Kural AAC Assistive Voice & Sign

A Visual Communication Platform with AAC, Speech Assistance and Sign-Language Support

Author

Ms. HARITHA M

Project Members **PARAMASIVAM A, PRIYAN S, RITHISH C, SHANMUGA IYYAPPAN S**

Sri Shakthi Institute of Engineering and Technology Department of Computer Science and Engineering
Second Year
Corresponding Author: paramasivam25cs@srishakthi.ac.in

Abstract

Augmentative and Alternative Communication (AAC) technologies support people who experience difficulty communicating through speech. This paper presents Kural AAC, an assistive communication platform designed to bring picture-based communication, sentence construction, speech assistance and sign-language support together in a unified interface. The prototype dashboard provides access to an AAC communication board, a sign-language assistant, a voice assistant, communication history and application settings. The AAC board organizes familiar symbols into categories such as needs, food and drink, feelings, people, activities, emergency messages and quick phrases. Users can select symbol cards to construct a message, view the sentence and request spoken output. English and Tamil are presented as language choices in the interface. Additional screens illustrate account registration, role selection, speech preferences and communication logs. The sign-language assistant screen provides a camera-based interaction area and a panel for recognized signs, while the voice assistant screen presents microphone dictation and typed-phrase speech controls. The paper describes the prototype screens, functional scope, logical architecture, data organization, interaction algorithms, accessibility considerations and a verification plan. The supplied screenshots provide evidence of the interface design; they do not by themselves establish the accuracy of sign recognition, speech recognition, backend persistence or completed user testing. Kural AAC is therefore presented as a project prototype and design contribution, not as a clinically validated communication intervention.

Keywords: Augmentative and Alternative Communication, AAC, assistive technology, picture-based communication, text-to-speech, speech recognition, sign-language assistance, Tamil language, accessibility.

1. Introduction

Communication enables people to express needs, preferences, emotions, questions and social messages. Some individuals experience barriers when relying on spoken language alone. AAC provides alternative or supplementary ways to communicate, including gestures, symbols, pictures, communication boards and speech-generating devices. A digital AAC platform can combine visual vocabulary with synthesized speech so that a selected message can be both seen and heard.

Kural AAC is designed as a multi-tool communication interface. Its central AAC board presents picture or symbol cards and a sentence builder. The wider platform also includes a sign-language assistant, a voice assistant, communication history and settings. The interface offers English and Tamil language options, reflecting the importance of local-language accessibility. The design aims to support communication in home, educational and social contexts while allowing the communication vocabulary and speech preferences to be configured.

The project focuses on the design and documentation of a working-style prototype interface. The screenshots

supplied for this paper show the dashboard and the principal feature screens. The paper distinguishes visible interface elements from behaviours that require separate functional verification. This distinction is important because a screen design alone cannot prove that camera recognition, microphone processing, account storage or speech output works reliably on a target device.

2. Problem Statement

People with speech or communication difficulties may not be able to communicate their needs quickly using speech alone. Paper-based boards can be useful but may be difficult to update, carry or personalize. Digital tools can offer flexible vocabulary, language selection and speech output, but users may face fragmented workflows when separate applications are required for picture communication, voice input and sign-language assistance. A further challenge is that language and speech-resource availability varies across devices.

The project addresses the need for a unified, visually clear communication interface that provides symbol-based message construction and access to complementary voice and sign-related tools. It also seeks to make common communication actions available through a consistent dashboard and to provide a history view for reviewing previously recorded communication events.

3. Objectives

- Design a clear dashboard that provides direct access to the main communication tools.
- Organize AAC symbols into meaningful categories and display English and Tamil labels where provided.
- Allow users to build a message by selecting multiple symbol cards and present the resulting sentence visibly.
- Provide controls for requesting speech output from a constructed or typed phrase.
- Present a camera-based sign-language assistance workflow with a recognized-sign result area.
- Provide a voice-assistant interface for typed phrases and microphone-based dictation.
- Present communication history with filters, favorites and export-oriented controls.
- Provide settings for language preference, speech voice, speed and volume.
- Design the interface so that later testing, personalization and accessibility improvements can be incorporated.

4. Scope of the Prototype

The supplied Kural AAC screens show a web-style dashboard with a left navigation panel and a main content area. The navigation items are Home, AAC Communication, Sign Language Assistant, Voice Assistant, Communication History, Settings and Help and Support. The top area includes English and Tamil language choices and a sign-in control. The home dashboard introduces the platform and presents quick access to the main communication tools.

The AAC communication screen includes a sentence builder, category filters, a symbol search field and a grid of communication cards. The visible category labels include All Symbols, Needs, Food & Drink, Feelings, People, Activities, Emergency and Quick Phrases. Sample cards shown include Water, Restroom, Sleep, Help Me, Medicine, Bath, Wheelchair, Fan / AC, I Want to Eat, Rice / Meals, Milk and Tea / Coffee. Tamil labels are displayed beneath several English labels.

The sign-language assistant screen presents a live camera area, a recognized-sign panel, an Auto Speak option, output-language controls and a list of supported sign examples. The voice assistant screen presents a microphone interaction area, a phrase-entry box, language choices and a Speak Text action. The history screen presents phrase counts, favorites, source filters and individual log entries. Settings include language preference, speech engine selection, speech speed and volume.

