

# Let's Chat with LEARN! An AI-Powered Chatbot for Growing Young Communicators

Suryani Atan<sup>1\*</sup>, Aw Guat Poh<sup>2</sup>, Kartini Anwar<sup>3</sup>, Dharmaraj Sundararaj<sup>4</sup>, Dharani Kanna Ravi<sup>5</sup>,  
Lim Kheng Yong<sup>6</sup>, Muhammad Hidir Mohd Ezat<sup>7</sup>

<sup>1,2,3,4,5,6,7</sup> Asian Languages and Cultures (ALC), National Institute of Education (NIE)–Nanyang Technological University (NTU), Singapore, Singapore

\*Corresponding Author: suryani.atan@nie.edu.sg

## Abstract

*This paper investigates the enhancement of oral proficiency among seven- and eight-year-old Mother Tongue Language (MTL) learners in Singapore primary schools. Linguistic data with a focus on reading aloud and speaking competencies – including pronunciation accuracy, grammatical usage, and lexical diversity – were collected to inform the design of an AI-powered educational tool. These insights guided the development of LEARN (Language automated Evaluation by generating Answers/questions from caRtoons), a conversational chatbot co-developed by the National Institute of Education (NIE) – Nanyang Technological University (NTU), and the Singapore Institute of Technology (SIT). LEARN engages students in short, scaffolded dialogues based on picture description tasks, providing an interactive and supportive environment for practicing oral skills in their respective mother tongues. Grounded in educational scaffolding principles, LEARN facilitates gradual language development through iterative conversational loops. This paper outlines the chatbot's core features, pedagogical affordances, and its potential to enhance oral communication in early language education.*

## Keywords

AI-Powered Chatbot, Learning Loops, Malay Language, Mother Tongue Languages, Oral Proficiency

## INTRODUCTION

This project addresses the challenge of enriching the Mother Tongue Language (MTL) – Chinese, Malay, Tamil – learning environment at home for Primary 1 and Primary 2 students in Singapore, with the objective of improving oral proficiency by 20% over a two-year period. To achieve this, we propose the integration of AI-powered conversational chatbots that engage young learners in short, structured dialogues based on the Picture Description Task. These chatbots provide an interactive platform for practicing spoken MTL, fostering language development through meaningful exchanges. Our innovative approach comprises several key components: (i) Automated generation of age-appropriate questions and answers derived from cartoon images; (ii) Development of conversational AI chatbots for both interaction and data collection; (iii) Automated speech recognition tailored to young children who often speak a blend of English and MTL; (iv) Automatic picture captioning techniques to extract and align key speech topics with visual content. To support system training and refinement, data was collected from approximately 659 Primary 1 and 2 students across ten primary schools. This dataset includes recordings of students reading aloud and engaging in picture-based conversations, forming the basis for iterative improvements to the chatbot's capabilities.

## LITERATURE REVIEW

The integration of artificial intelligence (AI) in language education has gained significant momentum, particularly in

early childhood and primary education contexts. AI-powered chatbots have emerged as promising tools for facilitating language learning through interactive, personalized, and engaging experiences. In the area of AI chatbots in language learning, recent systematic reviews highlight the growing use of AI chatbots in language teaching and learning across various age groups and linguistic contexts [1]. These tools are increasingly recognized for their ability to simulate conversational practice, provide immediate feedback, and support learner autonomy. Chatbots have been employed to teach both English as a Second Language (ESL) and foreign languages, demonstrating positive outcomes in vocabulary acquisition, pronunciation, and learner motivation [1]. In early language education, AI tools – including chatbots and social robots – are often grounded in developmental psychology and human – computer interaction theories. These tools are designed to personalize language input and simulate peer-like interactions, which are particularly effective for young learners [2]. Studies show that personalization, scaffolding, and multimodal engagement (e.g., visual prompts, speech recognition) are key factors influencing the success of AI interventions in early language development [2]. In the area of pedagogical affordances and limitations, chatbots offer several pedagogical affordances, such as: (i) scaffolder learning through guided conversations; (ii) immediate feedback on pronunciation and grammar; (iii) low-pressure practice environments that reduce anxiety; and (iv) data-driven insights for educators to monitor progress. However, limitations persist. These include challenges in

accurately recognizing children's speech, especially when it involves code-switching or mixed-language utterances, and the need for culturally and linguistically appropriate content [3]. Moreover, the novelty effect of AI tools may diminish over time, necessitating longitudinal studies to assess sustained impact [1].

