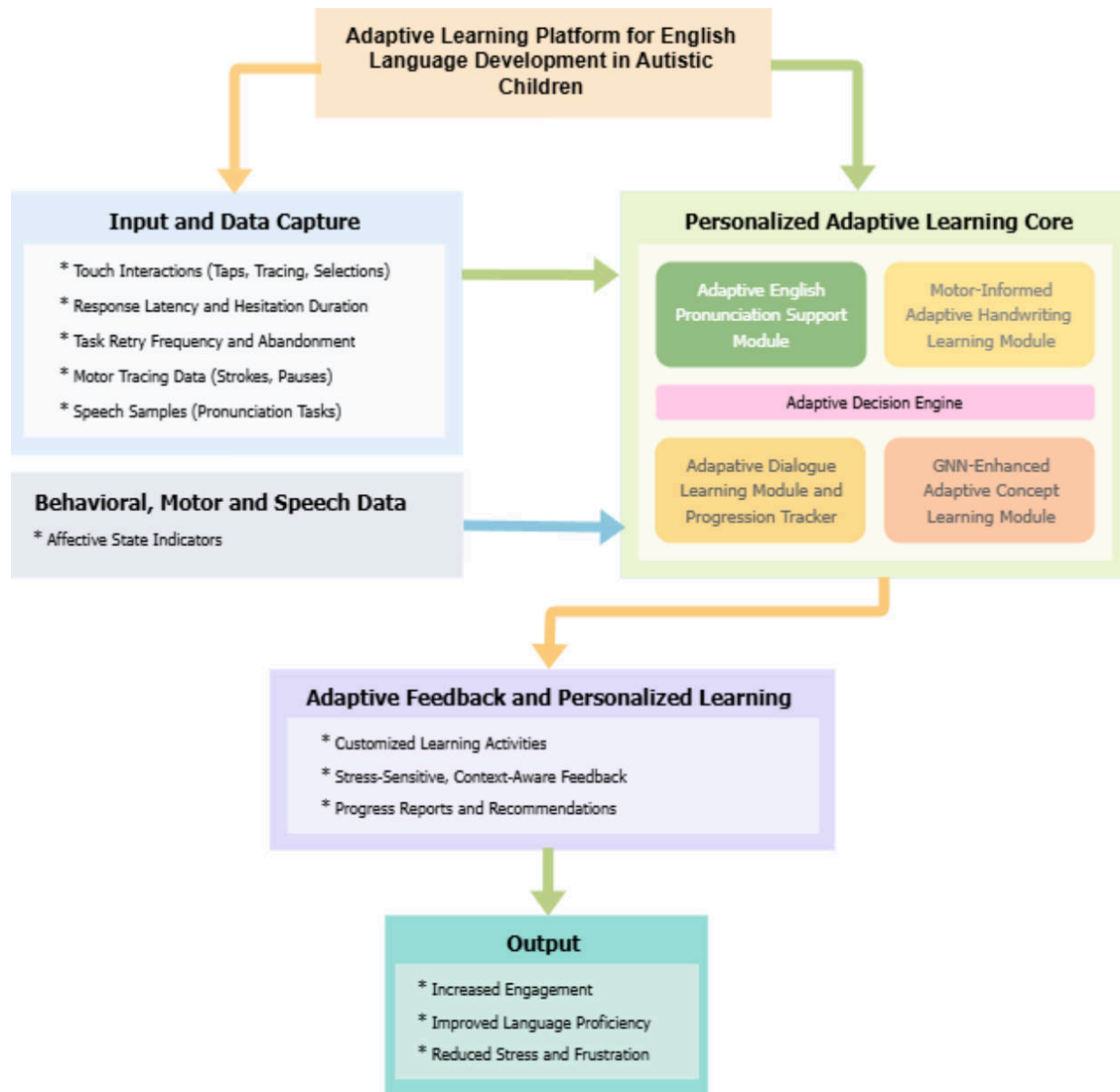


Figure 1: System overview

3.3 Component Overview Diagram.

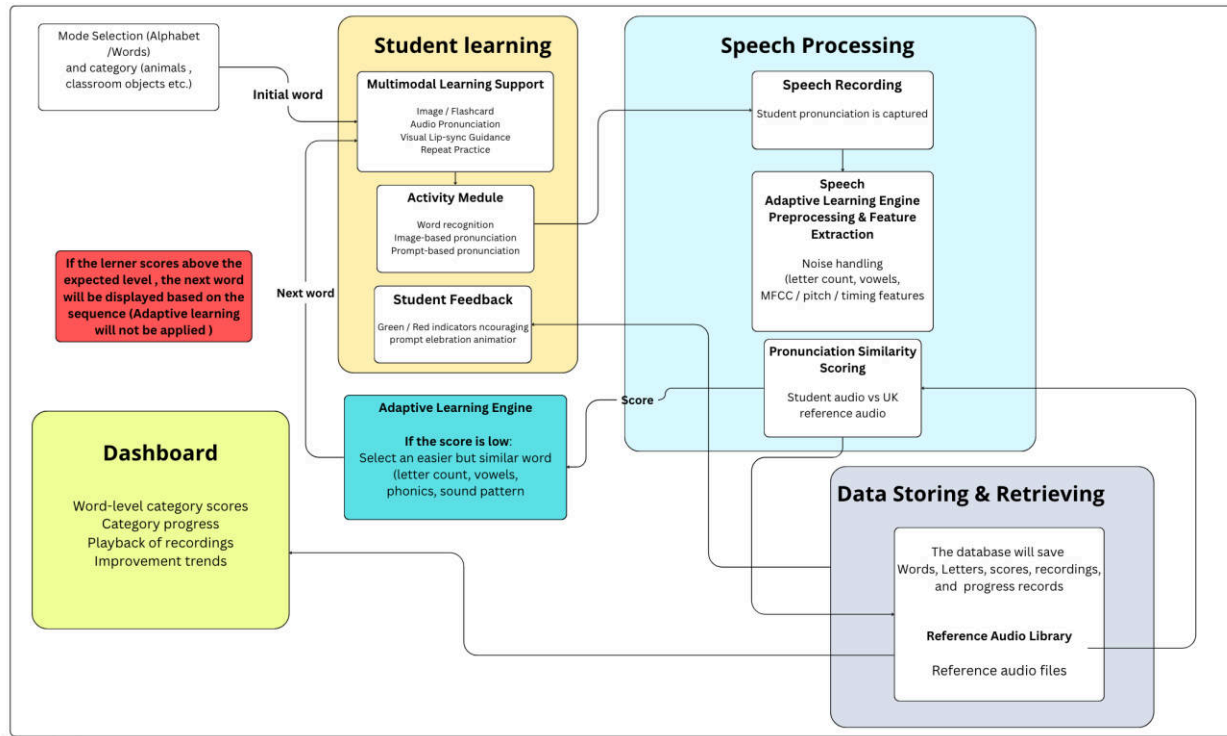


Figure 2:component overview

4. Work Breakdown Structure

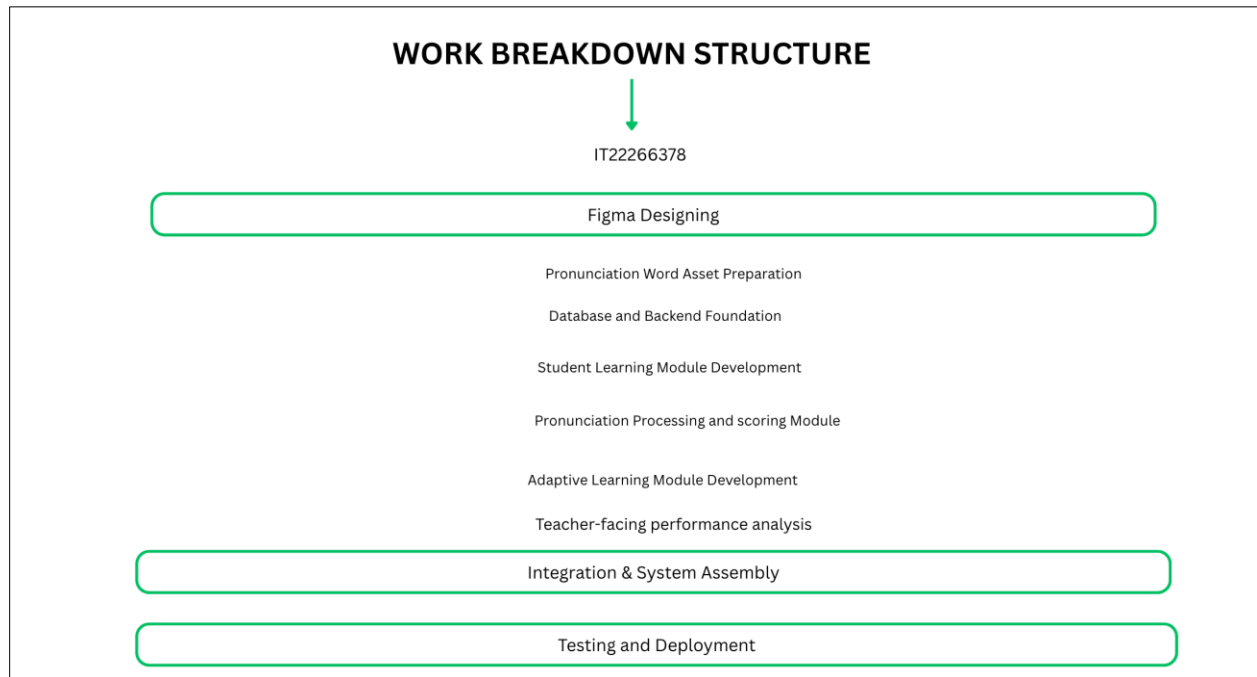


Figure 3:WBS

5. Gantt Chart

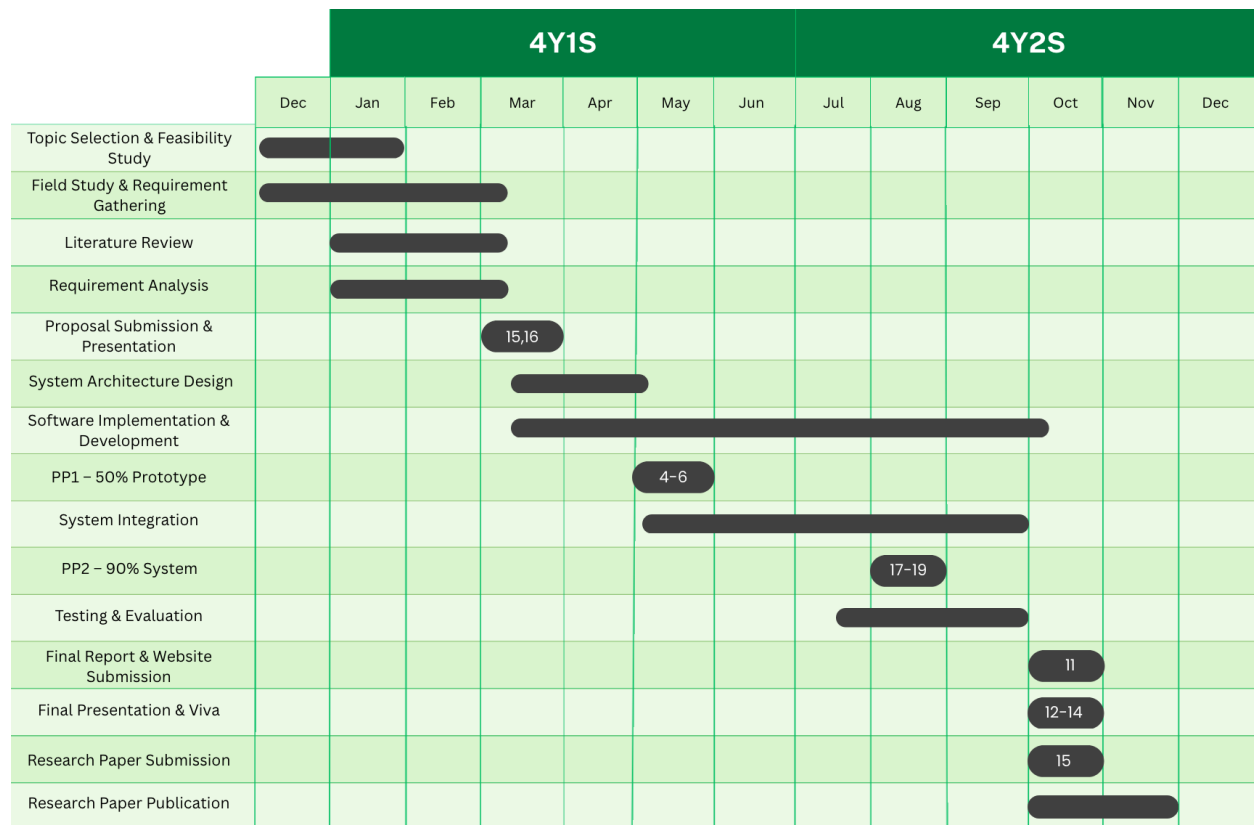


Figure 4: Gantt Chart

6. PROJECT REQUIREMENTS

6.1 Stakeholder identification

- **Autistic children** - Primary users who will directly interact with the system for learning English pronunciation through adaptive activities, guided practice, and reinforcement-based exercises.
- **Special education teachers** - Main facilitators who will supervise student usage, monitor pronunciation progress, identify learning difficulties, and use the system to provide more personalized support.
- **School principals** - Institutional stakeholders who are responsible for ensuring inclusive education practices, supporting the adoption of assistive learning technologies, and reviewing student outcome data for decision-making.
- **Parents / Guardians** - Secondary stakeholders who support children outside the classroom, encourage continued learning practice, and may benefit from understanding the child's progress and challenges.

6.2 Functional Requirements