The paper documents these visible prototype elements and the intended workflows. It does not claim that the

sign-recognition model, speech-recognition service, authentication backend, database, CSV export or every speech feature has passed functional testing. Such claims require execution evidence from the actual application.

5. Existing Approaches and Design Rationale

AAC approaches range from unaided communication methods, such as gestures and facial expressions, to aided methods, such as printed boards, symbol books and electronic speech-generating devices. Picture-based systems can make vocabulary more concrete and provide a visual reference for the intended message. Electronic systems can add voice output, flexible vocabulary and language options.

Speech synthesis converts text into audible speech, allowing a device to speak a word or sentence selected by the user. Speech recognition performs a different task: it attempts to convert spoken input into text. Sign-language recognition is also distinct; it attempts to interpret visual hand or body movements and may require camera input, image processing and a trained recognition model. These technologies have different accuracy, privacy and hardware requirements and should be evaluated independently.

Kural AAC adopts a modular interface so that symbol selection, voice input and sign-language assistance can be accessed from a common dashboard. The design rationale is to reduce navigation between separate tools while keeping the main AAC communication board prominent.

6. Prototype Screen Descriptions

6.1 Home Dashboard

The home dashboard introduces Kural AAC as an assistive voice and sign platform. Its headline, “Express Yourself Clearly with Assistive AAC & Sign to Voice,” communicates the overall purpose. The dashboard provides a Start Communicating action and a Try Sign Camera action. Below the introductory panel, the Core Communication Tools section presents quick-access cards. This screen acts as the navigation hub and helps users identify the primary functions without searching through nested menus.

6.2 Sign-in Screen

The sign-in dialog contains email-address and password fields, a Sign In action and a tab for switching to account registration. It also displays instant demo-profile choices for a communicator, therapist and caregiver. These visible elements indicate that the design considers different user roles. In a deployed system, authentication must be connected to a secure identity service, passwords must never be stored as plain text, and demo profiles must be clearly separated from real user accounts.

6.3 Account Registration Screen

The registration dialog provides fields for full name, optional Tamil name, email address and password. It offers role selection for Communicator, Caregiver and Therapist, together with avatar choices. This design can support role-aware navigation and personalization. Role selection in the interface should not be interpreted as professional verification; access to sensitive records and caregiver functions requires appropriate authorization controls.

6.4 AAC Communication Board

The AAC board is the central picture-based communication interface. At the top, a sentence builder displays the number of selected words and provides a visible area where the message is assembled. The Speak Sentence action requests spoken output, while Copy, Undo and Clear support message editing. Category filters and symbol search help users locate a desired card. Each card presents an icon, an English phrase and, where available, a Tamil equivalent. The visible cards cover immediate needs, daily routines, food, comfort and assistance.

A typical interaction begins when the user chooses a category, selects one or more symbol cards, checks the sentence shown in the sentence builder and activates Speak Sentence. The interface also provides a way to remove the last selection or clear the sentence. The design should preserve the user's ability to correct a message before speaking it.

6.5 Sign Language Assistant

The sign-language assistant screen is divided into a live camera feed and a recognized-sign panel. The camera area includes a Start Camera control and a standby state. The result panel displays the recognized sign, offers an Auto Speak option and provides Speak and Add to Sentence actions. A supported-sign list shows sample vocabulary such as Hello, Thank You, Help, Water, Food / Eat and Yes, with Tamil labels. The interface also offers English and Tamil output selection.

This screen describes the intended interaction design, not proof of recognition accuracy. Actual sign recognition depends on the sign vocabulary, camera quality, lighting, hand position, model training data and inference implementation. A production system should identify the supported sign set clearly and should not imply that it can translate unrestricted sign language unless that capability has been demonstrated.

6.6 Voice Assistant

The voice assistant screen is titled Speech & Voice Communicator. It includes English and Tamil choices, a Test Voice control, a microphone area for dictation, a phrase-entry text box and a Speak Text action. The screen therefore brings together two complementary workflows: entering text and requesting speech, and using a microphone to dictate speech into text. Quick assistive phrases are also indicated below the main interaction panel.

Speech output depends on the selected synthesis engine and installed voice resources. Microphone dictation depends on a speech-recognition service, microphone permission and, in some implementations, network connectivity. The two functions should be tested separately and should provide clear feedback when permission, language or service availability is missing.

6.7 Communication History

The history screen presents communication logs and summary cards for total phrases, favorites, sign gestures and voice inputs. Filters include All, Favorites, AAC Board, Sign and Voice. A search field helps locate a past message. Individual entries show the source, time, language and phrase, with actions such as Speak, favorite, add, copy and delete. An Export CSV control and Clear All action are also visible.

A history feature can help a user or caregiver revisit messages and review vocabulary that has been used. However, selection counts are descriptive records only. They must not be used as a diagnostic score or as proof of a child's language development, communication ability or clinical progress. History data should be protected and should be collected only with an understandable purpose and suitable caregiver controls.