In the context of Mother Tongue language learning, while most existing studies focus on English or widely taught foreign languages, there is a growing need to explore AI applications in mother tongue languages such as Chinese, Malay, and Tamil. The **LEARN** chatbot project addresses this gap by tailoring conversational AI to the linguistic and cultural context of young Chinese, Malay, and Tamil Language learners in Singapore. By incorporating cartoon-based prompts and automated question generation, LEARN aligns with best practices in early language education and AI design.

### OBJECTIVES

This research project aims to address the challenge of enhancing the Mother Tongue Language (MTL) learning environment at home for Primary 1 and Primary 2 students in Singapore. The central goal is to improve students' oral proficiency in MTL by 20% over a two-year period using AI-powered conversational tools.

To achieve this, the project proposes the implementation of AI conversational chatbots that engage students in short, structured dialogues based on the Picture Description Task. These chatbots are designed to facilitate spoken exchanges in MTL, enabling students to practice and develop their oral language skills in a realistic, engaging, coherent, and meaningful manner.

The innovative approach comprises several key components. Together, these components aim to create a robust and adaptive learning system that supports oral language development in young learners through interactive, AI-driven conversations:

- 1) Automatic generation of age-appropriate questions and answers derived from cartoon images.
- 2) Development of conversational AI chatbots for both interaction and data collection.
- 3) Automated speech recognition tailored to young children who often speak a mixture of English and MTL.
- 4) Automatic picture captioning techniques to extract and align key speech topics with visual reference content.

### METHODS

#### Alignment of Objectives and Methodology

The primary objective of this research is addressed through the development and deployment of the LEARN system – an AI-powered chatbot platform designed to facilitate

meaningful, picture-based conversations in MTL. To achieve this, the methodology incorporates the following aligned components:

*Objective 1: Enable realistic, engaging, and coherent short conversations in MTL*

**Methodological Alignment:** The Dialogue Engine (WP1b) and Question & Answer Generation (WP1a) modules generate age-appropriate prompts from cartoon images and ensure natural language understanding (NLU) and generation (NLG) for coherent dialogue flow.

*Objective 2: Automate question and answer generation from visual prompts*

**Methodological Alignment:** The system uses Visual Question Generation (VQG) and Visual Question Answering (VQA) techniques to create dynamic, image-based interactions that stimulate learner engagement.

*Objective 3: Support spoken exchanges in mixed-language contexts (English + MTL)*

**Methodological Alignment:** The Speech Engine (WP2) integrates Automatic Speech Recognition (ASR) and a Translation Manager (TM) to handle multilingual input and output, accommodating the linguistic realities of young learners.

*Objective 4: Provide automated assessment and feedback on oral proficiency*

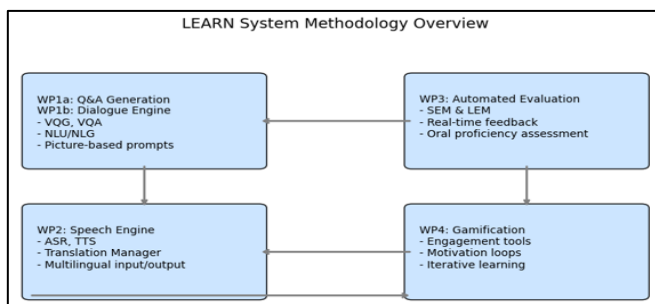
**Methodological Alignment:** The Automated Evaluation Engine (WP3) includes a Speech Evaluation Module (SEM) and a Language Evaluation Module (LEM) to assess pronunciation, grammar, and lexical diversity in real time.

