electronics

Article

SOUTY: A Voice Identity-Preserving Mobile Application for Arabic-Speaking Amyotrophic Lateral Sclerosis Patients Using Eye-Tracking and Speech Synthesis

Hessah A. Alsalamah [1,*], Leena Alhabrdi [1], May Alsebayel [1], Aljawhara Almisned [1], Deema Alhadlaq [1], Loody S. Albadrani [1], Seetah M. Alsalamah [2] and Shada AlSalamah [1,*]

[1] Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[2] Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
* Correspondence: halsalamah@ksu.edu.sa (H.A.A.); saalsalamah@ksu.edu.sa (S.A.)

Abstract

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disorder that progressively impairs motor and communication abilities. Globally, the prevalence of ALS was estimated at approximately 222,800 cases in 2015 and is projected to increase by nearly 70% to 376,700 cases by 2040, primarily driven by demographic shifts in aging populations, and the lifetime risk of developing ALS is 1 in 350–420. Despite international advancements in assistive technologies, a recent national survey in Saudi Arabia revealed that 100% of ALS care providers lack access to eye-tracking communication tools, and 92% reported communication aids as inconsistently available. While assistive technologies such as speech-generating devices and gaze-based control systems have made strides in recent decades, they primarily support English speakers, leaving Arabic-speaking ALS patients underserved. This paper presents SOUTY, a cost-effective, mobile-based application that empowers ALS patients to communicate using gaze-controlled interfaces combined with a text-to-speech (TTS) feature in Arabic language, which is one of the five most widely spoken languages in the world. SOUTY (i.e., "my voice") utilizes a personalized, pre-recorded voice bank of the ALS patient and integrated eye-tracking technology to support the formation and vocalization of custom phrases in Arabic. This study describes the full development life cycle of SOUTY from conceptualization and requirements gathering to system architecture, implementation, evaluation, and refinement. Validation included expert interviews with Human–Computer Interaction (HCI) expertise and speech pathology specialty, as well as a public survey assessing awareness and technological readiness. The results support SOUTY as a culturally and linguistically relevant innovation that enhances autonomy and quality of life for Arabic-speaking ALS patients. This approach may serve as a replicable model for developing inclusive Augmentative and Alternative Communication (AAC) tools in other underrepresented languages. The system achieved 100% task completion during internal walkthroughs, with mean phrase selection times under 5 s and audio playback latency below 0.3 s.

Keywords: amyotrophic lateral sclerosis; eye tracking; text-to-speech; Arabic language; assistive technology; mobile health; human–computer interaction; speech synthesis; gaze interaction; augmentative and alternative communication

check for updates

Academic Editor: Yuankai Qi

Received: 11 July 2025
Revised: 9 August 2025
Accepted: 12 August 2025
Published: 14 August 2025

Citation: Alsalamah, H.A.; Alhabrdi, L.; Alsebayel, M.; Almisned, A.; Alhadlaq, D.; Albadrani, L.S.; Alsalamah, S.M.; AlSalamah, S. SOUTY: A Voice Identity-Preserving Mobile Application for Arabic-Speaking Amyotrophic Lateral Sclerosis Patients Using Eye-Tracking and Speech Synthesis. *Electronics* **2025**, *14*, 3235. https://doi.org/10.3390/electronics14163235

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Amyotrophic Lateral Sclerosis (ALS) is a rapidly progressive neurodegenerative disease that leads to the degeneration of upper and lower motor neurons in the brain and spinal cord. As ALS advances, patients experience progressive muscle weakness and paralysis, eventually losing voluntary control of their limbs, facial muscles, and the ability to speak. Although cognitive functions often remain intact, the loss of verbal and physical expression significantly impairs autonomy, mental health, and quality of life [1].

ALS has a significant and growing public health impact. In the United States, the number of ALS cases was estimated at 32,893 in 2022 and is projected to exceed 36,300 by 2030 due to aging demographics [2]. Globally, prevalence was approximately 222,800 cases in 2015 and is expected to rise to 376,700 by 2040 [3]. The lifetime risk of developing ALS is estimated at 1 in 350–420 individuals [4]. These figures underscore the need for scalable, inclusive, and culturally adapted communication technologies.

Arabic is the native language of approximately 313–373 million people, with over 400 million using it daily, making it one of the five most widely spoken languages in the world [5]. In Saudi Arabia, the fourth-largest Arabic-speaking population globally [6], a 2020 national survey of ALS care providers reported that 100% lacked access to eye-tracking communication tools, and over 92% noted inconsistent availability of speech-generating devices [4]. Furthermore, nearly half of providers reported limited access to speech-language pathologists, and delays in respiratory intervention were common. These findings underscore systemic gaps in local ALS care, particularly in the availability of culturally appropriate Augmentative and Alternative Communication (AAC) technologies.

The inability to communicate effectively creates a profound emotional and psychological burden for both patients and caregivers. Communication loss isolates patients socially and limits their ability to make decisions regarding care, daily preferences, and emotional expression. Consequently, the development of assistive technologies that can bridge this gap is not only a technical challenge but a human imperative.

AAC systems offer pathways for restoring communication by enabling users to express themselves through non-verbal means. These systems range from low-tech communication boards to advanced digital solutions that use text-to-speech (TTS), gesture recognition, or brain–computer interfaces (BCIs). Among these, eye-tracking has emerged as one of the most viable modalities for late-stage ALS patients, who may retain eye movement long after losing control of other muscles [7,8].

Modern gaze-based AAC platforms allow users to select words or phrases on a digital interface using eye fixation, which are then vocalized via synthesized speech. These systems typically rely on dwell-time mechanisms, where maintaining gaze on a particular icon for a set duration triggers selection. However, despite the technological progress, the majority of AAC solutions remain focused on English-speaking users and Western cultural contexts.

This creates significant barriers for individuals from linguistically and culturally diverse backgrounds, particularly Arabic speakers. Arabic is a morphologically rich and scriptually complex language with numerous regional dialects. Existing AAC systems often do not support right-to-left (RTL) scripts or provide culturally resonant vocabulary. Moreover, most Arabic TTS systems rely on Modern Standard Arabic, which lacks the emotional nuance and familiarity of regional spoken dialects.

Another limitation of conventional AAC tools is the use of synthetic speech engines, which produce robotic and impersonal voices. For individuals who lose their voice due to ALS, this can create a sense of detachment from their own identity. Voice banking—recording a person's voice in advance and using it later through concatenated or parametric synthesis—has emerged as a valuable solution to preserve vocal identity. However, such

services are not widely available in Arabic and are rarely integrated into mobile, offline-capable platforms.

To address these critical gaps, we introduce SOUTY, a novel voice identity preserving mobile application designed to support Arabic-speaking patients with ALS. The name SOUTY is derived from the Arabic word for "my voice", symbolizing the application's mission to restore expressive communication for individuals who have lost the ability to speak. SOUTY is an iPhone Operating System (iOS)-based, offline-capable AAC system that combines eye-tracking with a pre-recorded Arabic voice bank to deliver a personalized, culturally relevant communication experience. The system allows users to select predefined phrases, navigate phrase categories, and construct custom messages using an on-screen Arabic keyboard—all controlled via gaze alone.

SOUTY emphasizes accessibility, dignity, and linguistic inclusion. It leverages existing mobile hardware (iPhones/iPads with front-facing cameras) and requires no external devices or network connection. The use of real voice recordings from an Arabic-speaking ALS patient enhances the emotional connection and cultural authenticity of the speech output. This paper presents the full lifecycle of SOUTY's development, from conceptual design to system implementation and evaluation. Expert interviews with speech-language pathologists and Human–Computer Interaction (HCI) professionals were conducted to validate its clinical applicability and usability. In addition, a public survey was conducted to assess awareness of ALS and the perceived value of eye-tracking-based AAC tools. The development was guided by user-centered design (UCD) principles and aligned with established HCI models that prioritize accessibility, feedback, and personalization.

