Langmates: Revolutionizing Language Education

Abir Hossain[1], Jessi Jha[1], Mostafa Afr[1,2], Sebastian Tseng[1], Terhi Nurminen[1,3], Yajjat Sekhsaria[1,2]
[1]*Electrical and Systems Engineering,* [2]*Wharton School,* [3]*Cognitive Science Program*
University of Pennsylvania, Philadelphia, United States of America
{abirh, jessijha, mostafaa, sebtseng, terhi, yajjat}@seas.upenn.edu

***Abstract*—Langmates is an AI-powered language learning platform that brings real-time, adaptive speaking and listening practice into K–12 classrooms while keeping students actively engaged. Using OpenAI's Realtime API and multilingual transcription models, Langmates enables students to hold natural conversations with animated AI characters in more than 40 languages. Each interaction is tailored to student interests and proficiency levels, featuring dynamic character personalities, gamified elements, and progress tracking aligned with the standards of the American Council on the Teaching of Foreign Languages (ACTFL). Teachers can upload syllabi to shape conversation topics, monitor student performance, and integrate Langmates seamlessly into existing instruction. Designed with pedagogical best practices and developmental psychology in mind, the platform emphasizes fluency, contextual learning, and intrinsic motivation. Early pilots in multiple schools demonstrated high engagement from both students and educators, positioning Langmates as a scalable and impactful tool for modern language education-all the while keeping learnign fun.**

***Index Terms*—Language learning, LLMs, Edtech, US Schools, K-12, Artificial Intelligence, Education.**

I. Introduction

OVER the past two years, advances in artificial intelligence have begun to transform traditional educational methods. In particular, large language models (LLMs) have created new opportunities for enhancing how students practice and acquire foreign languages. While many EdTech companies targeting the consumer market, such as Duolingo and Babbel, have incorporated AI to improve individual learning experiences, there has been little adoption of AI-driven tools within the U.S. K-12 education system. This gap is particularly concerning given the persistent shortcomings in American language education: many students graduate after years of study without achieving conversational proficiency, and students' oral skills generally lag behind their reading and listening skills [1].

At the same time, national organizations such as the U.S. Department of Education and the Center for Applied Linguistics have called for a renewed focus on developing students' speaking and listening skills [14] [16], emphasizing that early and immersive practice is critical for long-term success. However, resource constraints—including limited classroom time, a shortage of qualified teachers, and lack of access to native speakers—make it difficult for schools to offer enough real-time conversational practice.

Funded by the Department of Electrical and Systems Engineering at the University of Pennsylvania, School and Engineering and Applied Science
Advised by Chris Callison-Burch, ccb@upenn.edu

A. Inspiration

As a team of passionate language learners, we have experienced firsthand both the rewards and the challenges of learning new languages. Reflecting on our own experiences within the U.S. education system, we recognized a widespread problem: despite years of study, many students struggle to hold a conversation in their target language. The heavy emphasis on rote memorization of vocabulary and grammar at the expense of authentic speaking practice still prevalent in many classrooms leaves students disengaged and unprepared for real-world communication.

We believe that language learning should be rooted in meaningful interaction and cultural appreciation. Successful language acquisition requires more than textbook exercises; it demands opportunities for immersive, motivating conversation, tailored to each learner's needs and interests. We each found immense value in our lives by connecting with people speaking languages we enjoyed, and we wanted to bring that value to every child in the country. This inspired us to create Langmates: a platform that makes language learning both effective and fun.

B. Problem Statement

Language learning in the United States faces a crisis of engagement, effectiveness, and enthusiasm. Despite years of instruction, 40% of Americans who studied a foreign language report that they could not navigate daily life in a country where the language is spoken [13]. Further, enrollment in college-level language courses has decreased significantly in the past ten years [15]. These statistics suggest that American students are not only struggling to succeed in language learning but are increasingly opting out of it altogether.

C. Our Solution

Langmates is an AI-powered platform designed to make language learning engaging and personalized. Instead of drilling vocabulary or memorizing grammar rules, students practice real-world communication through conversations with AI characters that adapt to their interests, proficiency level, and goals.

Using OpenAI's Realtime API, Langmates enables natural, low-latency conversations in over 40 languages. Students can choose from a variety of character voices and personalities, creating a playful, immersive experience that mirrors real-world dialogue. Built-in progress tracking, aligned with ACTFL standards, gives students and teachers clear milestones

to measure their improvement and helps maintain motivation over time.

Teachers can upload lesson plans to customize conversation topics and learning goals for each class and individual student. They also have access to student conversations and can monitor student progress through a dedicated dashboard, ensuring that Langmates fits seamlessly into classroom instruction. This makes it easy to incorporate Langmates into existing curricula: there's no need to create new exercises or spend large amounts of additional time grading the conversations.

To ensure its educational effectiveness, Langmates is built on pedagogical, linguistic, and neurodevelopmental best practices. The platform emphasizes frequent speaking practice, contextual learning, and positive reinforcement—strategies shown to support stronger language acquisition [1]. Designed in consultation with educators and learning scientists, Langmates focuses on building communicative competence and intrinsic motivation, helping students develop the fluency, confidence, and cultural understanding that traditional methods often fail to achieve. By combining personalization, gamification, and real-time feedback, we transform language learning into an experience students genuinely want to stick with.

II. Related Work

Foreign language instruction in U.S. K–12 schools has faced persistent challenges. While some schools offer robust programs, many struggle with limited instructional time, a shortage of qualified teachers, and a lack of immersive, conversation-based practice. At the same time, the rise of consumer-focused language learning apps has made vocabulary acquisition more accessible, but often falls short in developing speaking and listening skills. This section reviews both the institutional context of language learning in U.S. schools and the capabilities and limitations of popular EdTech tools, providing the foundation for understanding where Langmates fits within the broader language education landscape.

A. Language Learning in US Schools

A 2008 national survey by the Center for Applied Linguistics reported that only 25% of elementary schools offered any form of foreign language education, and just 6% offered immersion programs—considered essential for achieving conversational fluency. Further, these programs are unevenly distributed, often correlating with socioeconomic disparities across school districts [4].

In addition to access gaps, national trends show a declining emphasis on language learning. From 2016 to 2021, enrollment in college-level foreign language courses declined by 16.6% [15]. By contrast, 92% of European students study at least one foreign language in school, compared to just 20% in the U.S. These trends reflect both structural barriers and a broader disengagement from language learning in the American educational system.

Recognizing the long-term consequences of this gap, major government organizations such as the U.S. Department of Education, the U.S. Department of Defense, and the Committee for Economic Development have called for an education system that equips U.S. students with the ability to "communicate effectively in languages other than English" [4], while the Center for Applied Linguistics has called for earlier and more immersive language instruction [16]. Across the board, these bodies emphasize the importance of expanding foreign language offerings, increasing instructional time, and integrating culturally responsive content into language curricula [4] [16].

B. Popular Language Learning Tools

While digital platforms such as Duolingo and Babbel have made language learning more accessible to individuals, research suggests that their focus on vocabulary retention and continuous repetition has limited effectiveness in developing speaking skills. A study from Michigan State University found that Babbel users made significant gains in vocabulary and grammar after 3 months of use but showed only modest improvement in oral fluency, as measured by standardized speaking assessments [2].

