Linguist AI - CS491 Senior Project

Specification Report

Group T2314:

- Yağız Can Aslan 22001943
- Kardelen Ceren 22003017
- Selim Can Güler 22002811
- İlkim Elif Kervan 22002223
- Tolga Özgün 22003850

Supervisor: Halil Altay Güvenir

Innovation Expert: Ahmet Eren Başak

Project URL: https://linguistai.app

Specification Report

November 17, 2023

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the CS491 course.

Table of Contents

1. Introduction 1
1.1. Description 1
1.2. High-Level System Architecture & Components of Proposed Solution 2
1.3. Constraints 3
1.3.1. Implementation Constraints 3
1.3.2. Economic Constraints 4
1.3.3. Ethical Constraints 4
1.3.4. Maintainability Constraints 4
1.3.5. Language Constraints 4
1.3.6. Privacy and Security Constraints 5
1.4. Professional and Ethical Issues 5
1.5. Standards 5
2. Design Requirements 10
2.1. Functional Requirements 10
2.1.1. User 10
2.1.2. Chatbot 10
2.2. Non-Functional Requirements 11
2.2.1. Usability 11
2.2.2. Reliability 11
2.2.3. Performance 12
2.2.4. Supportability 12
2.2.5. Scalability 12
2.2.6. Security & Privacy 13
3. Market Analysis 14
3.1. Target Users 14
3.2. Market Research 14
3.2.1. Learning Modes 14
3.2.2. End Users 14
3.2.3. Emerging Trends 14
3.2.4. Our Position in the Market 14
3.3. Competitive Analysis 15
3.3.1. Language Learning Competitors 15
3.3.2. AI Language Learning Competitors 15
4. Academic Analysis 16
4.1. Large Language Models 16
4.2. AI in Language Learning 16
4.3. Ethics of Language Learning Models 18
5. Mockups 19
References 25

1. Introduction

English is one of the most spoken languages in the world, with millions of native and non-native speakers [1]. It has become the default language of international events, businesses, technology, etc. Therefore, many people aim to learn English. For instance, in the Turkish education system, students are taught English as a second language for ten years starting from elementary school. That said, most students cannot speak English properly due to the ineffective traditional teaching methods, including grammar and vocabulary memorization using textbooks and flashcards. Students cannot incorporate the language into their everyday lives using the aforementioned methods and, therefore, forget the memorized rules of the language. Also, there is a lack of resources when it comes to writing and speaking materials, as the main focus of language teaching is based on grammar and vocabulary. However, language is not something to be memorized, it is a culture that should be experienced.

In addition to the traditional language learning techniques, there are applications that students can use to learn English, like Cambly and Duolingo. Duolingo is a mobile application that offers lessons for various skill levels [2]. These lessons aim to teach grammar and vocabulary using different activities such as filling in the blanks, translating sentences, speaking, and listening. On the other hand, Cambly is an online platform designed to connect students with native speakers [3]. While using Cambly, students have a chance to interact with a native speaker during the classes. However, the lessons in Duolingo are not customized to individuals and might get boring after a while. Also, they are designed for beginner and intermediate learners. Therefore, someone who wants to speak more advanced or learn academic vocabulary cannot use Duolingo. As for Cambly, it allows students to experience the language culture. However, it might not be affordable for all students. Also, students with self-esteem problems and shyness would be uncomfortable using this app.

LinguistAI is an online mobile application where users can learn vocabulary and practice their language skills using a chatbot. Our application detects the skill level of the user during the chat sessions and adapts to the user's level. Users can have a conversation with the bot about a topic or act on a scenario like buying a coffee or plane ticket. Our application will include gamification elements such as achievements, leaderboards, and daily streaks. All in all, Linguist AI will enable users to learn English tailored to their level by eliminating artificial learning items such as flashcards and textbooks.

The rest of this report will describe LinguistAI, its architecture and technologies, project constraints, functional and non-functional requirements, market research, academic analysis, and mockups.

1.1. Description

LinguistAI aims to help language learners practice vocabulary as well as writing and speaking skills via a mobile application featuring an AI chatbot. Learners will be expected to have basic English skills, reaching at least A2. Unlike other applications mentioned in the introduction,

LinguistAI will detect the user's skill level through vocabulary and grammar. Afterwards, the bot will adapt to the detected level and respond to the user in an easily understandable way while also integrating unfamiliar vocabulary. Users will be able to select the words they do not know and then view their meaning and usage in different contexts. Furthermore, the bot will incorporate the selected unknown words into its response messages so that the user can comprehend the new vocabulary easily. In addition to these, users will be able to speak to the bot. LinguistAI will detect the user's words using an external API and create a relevant response. Also, the chatbot can respond with audio if the user requests, enhancing the user's listening abilities. Furthermore, while chatting, the bot will create a user profile, including the user's demographic data, their likes and dislikes, and suggest new topics that might interest the user. The standout feature of LinguistAI is its ability to engage with users in a conversational format, allowing it to seamlessly track and maintain the flow of conversation for continued responses.

1.2. High-Level System Architecture & Components of Proposed Solution

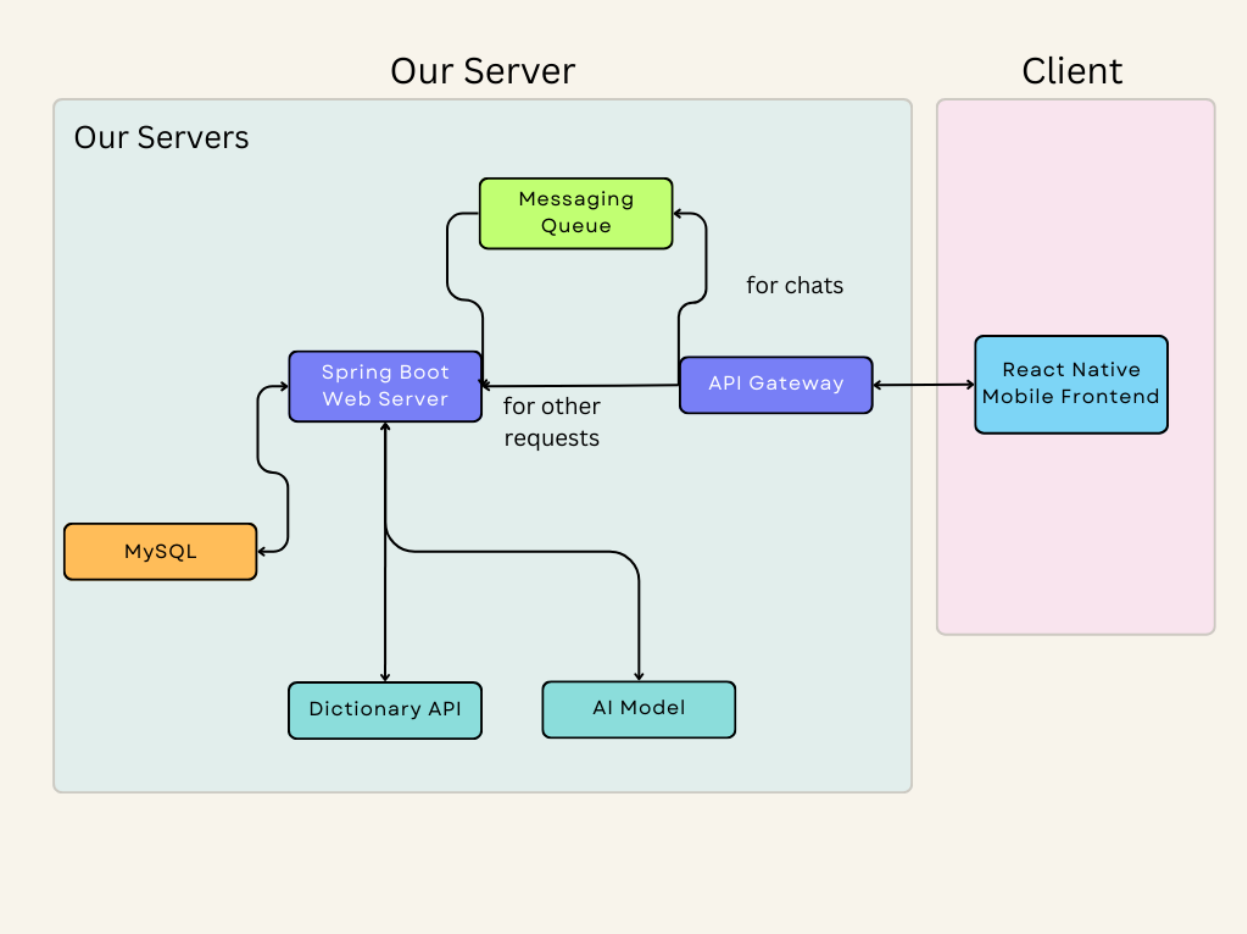

Fig. 1. LinguistAI High-Level System Architecture

We have adopted the microservice pattern as we will have many services in our application, such as the dictionary service and machine learning (ML) service. We will utilize Java Spring Boot for our API Gateway and one of the web services responsible for chat and user-related requests. To facilitate asynchronous communication between the bot and users in chats, we decided to use messaging queues. Furthermore, after careful consideration, we decided to use MySQL as our database for chats and user-related information. Also, we will be using TypeScript for the dictionary service and Python for the ML service. Lastly, React Native will be used for mobile development.