6.8 Settings and Preferences

The settings screen contains primary language preference cards for English and Tamil. A Speech Engine & Voice section displays a synthesizer voice selector, Test Speech control, speech-speed slider and volume slider. These controls provide a place to adjust speech presentation to user preference. The actual list of voices and the availability of Tamil speech depend on the operating system, browser or speech service installed on the device.

7. Functional Requirements

- FR-01: The application shall provide navigation to Home, AAC Communication, Sign Language Assistant, Voice Assistant, Communication History and Settings.
- FR-02: The AAC board shall display symbol cards grouped by category.
- FR-03: Each available symbol card shall have a linked message label.
- FR-04: Selecting a symbol shall add its message to the sentence builder.
- FR-05: The user shall be able to undo the latest selection and clear the current sentence.
- FR-06: The application shall provide a control to request speech output for the constructed sentence.

- FR-07: The voice assistant shall provide a text-entry field and a speech-output action.
- FR-08: Where speech recognition is implemented, the user shall be able to initiate microphone dictation after granting permission.
- FR-09: The sign assistant shall provide a camera start control and a result area for recognized signs.
- FR-10: The user shall be able to choose an available output language.
- FR-11: The history module shall display stored communication records and provide filtering and search controls.
- FR-12: The settings module shall expose available speech voice, speed and volume preferences.
- FR-13: Account and role functions shall restrict protected settings and records to authorized users.

8. Non-Functional Requirements

- **Usability:** Main actions should be easy to locate, labels should be understandable and navigation should remain consistent.
- **Accessibility:** Symbol cards should have sufficiently large touch targets, clear labels, adequate contrast and spacing. Colour should not be the only way to communicate meaning.
- **Responsiveness:** Selecting a symbol should update the sentence area without unnecessary delay.
- **Reliability:** The application should handle unavailable speech services, denied camera or microphone permissions, missing images and unsupported languages without crashing.
- **Privacy and security:** Personal details, communication history and camera or microphone data should be protected and processed only for stated purposes.
- **Portability:** The layout should adapt to supported desktop, tablet and mobile screen sizes.
- **Maintainability:** Vocabulary, translations, categories and supported signs should be managed as structured data.
- **Language support:** English and Tamil options should be shown only where suitable text and speech resources are available.

9. System Architecture

Kural AAC can be represented as a layered system. The presentation layer contains the dashboard, AAC board, sign assistant, voice assistant, history and settings screens. The application-logic layer manages navigation, symbol selection, sentence construction, permissions, language preferences and history operations. The data layer stores user profiles, vocabulary, localized labels, sign vocabulary, settings and communication records. The speech and recognition services layer connects to text-to-speech, speech recognition and the sign-recognition component.

Logical processing flow

```
User → Presentation Interface → Application Logic → (Vocabulary / User Data / History
Storage) → Speech or Recognition Service → Result and User Feedback
```

For AAC symbol selection, the application retrieves the selected symbol's label, appends it to the sentence builder and displays the updated message. When the user requests speech, the speech manager checks that the engine is ready and that the selected language is available before sending the text for synthesis. For voice dictation, the application requests microphone permission and sends captured audio to the configured recognition service. For sign assistance, camera frames are passed to the configured recognition component and the returned label is displayed for user confirmation.

The architecture is logical and technology-independent because the supplied project information does not specify the final programming framework, database or deployment environment. The implementation report should describe the actual technologies used in the source code rather than assume a particular stack.

10. Data Design

A structured data model can keep vocabulary and user preferences consistent across the different screens. The following entities describe the information required by the prototype design.

Entity	Important fields	Purpose
UserProfile	user_id, name, email, role, avatar	Stores account identity and role information.
Category	category_id, category_name, display_order	Groups AAC symbols into familiar topics.
SymbolItem	symbol_id, category_id, image_ref, label_en, label_ta	Stores a symbol and its language labels.
SignVocabulary	sign_id, sign_label, language_labels, model_label	Defines the sign labels supported by the sign module.
CommunicationLog	log_id, user_id, source, message, language, timestamp, favorite	Stores communication events when history is enabled.
AppSettings	user_id, language, voice_id, speech_rate, volume	Stores user speech and display preferences.

The exact database technology and storage location are implementation decisions. If communication logs are stored, the application should provide clear controls for viewing and deleting them. Only the minimum information needed for the communication purpose should be retained.

11. Core Algorithms and Interaction Workflows

11.1 AAC sentence construction

1. Load the available categories and symbol records.
2. Display the selected category or all available symbols.
3. When a symbol is selected, retrieve its linked label in the chosen language.
4. Append the label to the current sentence and refresh the visible sentence builder.
5. If Undo is selected, remove the most recently added item; if Clear is selected, empty the sentence.
6. If Speak Sentence is selected, pass the current sentence to the speech manager.
7. Display an understandable status if the sentence is empty or speech is unavailable.

11.2 Text-to-speech workflow

1. Receive the text to be spoken and the selected language.
2. Check whether a speech engine is available and initialized.
3. Check whether the selected language and voice are supported.
4. Apply the selected speech rate and volume where supported.
5. Submit the text to the speech engine.
6. Report completion or an error without removing the visible text.