Similarly, a study comparing Duolingo learners with students receiving direct instruction found that the latter significantly outperformed the app users in pronunciation accuracy. The German speech recognition technology the researchers used was able to identify only 17.3% of German words spoken by learners, compared to 49.7% accuracy for students who received classroom-based instruction [3]. These results highlight the limitations of app-based learning, particularly in fostering real-time verbal communication skills.

These gaps present an opportunity for more conversation-focused platforms to enter the space. Langmates addresses this need by delivering adaptive, interactive speaking practice through real-time AI conversations tailored to individual student proficiency. By focusing on oral proficiency, integration with classroom instruction, and accessibility for K–12 environments, Langmates offers a distinct alternative to consumer-oriented apps that emphasize memorization and repetition.

III. Design and Methodology

The development of Langmates followed a structured engineering design process to meet functional, educational, and technical requirements. Much of the innovation of Langmates lies in the system design element, as we strived to put together a system that would be fun and genuinely elevate language learning for students. This section outlines the specifications that guided the project, the constraints encountered, and the methodology used to design and refine the system, with each subsection discussing a different aspect of the design.

A. Pedagogic Considerations

Early in the design process, we consulted with experts in language acquisition and educational psychology to ensure that Langmates would support effective learning for school-aged users. Dr. Yuko Butler, a professor of linguistics at Penn with a focus in educational psychology and foreign language development in children, offered guidance on how children interact with language learning tools. We were initially worried that students could be missing on important body language

cues by learning from an AI companion. However, Dr. Butler affirmed that AI can be an effective language learning tool, even to the extent that some children might be more comfortable practicing their speaking skills with a virtual companion. Based on her input, we implemented features such as age-based interaction styles, adjustable correction frequency, and animated characters with eye contact-like behavior to support younger learners. Dr. Butler also advised us that younger students focus more on communicating meaning, rather than speaking with perfect grammar, leading us to modulate the way the AI interacts with different age groups more drastically than we had initially planned.

We also consulted with Dr. Amy Wrzesniewski, a psychologist and professor at the Wharton School, who emphasized the importance of student motivation and the psychological impact of visible progress. In response, we incorporated a progress bar and light gamification elements, both aimed at reinforcing a sense of achievement and keeping students engaged over time. These features are still in development, with achievement stickers, completion markers, and more detailed progress bars planned for future iterations.

In addition to external guidance, We applied insights from our own coursework in language acquisition, language processing, and human systems engineering to inform the user experience. In particular, the concept of a critical period in language learning was central to our design. While the precise age cutoff is still debated, research clearly shows that children and adolescents learn languages most effectively through exposure, acquiring both vocabulary and grammar implicitly rather than through formal instruction. This learning is most efficient when new material is embedded within familiar content—a principle known as comprehensible input. This informed our approach to conversational design, ensuring that interactions adapt dynamically to the student's proficiency level. To preserve attention and reduce cognitive load, we also limited AI response length and enabled students to interrupt the AI at any time, supporting more natural turn-taking and responsiveness, among other more subtle design choices for reducing learners' cognitive load.

B. Specifications and Constraints

The design of Langmates was guided by a set of key specifications aimed at creating a language learning platform that is effective, engaging, and suitable for deployment in K–12 classrooms. Our main focus areas were ensuring usability and conversation flow while keeping the product fun and secure to use, with a final goal of delivering a fully functional product capable of providing comprehensive listening and speaking practice in 10 languages through dynamic interactions with unique characters.

More specifically, to be an appropriate tool for modern classrooms, Langmates was required to support the most common foreign languages taught in US schools and provide useful feedback for students and teachers while keeping the user experience intuitive and aligned with data security practices. The minimal set of target languages was defined as those languages in which an Advanced Placement (AP) exam is currently offered[1] as well as the official UN languages not currently part of the AP program[2], providing access to as many learners as possible. For providing useful feedback, Langmates needed to be able to evaluate student performance in line with a standardized quantitative framework. Accessibility and data privacy specifications include compliance with appropriate legal requirements and design standards.

A key design priority was to create an engaging and intuitive learning experience, with the goal of improving educational outcomes through sustained student interaction with the platform. To achieve this, Langmates needed to be prompt-engineered to tailor its conversational responses to individual student interests, maintaining engagement through a variety of interactive activities. Students should also be able to select among different characters, each featuring distinct personalities, voices, and animations.

To properly support the development of oral expression and listening comprehension skills, conversations with Langmates needed to feel smooth, responsive, and natural. From a quantitative perspective, the system was designed to achieve a conversational latency of less than one second. On the qualitative side, the platform needed to deliver content with context-appropriate word and grammar choices, natural tone, and emotionally expressive responses, ensuring that conversations resembled authentic human interactions as closely as possible.

Several constraints shaped the development process. The technical specifications of the APIs used limited the number of languages we can support, though the number greatly exceeds our target. Latency presented another major technical challenge, as we are limited by the APIs' response and processing times as well as relay speed. Further, cost constraints related to API usage, particularly from OpenAI's Realtime API and ElevenLabs' voice synthesis service, necessitated efficient design choices. Legal and financial barriers associated with cybersecurity concerns and the licensing of recognizable characters also introduced additional considerations.

Due to ongoing latency challenges, full character voice functionality remains under development; however, animated character interactions have been successfully implemented. Overall, the core specifications for Langmates have been met or exceeded. In addition, several new specifications were introduced midway through the semester as we realigned the platform's development goals based on evolving stakeholder feedback and classroom testing results.

C. Design Procedure and Iterations

The design of Langmates followed a user-centered, iterative development approach, grounded in educational psychology research and informed by continuous feedback from learning researchers, students, teachers, and advisors. While specifications evolved throughout the year and the product underwent several major shifts in focus, the core vision—instilling a love of language learning in children through fun, engaging conversations with cartoon characters—remained constant. Overall,

[1]Chinese, French, German, Italian, Japanese, Spanish; excluding Latin
[2]Arabic, Russian

the design process balanced long-term planning and modular system architecture with responsiveness to rapid developments in AI technology and the feedback we received about classroom needs.

Langmates began as a program focused on delivering comprehensive real-time language tutoring through animated characters, with explicit language instruction embedded directly into the product. From the outset, the platform was intended for school-aged children, based on research showing that early immersion is critical both for developing fluency and fostering a lifelong love of language learning.

From the earliest parts of the fall semester, we didn't want to make Langmates in a vacuum, implementing the shiniest and newest features that we could think of as engineers. We contacted the University assisted schools through the Netter Center, as well as some schools we had personal contact with. We worked with these schools throughout the process, asking for their feedback on features we wanted to implement. In the end, we ended up working with seven schools in an attempt to build the best integrated system. While we sometimes would build a bit in one direction, such as with our idea of a good UI, we would often get input from teachers that made us realize we need to be building this in tandem with them, adding in the features they cared about.

Through interviews with teachers, we discovered that the greatest unmet need in classrooms was not explicit instruction, but rather opportunities for students to practice using the vocabulary and grammar they had already been taught to communicate real ideas. In response, we shifted Langmates' focus toward becoming a practice companion: a partner for students on their language learning journey, and a complement to the expertise of trained teachers. We also explored how the product could support ESL learners both in and outside the classroom, but ultimately concluded that while the need is significant, addressing it effectively would require a substantially different design than what Langmates is intended to offer.