1.3. Constraints

Constraints refer to the specific limitations and conditions that influence the development and operation of the LinguistAI project. These constraints can be categorized into several key areas, each with its own unique set of requirements and considerations. In this section, a detailed overview of these constraints will be covered under Implementation Constraints, Economic Constraints, Ethical Constraints, Maintainability Constraints, Language Constraints, Privacy and Safety Constraints.

1.3.1. Implementation Constraints

- Github and Jira will be used to maintain the development process of our project.
- React Native will be used as a framework to develop our application for the most common mobile operating systems, such as iOS and Android, with a single codebase.
- Expo will be used to maintain the project architecture for the client side.
- React Native Platform will be used to create responsive layouts for different device sizes and to create layouts that are adaptive to landscape and portrait modes.
- React Navigation will be used to create Stack-based routing and bottom tab routing and navigation for the entire application.
- Axios will be used to send requests over the Internet and intercept them to prevent unwanted behavior when needed.
- Expo ImagePicker will be used when users want to add a profile picture to their profiles.
- Expo Notification will be used to push notifications to the customer.
- Backend will be developed using Spring Boot with Java.
- Slf4j will be used for logging server activity.
- Relational MySQL database will be the main point of storing data.
- Amazon AWS S3 will be used as a secondary database for non-relational data such as files.
- Java Persistence API (JPA) and Hibernate ORM will be used to ease database CRUD operations.
- The backend will be deployed to a Virtual Desktop Server serving Ubuntu hosted at GoDaddy [6].
- Pre-trained Large Language Models (LLMs) such as Llama2, GPT-4, Gemini will be fine-tuned with publicly available conversational datasets to fit the context of the application.
- New machine learning models will be trained with publicly available datasets for tasks such as text filtering, text analysis and text summarization.
- Large Language Models (LLMs) will be deployed on a remote Ubuntu server, running with respective quantization methods to support CPU processes.

- LLM services will be served leveraging microservice architecture, and a Flask application will be connected to an API Gateway.
- The API Gateway will also be hosted in the same Ubuntu machine and will be run on Spring Boot architecture, handling the redirection of API calls and authentication.

1.3.2. Economic Constraints

- Backend and database hosting will be free as it is stored in GoDaddy where the service has been previously paid until the end of 2025.
- Training and deploying machine learning models will be the biggest cost of this project at this point; however, an exact cost could not be calculated due to not knowing how much training is required and potential cost reductions with brand deals and sponsorships.
- For the release of the mobile application on the Apple App Store for iOS devices, creating a new developer account for LinguistAI will cost a yearly subscription.
- For the release of the mobile application on the Google Play Store for Android devices, creating a new developer account for LinguistAI will cost a one-time fee.
- Purchasing a domain will have a cost.

1.3.3. Ethical Constraints

- In order to combat the hallucinations, i.e., inaccurate, fictional, or misleading information generated by the model, and dangerous text outputs such as inappropriate content, harmful instructions, or biased and discriminatory language, methods such as prompt engineering, fine-tuning various text-analysis models and creating an ensemble model will be used.
- User inputs will be sanitized with a similar approach given to hallucinations and dangerous text outputs.

1.3.4. Maintainability Constraints

- The system will be maintained on a regular basis.
- Any issues and bugs will be fixed in the next updates.
- Users will be able to report issues and bugs.
- Application crashes will be tracked via specialized tools and will be reported with the users' consent.

1.3.5. Language Constraints

- The application will support English as a learning language.
- The mobile application interface will initially support English; however, it may support more languages in the later stages of development.
- The application users are expected to know English in a conversational fluency or A2 level in CEFR rating.

1.3.6. Privacy and Security Constraints

- The user is expected to provide a valid email, username, and password in order to use the application.

- All messages sent by the user will be stored in the application servers; therefore, the necessary policies will be provided to the users with their informed consent.
- Users will be able to download their stored data with respect to data privacy laws.
- Users will be able to remove their accounts in compliance with data privacy laws.
- Users will be provided with Terms of Service and similar privacy information before signing up for the application.
- The solutions for storing user data will be secured up to industry standards, such as the passwords will be stored with SHA-256 bit encryption and salting procedures.

1.4. Professional and Ethical Issues

- Transparency will be maintained in how user data is collected, stored, and used within the application.
- All users will be informed on which data is collected, and their informed consent will be held.
- While developing the machine learning models, the models will be monitored so as not to exhibit bias or offensive behavior.
- Transparency will be maintained in the process on how the training is conducted on the models, as well as what steps have been taken to ensure ethical use of the models.
- Both internal and external testing will be conducted on as diverse groups as possible to ensure user accommodation.

1.5. Standards

We use industry-proven and widely integrated methodologies, tools, and software in our development process. In all parts of our tech stack, floating point numbers are stored based on the IEEE 754 standard. To track project related technical and non-technical tasks we utilize Jira, Confluence and GitHub issues. As our version control system we use Git, the most widely used version control tool in the industry. We host our repositories on GitHub. Our development team adheres to 2-week-long Scrum cycles, where we hold regular “stand up” meetings to discuss our progress, problems and plans. We hold Sprint review, retrospective and planning meetings at the end of each Sprint. Moreover, in order to streamline our development process, we have set up tailored pull request and commit conventions (see the figures below for the Confluence pages of the conventions).