Student-side functions

- The system shall allow the student to choose between alphabet pronunciation and word pronunciation.
- The system shall support the alphabet pronunciation module as an optional learning pathway without adaptive progression.
- The system shall present the selected word to the student using multimodal learning supports, including images or flashcards, audio pronunciation, and visual pronunciation guidance.
- The system shall allow the student to repeat the learning supports and practice the selected word multiple times before proceeding to reinforcement activities.
- The system shall provide two to three reinforcement activities related to the selected word.
- The system shall capture the student's speech during pronunciation-based reinforcement activities.
- The system shall preprocess the captured speech before pronunciation analysis.
- The system shall compare the student's captured pronunciation with the stored reference pronunciation.
- The system shall generate a similarity-based pronunciation score for the spoken word.
- The system shall provide simple child-friendly feedback to the student, such as correct/incorrect indicators, visual cues, colors, and celebratory animations.
- The system shall not display detailed pronunciation scores or analysis directly to the student.
- The system shall determine the next learning word based on the student's pronunciation performance.
- The system shall provide an easier but pronunciation-related word when the student performs below the expected threshold.

- The system shall continue to the next word in the planned progression when the student performs at or above the expected threshold.

Teacher-side function

- The system shall allow the teacher to select the initial word category for the student.
- The system shall allow the teacher to select the initial target word for the student.
- The system shall allow the teacher to view word-level pronunciation scores for each student.
- The system shall allow the teacher to view category-wise performance percentages for each student.
- The system shall allow the teacher to monitor student progress over time.
- The system shall allow the teacher to view practice frequency and completion progress for each student.
- The system shall allow the teacher to play back the student's recorded pronunciation attempts from the dashboard.
- The system shall maintain student-wise pronunciation records for future review.
- The system shall allow the teacher to identify areas in which the student shows repeated pronunciation difficulty.
- The system shall support secure teacher login for access to the dashboard.

6.3 Non-Functional Requirements

1. The system shall provide a simple, clear, and user-friendly interface suitable for autistic children and teachers.
2. The system shall record, process, and display pronunciation results within an acceptable response time.
3. The system shall store student recordings, scores, and progress data accurately and consistently without data loss.
4. The system shall provide consistent and reliable pronunciation similarity scores using a standard UK pronunciation reference.
5. The system shall ensure that only authorized teachers can access student records and dashboard data.
6. The system shall be easy to update, modify, and extend with new words, activities, and pronunciation data.
7. The system shall support the future addition of more students, words, categories, and learning activities.
8. The system shall work properly on the intended devices such as tablets, mobile devices.
9. The system shall use an autism-friendly design with minimal distractions, clear visuals, and supportive feedback.
10. The system shall be available for use during learning sessions and allow teachers to review stored records when needed.