The remainder of this paper is organized as follows. Section 2 provides a comprehensive review of existing AAC technologies, with a particular focus on the limitations of Arabic-language assistive tools. Section 3 outlines the methodology adopted for the development and validation of the system. Section 4 presents the system architecture and interaction design of SOUTY. Section 5 details the implementation process, including key modules and software tools. Sections 6 and 7 report the results of technical testing and user-centered evaluation, respectively. Section 8 offers a critical discussion of design implications, limitations, and opportunities for future enhancement. Finally, Section 9 concludes the paper and summarizes key contributions.

2. Related Work

The development of assistive technologies for individuals with ALS has accelerated in recent years, driven by advances in eye-tracking, speech recognition, brain–computer interface (BCI), and artificial intelligence (AI). These technologies seek to address the severe motor and speech impairments caused by ALS by enabling alternative forms of communication and control. This section provides an overview of the most relevant research on assistive input modalities, multimodal systems, machine learning-enhanced communication platforms, and the critical lack of support for Arabic-speaking users. The final subsection discusses the design motivations behind SOUTY in light of the observed gaps.

2.1. Eye-Tracking and Gaze-Based Control

Eye-tracking systems are widely used in assistive technologies due to their non-invasive nature and ability to serve users with limited or no limb movement. Edughele et al. [7] reviewed eye-tracking solutions tailored for ALS patients and highlighted their application in communication, gaming, and environmental control. They noted that integrating artificial intelligence improves gaze estimation accuracy and reduces calibration time, making systems more usable for individuals with advanced motor impairments.

Fischer-Janzen et al. [9] conducted a comprehensive scoping review of gaze-controlled robotic arms, classifying systems based on input modality (Video Oculography, Infrared Oculography, Electrooculography) and system complexity. Their findings reinforce the need for robust, user-friendly gaze interaction systems that are adaptable to real-world environments and user fatigue.

2.2. Speech Recognition and Audiovisual Monitoring

Speech recognition remains a viable input modality for ALS patients in the early stages of bulbar involvement. Cave and Bloch [10] found that speech interfaces work well for individuals with mild dysarthria but degrade significantly as speech impairments progress. This supports the use of speech as a supplementary modality rather than a standalone solution in AAC systems.

Neumann et al. [11] introduced NEMSI, a scalable platform for ALS symptom monitoring using audiovisual signals such as pitch, speaking rate, and facial motion asymmetry. Their work shows how communication tools can double as clinical monitoring platforms, enabling longitudinal tracking of disease progression through natural interaction.

Hoyle et al. [12] proposed a novel input modality called EarSwitch, which uses voluntary contractions of the tensor tympani muscle to generate ear rumbling as a signal. Their large-scale study demonstrated that ear-based control could complement or substitute for gaze tracking, especially for users experiencing eye fatigue or blinking difficulties.

2.3. Multimodal Assistive Systems and BCI Integration

Combining multiple input modalities can improve accessibility and system resilience. Bonanno et al. [8] emphasized the need for multimodal systems that integrate gaze, speech, BCI, and haptics to meet the evolving needs of people with neurological conditions. They called for user-centered, AI-enhanced assistive technologies that support dynamic adaptation and personalization.

Kew et al. [13] conducted a systematic review on the application of machine learning and BCIs in ALS, identifying Random Forests and other models as effective predictors of disease progression. Their work supports the development of intelligent AAC systems that respond in real time to physiological signals or intent prediction.

Zhou et al. [14] introduced the Augmented Body Communicator, a system that combines robotic augmentation and large language models (LLMs) to support gesture-driven communication for individuals with upper-limb limitations. This approach highlights the convergence of multimodal control and AI for expressive interaction beyond conventional AAC.

2.4. Arabic-Language Assistive Technologies

Despite advances in assistive technologies, most commercially available AAC tools are designed for English-speaking users. Tools such as Hawkeye Access [15], Vocable AAC [16], and EyeTech [17] provide gaze-based interaction but lack Arabic language support, localized phrase sets, and cultural tailoring. Similarly, while Tobii provides high-quality gaze hardware [18], it is primarily configured for Western languages and interfaces.

Attempts to build Arabic-compatible systems such as iWriter [19] demonstrate the feasibility of Arabic gaze typing but are limited by slow interaction speeds and a lack of dialectal voice synthesis. Most Arabic TTS engines rely on Modern Standard Arabic, which lacks the emotional nuance and familiarity of regional dialects.

Recent reviews of TTS technologies based on deep learning [20] show great promise in creating expressive synthetic voices. However, these systems typically require large datasets and significant computing resources, which are unavailable for Arabic dialects. This limits their utility in building culturally resonant AAC tools.

2.5. Design Motivation: Addressing Gaps Through SOUTY

Londral [1] emphasized that assistive technologies empower patients not only to communicate but also to engage with caregivers and participate in healthcare decisions. However, current AAC systems face several persistent limitations:

- Predominant focus on English and Western language users.
- Lack of dialect-specific Arabic voice banks.
- Dependence on synthetic speech, which reduces emotional connection.
- Limited offline operation, making tools less accessible in low-connectivity settings.

SOUTY was developed to address these gaps through

- A fully Arabic-language interface and phrase set designed for Saudi users;
- A pre-recorded Arabic voice bank from an ALS patient, enabling personalized and culturally resonant speech;
- Offline functionality that allows uninterrupted use without internet access;
- A lightweight, gaze-only interface deployable on mainstream iOS devices.

SOUTY contributes to the growing field of inclusive AAC design by offering a replicable model for developing gaze-based communication tools in underrepresented languages and cultures. Its integration of eye-tracking, natural speech playback, and mobile deployment reflects current best practices in user-centered, AI-aligned assistive technology. Moreover, the relevance of accessibility and interaction standards such as ISO 9241 [21] in shaping the usability dimensions of AAC systems is acknowledged.

3. Methodology

The development of SOUTY followed a UCD methodology grounded in iterative development and interdisciplinary collaboration. This approach ensured that both technical and clinical perspectives were integrated throughout the lifecycle, from conceptualization to evaluation.

3.1. Methodological Framework

The project was structured in four key phases:

1. Requirement Analysis: We conducted a focused literature review and consulted with domain experts—including speech-language pathologists and HCI specialists—to understand the needs of Arabic-speaking ALS patients and the limitations of existing AAC systems.
2. System Design and Prototyping: Based on initial requirements, we developed interaction flows, a modular system architecture, and interface wireframes. Design emphasis was placed on gaze-based interaction, cultural and linguistic relevance, offline operability, and personalization via voice banking.
3. Implementation: The prototype was developed using Swift and Xcode for iOS, employing Apple's ARKit and AVFoundation frameworks. All system components (eye-tracking, phrase selector, audio output, keyboard input, and database management) were built modularly for extensibility and testability.
4. Validation and Iteration: SOUTY was evaluated through qualitative expert interviews, an online public awareness survey, and internal usability walkthroughs. The feedback from these evaluations informed refinements to the UI, audio engine, phrase bank, and calibration process.The sample size for expert interviews (n = 2) was deemed sufficient for initial feedback as both individuals represent key domains (speech pathology and HCI). The public survey sample (n= 236) was collected via social media and academic channels using convenience sampling. The questionnaire underwent

internal pilot testing with 10 participants to ensure clarity, and descriptive statistics were used to summarize readiness and awareness.

Given the exploratory scope and the use of non-probability convenience sampling, the survey results were analyzed using descriptive statistics only. The primary objective was to establish baseline awareness and readiness levels to inform subsequent hypothesis-driven studies. Inferential statistical tests (e.g., chi-square) were not applied, as the sample was not designed to be representative of the broader population and such analyses would have limited generalizability conclusions at this stage.