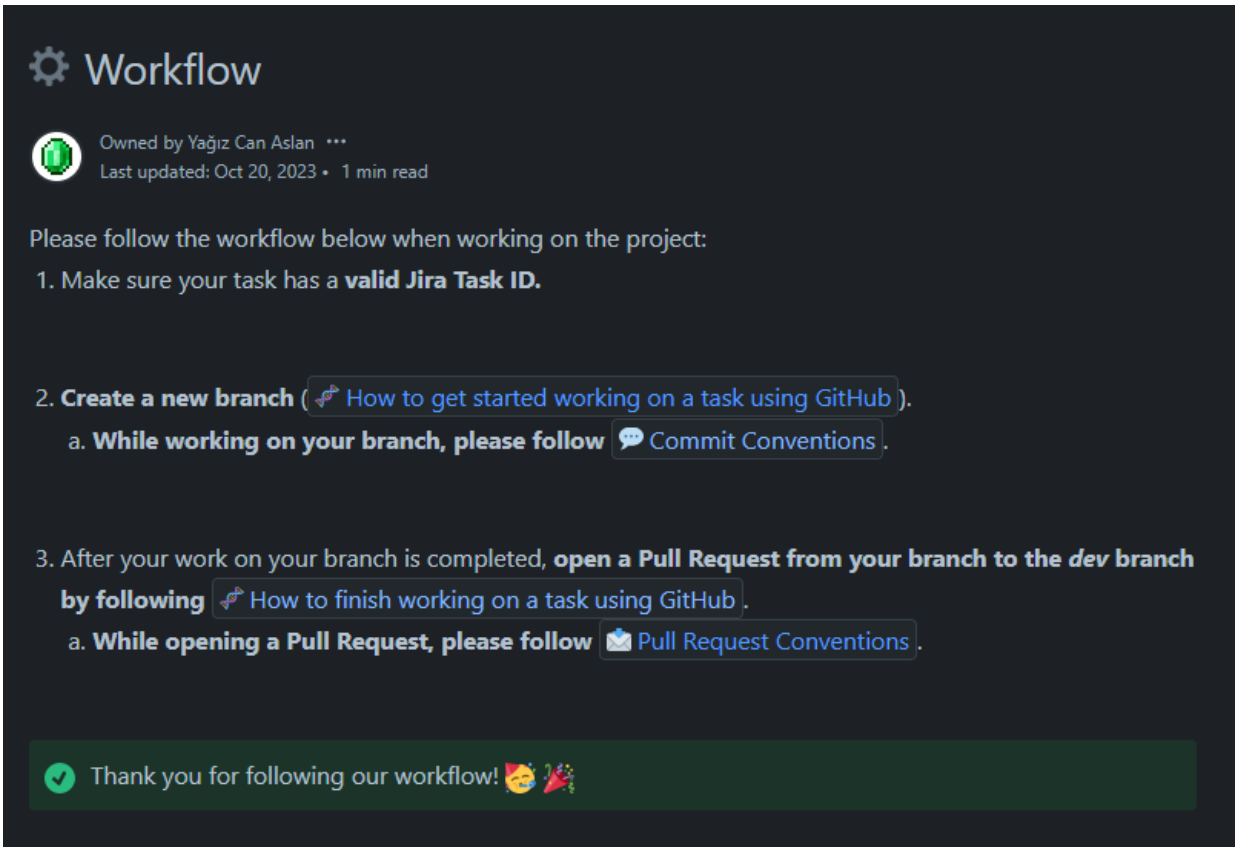

Fig. 2. LinguistAI Workflow

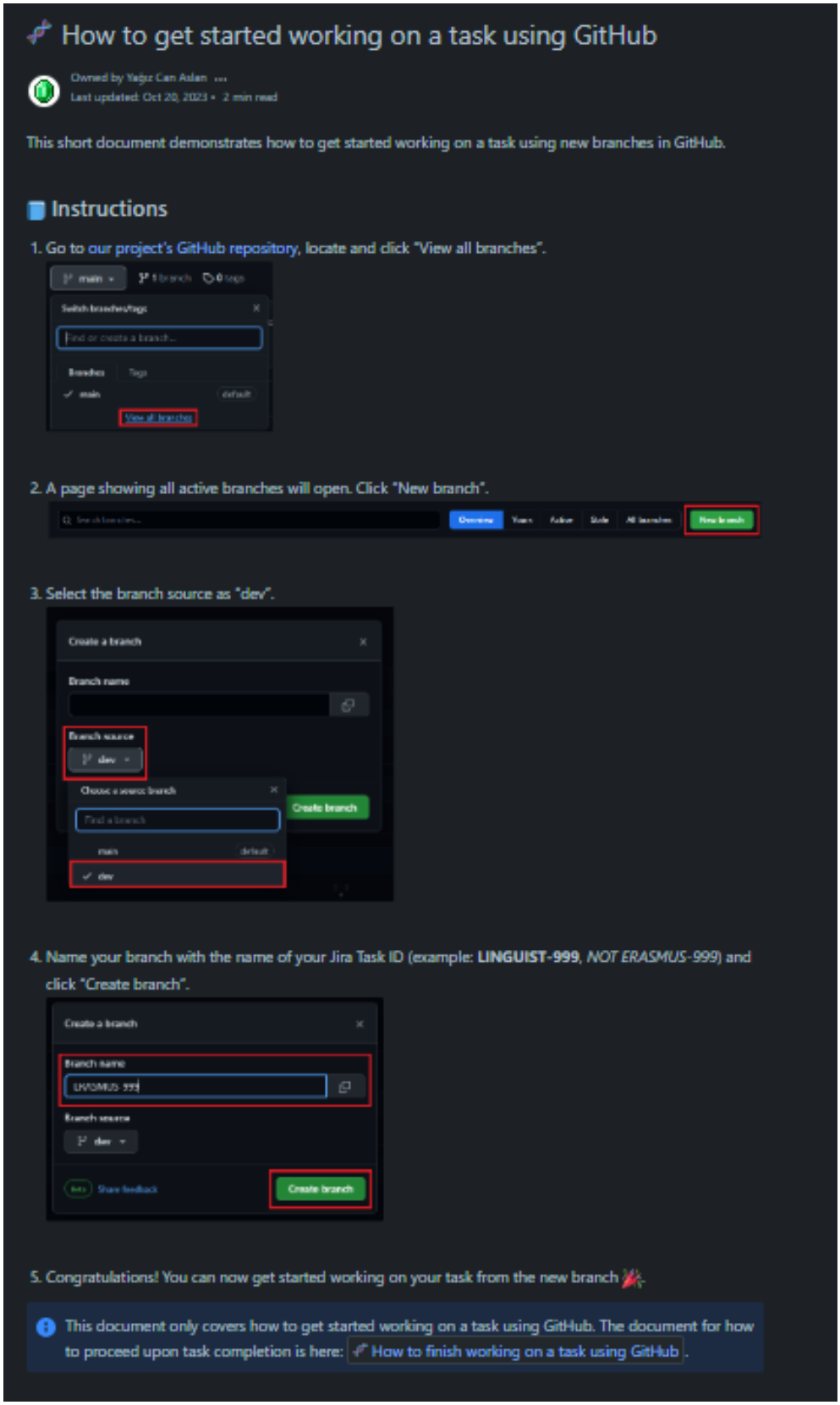

Fig. 3. LinguistAI Branch Creation Instructions

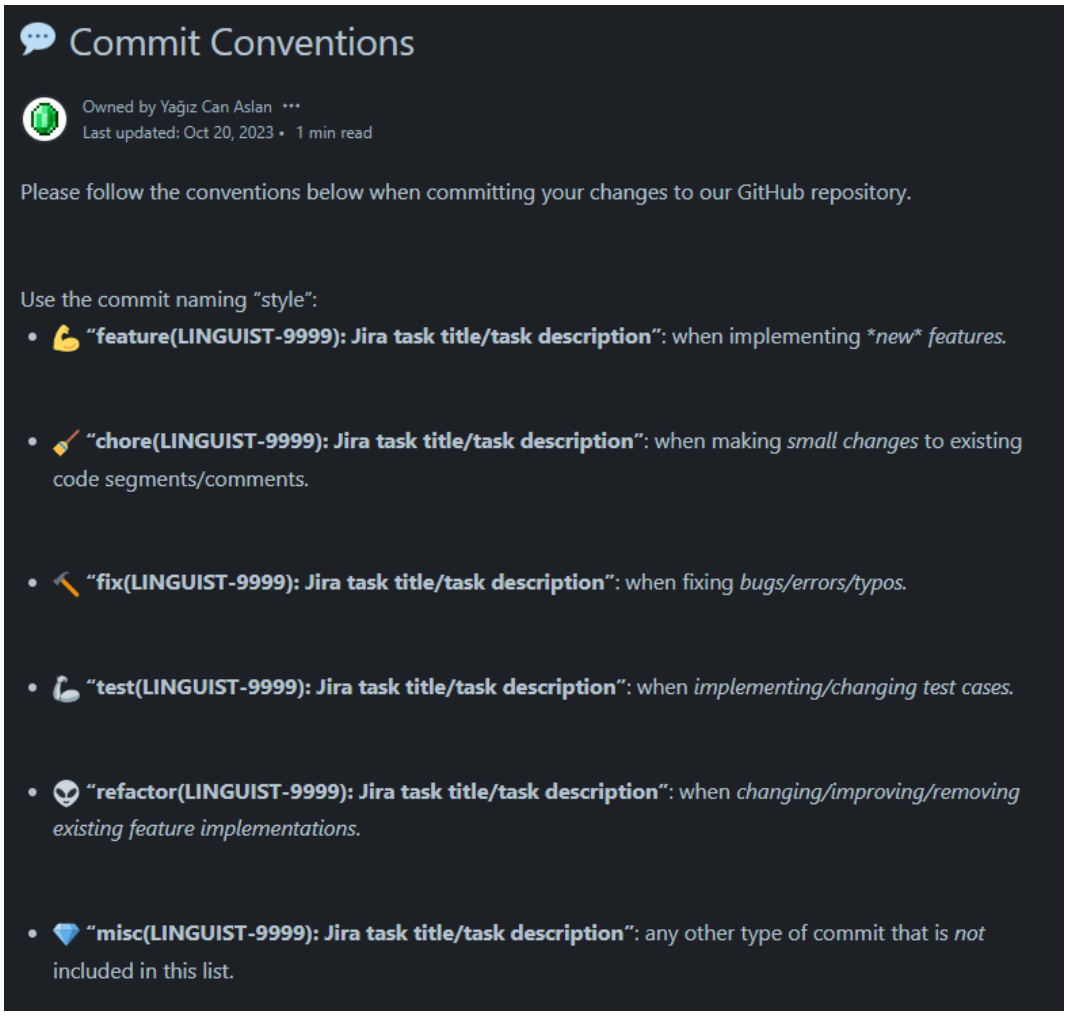

Fig. 4. LinguistAI Commit Conventions

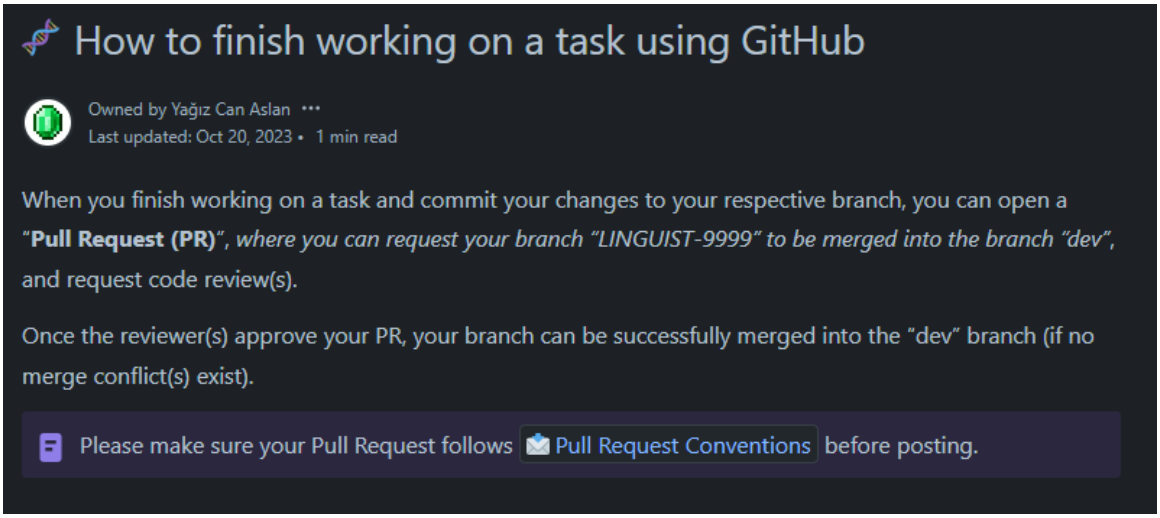

Fig. 5. LinguistAI Task Finishing Instructions

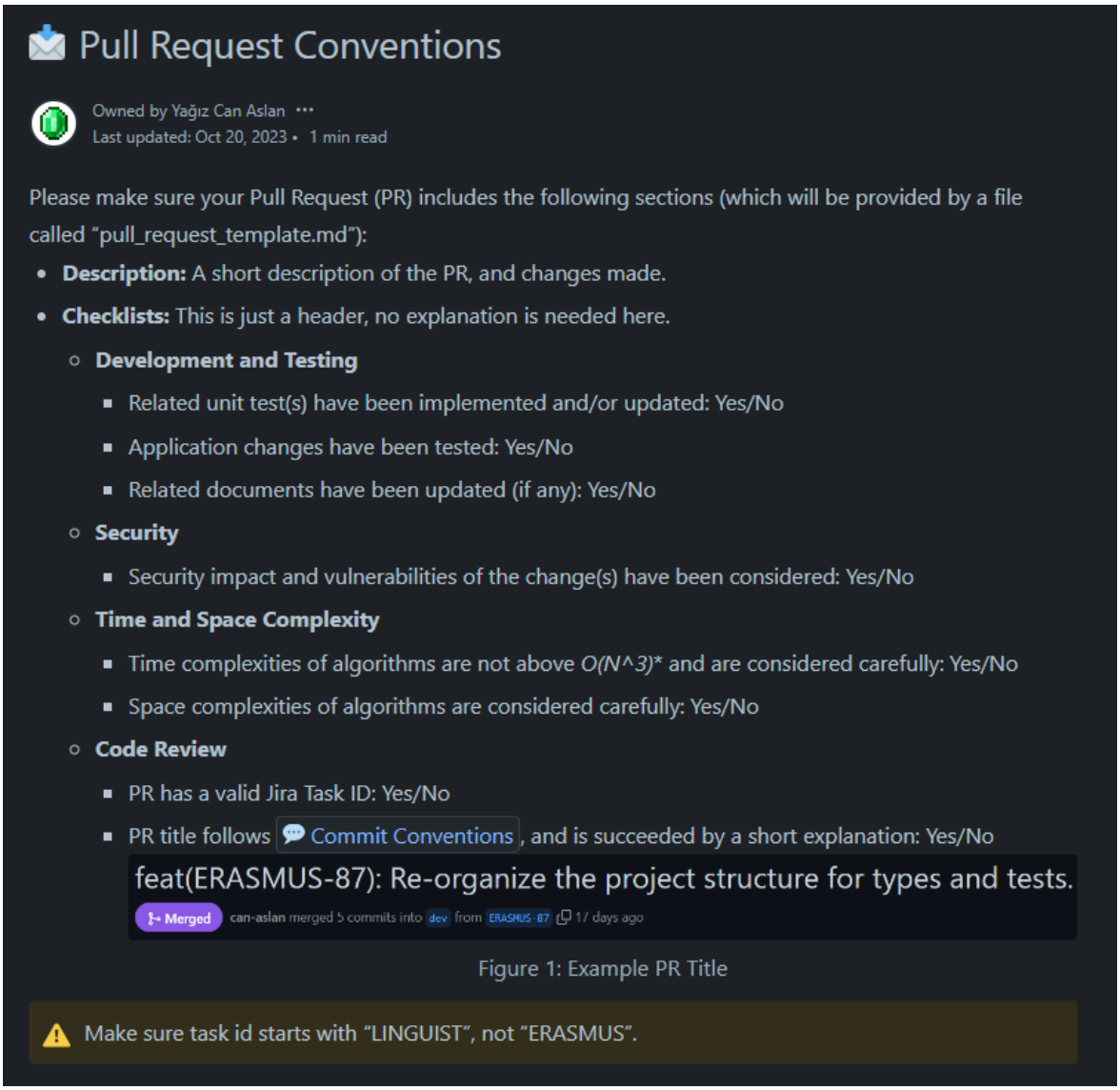

Fig. 6. LinguistAI Pull Request Conventions

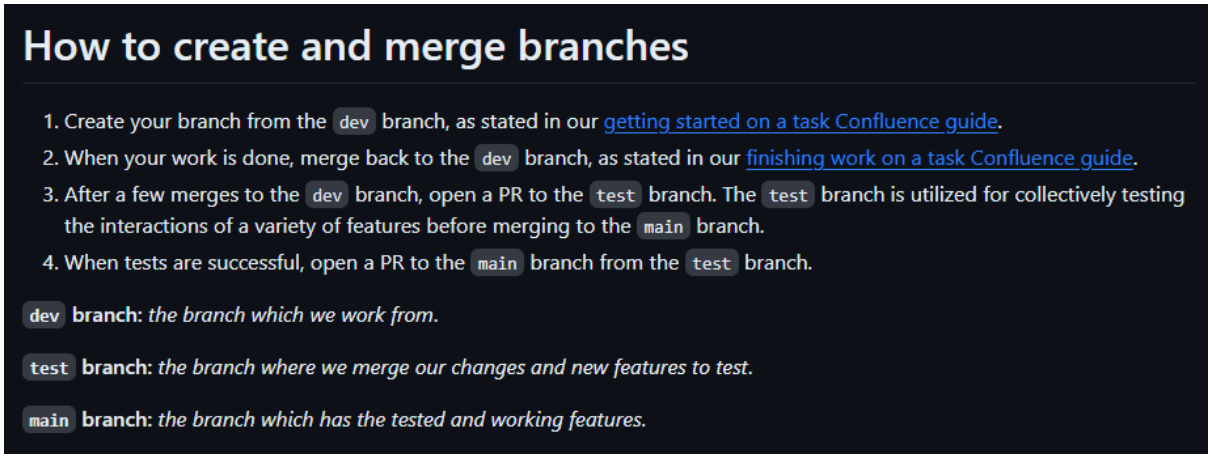

Fig. 7. LinguistAI Branch Instructions

2. Design Requirements

2.1. Functional Requirements

2.1.1. User

- The user can register/login by using their email.
- The user can reset their password using their email.
- The user can have a conversation with the chatbot by texting with the chatbot.
- The user can also choose to have a conversation with the chatbot by speaking.
- The user can listen to the pronunciation of the chatbot's texts.
- During their conversation with the chatbot, the user will be able to select a word they do not know and add it to their unknown word storage.
- The user can see the meaning of the words they select, along with an example sentence including the word.
- The user can see the list of unknown words they saved so far, with their meanings and example sentences available on click.
- The user will be presented with weekly tests to determine how well they know unknown vocabulary.
- The user can access their profile to see and edit their personal information.
- The user can see other users' profiles with their basic information and follow them.
- The user can view the leaderboard to compare themselves with their friends in terms of the number of words they learned in a particular time span.
- The user can be notified on a daily basis to remind them of their daily practice.
- The user can be notified of the progress of the other users they follow to encourage them to use the application.
- The user can see their daily streak on a pop-up when they initiate a chat with the bot for the first time in the day.
- The user can win awards and badges based on their learning statistics that are based on time spent on the app, number of words learned, language level-ups, etc.
- The user's proficiency level can be tested.
- The user will be presented with predefined role-play scenarios as prompts.
- The users will be able to opt-out of model personalization using user's personal information.

2.1.2. Chatbot

- The chatbot will initiate the conversation when the user opens up the chat screen.
- The chatbot can adapt its texts to be similar to the user's proficiency level.
- The chatbot will use words from the unknown word list more frequently.
- The chatbot will be given a summary of chat history and will adapt to it for further responses.

2.2. Non-Functional Requirements