11.3 Voice dictation workflow

1. The user activates the microphone control.
2. The application requests microphone permission if it has not already been granted.
3. Audio is captured and passed to the configured speech-recognition service.
4. The recognized text is shown for review and correction.
5. The user may request speech output or use the recognized text in the communication workflow.

11.4 Sign-language assistance workflow

1. The user starts the camera and grants camera permission.
2. The camera supplies frames to the sign-recognition component.
3. The recognition component compares the observed gesture with its supported vocabulary.
4. A candidate label is displayed in the recognized-sign panel.
5. The user reviews the label and chooses Speak or Add to Sentence.
6. If confidence is insufficient or no supported sign is detected, the interface should communicate that no reliable result is available.

12. Prototype Evaluation and Verification Plan

The supplied screenshots support a review of the screen layout and the presence of the listed interface controls. They do not provide sufficient evidence to confirm backend operations, speech quality, recognition accuracy, security or persistence. The following verification cases define how the prototype should be tested on its target device.

ID	Test	Expected observation
T01	Open the dashboard.	Dashboard and navigation items are displayed.
T02	Open the AAC board and choose a category.	Relevant symbol cards are displayed.
T03	Select multiple symbol cards.	Labels are appended in order to the sentence builder.
T04	Use Undo and Clear.	The sentence is updated or emptied correctly.
T05	Request speech for a sentence.	The selected voice speaks the text, or a clear error is shown.
T06	Change language to Tamil and test speech.	Tamil speech works if a compatible voice is installed; otherwise availability is reported.
ID	**Test**	**Expected observation**
T07	Start sign camera and test known supported signs.	The system displays a candidate label only when recognition returns a result.
T08	Use microphone dictation.	Recognized text is displayed, or permission/service errors are handled.
T09	Search and filter communication history.	Matching records and selected filters are displayed.
T10	Change speech rate and volume.	Settings are applied where supported by the speech engine.
T11	Sign out and access a protected area.	Protected information is not available without authorization.

Evaluation should record the device, operating-system or browser version, speech engine, language resources, permissions, test inputs and observed outcomes. Sign recognition should be evaluated using a defined set of signs and multiple recording conditions. Speech recognition should be evaluated separately from text-to-speech. No numerical

accuracy, user satisfaction score or clinical outcome is claimed in this paper.

13. Accessibility, Ethics and Privacy

AAC vocabulary should reflect the communication needs and preferences of the person using it. It should include more than requests for objects and should make room for social interaction, refusal, questions, feelings, help and personal choices. Symbols should be familiar to the individual; generic images may need to be replaced with photographs or culturally relevant symbols. The interface should allow sufficient time for selection and should avoid relying on colour alone.

Camera and microphone functions require clear permission prompts and visible indicators. Where possible, processing should occur locally; if data is sent to a remote service, the user should be informed about what is transmitted and why. Communication history can contain sensitive information and should be protected with access controls, retention limits and deletion options. Demo accounts should not contain real personal information.

The system should support, rather than replace, a person's existing communication methods. It should not be used to diagnose a condition or infer developmental progress from the number of symbol selections. Any study involving children or other participants should follow institutional ethics requirements and obtain appropriate consent and assent.

14. Limitations

- The screenshots show the prototype interface but do not establish the reliability or completeness of the underlying services.
- Sign-language recognition is limited by the defined sign vocabulary, training data, camera conditions and recognition model.
- Speech recognition and text-to-speech depend on permissions, operating-system support, installed voices and service availability.
- Tamil speech output may not be available on every device even when Tamil text is displayed correctly.
- The visible history and account screens do not independently confirm secure authentication, persistent storage or export behaviour.
- The current paper does not claim clinical validation or completed evaluation with AAC users.

15. Future Enhancements

Future development can expand the symbol library with personalized photographs, topic-specific vocabulary and carefully reviewed Tamil translations. The sentence builder can be improved with phrase prediction that remains user-controlled. Alternative access methods such as switch scanning, adjustable touch targets and partner-assisted selection can make the interface more flexible. The sign assistant can be improved through a clearly defined sign vocabulary, representative datasets and evaluation under varied lighting and camera positions. Voice features can be enhanced with offline speech resources where available, clear permission handling and user-adjustable speech settings. Further work should include participatory design with AAC users, caregivers, educators and speech-language professionals.

16. Conclusion

This paper has documented Kural AAC, a multi-module assistive communication prototype that brings together a picture-based AAC board, sentence construction, sign-language assistance, voice assistance, communication history and language and speech settings. The supplied interface screens show how users can navigate the platform and access its main communication tools. The design emphasizes visible messages, category-based vocabulary, English and Tamil presentation, and complementary voice and sign-related workflows. The prototype provides a foundation for implementation refinement and systematic evaluation. Functional reliability, sign-recognition performance, speech-service availability, privacy safeguards and usability with intended users must be verified before the

platform is presented as a validated assistive solution.