3.2. Ethical Considerations

Voice recordings were collected with informed consent from an ALS patient under supervision from the clinical team. No personal data was collected during development or evaluation, and all design elements were reviewed for cultural sensitivity and accessibility by qualified professionals.

4. System Design and Architecture

SOUTY was designed with a user-centered philosophy to meet the specific needs of Arabic-speaking individuals with ALS. The system aims to provide an intuitive, accessible, and offline-capable communication aid that leverages eye-tracking and real voice playback. This section outlines the design goals, system architecture, interface layout, and core components that enable SOUTY to function effectively across technical and usability dimensions.

4.1. Design Objectives

The design of SOUTY was driven by the following objectives:

- Provide a gaze-controlled communication platform for ALS patients who have lost the ability to speak.
- Support Arabic as the primary language with culturally relevant phrase sets.
- Enable speech playback using a real recorded Arabic voice instead of synthetic speech.
- Run offline on iOS devices without requiring network access or external hardware.
- Offer both predefined phrases and custom sentence construction through a gaze-controlled keyboard.
- Prioritize high-utility phrases (e.g., caregiver interactions, daily needs, pain, or discomfort) to reduce time-to-speech and ensure rapid communication in common scenarios.

4.2. System Architecture

Figure 1 illustrates the overall architecture of SOUTY. The system is composed of the following interconnected modules:

- Eye-Tracking Module: Utilizes the front-facing camera and ARKit to track gaze and detect dwell-time-based selections.
- Phrase Selector: Displays categorized, gaze-selectable Arabic phrases.
- Custom Keyboard: Allows the user to construct new phrases using an on-screen Arabic keyboard, also controlled by gaze.
- Audio Output Module: Maps selected phrases to audio files and plays them using AVFoundation.
- Voice Bank Repository: Stores pre-recorded WAV files, enabling natural speech output based on phrase selection.
- Logic Controller: Manages interaction flow, calibrates gaze input, and routes user selections to the appropriate output handler.

The Logic Controller uses threshold-based dwell-time validation (2 s) and exception handling for incomplete input cycles. Future iterations are designed to support platform

migration through modular APIs, with Android compatibility under consideration. All processing is handled locally on the device, ensuring privacy and offline functionality. The architecture supports communication between the Logic Controller and the Voice Bank Repository, especially for validating phrase existence and retrieving file metadata.

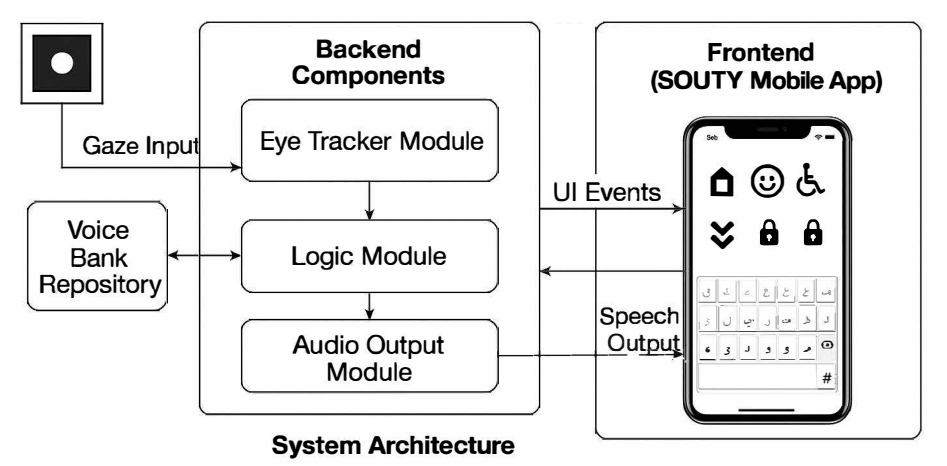


Figure 1. System architecture of SOUTY: This shows the interaction between the eye-tracking module, logic controller, phrase selector, and voice output, designed for Arabic-speaking ALS users.

4.3. User Interface (UI) Design

SOUTY's interface consists of a main dashboard with categorized phrase tiles and a navigation bar for switching to the custom keyboard. A gaze dwell-time of 2 s was established through internal calibration as optimal for accuracy without causing unintentional selections. All visual elements support RTL alignment for Arabic language users.

Icons were used in place of text on the phrase tiles to minimize reading effort and maximize clarity. When a phrase is selected, the corresponding voice recording is played instantly. Users can switch to the keyboard interface to spell out specific needs or names, with each character selected via gaze.

4.4. System Logic and Workflow

At startup, the system calibrates gaze zones and loads the phrase and audio banks into memory. During runtime, gaze coordinates are continuously monitored. If a user fixates on a UI element for the dwell threshold duration, a trigger is sent to the Logic Controller. Depending on the selected input (phrase or letter), the system routes the request either to the audio module or continues input collection for sentence construction.

The use case diagram (Figure 2) illustrates how different user intents are mapped to application actions.

This unified system design enables SOUTY to offer a consistent, low-friction communication experience tailored to the cultural and linguistic context of Arabic-speaking ALS users.

This architecture laid the foundation for the subsequent implementation of SOUTY, ensuring each component was modular, efficient, and aligned with the system's communication goals.

With the design goals and system architecture established, the next section details the implementation of each module, including software tools, calibration techniques, and audio integration.

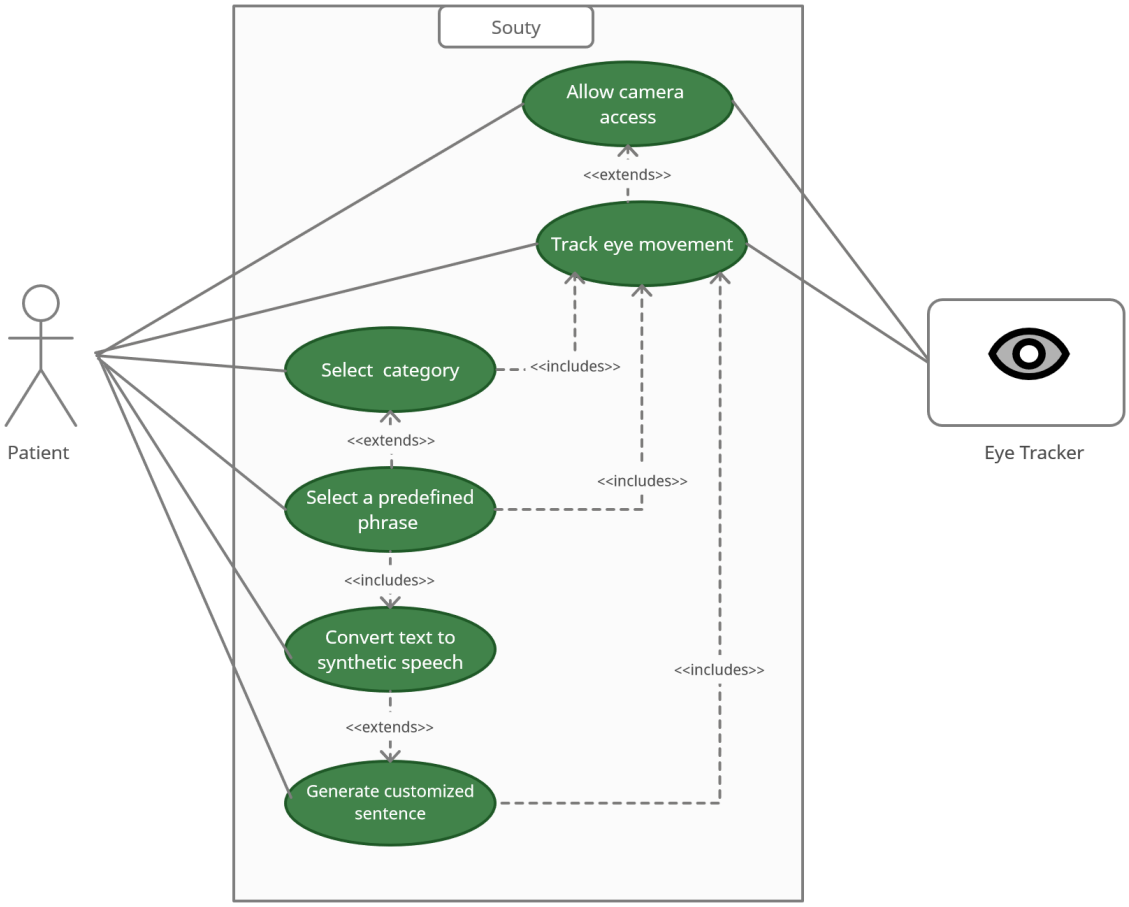


Figure 2. Use case diagram showing user interaction paths in SOUTY, including phrase selection, keyboard input, and audio playback via gaze control.