This section mentions the non-functional requirements under Usability, Reliability, Performance, Supportability, Scalability and Security.

2.2.1. Usability

Linguist AI will be a mobile application. Since Linguist AI will be a mobile application, the users will be able to use our application anywhere which can include while commuting on the bus, when waiting in line, while waiting for food, and many more examples. As such, it is vital that the application is accessible and easy to use. The users must be able to utilize our application without struggle. Therefore, the application must have an intuitive and simple user interface, ensure accessibility for all types of users (different types of color blind etc.), have a responsive design, maintain optimal performance, include methods for gathering user feedback and error reporting, be compatible across a variety of platforms and maintain a consistent design. The application also should have an easy-to-follow and informative user onboarding procedure upon first use, to maintain user retention and make the application easier to understand and use. In order to adhere to the aforementioned usability guidelines for our project, the users must be able to reach all parts of the application with at most 5 clicks. For ease of use, the navigation bar should be visible on all screens, and the back buttons must always be in the same place (for instance, top left). We have defined the following constraints to measure usability:

- Each main subsystem, such as chat and leaderboard, should be accessible within 2 clicks for every logged-in user.
- Each login step should take 4 steps at maximum with valid credentials.
- All back buttons should be placed on the top left of the screen, with 16 eDPI for left and top margins.
- Each main subsystem should have a visible bottom navigation.
- All network requests should be asynchronous so that users can interact with the front end while the backend response is being generated.
- All in-progress network requests should be clearly indicated to the user with a loading visual placed as the closest element to the target component.

2.2.2. Reliability

In order for our application to achieve its goals, the application must provide a consistent experience. All systems of our application (backend servers, databases, LLM integration, mobile application etc.) must cooperate smoothly without major errors. To provide a smooth experience to our users, all of our systems should have a 95% uptime. The user must be able to continuously use the application without encountering a major error in any step. Moreover, the responses of our conversational AI models must not be strange or unsettling, the users must know that whenever they use our application, they will be subject to the same experience.