References

[1] American Speech-Language-Hearing Association, "Augmentative and Alternative Communication (AAC)," Practice Portal. Available: https://www.asha.org/practice-portal/professional-issues/augmentative-and-alternative-communication/

[2] World Health Organization and UNICEF, "Global report on assistive technology," 2022. Available: https://www.who.int/publications/i/item/9789240049451

[3] Android Developers, "TextToSpeech," Android API Reference. Available: https://developer.android.com/reference/android/speech/tts/TextToSpeech

[4] World Wide Web Consortium (W3C), "Web Content Accessibility Guidelines (WCAG) 2.2," 2023. Available: https://www.w3.org/TR/WCAG22/

Appendix A. Prototype Interface Screens

The following figures reproduce the screenshots supplied for the Kural AAC prototype. Captions describe the visible interface and do not imply that every displayed control has been functionally verified.

Figure A1. Kural AAC home dashboard with navigation and core communication tools.

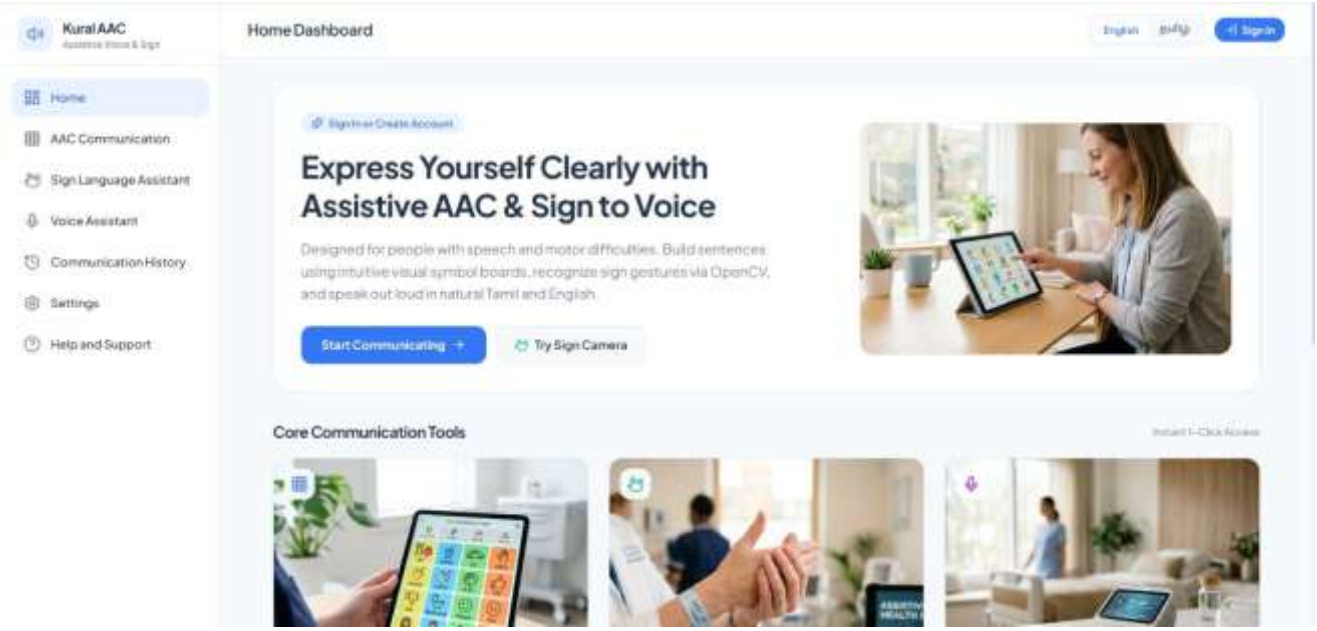


Figure A2. Sign-in dialog with email and password fields and demo profile choices.