7. Existing applications and their limitations

Application	Main focus	Key strengths	Limitations	How the proposed component defers
Read Along	Speech-based reading tutor for children	Provides real-time feedback while children read aloud, includes an in-app reading buddy, and offers a large set of stories and word games for early reading practice.	Mainly supports reading aloud and literacy practice, not structured pronunciation learning for autistic children. It does not clearly provide teacher-controlled word selection, pronunciation similarity scoring against a UK reference, or adaptive next-word selection based on pronunciation difficulty. This last point is an inference from its product focus	Proposed component focuses specifically on pronunciation learning, not general reading. It includes multimodal word practice, similarity-based pronunciation scoring, adaptive word progression, and teacher dashboard monitoring.
Proloquo2Go / Proloquo	AAC communication support	Designed for non-speaking users and AAC learners; highly customizable; includes Focus Mode to guide attention to key words without limiting vocabulary access.	Primarily supports communication through AAC, not pronunciation training. It emphasizes vocabulary access and communication support rather than English pronunciation scoring, activity-based pronunciation learning, or teacher-controlled adaptive word practice.	The proposed component is not an AAC communication board. It is a teacher-supported pronunciation learning component for Grade 3 autistic learners, with pronunciation evaluation and progress tracking.
Avaz AAC	AAC communication and learning support	Picture- and text-based AAC app; includes analytics and graphs for communication progress; supports children and adults with communication needs.	Strong for AAC-based communication but not specifically designed for pronunciation learning. It does not center on comparing student pronunciation with a UK reference audio, structured pronunciation activities, or adaptive next-word selection based on speech performance. This distinction is an inference from the official product description and feature set.	component targets foundational English pronunciation learning and combines multimodal practice, similarity-based scoring, adaptive progression, and teacher-side records.

Proposed Component	Adaptive pronunciation support for Grade 3 ASD learners	Multimodal practice using images, audio, flashcards, and visual pronunciation guidance; student speech recording; similarity-based pronunciation scoring with UK reference audio; adaptive next-word selection; teacher dashboard with scores, playback, and progress monitoring.	As a research prototype, it may initially rely on rule-based adaptation and word-level similarity scoring rather than highly advanced phoneme-level diagnosis.	The Proposed component It directly addresses the gap by combining pronunciation learning, adaptive progression, Similarity analysis and teacher-centered monitoring in one component tailored to ASD learners in the Sri Lankan ESL context.
---------------------------	---	---	--	---

Table 1: Existing systems and limitations

8. COMMERCIALIZATION PLAN

8.1 Value proposition

The proposed system offers a specialized and locally relevant solution for supporting English language learning among autistic children in Sri Lanka. Unlike general learning applications, this platform is specifically designed to support adaptive pronunciation learning, teacher monitoring, and curriculum-aligned intervention within special education settings.

The main value propositions of the system are as follows:

- Provides an adaptive English pronunciation learning platform tailored for autistic children.
- Supports personalized learning pathways based on student performance and progress.

- Help teachers monitor word-level pronunciation performance, learning trends, and student engagement.
- Reduces the difficulty of delivering the same content to all students by enabling individualized content progression.
- Aligns with the Sri Lankan educational context, making it more relevant than imported generic learning tools.
- Improves student confidence and participation through structured reinforcement-based activities.
- Assists schools in strengthening inclusive education practices through measurable student progress data.

8.2 Revenue model

- Freemium trial model - Schools and institutions can access a limited free trial version to evaluate the platform before making a commitment.
- Tier I institutional plan - Designed for small special education units with up to 30 students.
- Tier II institutional plan - Designed for mid-scale schools or larger special education units with 30 to 120 students.
- Tier III institutional plan - Designed for large institutions or multi-campus educational organizations with more than 120 students.

8.3 Cost estimation

Cost Category	Estimated Cost (LKR)	Description
Labor	20,400	Development effort for design, frontend, backend, adaptive learning logic, and dashboard implementation
Material / Learning assets	15,300	Preparation of pronunciation words, images, audio assets, and related educational content
Hosting	8,500	Backend hosting, database hosting, and deployment-related services
API / Tools	10,200	API testing tools, external libraries, and supporting technical services
Total Estimated Cost	34,000	Estimated budget for prototype-level implementation

Table 2: Cost estimation

8.4 Target market and market potential

The proposed system primarily targets the special education sector in Sri Lanka, with a focus on institutions that support autistic children and related learning needs.

The identified market includes:

- 27 special education schools
- 704 special education institutes

The main target customers are:

- Sri Lankan special education schools
- Special education units
- Autism support centers
- Inclusive education institutions

8.5 Competitive advantage

The proposed system has several competitive advantages compared to general English learning applications and static educational tools.

- It is specifically designed for autistic children, rather than for general learners.
- It supports adaptive pronunciation learning, which allows the next learning activity or word to change based on student performance.
- It includes teacher-facing monitoring and analytics, allowing teachers to track student progress more effectively.
- It is built for the Sri Lankan educational context, which strengthens relevance and local applicability.
- It combines student learning, adaptive logic, and teacher support in one platform instead of treating them as separate tools.
- It supports institutional use through a school-based licensing model, making adoption more realistic for the target environment.