*Objective 5: Sustain learner motivation through interactive design*

**Methodological Alignment:** The Gamification module (WP4) introduces game-like elements to maintain engagement and encourage repeated practice, supporting iterative learning loops.

*Objective 6: Train and refine the system using real student data*

**Methodological Alignment:** A dataset of 659 students from ten Singapore primary schools was collected in July 2023, including reading and picture conversation recordings. This data is used to train and fine-tune the AI models across all system components (Figure. 1).

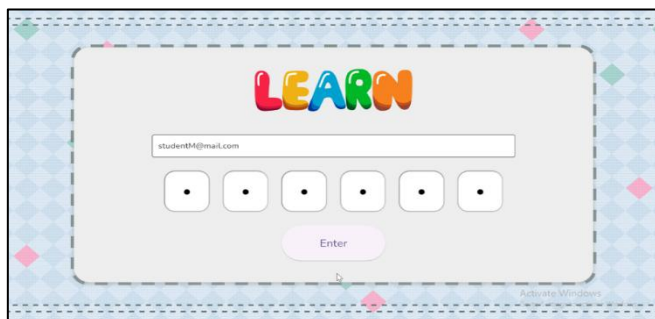


**Figure 1.** Components and Data Flow for AI-Powered Oral Proficiency Enhancement

## DEVELOPMENT OF THE LEARN PORTAL

### Description of the Portal

The LEARN portal is designed to facilitate oral language practice through interactive, cartoon-based scenarios. It supports three Mother Tongue languages – Chinese, Malay, and Tamil – providing multilingual flexibility for real-time conversations. By leveraging cartoon imagery, LEARN generates contextual dialogues that encourage both spontaneous and guided speech practice. The portal incorporates an adaptive scaffolding model that adjusts to varying student proficiency through tiered prompting and automated feedback. Gamification features are embedded to enhance learner motivation and promote sustained oral practice in home environments.



**Figure 2.** Homepage of LEARN Portal

Figure 2 illustrates the homepage of the LEARN portal. It is designed to be visually engaging and child-friendly, making it appealing for young learners:

- **Bright and Colourful Design:** The word "LEARN" is displayed in large, multicoloured letters at the top, immediately catching the eye and creating a playful tone.
- **Simple Login Interface:** It includes a clear email input field and six password/PIN boxes, making the login process straightforward.
- **Inviting "Enter" Button:** A large, pink "Enter" button adds a sense of fun and action.
- **Playful Background:** The light blue checkered background with small, colorful shapes enhances the cheerful and welcoming vibe.

This combination of vibrant visuals, simplicity, and playful elements encourages young users to engage with the platform confidently and enthusiastically.

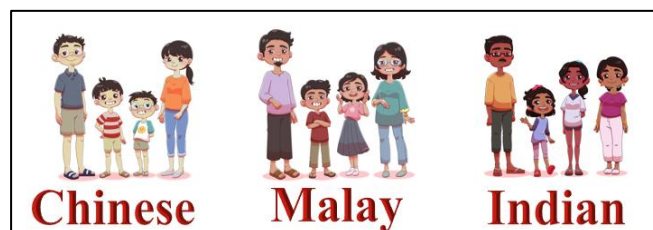
To reflect Singapore's multicultural society, LEARN features three representative families: a Chinese family, a Malay family, and an Indian family (see Figure. 3). These characters are embedded in cartoon-style scenarios depicting everyday life, serving two key purposes:

### Cultural Familiarity:

Students encounter culturally relevant content, fostering inclusivity and identity affirmation. The image showcases three cartoon-style families – Chinese, Malay, and Indian. This representation mirrors Singapore's multicultural society and helps students recognize and relate to familiar cultural identities. By embedding these characters in everyday scenarios, LEARN affirms students' backgrounds and promotes a sense of belonging.

### Visual Appeal:

The cartoon format is engaging and age-appropriate, increasing learner motivation and enjoyment. It is colourful, friendly, and age appropriate. It captures young learners' attention and makes the learning environment more inviting. The simplicity and warmth of the illustrations encourage engagement and make the platform feel approachable and fun.