During the fall semester, our primary focus was establishing the technical foundation for the platform. This included building a backend that could handle real-time conversation processing and integrate OpenAI's Realtime API and ElevenLabs' voice synthesis API into a single data pipeline. We also validated the LLM's multilingual capabilities and sought expert input in children's language acquisition to ensure consistent and high-quality conversational experiences across the languages offered on the platform. By the end of the semester, we had developed a minimalistic prototype of Langmates with the essential core features in place, although without significant system fine-tuning or user-facing polish.

The spring semester focused mainly on improving user experience and aligning the product with the practical and pedagogical needs of K–12 classrooms. Major updates included significant redesigns of the user interface to make the platform more intuitive and accessible for younger learners, the development of more granular system prompts to guide AI behavior, and the addition of new features such as session report generation and progress tracking. On the backend, we implemented a persistent database to store user conversations, enabling the platform to maintain memory across interactions and personalize the learning experience over time.

One major shift early in the spring was the decision to postpone full integration of voice synthesis. Due to significant latency challenges, real-time voice responses would have compromised the quality of user experience. Recognizing the rapid pace of advancement in AI tools, we chose to prioritize developing other features while monitoring the market for emerging voice services that better meet our latency and quality requirements. In exchange, we went through multiple iterations of our UI/UX where we completely revamped it across time to create an incredibly user-friendly design that meshes well with kids.

We have also expanded our testing into three schools this spring semester, through a paid pilot program. This has provided us valuable feedback on our system that we are working to improve on. One piece of feedback was adding an option to modulate the speed of speech for different learners. While we have this implicitly designed into our product based on the level of the student, a teacher wanted a method to do it manually so that she could have more control over her lessons. Additionally, we received some feedback on adding another layer of LLM checking for input to ensure that the language that is being spoken is correctly displayed. One issue we struggled with was the AI companion using the wrong script for languages that are similar, like Hindi and Urdu. Hindi uses the Sanskrit alphabet, whereas Urdu uses the Arabic one, and while the AI would usually catch on to what language you were speaking, the initial message may use the wrong script since it is guessing using the input words, since it is a transcription algorithm. We could add a restriction, likely using an additional check from an LLM that restricts the input text to only the target language and the students' native language, although this may remove some robustness for students that are bilingual (think comparing the target language to another language the student speaks such as comparing Spanish to Arabic).

Currently, Langmates is a functional speaking practice companion that can be used independently by students or integrated into classroom instruction by teachers. The product has gone through a number of iterations on both the frontend and backend, all of which were guided by the needs of students and educators. The next stages of product development will be guided by the results of our ongoing pilots but are likely to include enhancements to model memory and student progress tracking, the incorporation of additional gamification elements to sustain long-term engagement, and the expansion of character selection options to provide greater personalization.

D. System Architecture

Langmates is structured around a modular architecture designed for scalability, reliability, and performance, as shown in Figure 1. The system consists of three primary components: the frontend, the backend, and the data flow pipeline.

The frontend delivers a dynamic and responsive user experience for both students and teachers. It manages conversational interfaces, displays animated character interactions, and presents progress-tracking information through a visual

dashboard. The backend integrates OpenAI's Realtime API for multilingual conversation generation and ElevenLabs' API for voice synthesis, with voice features currently under refinement for future deployment. It is responsible for session management, language processing, user authentication, and ensuring data privacy. The backend also retrieves data from and saves it to Supabase databases to enable progress tracking and the utilization of conversational history to enhance current interactions.

In the data flow pipeline, student speech input is transcribed using a multilingual model and simultaneously sent to an LLM to generate context-appropriate responses. The LLM returns audio output, which is played alongside synchronized animations and a corresponding text transcript. Real-time updates to progress metrics are recorded and displayed for both students and teachers through the frontend dashboard. Conversations are stored in a database, enabling the creation of a personalized conversation history for each student. This history allows us to analyze speech patterns and generate long-term performance reports. It also supports the development of an AI "memory" that condenses each student's strengths, weaknesses, and personal preferences, allowing conversations to feel increasingly personalized over time. We are using our pilot deployments to gather at least a month's worth of data per student, which will enable us to implement these features in a way that robustly draws on long-term interaction history.

In addition to the main data pipeline, two other major functions rely on LLMs: the generation of session reports and the integration of teacher-uploaded syllabi into conversations. At any point, users can choose to generate a report of a conversation, which sends the full dialogue through OpenAI's API with formatting instructions and returns the information needed to produce a standardized summary. Similarly, when a teacher uploads a curriculum document, it is processed by an LLM that extracts key instructional content, which is then incorporated into the student's conversation prompts.

This modular structure ensures that Langmates can evolve with advancements in AI and speech technologies, while maintaining flexibility for future feature expansion and classroom customization.

IV. Implementation

The implementation phase of Langmates focused on translating the system design into a functional, scalable, and user-friendly platform. Emphasis was placed on creating a robust web application, integrating real-time conversational AI capabilities tailored for the pedagogic needs of K-12 classrooms, developing engaging frontend interactions, and establishing reliable systems for progress tracking and curriculum customization. Although initial plans included full character voice integration, the current build prioritizes a low-latency conversation experience enhanced by animation features, while voice synthesis remains under refinement.

1) Backend: The backend architecture of Langmates was built to support low-latency, real-time multilingual conversations while maintaining scalability and data security. The primary conversational engine is powered by OpenAI's Realtime API, which has the capability to directly interpret multiple languages in a single audio input, making it optimal for language learners who might code-switch mid sentence. The Realtime API then generates natural language responses in line with the instructions we provide in system prompts, and handles the flow of the conversation throughout the session. The conversations are transcribed by Open AI's Whisper software and displayed to the user along with the audio responses. The backend also manages session states for individual users, tracks student interactions, and maintains secure records of learning progress, including session duration, conversation content, and vocabulary milestones.

Integration with ElevenLabs' voice synthesis API was explored to generate character-specific voice outputs. However, due to ongoing latency and cost challenges, full deployment of synthesized character voices remains under development. Current system functionality prioritizes delivering smooth spoken dialogue with the default voices provided by the OpenAI Realtime API.

To optimize performance, backend services employ efficient request handling and minimize API call redundancies. Session data and user profiles are stored securely using Supabase services, ensuring compliance with data privacy standards such as FERPA and COPPA.

2) Frontend: The Langmates frontend was developed using Next.js and React, leveraging a modular component structure for scalability and maintainability. The ShadCN UI library was selected to provide a clean, accessible, and consistent visual design across both student and teacher interfaces.

For students, the frontend delivers an interactive language learning environment where the user engages in real-time conversation with animated, AI-driven characters. The animation system is built using lightweight, web-optimized libraries that synchronize facial movements and gestures to match AI-generated dialogue text. Feedback elements, including progress bars and session summaries, are embedded into the interface to provide immediate feedback and encourage sustained engagement.

We are currently in the process of developing a dedicated teacher-facing dashboard, with a working version of the features integrated directly into the chat interface. This allows teachers to upload lesson plans and instructional materials through a sidebar panel and generate progress reports. These lesson plans are parsed on the backend to customize conversation topics and vocabulary sets dynamically, while the report is created through analyzing and condensing the student conversation using an LLM. This initial integration lays the groundwork for a more robust set of tools to support instructional alignment and progress monitoring in future iterations.