8.6 Intellectual property considerations

- **Copyright protection** for the software code, user interface design, educational activity structure, and teacher dashboard design.
- **Protection of adaptive learning logic** used to determine personalized progression and pronunciation support.
- **Protection of pronunciation asset organization and database schema**, including structured word sets, categories, and progress tracking design.
- **Trademark registration** for the product name, branding, and visual identity if the system is commercialized in the future.
- **Protection of generated content and system documentation**, including design documents, workflow structure, and technical architecture.

9. BUDGET AND BUDGET JUSTIFICATION

Description	Total Amount (Rs.)
Cloud Database and Storage Services (Firebase / PostgreSQL) – 6 months	6000
Backend / Server Hosting for Dashboard and Adaptive Logic – 6 months	5500
Speech Processing and Development Tools (libraries, APIs, optional premium tools)	4000
Mobile / Tablet Application Development Environment and Testing Tools	3500
Internet and Utility Cost for Development and Testing	3000
Transport Cost for School Visits / User Testing / Data Collection	6000
Miscellaneous Expenses (printing, small equipment, replacements, contingencies)	3500
Total Budget	31,500

Table 3: Budget and Justification

9. REFERENCES

- [1] S. M. Witt and S. J. Young, "Phone-level pronunciation scoring and assessment for Interactive language learning," *Speech Communication*, vol. 30, no. 2–3, pp. 95–108, Feb. 2000. [Online].
- [2] Y. El Kheir, A. Ali, and S. A. Chowdhury, "Automatic pronunciation assessment – a review," in *Findings of the Assoc. for Computational Linguistics: EMNLP*, 2023, pp. 8304–8324.
- [3] Y. Getman, T. Grósz, M. Kurimo, and G. Salvi, "assessment challenge (NOCASA)," *arXiv*, Apr. 2025. "Non-native children's automatic speech
- [4] O. Grynszpan, P. L. Weiss, F. Perez-Diaz, and E. Gal, "Innovative technology-based interventions for autism spectrum disorders: A meta-analysis," *Autism*, vol. 18, no. 4, pp. 346–361, 2014.
- [5] D. M. Kagohara et al., "Using iPods® and iPads® in teaching programs for individuals with developmental disabilities: A systematic review," *Research in Developmental Disabilities*, vol. 34, no. 1, pp. 147–156, 2013.
- [6] J. Zhang, Z. Zhang, Y. Wang, Z. Yan, Q. Song, Y. Huang, K. Li, D. Povey, and Y. Wang, "speechocean762: An open-source non-native English speech corpus for pronunciation assessment," in *Proc. Interspeech 2021*, 2021.
- [7] H. Ryu, S. Kim, and M. Chung, "A joint model for pronunciation assessment and mispronunciation detection and diagnosis with multi-task learning," in *Proc. Interspeech 2023*, pp. 959–963, 2023.
- [8] B.-C. Yan, J.-T. Li, Y.-C. Wang, H.-W. Wang, T.-H. Lo, Y.-C. Hsu, W.-C. Chao, and B. Chen, "An effective pronunciation assessment approach leveraging hierarchical transformers and pre-training strategies," in *Proc. 62nd Annu. Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1737–1747, 2024.
- [9] P. M. Rogerson-Revell, "Computer-assisted pronunciation training (CAPT): Current issues and future directions," *RELC Journal*, vol. 52, pp. 189–205, 2021.
- [10] Y. Shen, A. Yasukagawa, D. Saito, N. Minematsu, and K. Saito, "Optimized prediction of fluency of L2 English based on interpretable network using quantity of phonation and quality of pronunciation," in *Proc. IEEE Spoken Language Technology Workshop (SLT)*, pp. 698–704, 2021.
- [11] P. W. S. Leung, S. X. Li, C. S. O. Tsang, B. L. C. Chow, and W. C. W. Wong, "Effectiveness of using mobile technology to improve cognitive and social skills among individuals with autism spectrum disorder: Systematic literature review," *JMIR Mental Health*, 2021.

[12] Y. Getman, T. Grósz, M. Kurimo, and G. Salvi, “Non-native children’s automatic speech assessment challenge (NOCASA),” in *2025 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2025.