5. System Implementation

The SOUTY application was implemented as a native iOS solution using the Swift programming language and the Xcode development environment. The modular software architecture translates key functional components—gaze tracking, phrase selection, custom input, and Arabic voice playback—into cohesive system logic. Implementation followed an iterative, test-driven development process.

5.1. Platform and Tools

Development was conducted using Xcode (v14) on macOS, with deployment targeting iPhones and iPads equipped with TrueDepth cameras. SOUTY requires iOS 14 or later and is optimized for iPhone X and newer models with TrueDepth cameras. The following iOS frameworks were used:

- ARKit: ARKit was used to track facial geometry through TrueDepth camera input. Specifically, ARKit's `ARFaceAnchor` and eye transform data were used to estimate the user's gaze direction and map it to on-screen coordinates. Gaze estimation was refined using device orientation and screen layout context to ensure robust interaction during natural head movement.
- AVFoundation: For precise playback of WAV files corresponding to selected phrases.
- UIKit: For rendering gaze-interactive RTL UI components and managing views.
- SQLite: For lightweight local database management of phrase categories and associated audio paths.

5.2. Modular Software Components

SOUTY is structured into five key modules:

1. Eye Tracking Module: Captures user gaze vectors using ARKit's face anchors. Coordinates are mapped to interactive screen zones.
2. UI Manager: Manages RTL interface layouts, tile highlighting, keyboard display, and gaze-based dwell timers.
3. Audio Output Module: Interfaces with AVFoundation to load and play voice recordings.
4. Phrase Manager: Interfaces with SQLite to retrieve phrases and audio mappings based on user selections.
5. Voice Bank Repository: Contains WAV files recorded in Saudi Arabic by a real ALS patient, mapped to corresponding semantic phrases.

Figure 3 shows the gaze-controlled phrase selection interface. Figure 4 illustrates the on-screen Arabic keyboard that supports custom message creation via gaze selection.

The Logic Controller manages the flow of user interaction, transitioning between phrase selection, keyboard input, and playback. It interfaces with all other modules and is responsible for coordinating input processing and error handling.

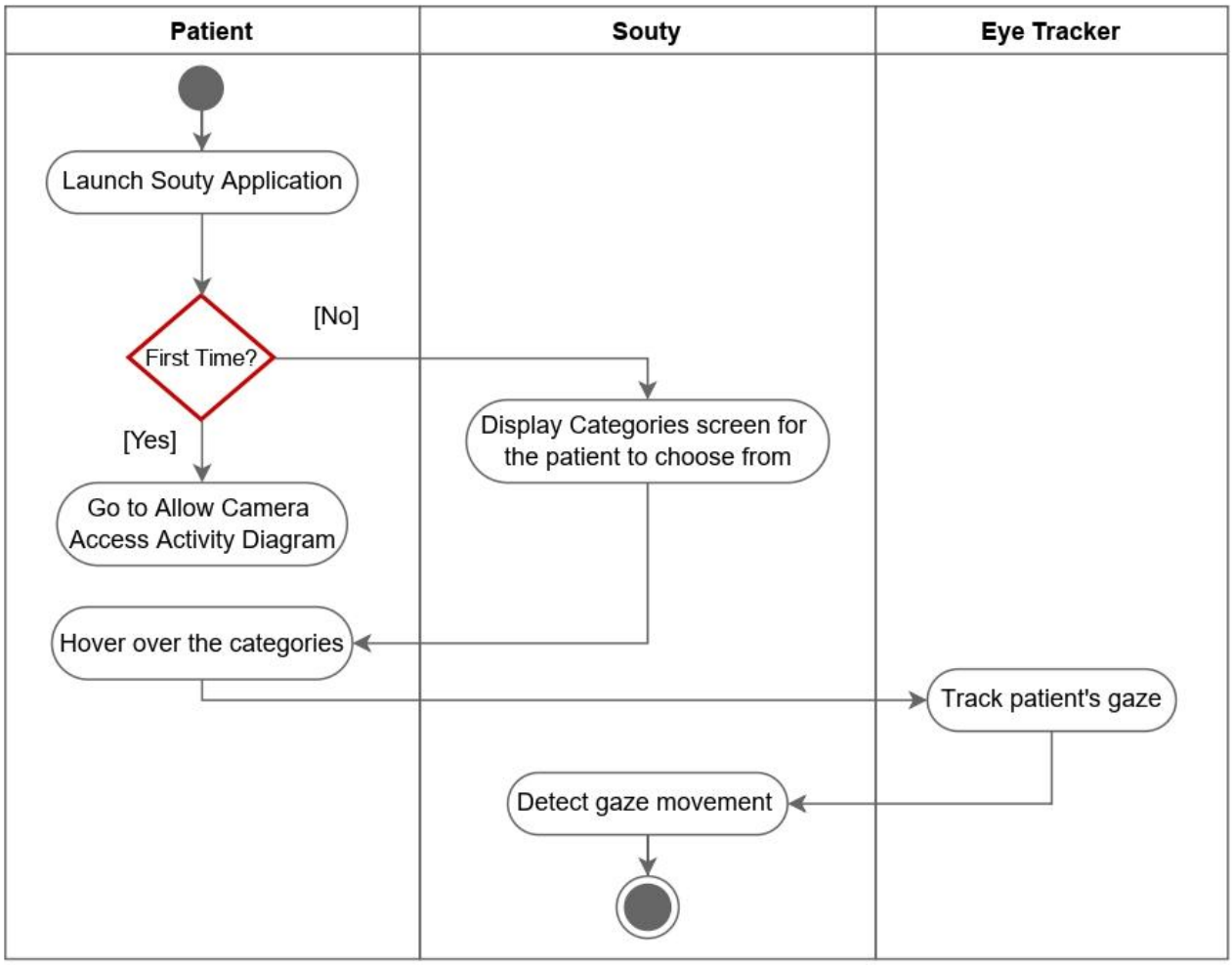


Figure 3. Phrase selection interface with categorized tiles and gaze highlighting.