As a browser-based application, performance and usability were major considerations. Through server-side rendering and static site generation, Next.js accelerates initial page loads, ensures a consistent user experience, and also enhances search engine optimization (SEO). Further, its integrated routing system and automatic code splitting optimize application efficiency. As a result, Langmates is fully functional on basic devices commonly used in K–12 classrooms.

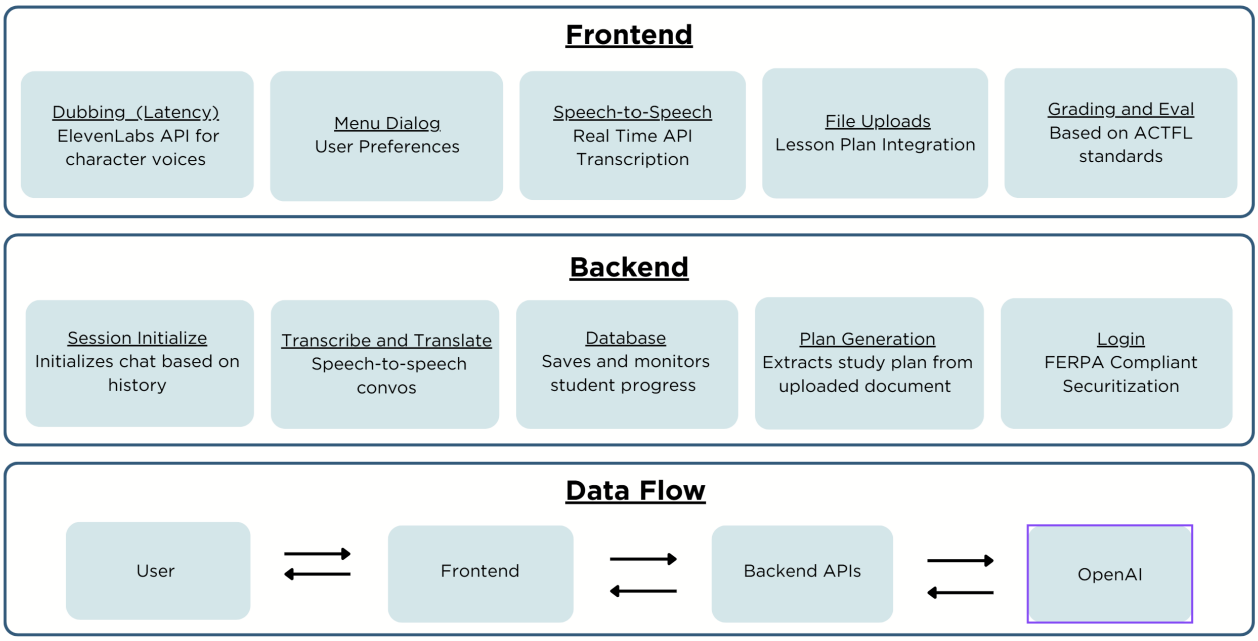

Fig. 1: System Design Flow Chart

A. *Deployment Environment*

Langmates was deployed using a cloud-native architecture to ensure high availability, scalability, and low-latency performance. The frontend is hosted on Vercel, providing global CDN distribution for fast load times and optimized content delivery. Backend services and database storage are hosted on Supabase for flexible scaling based on user demand.

All AI model requests and user session data are routed through secured endpoints with encrypted communication to maintain compliance with data protection regulations. Cloud infrastructure was chosen specifically to eliminate the need for specialized local hardware, making Langmates easily deployable in a wide range of school technology environments. End users require only a standard web-connected device with a microphone, such as a laptop, tablet, or Chromebook, to access the full functionality of the platform.

B. *User Interface and User Experience (UI/UX)*

User experience design for Langmates focused on maximizing student engagement while maintaining accessibility and ease of use. The student-facing interface utilizes intuitive visual cues that encourage exploration without overcrowding the layout, while the educator side focuses on structured, well-organized layouts that provide an easy overview of student progress. Color schemes, typography, and visual elements were selected to meet WCAG 2.1 accessibility standards, ensuring that the platform is accessible to users with a wide range of needs.

Key UI/UX features implemented include:

- **Animated Characters:** Characters visually respond to student inputs with facial expressions and gestures synchronized to dialogue content. This visual feedback makes the conversational experience feel more human and immersive, even before full voice capabilities are added.
- **Progress Tracking:** A visible progress bar and achievement system allow students to track their language practice over time. Metrics such as session length, frequency of conversation, and vocabulary milestones are stored to reinforce positive learning habits.
- **Customizable Conversations:** Settings allow students to adjust conversation difficulty based on their proficiency level, providing differentiated support for elementary, middle, and high school learners.
- **Intelligent AI responses:** Langmates stores all conversations the student has, allowing the AI to learn about the student and their interests. This way, every interaction is uniquely tailored for the specific student. The program is also tuned to detect signs of disengagement, with built-in mini-games and other activities to suggest to the student to keep them practicing their language skills.

The ShadCN UI library was selected to implement the UI for its minimalist design philosophy, as its emphasis on clarity and functional elegance aligns with our design goals. The library provides a collection of well-crafted and easily navigable components that maintain a clean aesthetic while providing comprehensive functionality.

The conversational experience is shaped by carefully engineered system prompts that guide interactions toward educational and age-appropriate content. Langmates adapts to each student's interests, encouraging them to elaborate and engage in extended dialogue. When relevant, it may suggest vocabulary-focused activities from its internal game library to keep conversations varied and purposeful. If a student attempts to steer the conversation off-topic, the system gently redirects

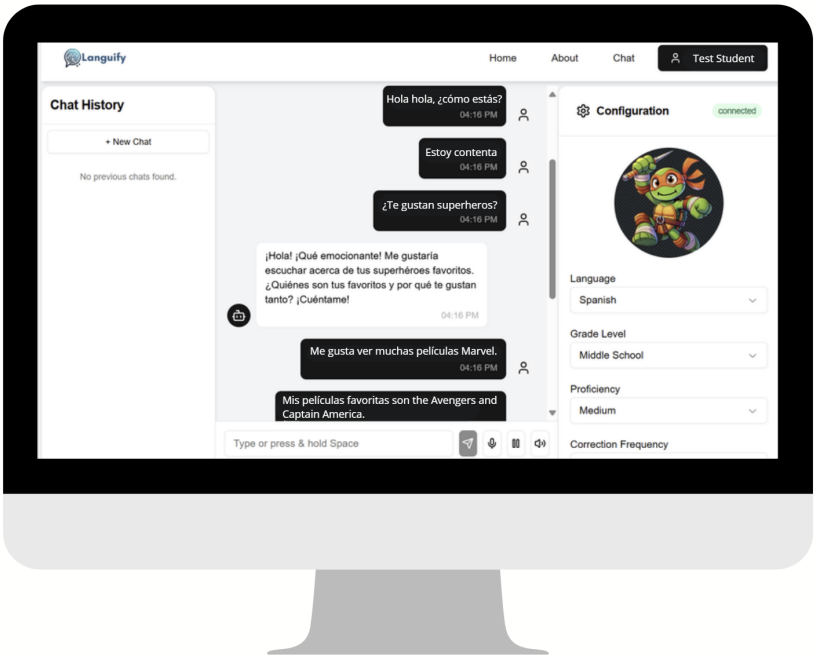

Fig. 2: User Interface Sample

them, and built-in moderation tools prevent engagement with inappropriate prompts. In the case of concerning disclosures, Langmates is designed to encourage the student to speak with a trusted adult and automatically notify the teacher of the interaction.