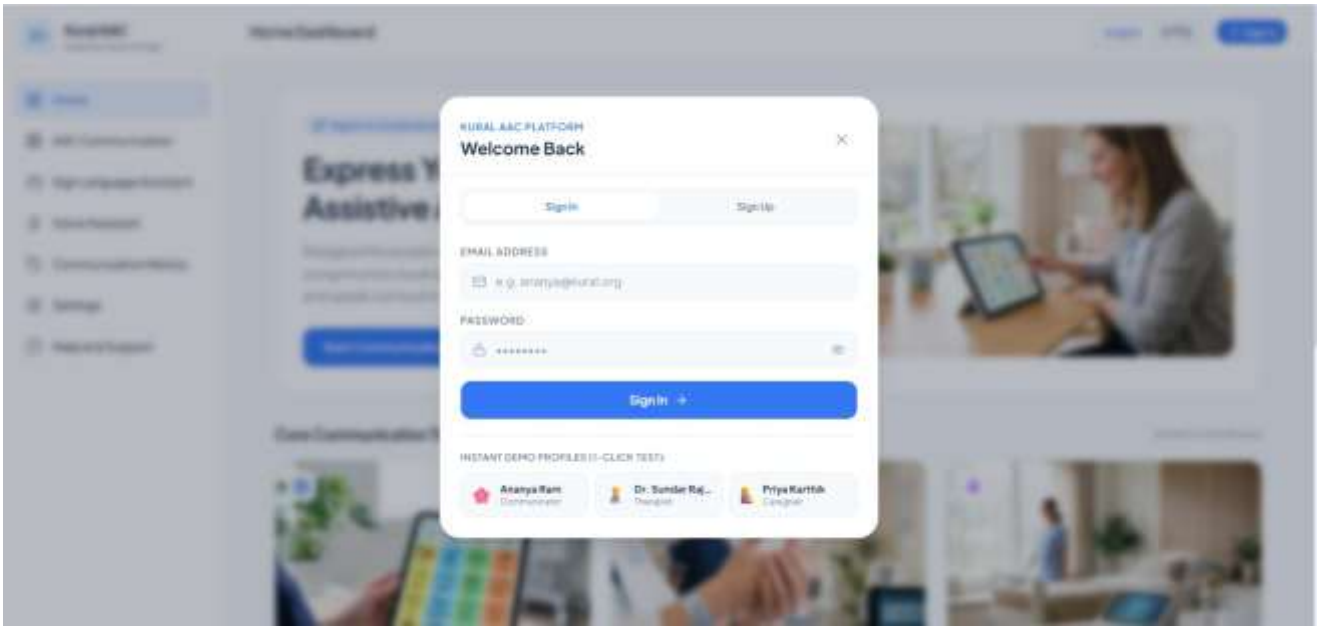


Figure A3. Account registration interface with role selection and avatar choices.

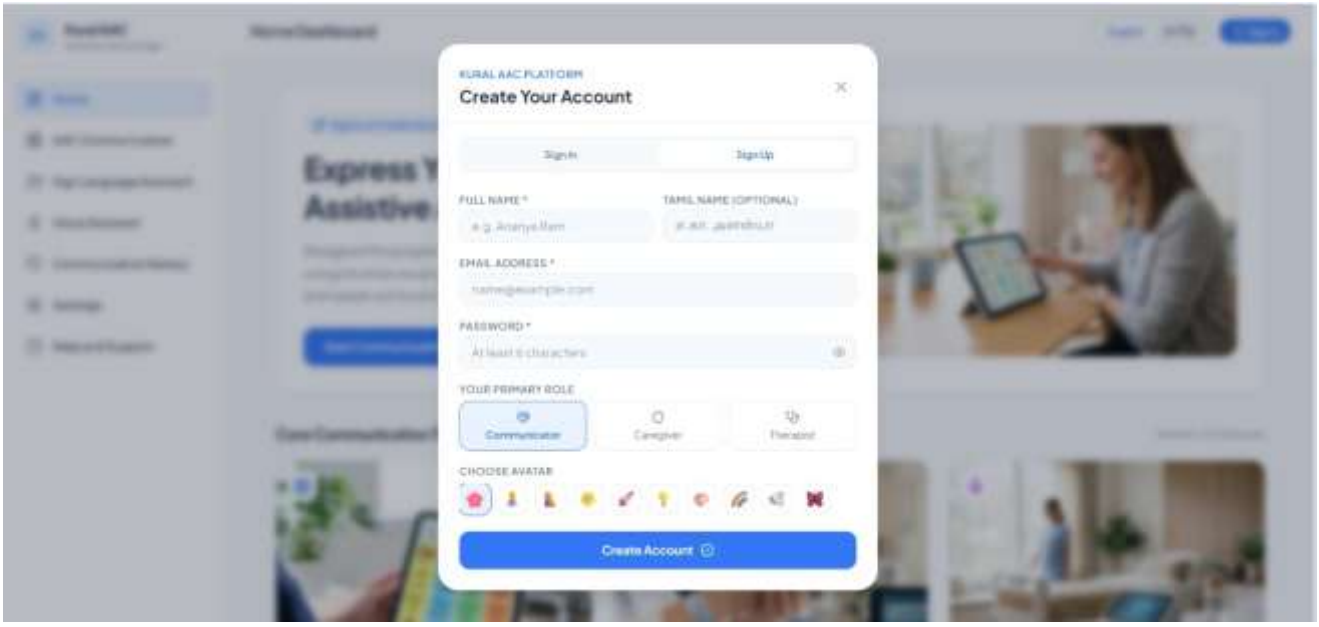


Figure A4. AAC communication board with sentence builder, category filters and symbol cards.