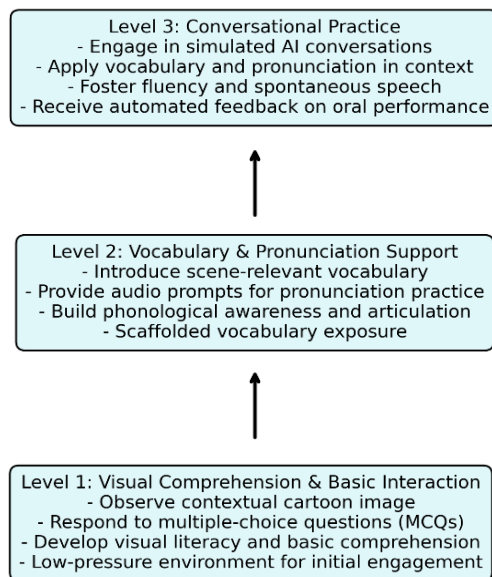
**Figure 3.** Multiracial Characters in LEARN

The accompanying image exemplifies LEARN's commitment to multicultural representation and visual engagement. By depicting Chinese, Malay, and Indian families in a cartoon format, it supports the platform in fostering cultural familiarity and learner motivation. These visuals serve as both educational tools and inclusive symbols, reinforcing the journal's emphasis on identity affirmation and age-appropriate design.

### Pedagogical Framework

LEARN is grounded in a structured pedagogical framework that scaffolds speaking skills through progressive learning stages. The system is divided into three levels, each designed to build upon the previous to support oral language development (see Figure. 4).

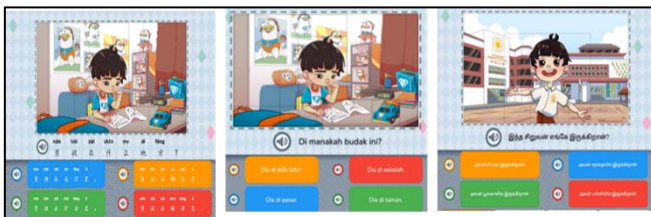




**Figure 4.** LEARN Learning Design and Scaffolding

#### Level 1: Visual Comprehension and Basic Interaction

Students begin by observing contextual images and responding to multiple-choice questions (MCQs). This stage promotes visual literacy and basic comprehension in a low-pressure setting (see Figure. 5).



**Figure 5.** Visual Comprehension and Basic Interaction

#### Level 2: Vocabulary and Pronunciation Support

Learners are introduced to scene-relevant vocabulary and receive pronunciation support via audio prompts. This level strengthens articulation and phonological awareness (see Figure. 6).



**Figure 6.** Vocabulary and Pronunciation Support

#### Level 3: Conversational Practice

Students engage in simulated conversations with the AI system (see Figure. 7), applying vocabulary and pronunciation skills in dynamic, communicative exchanges that foster fluency and spontaneity.



**Figure 7.** Malay Language Conversation in LEARN

The image in Figure. 7 effectively supports Level 3: Conversational Practice by providing a realistic and relatable scenario – a classroom cleaning activity – that students can discuss with the AI system. It aligns with the learning goals in the following ways:

#### 1) Contextual Vocabulary Use

Students are prompted to describe the scene using relevant vocabulary (e.g., *bilik darjah*, *mop*, *bersih*) – classroom, mop, clean – reinforcing word recognition and usage in context.

#### 2) Pronunciation Practice

The interface includes audio playback features, allowing learners to hear and repeat phrases, improving their pronunciation and listening skills.

#### 3) Interactive Dialogue

The AI initiates a conversation in Mother Tongue Language, for example, in Malay language, asking questions like "Bolehkah kamu beritahu saya di mana tempat ini?" (Can you tell me where this place is?), encouraging students to respond spontaneously, which fosters fluency and confidence.

#### 4) Visual Support

The cartoon-style illustration makes the activity engaging and age-appropriate, helping learners connect language with action visually.