Langmates incorporates accessibility features such as keyboard navigability, high-contrast visual modes, and screen reader compatibility to ensure inclusivity across a wide range of user needs. Usability testing was conducted with both university staff and high school students to evaluate system responsiveness, clarity of interface, and overall engagement. Feedback from these sessions directly informed design refinements, including improvements to layout structure, positioning of key controls, and phrasing of instructional prompts to ensure that the platform is as intuitive as functional.

V. Testing and Verification

Langmates was evaluated through a combination of internal testing and school-based pilot trials to assess language learning outcomes, usability, and overall user experience. Feedback from students, educators, and advisors provided critical input that shaped the refinement of the platform.

A. Preliminary Feedback

After the completion of our first prototype back in the fall, several students and staff from the University of Pennsylvania were asked to test and rate the chatbot on Language Accuracy, Phonetics, Realism, Engagement, and Ease of Use. Feedback across all five categories was largely positive. Most participants found the tutor highly realistic and accurate in conversation. However, several recommendations were made to enhance the chatbot's UI controls, making it more visually appealing and user-friendly.

B. High School Student Experiences

Following improvements to the UI, customization features, and chatbot conversational capabilities, we began evaluating the effectiveness and user experience of Langmates in a school setting. This feedback provided valuable insight into how the system performs in real-world educational environments.

A group of 13 high school students participated in a hands-on trial of the Langmates AI tutor. Students were asked to evaluate specific aspects of their experience:

- **Engagement**: How pleasant conversations are with the chatbot and the extent to which the user would want to keep using it.
- **Ease of use**: The overall simplicity and navigability of the UI and speech functions.
- **Naturalness**: The degree to which the tutor's responses feel like a legitimate human-to-human conversation.
- **Effectiveness**: How much a student feels that he/she genuinely learned and is able to speak the learned language.

The results were consistently positive, indicating strong potential for Langmates to support language development through meaningful, adaptive conversation. Many students highlighted the ability to interact with different character voices, the natural tone of conversations, and how well the system adapted to varying fluency levels.

For example, one student proficient in Bengali set the chatbot to high school level with high proficiency. Langmates responded primarily in Bengali, producing fluent, natural dialogue. The student described the experience as "talking to his own family in their native tongue."

At the other end of the spectrum, a beginner-level Portuguese student configured the chatbot to elementary level with low proficiency. In this case, Langmates adopted a nurturing tone, guiding the student through basic vocabulary

Overall Progress

The student is showing a positive attitude towards learning Arabic and is making incremental progress in vocabulary acquisition. Continued practice and exposure will aid in developing more complex language skills.

Additional Notes:

The student was playful and engaged, showing interest in learning through fun methods like mimicking Mickey Mouse. Encourage this creative approach to maintain enthusiasm. No tech issues were noted.

ACTFL Category Scores

Domain	Score	Level	Comment
Domain	Score	Level	Comment
Text Type	3	Novice High	The student used simple phrases and repeated requests effectively.
Comprehensibility	3	Novice High	The student was mostly clear but required repetition for understanding.
Language Control	2	Novice Mid	Basic control with frequent repetition needed for clarity.
Vocabulary Use	3	Novice High	Limited range, focused on a specific word but persistent in learning.
Listening Comprehension	3	Novice High	Understood basic responses, needed clarification for specific terms.
Communication Strategies	4	Intermediate Low	Used repetition and clarification requests to maintain conversation.
Fluency	2	Novice Mid	Hesitant, with frequent pauses and repetition.
Cohesion & Discourse	2	Novice Mid	Limited to short exchanges, little use of cohesive devices.

Fig. 3: Sample Generated Feedback Report

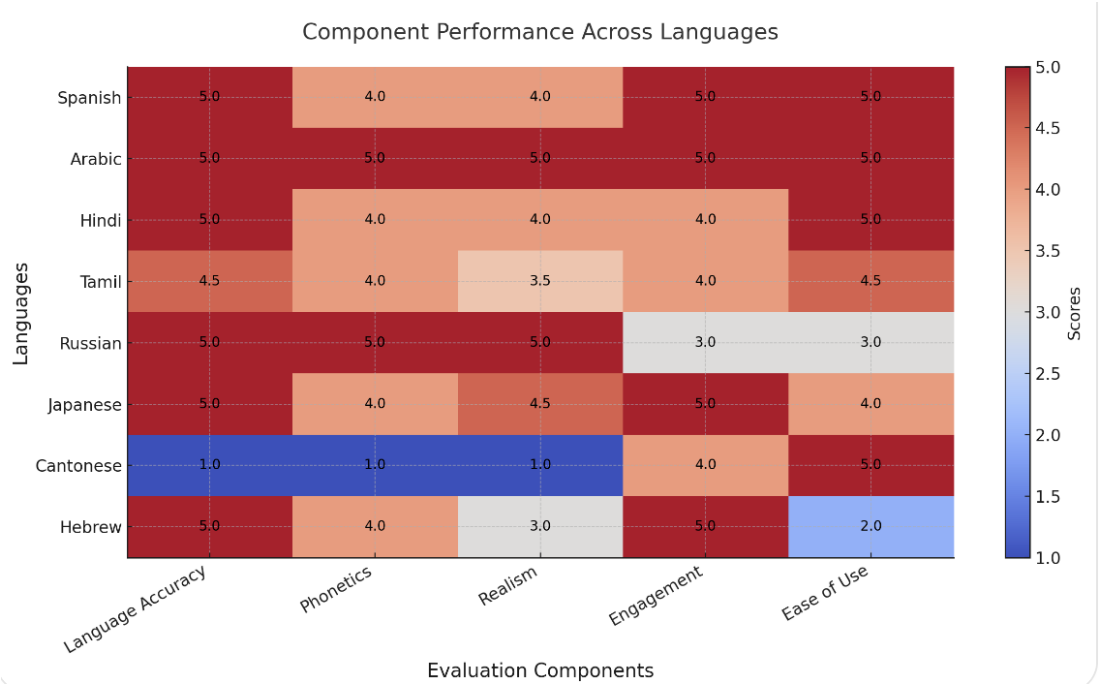

Fig. 4: Feedback from participants by language (scale of 1–5)

with patience and encouragement. The student noted that this approach felt more effective than traditional learning methods.

C. Educator Feedback

Educator feedback emphasized the importance of simple integration, instructional flexibility, and a robust user experience. A Spanish and French teacher, who works with students across a range of proficiency levels, tested Langmates and expressed enthusiasm about its potential—particularly the ability to tailor conversations to different grade levels.

One of the most requested features was a file upload option that would allow teachers to integrate existing teaching materials directly into the platform. Many educators possess a wealth of resources accumulated over years of experience, and the ability to align AI interactions with their current curricula was viewed as essential.