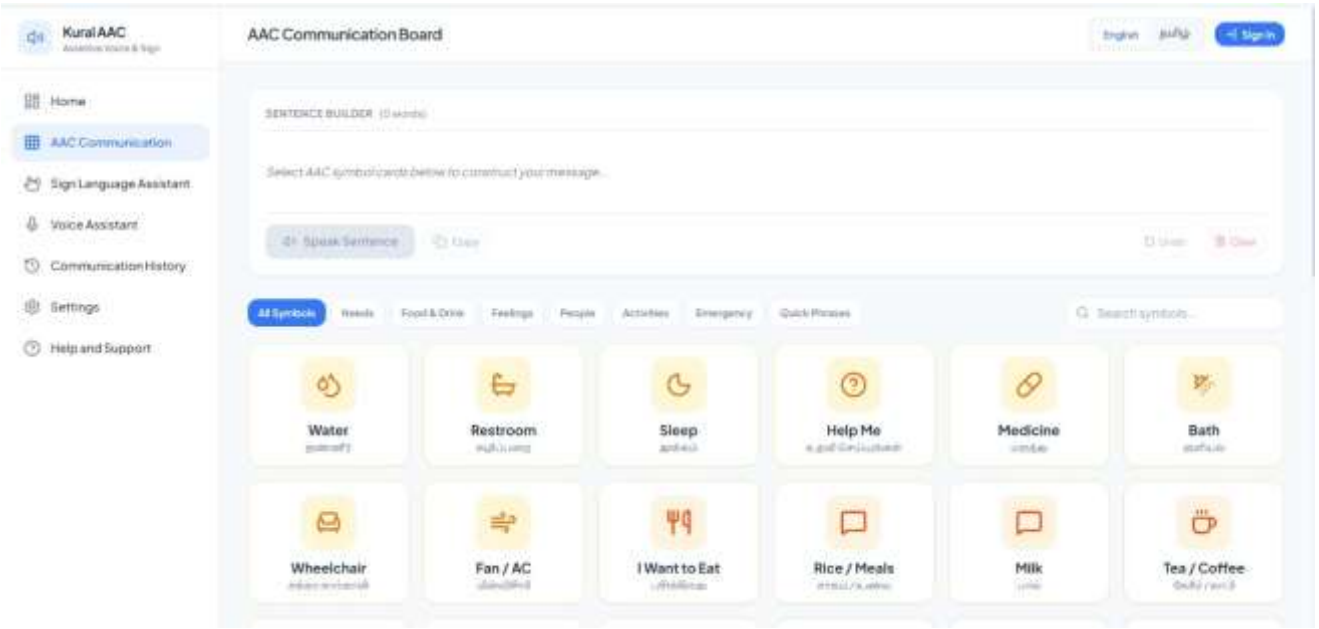


Figure A5. Sign-language assistant with camera area, recognized-sign panel and supported sign list.

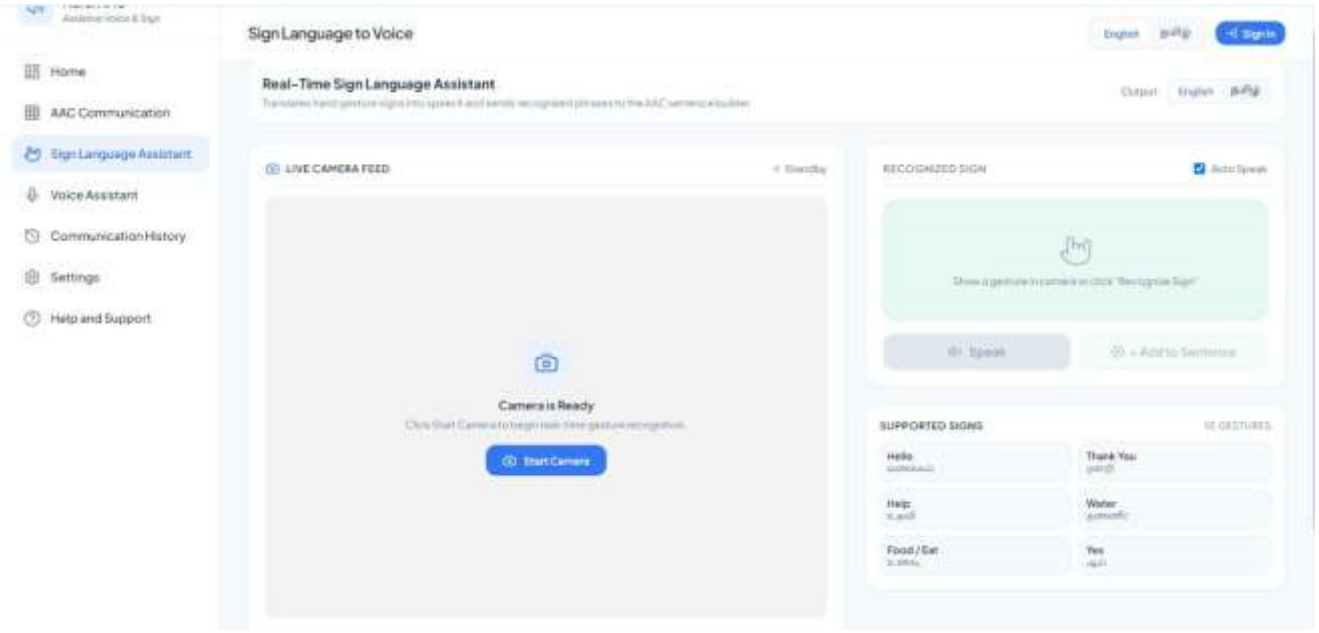


Figure A6. Voice assistant with microphone dictation, phrase entry and speech controls.