In summary, Table 1 shows the focus, activities, purpose and support features that LEARN has to achieve the learning goals at each level.

LEARN exemplifies the potential of digital learning environments to systematically support the development of oral language proficiency while embedding principles of cultural inclusivity. Its pedagogical design – characterised by structured progression, multimodal input, and culturally contextualized content – aligns with contemporary approaches to language education that emphasize communicative competence and learner engagement. Recent research highlights how digital tools, when thoughtfully implemented, can empower learners by fostering agency, supporting plurilingual identities, and promoting inclusive

pedagogies [4]. Moreover, studies have shown that platforms incorporating voice, video, and interactive elements significantly enhance vocabulary acquisition and oral fluency, particularly in diverse cultural contexts [5]. Therefore, LEARN serves as a compelling model for

integrating technology into linguistically diverse classrooms, prompting educators to consider how digital tools can be leveraged not only for skill acquisition but also for affirming students' cultural identities.

**Table 1.** Summary of LEARN Portal Learning Levels

Level	Focus	Activities	Purpose	Support Features
<b>Level 1: Visual Comprehension &amp; Basic Interaction</b>	Foundational engagement through visual stimuli	Observe cartoon-style images; Respond to MCQs	Develop visual literacy and basic comprehension; Provide low-pressure entry into oral practice	Age-appropriate prompts; Immediate feedback
<b>Level 2: Vocabulary &amp; Pronunciation Support</b>	Language building through vocabulary and phonological practice	Introducing scene-relevant vocabulary; Practice pronunciation via audio prompts	Strengthen articulation and phonological awareness; Expand expressive language capacity	Adaptive scaffolding; Multilingual support
<b>Level 3: Conversational Practice</b>	Application of language skills in AI-driven dialogue	Engage in simulated conversations with chatbot	Foster fluency, spontaneity, and confidence; Encourage contextual use of vocabulary and grammar	Automated speech and language evaluation; Feedback loops

## RESULTS AND DISCUSSION

The participant pool consisted of 275 male and 384 female students. Among them, 326 were Primary 1 students and 333 were Primary 2 students. Language group breakdowns included 332 Chinese students, 266 Malay students, and 61 Tamil students. The initial trial phases of LEARN yielded

promising outcomes, particularly in its capacity to engage young learners and support oral language development. Feedback was collected from Primary 1 and Primary 2 students across three language streams – Chinese, Malay, and Tamil – highlighting high levels of enjoyment and perceived ease of use (see Table 2).

**Table 2.** Student Feedback

Language	Enjoyed Very Much (%)	Enjoyed (%)	Did Not Enjoy (%)	Very Easy (%)	Not Easy (%)
<b>Chinese</b>	53.6	40.2	6.2	78.2	21.8
<b>Malay</b>	66.7	27.4	5.9	82.8	17.2
<b>Tamil</b>	80.0	20.0	0.0	80.0	20.0

### Student Engagement and Enjoyment

Across all language groups, most students reported enjoying the gamification component of LEARN. Notably, 80% of Tamil language students selected “I enjoyed it very much!”, followed by 66.7% of Malay and 53.6% of Chinese language students. These findings suggest that LEARN’s interactive and gamified elements are effective in capturing students’ interest and sustaining engagement.

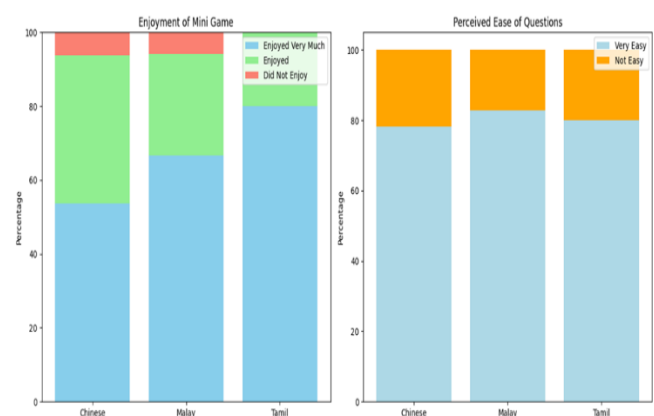
### Perceived Ease of Use

When asked whether the questions were easy, over 78% of students in each language group responded affirmatively. This reflects the effectiveness of LEARN’s scaffolded instructional design, which introduces vocabulary and conversational structures in a gradual and intuitive manner.