Ease of use was also consistently highlighted as a key priority. For instructors less familiar with AI-based tools, a streamlined interface with intuitive, single-click functionality was considered critical. Progress tracking tools—such as vi-

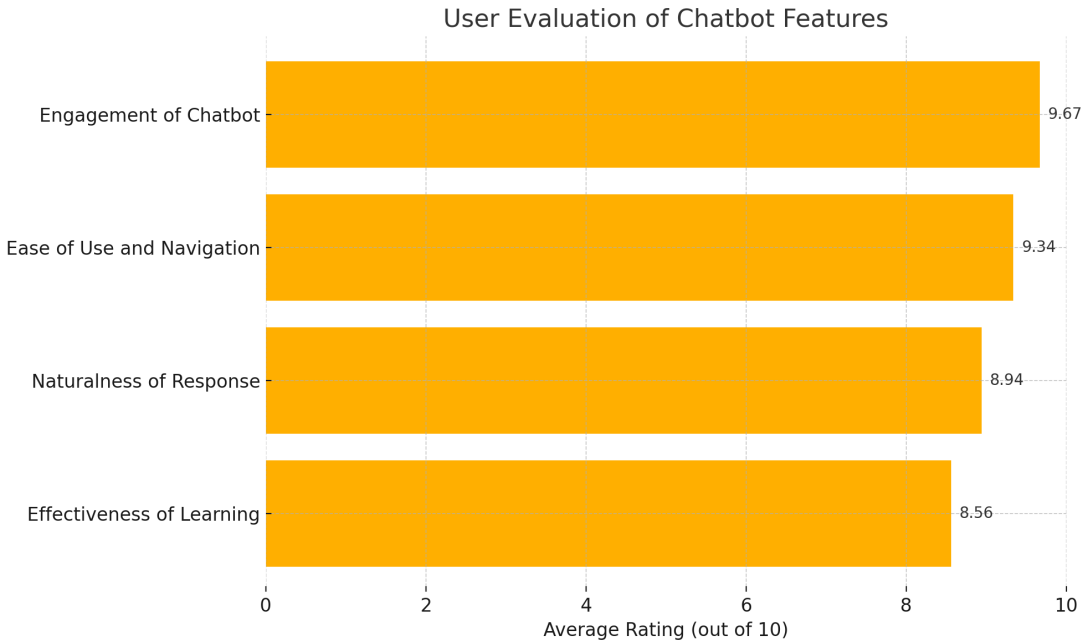

Fig. 5: Feedback from participants (scale of 1–5)

sual indicators, time logs, and conversation summaries—were identified as especially valuable for assessing student learning and adjusting instruction accordingly.

One teacher spoke particularly to the lack of speaking practice outside of class time, which he sees as the most effective way to internalize a new language. He noted that Langmates fills this gap, offering students the ability to practice spoken language in a low-stakes, conversational format. The flexibility to serve multiple grade levels was also highlighted as a major strength of the platform.

As described in the Design and Methodology section, we launched paid pilot programs in three schools as of April. These pilots have provided valuable feedback that continues to shape development. Teachers have praised the language model's ability to adapt fluidly to conversations, while students often exceeded expected usage time—validating our dual goals of educational value and engagement. While we do not yet have hard data on learning outcomes such as test scores, we anticipate these results as pilots progress.

Feedback has also helped surface areas for refinement and edge cases we had not initially anticipated. Teachers suggested adjustments to how grade reports are presented, refinements to control options, and general improvements to teacher-facing tools. Some feedback also highlighted technical challenges, such as handling pronunciation errors that are currently difficult for the system to detect. These insights are helping guide improvements for future iterations of Langmates.

VI. Applicable Standards

Langmates was developed in alignment with a range of engineering, educational, data privacy, accessibility, and software development standards to ensure a reliable, ethical, and effective learning platform. In particular, the platform's use of artificial intelligence adheres to IEEE P7001 guidelines for transparency and ISO/IEC 23894 standards for AI risk management, minimizing bias and ensuring responsible AI behavior. The ISTE Standards for Technology in Education further shaped the design by promoting equitable and personalized learning experiences.

Data protection and user privacy were prioritized in compliance with FERPA and COPPA, safeguarding sensitive student information and ensuring appropriate consent procedures. ISO/IEC 27001 principles for information security management informed backend development, and best practices from the KidSAFE Seal Program were incorporated to promote safe interaction for younger users.

Accessibility was addressed through adherence to WCAG 2.1 guidelines and Section 508 of the Rehabilitation Act, ensuring that Langmates is usable by students with disabilities through screen reader compatibility, keyboard navigation, and accessible design features.

The educational framework of the platform reflects IMS Global LTI standards for learning tool interoperability and SCORM standards for sharable content, facilitating integration with school learning management systems. IEEE Learning Technology Standards guided instructional content structuring, while the ACTFL Proficiency Guidelines informed the platform's student progress tracking and reporting systems.

User interface design followed ISO 9241 ergonomic principles, Nielsen's Usability Heuristics, and ISO/IEC 29110 standards for child-centered design, ensuring that the system is intuitive, engaging, and developmentally appropriate. The software development process was structured according to ISO/IEC 25010 for quality requirements, IEEE 829 for testing documentation, and ISO/IEC 12207 for software life cycle management.

By grounding its design in these standards, Langmates ensures a platform that meets high expectations for security,

accessibility, educational impact, and technical reliability.

VII. Ethical Considerations

The development of Langmates raised several important ethical considerations centered around user privacy, data security, responsible use of AI, and protection of minors. From the outset, the platform was designed to comply with FERPA and COPPA regulations, ensuring that all student data is securely stored, access-controlled, and collected with appropriate consent for users under the age of 13. Data transmission and storage practices follow recognized information security standards to protect sensitive personal information.

Transparency was another priority. Langmates clearly identifies that users are interacting with an AI-powered system, helping to prevent misunderstandings about the nature of conversations. This transparency is especially important in educational settings, where trust and informed participation are critical.

Potential biases in AI-generated responses are another serious risk. While LLMs cannot be guaranteed to never produce unwanted outputs, the prompts are carefully tuned to control system behaviors and restrict conversational domains to those appropriate for school-aged children. Language outputs are monitored to ensure that conversations remain appropriate, inclusive, and aligned with educational goals.

Finally, the platform was designed with special consideration for its youngest users. Interfaces, conversational topics, and engagement features were built to be developmentally appropriate, promoting safe and positive language learning experiences for children and adolescents.

By addressing these ethical challenges early and consistently throughout the design process, Langmates aims to provide a trustworthy, safe, and supportive environment for students and educators alike.

VIII. Environmental, Social, and Economic Impact

The deployment of Langmates brings important environmental, social, and economic considerations. On the environmental side, the use of large AI models and cloud infrastructure contributes to energy consumption and associated carbon emissions. To mitigate this, cloud hosting providers with commitments to renewable energy were prioritized and backend systems were designed to minimize unnecessary API calls and data processing. Future improvements will continue to focus on optimizing computational efficiency and reducing the platform's overall environmental footprint.

Socially, Langmates has the potential to expand access to high-quality language education, particularly for students in under-resourced schools where access to native speakers and personalized instruction is limited. By making interactive, personalized language practice broadly available, the platform supports greater equity in educational opportunities. In addition, exposure to multiple languages and cultural contexts through conversational practice encourages cross-cultural understanding, empathy, and global citizenship among students.

Economically, Langmates offers a scalable alternative to traditional language instruction models, reducing the need for costly tutoring or supplemental programs. The platform's cloud-based design allows it to serve large numbers of students at low marginal cost, making it a practical solution for schools and districts with limited budgets.

By addressing environmental sustainability, promoting educational equity, and offering a cost-effective solution, Langmates aims to create positive and lasting impacts beyond the immediate goals of language learning.