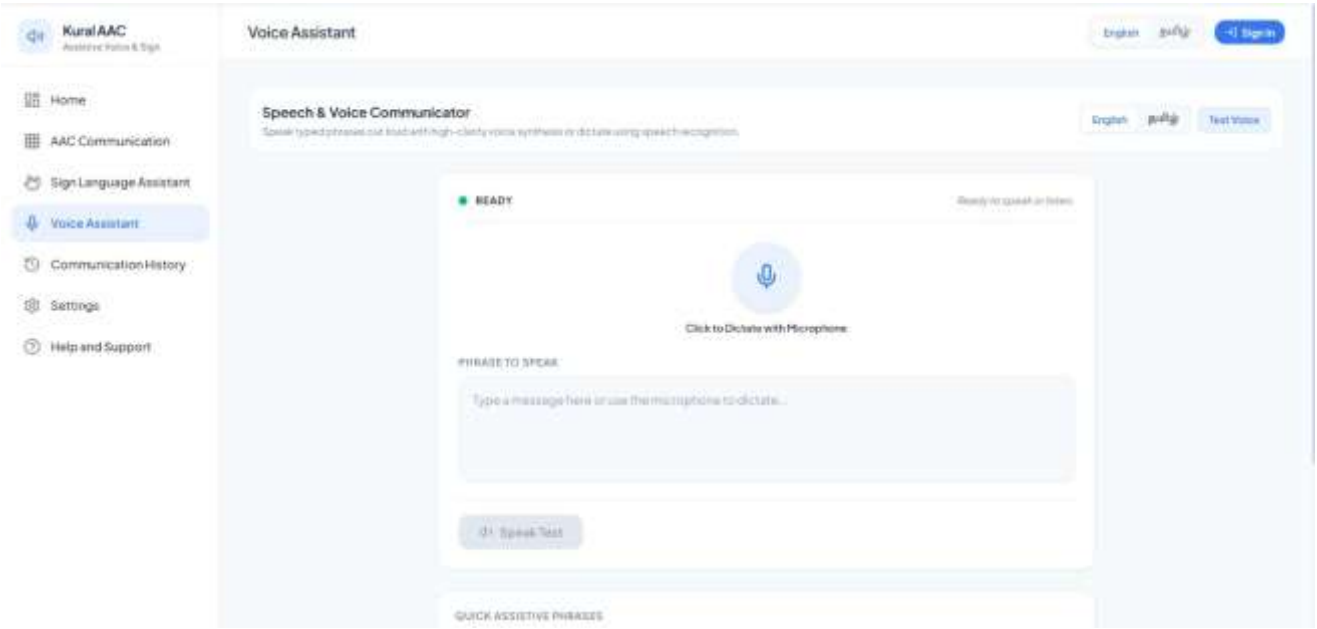


Figure A7. Communication history with phrase summaries, filters and individual log entries.

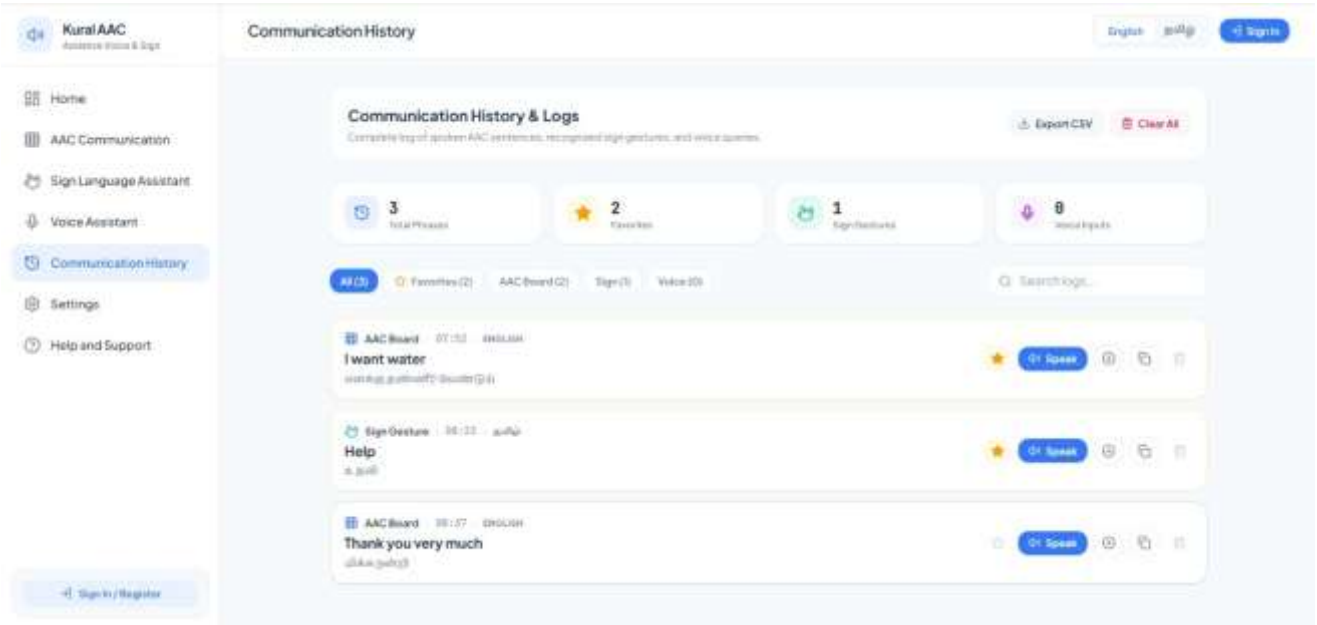


Figure A8. Settings screen with language preference and speech voice, speed and volume controls.