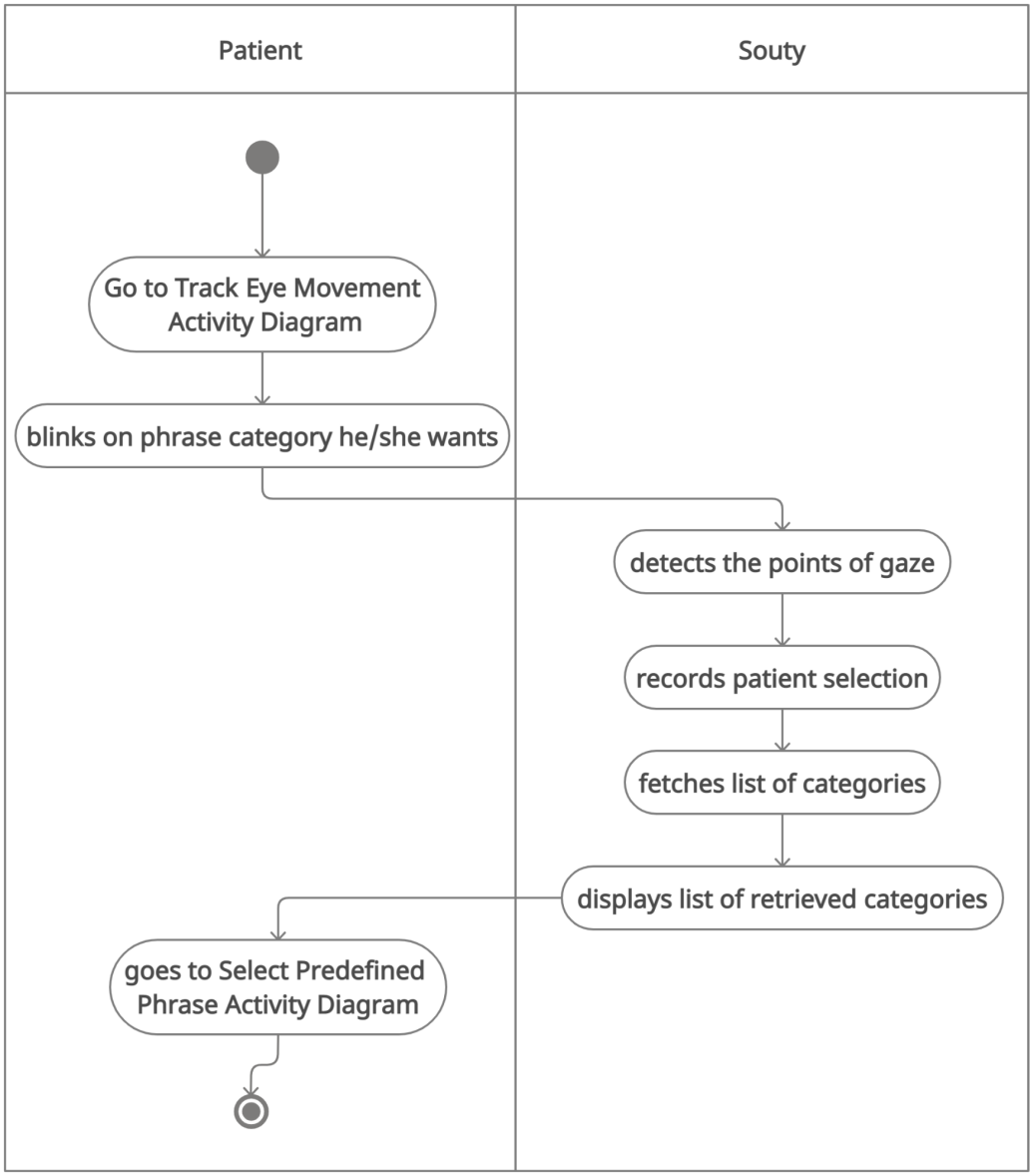


Figure 4. Custom Arabic keyboard screen used for free-form phrase input.

5.3. Calibration and Interaction Flow

SOUTY's eye-tracking functionality relies on Apple's ARKit framework, which uses facial tracking to estimate the direction of the user's gaze in real time. Upon app launch, a five-point calibration sequence maps estimated gaze vectors to defined screen coordinates. The system then monitors continuous gaze input and identifies a target selection when the user maintains their gaze on a UI element for at least 2 s (dwell-time threshold). Hover highlighting provides visual feedback during this dwell period, and once completed, the selection is passed to the Logic Controller for phrase retrieval and playback.

Upon launch, the system enters a gaze calibration routine, prompting the user to focus on five sequential screen points. Coordinates are validated based on gaze stability (~500 ms) and spatial accuracy (~±1.5 cm deviation). Successful calibration triggers UI activation.

During operation, a 2-second dwell time threshold is used for selection. Hovered UI components are highlighted in real time. Once activated, the selected phrase is passed to the Audio Output Module.

5.4. Database and Audio Playback Integration

An SQLite database maintains mappings between phrase IDs, categories, and associated WAV file paths. AVFoundation handles low-latency audio playback. Audio files were pre-recorded by a native speaker and stored locally. Figure 5 presents the simplified Entity Relationship Diagram (ERD) used.

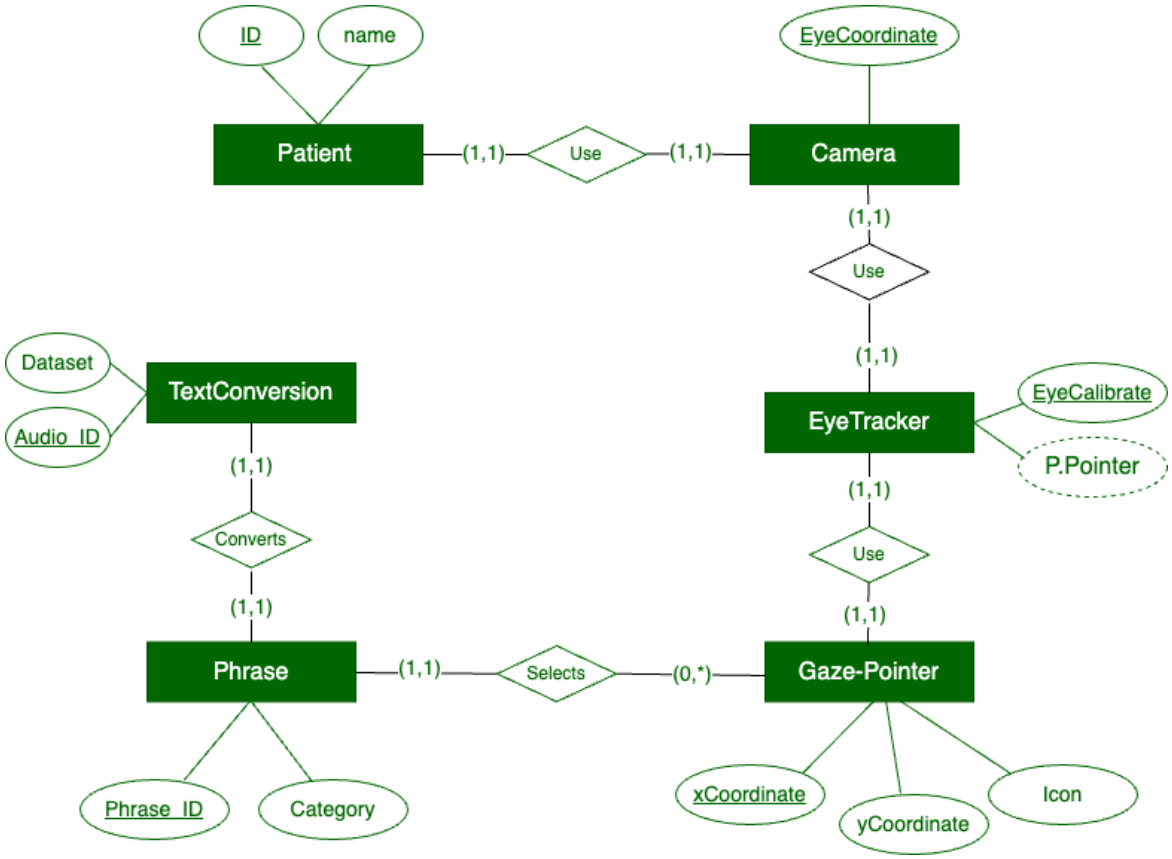


Figure 5. Simplified ERD illustrating SOUTY's internal data structure, including patients, input devices (camera, eye-tracker, gaze pointer), phrase selection flow, and audio playback mapping. The relationships enforce data consistency and support real-time interaction for gaze-based phrase selection (* denotes "many" cardinality).