10. APPENDICES

10.1 Risk Analysis

Risk	Description	Impact	Mitigation Strategy
Limited participant accessibility	Difficulty in obtaining sufficient participation from autistic children within the intended age group and school setting	High	Conduct early coordination with schools, special education teachers, and parents/guardians, and use a manageable pilot sample
Variability in speech patterns	Pronunciation attempts may vary due to differences in communication style, speech clarity, hesitation, or sensory factors	High	Use flexible scoring thresholds, repeated practice opportunities, and teacher-reviewed progress interpretation rather than relying only on a single attempt
Audio capture quality issues	Background noise, microphone limitations, and inconsistent recording environments may affect pronunciation analysis accuracy	High	Use noise-reduction preprocessing, basic recording quality checks, and testing in a quiet classroom environment
Pronunciation scoring inaccuracies	The system may not always accurately reflect the child’s actual pronunciation performance in early stages	Medium	Validate scoring logic with sample recordings, compare results with teacher observations, and refine similarity thresholds iteratively
Adaptive logic mismatch	The adaptive learning flow may recommend content that is too easy, too difficult, or not suitable for the learner’s current ability	Medium	Apply rule-based progression with teacher oversight, minimum exposure rules, and continuous adjustment based on learner performance
Usability and engagement issues	Some autistic children may face difficulty interacting with the interface due to sensory,	High	Design a simple and predictable UI with clear visuals, limited distractions, large touch targets, and reinforcement-based activities

	cognitive, or attention-related challenges		
Integration complexity	Mobile frontend, backend APIs, database, and audio processing modules may be difficult to integrate smoothly	Medium	Use modular system architecture, well-defined REST APIs, and incremental integration testing across components
Data privacy and child protection concerns	Student recordings, progress records, and personal learning data are sensitive and must be protected	High	Use authentication, secure storage, restricted access, anonymized datasets for analysis, and avoid unnecessary personal data collection
Teacher adoption challenges	Teachers may find the system difficult to use if the dashboard or workflow is too complex	Medium	Keep teacher-facing features simple, provide clear progress summaries, and conduct basic usability evaluation with teachers
Time and resource limitations	The scope of audio analysis, adaptive learning, and dashboard development may be difficult to complete within the project period	Medium	Prioritize core MVP features, use existing libraries and frameworks, and limit the first version to feasible functions

Table 4: Risk Analysis

10.2 Ethical Considerations and Consent Documentation

This research involves the design and development of an adaptive English pronunciation support component for Grade 3 children with] ASD. As the study involves children, student learning data, and recorded speech attempts collected during reinforcement-based activities, ethical considerations are essential throughout the research process to protect participants' rights, dignity, privacy, and well-being.

The study will be conducted with special attention to the needs of autistic children, ensuring that the research process and system interactions are safe, respectful, and non-harmful. Participation in the study will be entirely voluntary. Since the primary participants are children, informed consent will be obtained from parents or legal guardians before any data collection or system interaction takes place. Where appropriate, assent from the child will also be considered in a simple and understandable manner.

Participants and their guardians will be clearly informed about the purpose of the research, the type of data collected, how the data will be used, and the measures taken to ensure confidentiality. The study will avoid any form of coercion, and participants will have the right to withdraw from the research at any stage without any negative consequences.

The following ethical principles will be followed during the research:

- Participation in the study will be voluntary.

- Informed consent will be obtained from parents or legal guardians before involving any child in the study.
- Participants and guardians will be informed about the purpose and procedures of the research before data collection begins.
- Student identities will be kept confidential, and personally identifiable information will not be disclosed in the research findings.
- Audio recordings and performance records collected through the system will be used only for academic and research purposes.
- Only the minimum necessary data required for system evaluation and improvement will be collected.
- Participants will have the right to withdraw from the study at any time without providing a reason.
- The system design and evaluation process will avoid causing emotional distress, sensory discomfort, or pressure on the child.
- Teachers and parents/guardians will be able to provide observations if a child becomes uncomfortable during participation.
- All collected data will be securely stored and accessed only by the research team.

Special care will be taken when handling recorded speech data. Since pronunciation attempts may indirectly reveal a child's communication characteristics, recordings will be stored securely and, where possible, anonymized before analysis. Any presentation of results in the final report will use aggregated findings rather than identifiable student-level disclosure.

The research will also ensure that the proposed component is used as a supportive educational tool rather than a diagnostic or clinical assessment system. Performance outputs from the system will not be interpreted as a medical judgment, but as learning support indicators intended to assist teachers in monitoring progress and adapting instruction.