2.2.3. Performance

As stated in section 2.2.1, our users can use our application almost anywhere. As such, it is important for our application to be responsive. For this reason, the performance of the application will be measured through various metrics such as time to load (TTL), mean time to respond (MTTR) and time to first byte (TTFB). An upper bound is declared for all of these metrics in various contexts. The constraints considered for such metrics are as follows:

- The time to the first byte of the chat system should be strictly under 10 seconds. This also applies to a potential version of the chat which uses text streaming via sockets.
- Mean response time of the authentication service should be less than 5 seconds.
- Average time to load a page on the mobile application should be less than 3 seconds.
- Adding new words to the word list should take less than 2 seconds.
- Loading user profiles should take less than 4 seconds.
- Mean response time of the main subsystems besides chat and authentication should be less than 4 seconds.

If these conditions are not met, the users will have to wait a considerable amount of time to utilize our features which will decrease our application's usage and language teaching effectiveness significantly. With these metrics in mind, our expectation is having a UI that is performant and responsive, so the application does not feel clunky, slow or old, and the connections between the subsystems of our project such as backend servers, databases, LLM integration and mobile application happening relatively fast to enable the aforementioned smooth and responsive experiences.

2.2.4. Supportability

In order to reach a broader user base, our mobile application must be available for a variety of platforms, including Android and iOS. As such, we must maintain the application for different operating systems and mobile devices. Our application must be able to be maintained and expanded upon easily. This can be addressed by following good software development practices and adhering to tested and established design paradigms such as Object Oriented Programming. In order to adhere to these values, we established the following constraints:

- The minimum Android SDK version should be 24.0 which encompasses more than 98% of the active Android user base, and is the recommended minimum by Google.
- The minimum iOS version should be 15 which is the version recommended by Apple. 94% users use it.

2.2.5. Scalability

Since our application is a mobile application, it will be easy for new users to download and start using our application. Therefore, our systems (backend servers, databases, LLM integration, mobile application etc.) must be prepared and ready to handle a surge in users and server loads. Our system must be able to handle at least 1000 concurrent users. We must also develop our

systems in such a manner that migration, adding new features or distributing our application to new vendors (for mobile applications, like App Store, Google Play Store etc.) will not cause significant problems.

2.2.6. Security & Privacy

Linguist AI will store user data to provide a specialized learning experience. The stored data may include names, interests, passwords, emails, conversation history and many more types of information. Henceforth, it is crucial that our systems are as secure as possible to protect sensitive data and that we handle such data with extreme care. To achieve this, we can utilize complex encryption algorithms and various other tried and tested security protocols. Moreover, we will follow GDPR (General Data Protection Regulation) and KVKK (Kişisel Verilerin Korunması Kanunu) guidelines closely and will seek user consent whenever necessary, to strike a balance between personalization and privacy. We are implementing the following procedures for our security purposes:

- All passwords must be hashed with SHA-256 and salted.
- All sessions must be given a session token by JWT and all requests should have a valid session token.
- All users should be provided with a refresh token to regenerate the session token when expired.
- Session token should be valid for 10 hours.
- Refresh token should be valid for 24 hours.

3. Market Analysis

3.1. Target Users

Our target users can come from any native language background since our chats will be exclusively in English, and any potential translation tools will be offered in all languages. With approximately 1.5 billion English language learners worldwide [4], this ensures a wide range of users. However, our specific focus will be on people of all ages who are comfortable using mobile apps and have achieved at least an A2 level in English proficiency. This targeting helps to minimize frustration when engaging in conversations, especially for those who are beginners in the language.

3.2. Market Research

The English language learning market, estimated at 21.3 billion USD in 2022 [5], presents robust growth prospects driven by distinct trends and user preferences.

3.2.1. Learning Modes

The market is categorized into two segments: self-learning and tutoring. Currently, tutoring commands the largest market share. This is attributed to the growing need for customized language education, providing real-time feedback, and the convenience of in-person interaction. Despite tutoring's dominance, self-learning apps are expected to show remarkable growth with the highest projected compound annual growth rate (CAGR). Several factors fuel this trend, including the appeal of cost-effective tech-driven solutions, new online learning platforms, digital content, and the desire for customizable language learning experiences [6].

3.2.2. End Users

The market's end users include individual learners, educational institutes, government bodies, and corporate learners. In 2023, individual learners are at the forefront, claiming both the largest market share and the highest CAGR. Their prominence is influenced by the adoption of smart devices, faster internet, a growing appetite for live online content, and the search for professional growth and personal development. These learners expect their language learning journey to be flexible and tailored to their specific needs [6].

3.2.3. Emerging Trends

Incorporating Artificial Intelligence (AI) and Natural Language Processing (NLP) technologies for personalized learning, automated feedback, and intelligent content recommendations is a prominent trend [7]. Additionally, gamification, featuring interactive quizzes, rewards, and challenges, emerges as an effective technique to enhance learner motivation and retention.

3.2.4. Our Position in the Market

Given these market insights, it becomes evident that our language learning mobile app is well-positioned for success. By leveraging emerging trends such as AI and NLP technologies, we offer personalized learning experiences where self-learners can learn and practice at their own pace

without fear of judgment for their mistakes. It is convenient, relatively affordable, and offers gamification to motivate and retain users. Furthermore, it aligns with the broader market trends prioritizing flexibility, customization, and technology-driven learning. As a result, we anticipate that our app will tap into the vast potential of this dynamic and evolving market.

3.3. Competitive Analysis

3.3.1. Language Learning Competitors

Several companies, including Wall Street English and Berlitz Languages, provide structured language classes, available online and in physical schools, supplemented with web or app-based exercises. Both Cambly and Italki offer one-on-one or group chat sessions with an experienced tutor. In contrast, Rosetta Stone and Memrise specialize in interactive learning through flashcards reinforced with native audio. Additionally, there is a range of web and mobile applications such as Babbel, Busuu, Mondly by Pearson, and Duolingo, which offer interactive, bite-sized lessons focusing on vocabulary and grammar. Babbel extends its features to include podcasts, videos, and optional live virtual classes. Busuu offers a correction feature by voluntary native speakers, while Mondly's app caters to language learning for businesses, featuring employee tracking. Meanwhile, Duolingo stands out for its gamified approach to learning, complemented by progress tracking, allowing users to connect with friends and earn points as they progress in their language learning journey.

In summary, language learning applications often fall into two main categories: those that prioritize live tutor instruction and those offering structured online lessons. In contrast, our platform's primary feature revolves around an AI-driven chatbot. We believe this chatbot will offer an affordable, non-judgmental, and 24/7 accessible practice environment, effectively bridging the gaps left by traditional tutor-based approaches. Furthermore, we place a strong emphasis on enhancing both writing and speaking skills, aspects that online lessons may sometimes struggle to address adequately. In the following section, we explore language learning companies that use AI more extensively.

3.3.2. AI Language Learning Competitors

In AI-driven language learning, competitors introduce innovative features to enhance their existing apps or offer standalone solutions. Duolingo Max offers a beta feature that uses AI to explain users' mistakes in lessons and to engage with users in certain roleplay scenarios, such as ordering coffee. Mondly Chatbot is another app where users can chat with AI. Their chatbot's range of topics is also limited; however, in October 2023, Mondly announced they were working on a more advanced chatbot with a human-like appearance named LUNA [8]. Memrise Membot allows its users to discuss a wider range of predetermined topics with AI, from everyday scenarios like getting a haircut to more profound topics like talking about the impact of social media. Talkpal AI is a web application with an open-ended AI chat feature, real-time feedback, and optional roleplay scenarios. Notably, it avoids gamification elements in its approach. Lastly, Character AI is a web application known for its chatbots with distinct personalities. These chatbots can engage users in conversations; however, the app is not explicitly designed for language learning.

Our app distinguishes itself in several ways. Unlike certain competitors, our chatbot engages in open-ended conversations, not restricted to predetermined topics. It is uniquely designed for language learners, providing conversations that adapt to the user's proficiency and focusing on unfamiliar words to enhance vocabulary. Furthermore, we have incorporated gamification features into our AI to motivate and inspire users on their language learning journey.

4. Academic Analysis

4.1. Large Language Models

Large Language Models (LLMs), particularly represented by Transformer-based models, are complex artificial intelligence systems that can understand and generate human-like text. These models are trained to predict the next word or fill in masked words in a sentence to understand language patterns, semantics, and context. Additionally, attention mechanisms allow LLMs to focus on relevant input parts, mimicking a type of reasoning similar to human language understanding. As a result, LLMs produce coherent and contextually relevant text, making them useful for tasks such as chatbot conversations, content production, and language translation [9]. Most notable LLMs include OpenAI's ChatGPT, Meta's Llama, Google's BERT and PaLM.

In LLMs, the term "large" refers to the use of massive amounts of data and parameter counts that can reach billions. Beyond a certain scale threshold, these models exhibit emergent abilities not found in smaller counterparts. Three notable emergent abilities are in-context learning, where the model generates expected output without additional training; instruction following, where the model can perform well on unseen tasks through instructions; and step-by-step reasoning, which allows LLMs to solve complex tasks involving multiple reasoning steps [9,10].

LLMs present challenges from both academic and industrial perspectives. Academically, there is a lack of understanding of how these previously mentioned emergent abilities occur. Two additional challenges arise from an industrial standpoint. First, training capable LLMs proves difficult due to the high demand for computation resources, making repeated studies on various training strategies costly. The primarily industry-driven training of LLMs, with undisclosed details such as data collection and cleaning, adds to the complexity. Second, aligning LLMs with human values can be challenging because these models may generate toxic or harmful content, requiring effective risk-mitigation strategies [10].