The ERD defines the core entities, their attributes, and the relationships that ensure data integrity and enable optimized information retrieval. The ERD models the core components of SOUTY's development ecosystem, including the patient, camera, eye-tracker, gaze pointer, and text conversion engine. Several one-to-one (1:1) relationships exist; for example, a patient interacts with a single camera, which is linked to one eye-tracker. The eye-tracker, in turn, is associated with a single gaze pointer that selects a specific phrase. In contrast, the relationship between phrases and the gaze pointer is zero-to-many (0:*), as each phrase can be selected zero or multiple times.

5.5. System Screen Flow

Figure 6 illustrates the screen flow of the SOUTY application, outlining the key user interaction stages from initial launch to voice playback.

Upon launching the application, users are prompted to grant camera access and complete an eye gaze calibration routine. Once calibrated, the main interface presents multiple phrase categories, each containing commonly used expressions. The user navigates through these options using gaze fixation, selecting a phrase or navigating to a new category.

When a phrase is selected, the application plays back a pre-recorded Arabic audio clip corresponding to that phrase.

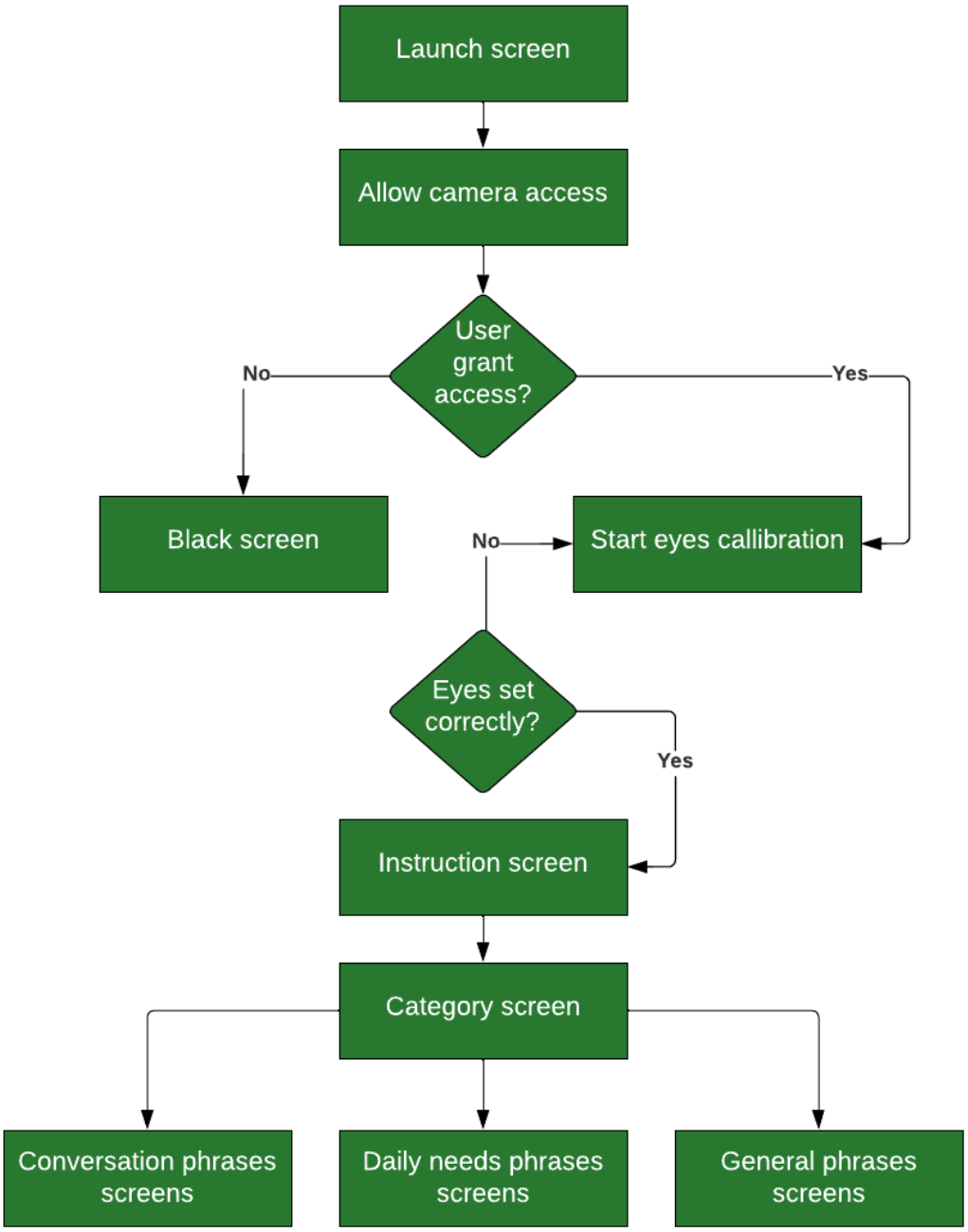


Figure 6. Screen flow of the SOUTY application, illustrating the sequential interaction states from application launch and gaze calibration to category selection, phrase playback, and custom phrase construction via the gaze-controlled Arabic keyboard.

If the user chooses the keyboard mode, the screen transitions to an on-screen Arabic keyboard that can be operated entirely by gaze. Characters are entered using dwell-time selection, enabling users to construct novel phrases beyond the predefined options.

This modular and accessible screen flow was designed to accommodate users with severe motor impairments, offering both quick communication through ready-made phrases and expressive flexibility via custom input. The following sections present results from testing and early evaluations of SOUTY's usability, reliability, and user acceptance.

6. System Testing

This section evaluates the technical reliability of SOUTY through developer-led testing strategies, focusing on internal correctness, performance under controlled conditions, and compliance with expected functional behavior. Tests were conducted using manual execution and Xcode's profiling tools on iOS devices. These methods were selected to ensure that SOUTY could deliver consistent functionality under real-world usage conditions typical for ALS patients.

6.1. Unit Testing

Each core module of the system was tested in isolation to validate individual behavior. The Eye Tracking Engine, Phrase Selector, Audio Playback Handler, Keyboard Input Manager, and SQLite Phrase Bank were tested using white-box techniques. Table 1 summarizes the results.

Table 1. Sample unit testing results.

Module	Test Action	Result
Eye Tracking Engine	Eye coordinates detection	Pass
Phrase Selector	Gaze selection and highlight	Pass
Audio Playback Handler	WAV playback via AVFoundation	Pass
Custom Keyboard	Dwell-based character input	Pass
Database Access	Phrase retrieval via SQLite	Pass

6.2. Integration Testing

Integration testing followed a bottom–up approach. Modules were incrementally combined, starting from camera access and gaze tracking, followed by database connections, UI interaction, and audio playback. Each integration point was validated on physical iOS devices (iPhone 12 and iPhone 13), confirming consistent data flow and synchronized behavior.

6.3. Performance Testing

Performance was profiled using Xcode Instruments under typical usage scenarios. Table 2 outlines key performance indicators.

Table 2. Performance metrics for SOUTY averaged over 10 runs under controlled lighting. Latency SD = 0.05 s; memory SD = 8 MB. Testing in variable lighting is part of ongoing work.

Metric	Observed Value
CPU Usage	28% (steady during gaze tracking and playback)
Memory Consumption	132 MB
App Launch Time	1.2 s
Audio Playback Latency	<0.3 s

6.4. Summary of Technical Testing

All modules passed unit and integration testing without critical issues. System resource usage was within acceptable bounds for mobile environments. Performance remained consistent across test devices. Future testing will incorporate additional metrics, such as frame rendering latency and error tolerance under poor lighting or camera obstruction. Future testing will incorporate additional metrics, including performance under low-light conditions and minor head movements, to better simulate real-world home use.

7. Evaluation and Results

This section evaluates SOUTY's usability, user satisfaction, and contextual appropriateness using expert validation, public awareness surveys, and structured usability walkthroughs.