### Visual Representation of Feedback

To better illustrate these findings, the following chart (see Figure. 8) presents a comparative view of student responses

across the three language groups:



**Figure. 8.** Visual Representation of Feedback

### Interpretation of Results

The positive feedback can be attributed to several key design features of LEARN:

### *Scaffolded Learning Design*

LEARN's tiered structure ensures that learners are not overwhelmed. Each level builds upon the previous, allowing students to progress at a comfortable pace.

### *Multimodal Interaction*

By integrating speech, image, and text, LEARN creates a rich, immersive learning environment that mirrors real-life communication scenarios.

### *Cultural Relevance*

The inclusion of cartoon families representing Singapore's major ethnic groups adds familiarity and relatability, which likely contributes to students' enjoyment and motivation.

### **Broader Implications**

LEARN demonstrates how AI can be meaningfully integrated into early education environments, particularly for oral language development. The use of multimodal data – speech, image, and text – enables realistic and engaging conversations. Its design also supports scalable deployment across diverse linguistic and socio-cultural contexts, making it adaptable beyond the Singaporean classroom.

## **CONCLUSION**

The LEARN portal demonstrates strong potential in addressing the persistent challenges of oral language acquisition among young learners of Mother Tongue Languages (MTL). By integrating conversational AI, culturally resonant cartoon imagery, and gamified learning experiences, LEARN offers a dynamic and engaging platform for developing speaking skills. Its design is rooted in sound pedagogical principles – scaffolding, multimodal engagement, and contextual learning – that collectively foster structured yet enjoyable language practice.

Positive feedback from students across the Chinese, Malay, and Tamil language streams underscores LEARN's versatility and inclusivity. High levels of engagement, enjoyment, and perceived ease of use validate the system's instructional approach and affirm its relevance in diverse classroom settings.

Beyond its immediate educational impact, LEARN presents a scalable and adaptable solution for multilingual education. Its modular framework allows for seamless customization to suit varied linguistic and cultural contexts, making it suitable for global deployment. The integration of AI-driven dialogue and real-time feedback further positions LEARN as a forward-thinking educational technology with strong potential for both formal and informal learning environments.

In conclusion, LEARN is more than a digital tool. It is a transformative platform that reimagines how young learners

interact with language. By blending technology with culturally responsive pedagogy, LEARN paves the way for more inclusive, effective, and engaging language education in the 21st century.

### **Acknowledgements**

Thank you to the funding agency, AI Singapore (AI.SG), National Research Foundation, and to our research partner and host institution, Singapore Institute of Technology (SIT).

## **REFERENCES**

- [1]. Chua, H. W., & Annamalai, N. (2025). A bibliometric analysis and systematic review in AI chatbots in language teaching and learning. *International Journal of Technology in Education (IJTE)*, 8(2), 274–331. <https://doi.org/10.46328/ijte.1035>
- [2]. Sun, H., Tan, J., & Lim, M. C. (2025). AI and early language learning: A scoping review. *AI Brain Child*, 1(4). <https://doi.org/10.1007/s44436-025-00005-3>
- [3]. Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- [4]. Galante, A., Piccardo, E., Marcel, F., Zeaiter, L. F., Cruz, J. W., & Barise, A. (2024). Decolonizing language learning in digital environments through the voices of plurilingual learners in the Global South. *Applied Linguistics*, amae090, 1–20. <https://doi.org/10.1093/applin/amae090>
- [5]. Pikhart, M., Klimova, B., & Al-Obaydi, L. H. (2024). Exploring university students' preferences and satisfaction in utilizing digital tools for foreign language learning. *Frontiers in Education*, 9, 1412377. <https://doi.org/10.3389/feduc.2024.1412377>