4.2. AI in Language Learning

The incorporation of chatbots into language learning has caught the interest of educators and researchers, who see these digital entities as valuable conversation partners for students [11]. Notably, research emphasizes the distinct benefits provided by chatbots, such as their availability at any time and adaptability to diverse learner needs and proficiencies [12]. Furthermore, chatbots' ability to engage in extensive practice allows learners to experiment with new language aspects, reinforcing recently acquired vocabulary and grammar [11]. Contrary to assumptions, research shows users have no less confidence or discomfort when conversing with chatbots than

with human counterparts [13]. Furthermore, demonstrated efficacy in improving overall foreign language proficiency [14] establishes chatbots as valuable tools in language education. An 8-week study with Korean English-as-a-Foreign-Language (EFL) students, for example, found that students who used chatbots improved their vocabulary skills more than the control group [15].

Another significance of AI in language learning is how research emphasizes the importance of learners' need for significant output practice and interaction in order to feel satisfied and confident in their language skills [16]. While immersion programs increase learners' willingness to converse, the majority of language learners do not have access to such immersive environments. AI chatbots emerge as potential substitutes for human interaction in this context, creating an immersive learning environment for students [17].

The motivation for learners to practice with chatbots rather than human partners often stems from a fear of making mistakes [11] and a preference for low-pressure opportunities to reinforce vocabulary and grammar skills. This tendency suggests that chatbots can play a unique role in providing a supportive and non-judgmental learning environment. A 16-week study of Korean EFL students revealed positive perceptions of chatbot interactions, attributing this positivity to the chatbots' role in alleviating social anxiety and peer pressure, fostering a learning environment marked by reduced stress and increased engagement [18]. Similar studies have found that learners' motivation and confidence have increased [19, 15].

Moving on to voice-based chatbots, a study involving Taiwanese EFL university students interacting with a speaking chatbot revealed that those in the experimental group experienced significant improvements in oral English proficiency [20]. Similarly, a 10-week study with Korean EFL students revealed improvements in speaking competence, pronunciation, and fluency, with students displaying more accurate vocabulary and grammar after chatbot interactions when compared to the control group [21]. A separate three-month study with Chinese EFL students revealed an important increase in students' speaking skills [17]. It should also be noted that studies indicate that people learn better from a human voice than from a machine voice [22].

Despite the potential benefits of chatbots in language learning, certain considerations should be made. According to research, beginner EFL students may struggle to initiate conversations with chatbots in the target language due to limited proficiency [23]. As a result, it is possible that targeting learners with a conversational level of English may be a more effective strategy. Furthermore, studies highlight the risk of boredom if chatbots engage in long monologues [24], emphasizing the importance of interactive and conversational-style [22] dialogues in engaging learners and making language learning more enjoyable and effective.

In conclusion, the introduction of chatbots into language learning has proven promising, providing benefits such as accessibility at all times and adaptability. Numerous studies show positive outcomes, ranging from improved vocabulary skills to greater oral proficiency to increased student motivation. However, it is critical to consider learners' proficiency levels, as beginners may face difficulties. It is essential to maintain interactive and engaging dialogue formats in order to avoid

boredom and maintain interest. As technology advances, strategic integration of chatbots holds tremendous potential for creating effective and immersive language learning experiences.

4.3. Ethics of Language Learning Models

Ethics has been a concern for humans for a long time on topics such as distinguishing right and wrong, justice and injustice, etc. With the development of new AI technologies, especially with the recent rapid rise in LLMs, a framework on the ethics of AI, which should adhere to the current laws and ethical ideas is needed.

Firstly, AI systems must be designed transparently. This means that any information regarding the AI's design, the training data, its outcomes, and its use of user's data must be disclosed and displayed transparently. This would make it accessible for the users and professionals to evaluate whether the model is adhering to ethical standards.

Secondly, the model should not exhibit discriminative behavior that would favor a certain group of people, whether it be ethnic or religious or any other kind. Any biases should be examined thoroughly, and its reasons should be investigated within the model's architecture or the training data's nature. Also, it is important to obtain the user's consent if the model uses any of their data. Furthermore, personalization should not be taken too far to a point that will threaten the privacy of the users. To achieve this, the data storage methods of the system can be publicly disclosed for evaluation. Lastly, who is accountable for the actions of the AI system must be determined. Moreover, it is important to be precisely clear about these points and avoid any shortcuts and generalized promises such as "the AI system adheres to transparency principles" and be more open about the system behind the model [25].