7.1. Expert Interviews

Two semi-structured interviews were conducted with domain experts. Both interviews were held in person and lasted approximately 45–60 min each. One interviewee was a researcher specializing in AI and HCI, and the other was a licensed speech-language pathologist. Discussions explored interface usability, phrase structure, cultural relevance, and the clinical impact of personalized voice playback.

Feedback from the subject-matter experts emphasized the value of non-synthetic, culturally resonant voice output; intuitive navigation through gaze tracking; and culturally resonant vocabulary for Arabic-speaking users with ALS. Both experts validated the clinical and technical relevance of the system and endorsed the integration of voice banking as a means of preserving user identity.

7.2. Public Awareness Survey

A cross-sectional online survey (n = 236) was distributed via social media and academic networks to assess public awareness of ALS, familiarity with gaze-tracking technologies, and perceptions of Arabic-language AAC tools.

Out of 236 respondents, 91.1% acknowledged communication challenges faced by people with disabilities, while 64% were previously aware of ALS. However, 96.2% had never used a gaze-based interaction system, and only 54.7% were familiar with TTS tools, highlighting a significant technological readiness gap among Arabic-speaking populations.

7.3. Usability Walkthroughs

Five internal participants (not involved in core implementation) conducted structured usability walkthroughs simulating ALS communication scenarios. Each participant followed a scripted sequence of interaction tasks designed to validate the system's end-to-end flow (see Figure 7), including the following:

- Granting camera access and completing gaze calibration;
- Navigating to and selecting a predefined phrase category;
- Playing a pre-recorded audio phrase;
- Constructing a custom message using the on-screen Arabic keyboard.

Overall feedback confirmed that the interface was intuitive and accessible. Participants suggested improvements such as adding a replay function, expanding phrase categories, and including shortcuts for caregiver interactions.

Table 3 outlines the task definitions; Table 4 summarizes the outcomes.

Table 3. Walkthrough task scenarios.

Task No.	Description
1	Launch app and allow camera access
2	Complete gaze calibration
3	Navigate to phrase category
4	Select and play phrase
5	Compose message with keyboard

Table 4. Walkthrough results summary.

Participant	All Tasks Completed	Notes
P1	Yes	Smooth tracking
P2	Yes	Needed calibration repeat
P3	Yes	Natural voice output appreciated
P4	Yes	Suggested more phrase icons
P5	Yes	Gaze accuracy was high

Phrase Entry Timing. During the usability walkthroughs, participants were observed to complete predefined phrase selection and playback in less than 5 s on average. This includes gaze calibration, category navigation, dwell-time selection (2 s), and audio output. When using the gaze-controlled Arabic keyboard to construct short custom phrases (2–4 words), task durations ranged between 12 and 20 s depending on familiarity and precision of gaze input. These timing results underscore the effectiveness of prioritizing essential phrases in reducing communication effort.

The evaluation confirmed the practical utility of SOUTY across both technical and user-centered dimensions. Expert validation underscored the system's cultural and clinical relevance. Public responses revealed both awareness and knowledge gaps regarding AAC technologies in Arabic. Walkthrough feedback was positive and directly informed future design refinements. Future validation with ALS patients and the inclusion of formal usability tools (e.g., System Usability Scale (SUS)) will enhance the robustness of user-centered evaluation.

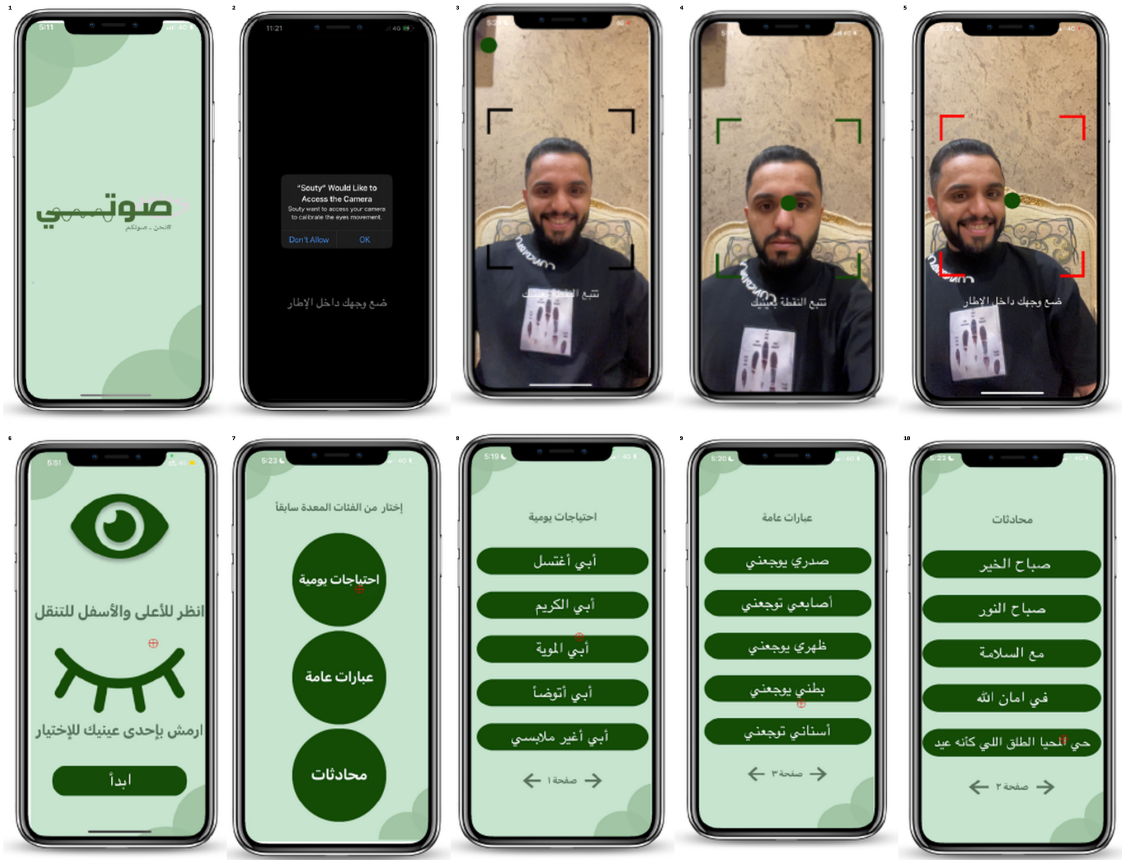


Figure 7. Sequential screenshots showing calibration, category selection, and phrase playback within SOUTY: (1) Launch screen, (2) Camera permission, (3) Face alignment, (4) Calibration (focused gaze), (5) Calibration correction, (6) Navigation instructions, (7) Category selection, (8) Daily needs phrases, (9) General phrases, (10) Conversational phrases.

7.4. Qualitative Effectiveness Evaluation (Internal Walkthroughs)

While this initial study did not include clinical deployment with ALS patients, internal walkthroughs allowed for qualitative observation of SOUTY's potential effectiveness. All five participants completed a structured interaction script using only gaze-based control, confirming system usability and reliability. In addition to task completion, participants reported high confidence in the system's responsiveness and low error incidence during phrase selection and audio playback. Informal timing measurements showed that urgent phrases could be selected and vocalized in under 5 s, while personalized phrases using the gaze keyboard ranged between 12 and 20 s, depending on the phrase length. These observations suggest that SOUTY enables basic communication within reasonable latency thresholds for mobile AAC tools. Future evaluation will include clinical trials and validated scoring tools such as the SUS and Communication Effectiveness Index (CETI).