IX. Conclusions and Reflections

Over the course of the year, our team designed and implemented Langmates, a web-based platform that brings realtime, personalized language conversation practice into classrooms. We successfully built a functional system that enables students to engage with AI-driven animated characters in over 40 languages and receive feedback on their progress. Teachers have the capability to upload syllabi and exercises into the system to align their students' interactions with classroom goals and activities, making it easy to integrate Langmates into existing courses. Continuous testing with students and teachers, culminating in a month-long pilot test in school environments, demonstrated high student engagement and satisfaction with the user experience.

The positive feedback we received from our test users, along with the launch of paid pilot programs, is a testament to the demand for innovative language instruction. These results reflect both the value of integrating AI meaningfully into the classroom and the enthusiasm students bring to language learning when instruction is interactive, personalized, and fun. Our experience confirmed that when students are given space to explore language through dynamic conversation—rather than rigid memorization—they become more engaged, confident, and motivated learners.

For us, this project offered valuable lessons in balancing technical ambition with real-world usability. From licensing constraints to latency concerns, several features we originally envisioned as core to the product—such as full character voice integration—had to be scaled back or postponed to maintain a high-quality user experience. At the same time, the rapid pace of AI tool development reshaped our technical roadmap. After spending a month in the beginning trying to engineer a custom solution for the realtime responses from the AI, OpenAI released a brand new API that provided that functionality out of the box, allowing us to redirect out focus toward other features almost overnight. Time and again we came upon situations like this, where we switched APIs and rebuilt our project in a differnt form nearly overnight. Due to the nature of AI, there were many times throughout the project we had to move flexibly in fast-moving technical landscapes, where adaptability can matter more than upfront completeness.

In addition to the diverse technical and subject matter expertise of our team, one of the most important contributors to our success was how well we worked together. We trusted one another to take ownership of deliverables, assigned tasks based on each other's strengths, and created an environment where we not only produced high-quality work but also genuinely enjoyed the process. As a team, we did our best

to accommodate one another and offer grace when it was needed—and in return, we could rely on each other to step up when someone had a rough midterm week or a family commitment. Although we did not always meet every deadline or check every box we had hoped to, we continued to pursue ideas that excited us and allowed the project to become what it is today. The difference between a project that struggles and one that thrives is often not just the idea or the code, but the trust and mutual respect among the people building it.

Our product also greatly benefited from the feedback and perspectives of classmates, professors, advisors, and test users in our partner schools. When working closely on a project, it's easy to become so immersed in the internal logic and technical architecture that you lose sight of the user's experience. External feedback reminded us to prioritize supporting features the users wanted rather than what we thought was important. For example, a more robust approach to remembering the user's conversations was given priority over the voice changing functionality due to feedback from teachers, and the current implementation of the mute functionality emerged directly from classroom testing sessions.

At the same time, we found that building a good product requires balancing responsiveness to feedback with a commitment to thoughtful design decisions. While we received suggestions to expand Langmates for English as a Second Language (ESL) learners, we ultimately recognized that our platform, as it stands, is not sufficient to address the needs of these students. Similarly, many of our adult mentors asked for features like explicit grammar correction during conversation, but we decided to prioritize modeling correct language use instead—aligning with educational research that shows learners benefit more from hearing corrected forms in context than from direct interruption as per our advice from Dr. Butler. These decisions reflect the importance of maintaining ownership over the product's vision, even as we remain open to continuous input and improvement.

Our goal was to build a product that adapts to the needs of individual students while remaining scalable and flexible enough for diverse classroom environments. But the broader vision extends beyond simply teaching languages: we aspire to make language learning in the United States both engaging and equitable. We believe in harnessing AI to bridge cultural and educational gaps through affordable, culturally informed learning environments—just as previous generations of online resources have transformed access to information. It has been exciting to contribute to this emerging space within language instruction, and we hope that the continued democratization of access will spark a lifelong passion for language learning and foster cross-cultural understanding among students across the United States and beyond.

We are currently exploring the transformation of Langmates into a startup, building on the momentum and technical progress made over the past year. As we make this transition, our priorities include ensuring full legal compliance with intellectual property and data privacy regulations, as well as strengthening the system's stability and reliability for long-term use. On the feature side, we aim to validate our progress tracking tools at scale, refine model behavior through continued prompt engineering, and implement subtle interface improvements to enhance usability. Perhaps most importantly, we are now moving beyond the exploratory phase of development-the time in the sandbox is over. As we enter more formal relationships with schools, we must define a clear product promise and deliver a system that is predictable, well-scoped, and aligned with the structured needs of educational institutions.

Looking back, this has been quite a year. For a team and a project that came together during the first week of class—without an existing idea or early-stage prototype—Langmates reflects what the senior design course can offer at its best: the space to build something meaningful from the ground up, shaped by collaboration, iteration, and ambition. [NEED FINAL SENTENCE]

X. Abridged Business Writeup

While we have a longer, more formal business plan and writeup, we will detail the basic strategy here.

1) Executive Summary: Langmates transforms traditional language education into an engaging, scalable enterprise solution by enabling students to converse with AI in the personas of beloved characters. This immersive approach addresses the engagement gap in foreign-language programs and integrates seamlessly with existing curricula, providing real-time progress tracking and data-driven insights for educators. From a business standpoint, our dual revenue model—focused on per-learner enterprise subscriptions with a secondary consumer offering—ensures recurring, stable cash flows while positioning Langmates as a leading contender in the burgeoning EdTech market. [7]

2) Value Proposition:

- **Enhanced Engagement:** Students practice with familiar, entertaining characters, making language learning fun and memorable.
- **Personalized Learning:** Educators tailor conversation topics by uploading existing syllabi, aligning AI interactions with classroom objectives.
- **Equitable Access:** Affordable, scalable pricing bridges resource disparities, granting all schools access to high-quality language practice.
- **Comprehensive Assessment:** Real-time analytics empower teachers to monitor progress and intervene where needed.
- **Institutional Scalability:** A per-learner subscription model simplifies budgeting and drives predictable, recurring revenue.

3) Stakeholders:

- **Students:** Primary beneficiaries who gain authentic speaking and listening practice in a risk-free, gamified setting.
- **Parents:** Participants in home learning see measurable gains in children's language skills alongside entertaining, character-driven interactions.
- **Educators:** Teachers leverage customizable content and robust tracking to deliver differentiated instruction at scale.

- **Educational Institutions:** Schools and districts modernize language programs affordably, improving outcomes and equity.
- **Technology Partners:** Collaborations with AI and cloud providers ensure platform reliability and continuous innovation.

4) Market Research: The U.S. EdTech sector was valued at $80.27 billion in 2023 with an 11%+ CAGR projected through 2028. [7] Globalization and workforce demands are increasing the need for multilingual skills, while under-resourced schools seek cost-effective, scalable solutions to close the language-learning gap. Early pilots in three Philadelphia schools and the Arizona Cultural Academy have demonstrated strong engagement and learning gains, validating school-district interest and partnership potential. Additionally, the North American online language-learning market is projected to reach $9.78 billion by 2031 (19.3% CAGR). [8] Moreover, adoption of generative AI in EdTech grew by 15% in 2023. [9]

5) Target Customer Segment:

- **Enterprise (Primary):** Public and private K–12 schools and districts purchasing per-learner subscriptions at approximately $200 per student per academic year (or $20/month), yielding predictable, high-margin revenue.
- **Individual Learners (Secondary):** Language enthusiasts subscribing at $24 per month, representing a high-volume but more competitive segment that we will explore selectively.