5. Mockups

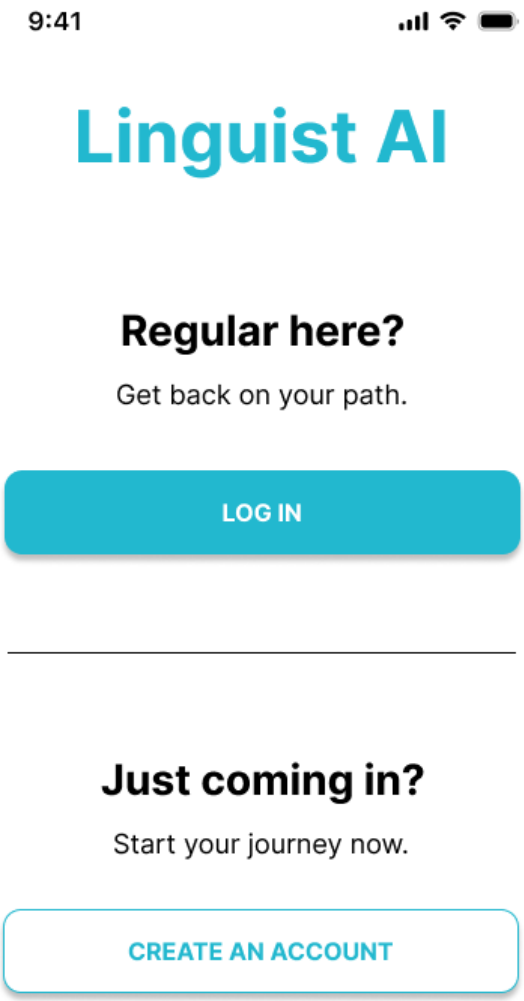

Fig. 8. Landing Screen

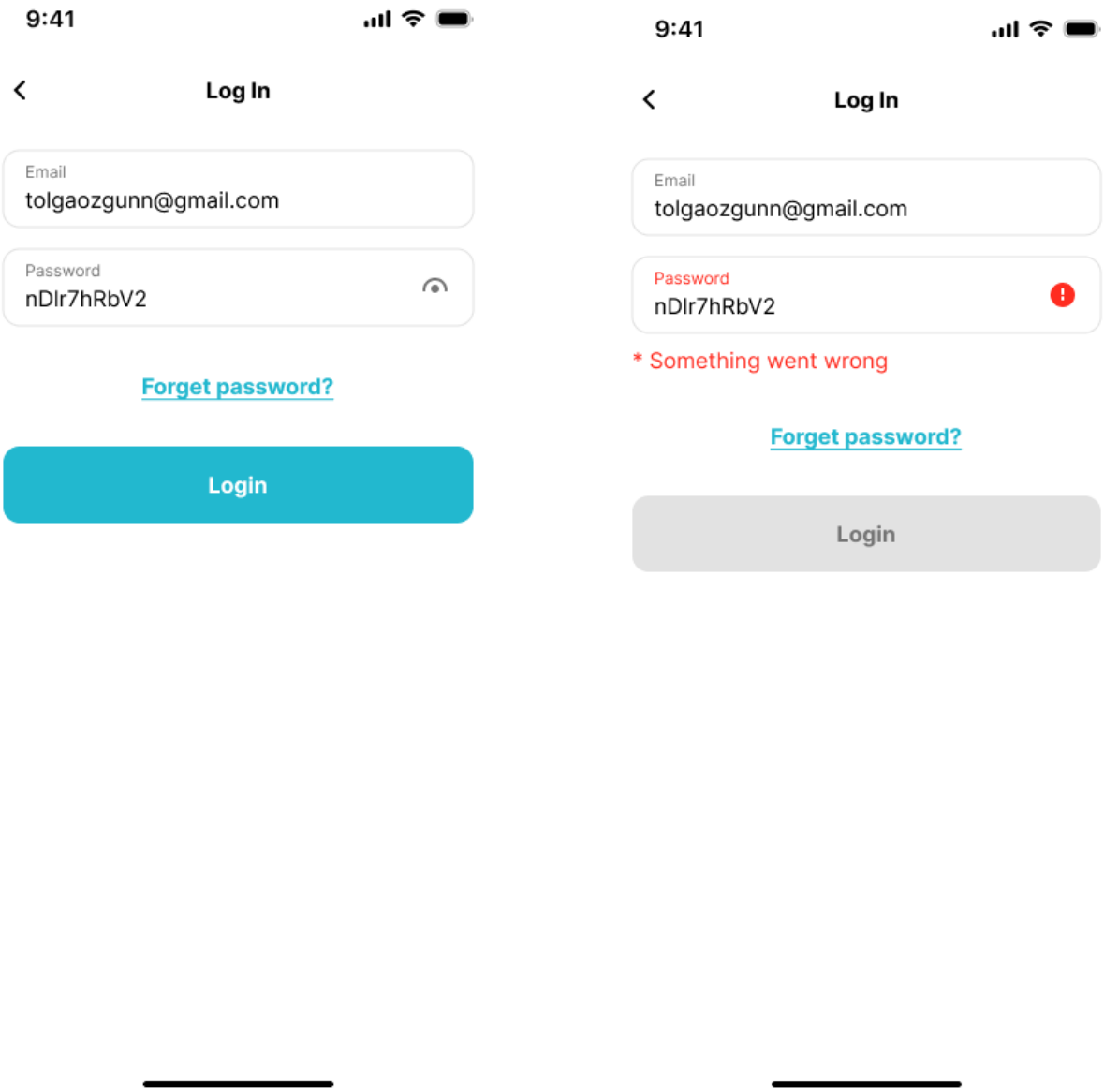

Fig. 9. Login Screens

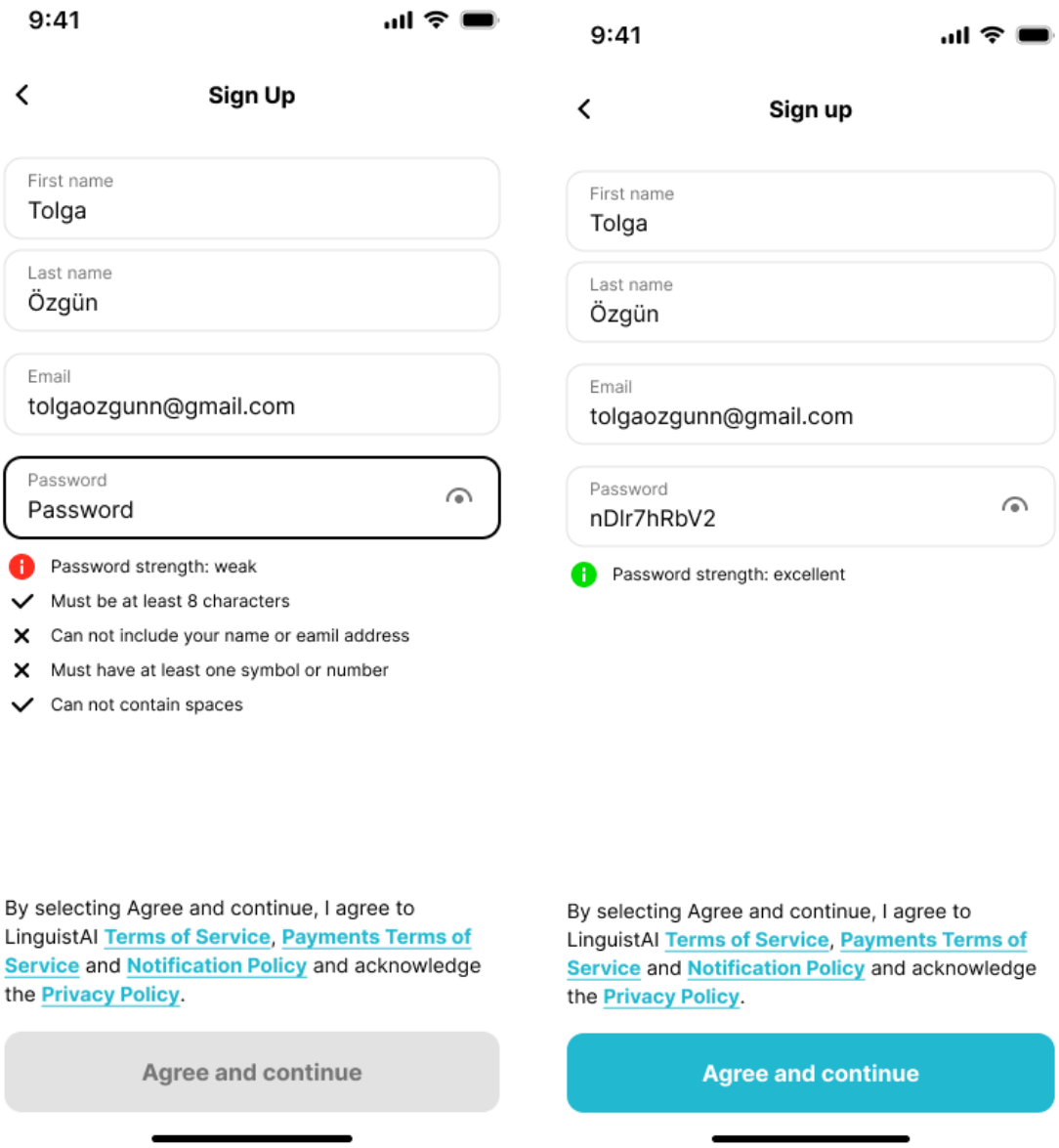

Fig. 10. Sign-up Screens

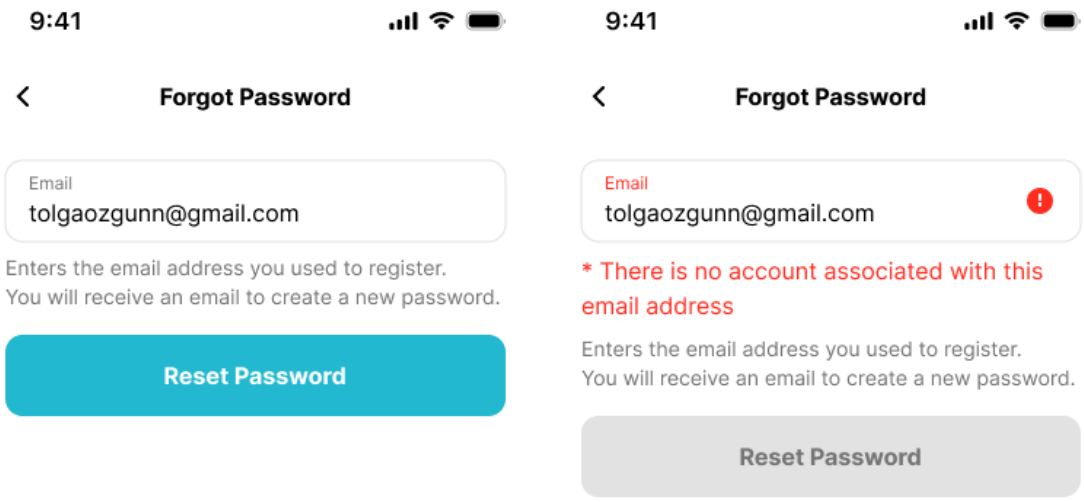

Fig. 11. Forgot Password Screens

Fig. 12. Chatting Screen

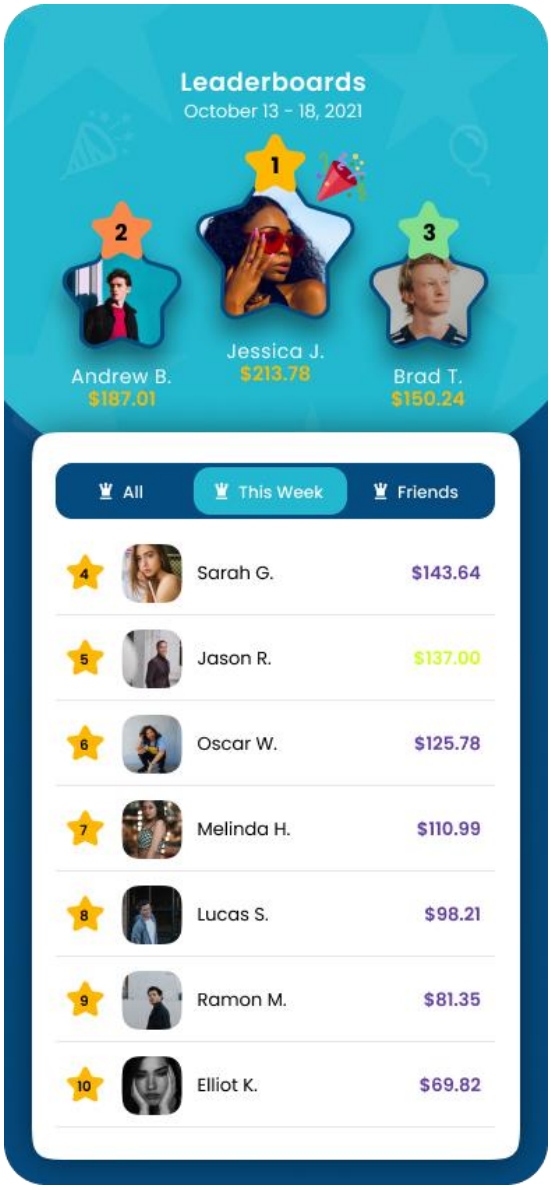

Fig. 13. Leaderboard Screen

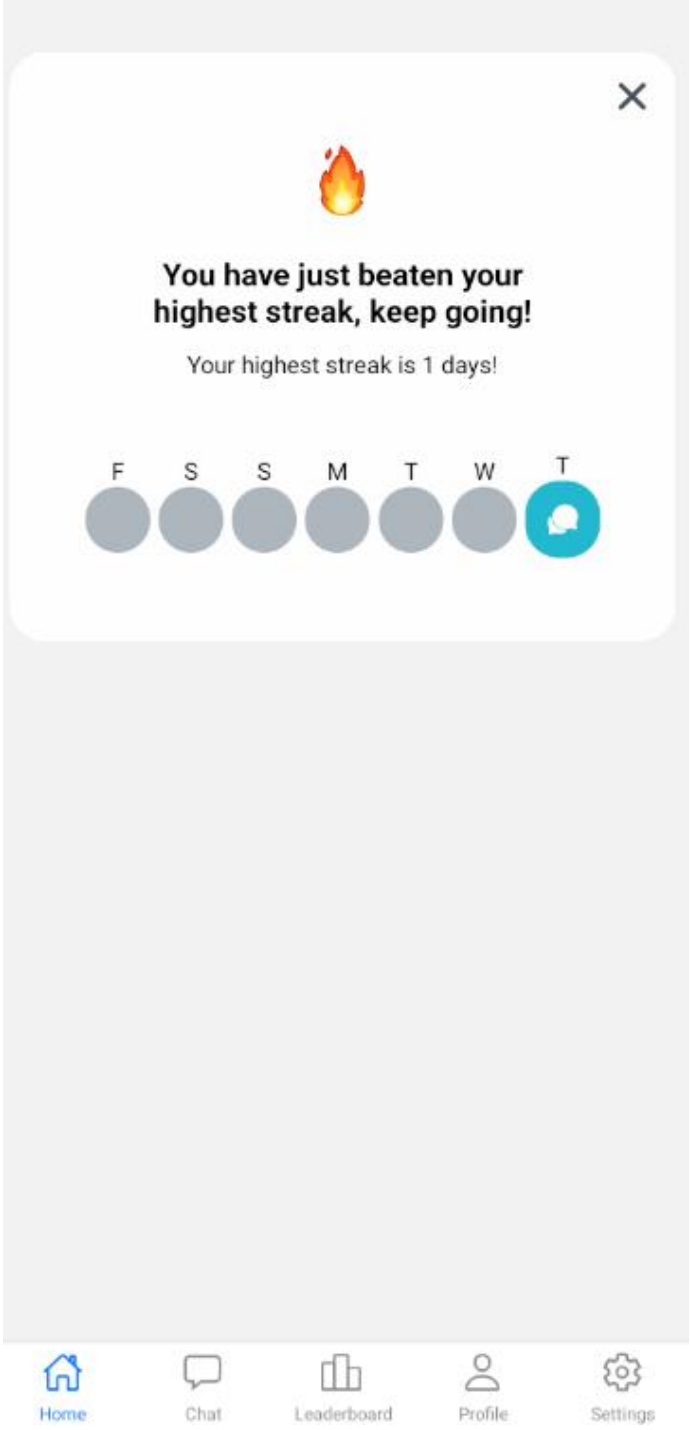

Fig. 14. Streak Screen

References

[1] E. H. Dyvik, “The most spoken languages worldwide 2023,” Statista. [Online]. Available: https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/ [Accessed: Nov 14, 2023].

[2] “Learn a language for free,” Duolingo. [Online]. Available: https://www.duolingo.com/ [Accessed: Oct 27, 2023].

[3] “Online English classes with native speakers,” Cambly. [Online]. Available: https://www.cambly.com/english?lang=en [Accessed: Oct 27, 2023].

[4] J. Bentley, “Report from TESOL 2014: 1.5 billion English learners worldwide,” International TEFL Academy. [Online]. Available: http://www.internationalteflacademy.com/blog/report-from-tesol-2-billion-english-learners-worldwide [Accessed: Oct 20, 2023].

[5] "English Language Learning: Global Strategic Business Report 2023-2030." Research and Markets. [Online]. Available: https://www.researchandmarkets.com/reports/5140838/english-language-learning-global-strategic. [Accessed: Oct 21, 2023].

[6] "English Language Learning Market." Meticulous Research. [Online]. Available: https://www.meticulousresearch.com/product/english-language-learning-market-5127. [Accessed: Oct 20, 2023].

[7] “Language Learning Market.” Global Market Insights. [Online]. Available: https://www.gminsights.com/industry-analysis/language-learning-market . [Accessed: Oct 21, 2023].

[8] "Get Ready to Meet Your New AI Conversation Partner." Mondly. [Online]. Available: www.mondly.com/blog/new-ai-conversation-partner/. [Accessed: Oct 21, 2023].

[9] E. Kasneci et al., “ChatGPT for good? On opportunities and challenges of large language models for education,” *Learning and Individual Differences*, vol. 103, p. 102274, 2023. [Online]. Available: https://doi.org/10.1016/j.lindif.2023.102274. [Accessed: Nov 11, 2023].

[10] W. X. Zhao et al., “A survey of large language models,” arXiv preprint. [Online]. Available: https://doi.org/10.48550/arXiv.2303.18223. [Accessed: Nov 11, 2023].

[11] L. K. Fryer et al., “Bots for language learning now: Current and future directions,” *Language Learning & Technology*, vol. 24, no.2, pp. 8-22, Jun. 2020. [Online]. Available: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/950396a1-e7a1-4eac-bd27-c3d194ce77e2/content. [Accessed: Nov 16, 2023].

[12] E. Ayedoun, Y. Hayashi, and K. Seta, “A conversational agent to encourage willingness to communicate in the context of English as a foreign language,” *Procedia Computer Science*,

vol. 60, pp. 1433–1442, 2015. [Online]. Available: https://doi.org/10.1016/j.procs.2015.08.219. [Accessed: Nov 16, 2023].

[13] J. Hill, W. Randolph Ford, and I. G. Farreras, "Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations," *Computers in Human Behavior*, vol. 49, pp. 245–250, 2015. [Online]. Available: https://doi.org/10.1016/j.chb.2015.02.026. [Accessed: Nov 16, 2023].

[14] S. Bibauw, T. François, and P. Desmet, "Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based call," *Computer Assisted Language Learning*, vol. 32, no. 8, pp. 827–877, 2019. [Online]. Available: https://doi.org/10.1080/09588221.2018.1535508. [Accessed: Nov 16, 2023].

[15] N.-Y. Kim, "Chatbots and Korean EFL Students' English Vocabulary Learning," *Journal of Digital Convergence*, vol. 16, no. 2, pp. 1–7, 2018. [Online]. Available: https://doi.org/10.14400/JDC.2018.16.2.001. [Accessed: Nov 16, 2023].

[16] T.-Y. Tai and H. H.-J. Chen, "The impact of google assistant on adolescent EFL learners' willingness to communicate," *Interactive Learning Environments*, vol. 31, no. 3, pp. 1485–1502, 2020. [Online]. Available: https://doi.org/10.1080/10494820.2020.1841801. [Accessed: Nov 16, 2023].

[17] Y. Yuan, "An empirical study of the efficacy of AI Chatbots for English as a foreign language learning in primary education," Interactive Learning Environments, pp. 1–16, 2023. [Online]. Available: https://doi.org/10.1080/10494820.2023.2282112 [Accessed: Nov 16, 2023].

[18] J. Jeon, "Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives," *Computer Assisted Language Learning*, pp. 1–26, 2022. [Online]. Available: https://doi.org/10.1080/09588221.2021.2021241. [Accessed: Nov 16, 2023].

[19] S. Lee *et al.*, "On the effectiveness of robot-assisted language learning," *ReCALL*, vol. 23, no. 1, pp. 25–58, 2011. [Online]. Available: https://doi.org/10.1017/s0958344010000273. [Accessed: Nov 16, 2023].

[20] M.-H. Hsu, P.-S. Chen, and C.-S. Yu, "Proposing a task-oriented chatbot system for EFL learners speaking practice," *Interactive Learning Environments*, vol. 31, no. 7, pp. 4297–4308, 2021. [Online]. Available: https://doi.org/10.1080/10494820.2021.1960864. [Accessed: Nov 16, 2023].

[21] D.-E. Han, "The effects of voice-based AI Chatbots on Korean EFL middle school students' speaking competence and affective domains," *Asia-pacific Journal of Convergent Research Interchange*, vol. 6, no. 7, pp. 71–80, 2020. [Online]. Available: http://dx.doi.org/10.47116/apjcri.2020.07.07. [Accessed: Nov 16, 2023].

[22] R. E. Mayer, “Using multimedia for e-learning,” *Journal of Computer Assisted Learning*, vol. 33, no. 5, pp. 403–423, 2017. [Online]. Available: https://doi.org/10.1111/jcal.12197. [Accessed: Nov 16, 2023].

[23] R. R. Divekar et al., “Foreign Language Acquisition via Artificial Intelligence and extended reality: Design and evaluation,” *Computer Assisted Language Learning*, vol. 35, no. 9, pp. 2332–2360, 2021. [Online]. Available: https://doi.org/10.1080/09588221.2021.1879162. [Accessed: Nov 16, 2023].

[24] S. Matsuura and R. Ishimura, “Chatbot and dialogue demonstration with a humanoid robot in the lecture class,” *Universal Access in Human–Computer Interaction. Human and Technological Environments*, pp. 233–246, 2017. [Online]. Available: https://doi.org/10.1007/978-3-319-58700-4_20. [Accessed: Nov 16, 2023].

[25] J. Bang, B.-T. Lee, and P. Park, “Examination of ethical principles for LLM-based recommendations in conversational AI,” *2023 International Conference on Platform Technology and Service (PlatCon)*, 2023. [Online]. Available: https://doi.org/10.1109/platcon60102.2023.10255221. [Accessed: Nov 16, 2023].