6) Competition: Key competitors include Duolingo [10], known for consumer gamification but lacking classroom integration, and ISSEN [11], which offers personalized AI tutoring without strong motivation mechanics. Traditional substitutes (private tutoring and teacher-training programs) are effective but expensive and difficult to scale. A Porter's Five Forces analysis indicates low threat of new entrants (high regulatory and training costs), moderate buyer power (tight school budgets mitigated by high switching costs), moderate supplier power (reliance on AI and voice-synthesis providers), and moderate threat of substitutes, with overall low-to-moderate competitive rivalry in the AI-for-schools segment.

7) Cost: Current AI-voice synthesis costs are approximately $0.12 per minute, [12] with projected reductions to $0.04 per minute over the next year. Assuming an average usage of two hours per week, this results in an annual cost of roughly $96 per user. Cloud infrastructure and licensing leverage economies of scale, further driving down per-user expenses.

8) Revenue Model:

- **Enterprise Subscriptions:** $200 per student per academic year (or $25/month) for schools, yielding $30,000 annual revenue and $15,600 gross profit per average private school (150 users).
- **Consumer Subscriptions:** $20 per month per individual user, capturing the broader online language-learning market on a cautious, secondary basis.

9) IP: Langmates' intellectual property portfolio will encompass:

- **Patents:** Applications for our real-time voice modulation techniques and dynamic syllabus-integration methods.
- **Trademarks:** "Langmates" brand identity and character-persona names.
- **Copyrights:** Software codebase, character animations, and curriculum content.
- **Trade Secrets:** Proprietary AI fine-tuning processes and data-driven assessment algorithms.

ACKNOWLEDGMENTS

We would like to begin by expressing our sincere gratitude to our advisor, Professor Chris Callison-Burch, for his invaluable insights. His guidance on incorporating AI into education and navigating the ethical considerations of this work was instrumental in shaping the direction and depth of our project. We also thank Professors Deliwala and Van der Spiegel for their instruction and support throughout the year.

A special shout-out goes to our test users—both at the University of Pennsylvania and at our partner pilot schools—whose thoughtful feedback helped us refine Langmates into a more effective and engaging tool. Their participation was essential to validating our platform in real-world educational settings.

Additionally, we thank the School of Engineering and Applied Science at the University of Pennsylvania for bringing our team together and providing the funding, resources, and infrastructure that made this project possible.

Finally, we are honored to have received the Norman Gross Senior Design Project Award from the Department of Electrical and Systems Engineering. This recognition affirms our belief in the importance of accessible, engaging language education and reflects the collaborative effort that went into building Langmates.

REFERENCES

[1] K. van Gorp and M. Coss, "What proficiency levels do K–16 world language learners achieve?," *American Council for Teaching Foreign Languages*, n.d. [Online]. Available: https://www.actfl.org/research/research-briefs/proficiency-levels-for-k16. [Accessed: Apr. 30, 2025].

[2] S. Loewen, D. Isbell, and Z. Sporn, "The effectiveness of app-based language instruction for developing receptive linguistic knowledge and oral communicative ability," *Foreign Language Annals*, vol. 53, no. 2, pp. 189–206, Jun. 2020, doi: 10.1111/flan.12454.

[3] C. Taylor and C.-L. Huang, "How effective is Duolingo at promoting implicit pronunciation learning?" *English Australia Journal*, 2024. [Online]. Available: https://www.englishaustralia.com.au/documents/item/2403.

[4] I. Pufahl and N. C. Rhodes, "Foreign language instruction in U.S. schools: Results of a national survey of elementary and secondary schools," *Foreign Language Annals*, vol. 44, no. 2, pp. 258–288, 2011, doi: 10.1111/j.1944-9720.2011.01130.x.

[5] L. Barrow and L. Markman-Pithers, "Supporting young English learners in the United States," *The Future of Children*, vol. 26, no. 2, pp. 159-183, 2016. Available: http://www.jstor.org/stable/43940586. [Accessed: Dec. 12, 2024].

[6] K. Sánchez, "New York City is trying to support immigrant students and asylum seekers, but resources are scarce," *Chalkbeat*, May 9, 2023. [Online]. Available: https://www.chalkbeat.org/newyork/2023/5/9/23716167/nyc-immigrant-students-asylum-seekers-support-english-learners. [Accessed: Dec. 12, 2024].

[7] GlobalData, "United States (US) EdTech Market, 2023–2028," *US EdTech Market Size, Share, Trends, Report Forecast to 2028*, Jan. 2024. [Online]. Available: https://www.globaldata.com/store/report/usa-edtech-market-analysis/. [Accessed: Apr. 30, 2025].

[8] GlobeNewswire, “North America Online Language Learning Market Projected to Reach $9.78 Billion by 2031,” *GlobeNewswire*, Dec. 6, 2024. [Online]. Available: https://www.globenewswire.com/news-release/2024/12/06/2993114/28124/en/North-America-Online-Language-Learning-Market-Analysis-by-Learning-Mode-Age-Group-Language-End-User-and-Country-Forecast-to-2031.html. [Accessed: Apr. 30, 2025].

[9] Market.US, “Generative AI in Edtech Market Size, Trends — CAGR of 41%,” Market.US, 2024. [Online]. Available: https://market.us/report/generative-ai-in-edtech-market/. [Accessed: Apr. 30, 2025].

[10] Duolingo Inc., “Duolingo,” Duolingo, 2025. [Online]. Available: https://www.duolingo.com. [Accessed: Apr. 30, 2025].

[11] ISSEN, “ISSEN AI Tutoring Platform,” ISSEN, 2025. [Online]. Available: https://issen.ai. [Accessed: Apr. 30, 2025].

[12] ElevenLabs, “ElevenLabs Pricing,” ElevenLabs, 2025. [Online]. Available: https://elevenlabs.io/pricing. [Accessed: Apr. 30, 2025].

[13] N. Mykhalevych, “Foreign Language Education Statistics: Second Languages in the US,” *Preply Blog*, Oct. 22, 2024. [Online]. Available: https://preply.com/en/blog/americans-foreign-language-survey/. [Accessed: Apr. 30, 2025].

[14] S. Morrison, “Resources for Elementary School Foreign Language Programs,” CAL Resource Guides Online, n.d. [Online], Available https://www.cal.org/twi/rgos/fles.html#search. [Accessed: April 30, 2025]

[15] N. Lusin, T. Peterson, C. Sulewski, and R. Zafer, “Enrollments in Languages Other Than English in US Institutions of Higher Education, Fall 2021” Modern Language Association, May 2023. [Online]. Available: https://www.mla.org/content/download/191324/file/Enrollments-in-Languages-Other-Than-English-in-US-Institutions-of-Higher-Education-Fall-2021.pdf. [Accessed: Apr. 30, 2025].

[16] ACTFL, AATG, AATI, AATSP, ACL/APA, ACTR, CLASS/CLTA, and NCSTJ/ATJ, “Standards for Foreign Language Learning,” n.d. Available: https://www.actfl.org/uploads/files/general/Documents/AppendixAStandardsforFLLexecsumm.pdf. [Accessed: April 30, 2025]