8. Discussion

The development and evaluation of SOUTY highlight several key contributions to the field of assistive technology, particularly within Arabic-speaking communities. As part of the broader research roadmap, future phases will involve pilot testing with Arabic-speaking ALS patients and the collection of larger, representative datasets. These data will support the application of appropriate inferential statistical methods (e.g., chi-square tests, *t*-tests) to examine associations between participant characteristics and technology acceptance, thereby strengthening the generalizability and clinical relevance of our findings. This section discusses SOUTY's comparative advantages, the design trade-offs encountered, and areas for future improvement.

8.1. Novelty and Differentiators

SOUTY offers a number of unique features that distinguish it from existing AAC systems:

- Arabic-first design: Unlike most gaze-based AAC tools that support English or generic Latin scripts, SOUTY was developed entirely in Arabic, featuring a user interface, a keyboard, and phrase categories specifically tailored to the Arabic script and cultural context. This design addresses the needs of underserved speakers of one of the most widely spoken languages in the world.
- Real voice bank: SOUTY uses a dataset of pre-recorded Arabic phrases by an ALS patient, preserving speaker identity and emotional expressiveness. This approach offers a more natural alternative to synthesized voices and resonates with users on a personal level.
- Offline operation: All functionalities—gaze tracking, phrase navigation, and audio playback—work without requiring internet access. This enhances privacy, lowers cost, and makes the app usable in home care and rural settings.
- Localized phrase design: Phrase categories were curated based on needs identified by medical experts and families of ALS patients. Categories such as "Daily Needs," "General Phrases," and "Conversations" reflect everyday communication scenarios. Internal walkthroughs confirmed that frequently used phrases—such as "I need help," "I am in pain," and "Please call someone"—could be selected and vocalized in under 5 s, validating the system's suitability for urgent or high-frequency communication needs.

8.2. Design Trade-Offs

During development, several important trade-offs were made:

- Voice customization vs. scalability: Using pre-recorded audio provides authenticity but limits the ability to scale across dialects or genders. A future version may require hybrid models that combine recorded anchors with synthetic speech.

- Hardware dependence: SOUTY currently supports only iOS devices with front-facing cameras. Expanding to Android will require platform-specific calibration and UI adaptations.
- User fatigue: Extended use of eye-tracking can be tiring for some users. While dwell-time tuning helps, additional input modes (e.g., blink or facial gesture recognition) may be beneficial in future versions.

While SOUTY currently functions as a standalone, offline-capable mobile application intended for single-user operation, we recognize the importance of system scalability for future deployments, particularly in clinical settings. Since all processing is handled locally and no server-based operations are involved, system performance remains unaffected by the number of users or phrase sets. However, future versions may include secure multi-profile support to allow shared devices in care facilities or multi-user environments. Similarly, expanding the voice bank with additional speakers or dialects may require lightweight cloud integration, which is under consideration for future iterations.

8.3. Comparison with Commercial Applications

Table 5 summarizes key differences between SOUTY and existing AAC tools that support eye tracking.

Table 5. Comparison of SOUTY with existing gaze-controlled AAC applications.

Feature	Hawkeye Access	Vocable AAC	SOUTY
Language Support	English only	English only	Arabic (Saudi dialect)
Offline Mode	Partial	Yes	Fully supported
TTS Source	Synthetic voice	Synthetic voice	Human voice bank
Device Platform	iOS only	iOS only	iOS only
Phrase Customization	Limited	Moderate	Predefined + Keyboard
User Interface	Western-centric	English phrases	Arabic UI and flow

8.4. Ethical and Cultural Considerations

The lack of eye-tracking AAC solutions in Saudi Arabia, as reported by Abuzinadah et al. [4], reinforces SOUTY's role as a foundational step toward localizing assistive communication for Arabic-speaking ALS patients. Moreover, personal voice preservation is increasingly recognized as a dignity-enhancing intervention for patients with degenerative diseases. SOUTY addresses this by incorporating a donated voice from a Saudi ALS patient with informed consent. The team took care to ensure that phrases were appropriate, respectful, and reflected social norms. Moreover, no user data is collected or transmitted, preserving patient privacy.

8.5. Lessons Learned

Throughout the development and early evaluation of SOUTY, several important lessons were identified that informed both design choices and future improvement opportunities. First, early engagement with domain experts—particularly licensed speech-language pathologists—proved invaluable. Their clinical insights guided the selection and structuring of phrase categories, ensuring that the vocabulary addressed practical communication needs in culturally and medically appropriate ways. Second, the project reinforced the viability of eye-tracking as a primary input modality for mobile-based assistive communication. However, its effectiveness depended heavily on careful interface design, including the use of real-time gaze feedback mechanisms such as hover animations, clear selection zones, and visual cues that reinforced interaction states. Third, user testing revealed a clear preference for short, contextually relevant phrases presented in categorized groups rather

than unconstrained text entries. This feedback validated SOUTY's modular and structured interaction model, which reduces cognitive load and supports faster communication—especially critical for users with limited energy or mobility. These lessons will guide future iterations and may serve as practical recommendations for others designing culturally inclusive AAC systems. Future evaluations will include statistical usability assessments (e.g., SUS scores), which were beyond the scope of this initial development phase.

9. Conclusions and Future Work

This study presents SOUTY, an Arabic-language AAC application for ALS patients using eye-tracking and voice identity preservation. The system was developed and evaluated through iterative design, internal performance testing, and expert validation. While the core system modules (gaze tracking, phrase playback, and UI interaction) demonstrated reliable technical performance, effectiveness was also observed qualitatively through user walkthroughs that confirmed usability, accessibility, and fast phrase selection. The conclusion is that SOUTY is a technically viable and culturally relevant AAC system that now requires clinical evaluation to determine its impact on user communication outcomes and quality of life.

The system combines multiple strengths: a user interface tailored for Arabic text and phrase categories, a pre-recorded voice bank that ensures natural and personalized output, and a robust eye-tracking mechanism that requires no external hardware. Through structured expert interviews and survey-based user validation, we demonstrated SOUTY's feasibility, acceptability, and impact potential for Arabic-speaking patients who have lost their ability to speak.

Future Directions

SOUTY lays the foundation for a broader research and development agenda in inclusive Arabic HCI design. This initial development phase did not include inferential statistical analyses or pilot testing with ALS patients. Future work will incorporate hypothesis-driven study designs with larger, representative samples and will apply appropriate inferential statistical methods (e.g., chi-square tests, t-tests) to examine associations between participant characteristics and acceptance or performance metrics. Pilot testing with ALS patients is also planned, enabling the collection of clinical and usability data that will undergo both quantitative and qualitative analyses. Planned improvements and future work include the following:

- Cross-platform support: Extend SOUTY to Android and web-based devices to reach a wider audience, including low-cost smartphones and tablets.
- Dynamic TTS integration: Explore hybrid models that combine pre-recorded anchors with Arabic deep learning-based TTS to offer real-time sentence flexibility while preserving voice identity.
- Voice diversity: Expand the voice bank to include female and child voices, allowing users to select or preserve their own voice during early ALS diagnosis.
- Custom phrase creation: Enable users or caregivers to add custom phrases, favorite phrases, and shortcut screens (e.g., for emergencies).
- Clinical validation: Conduct longitudinal trials with ALS patients and caregivers to evaluate performance, satisfaction, and real-world health outcomes.
- Internationalization: Collaborate with global accessibility networks and health organizations such as WHO to localize SOUTY for other Arabic dialects and regions.
- Scalable deployment: Extend SOUTY to support secure multi-user profiles on shared devices (e.g., in clinics or long-term care facilities), with optional cloud-based synchronization for larger voice datasets.

While gaze calibration proved reliable under normal lighting, its performance can degrade in low-light or high-motion environments. Camera access is mandatory; if permission is denied, the app exits gracefully. No alternate input method (e.g., voice or touch) was implemented, limiting usability to users with preserved ocular control.

Ultimately, SOUTY is a step toward more inclusive technology ecosystems where linguistic and cultural identity are embedded within assistive innovation. Its model can serve as a blueprint for future communication tools built not only for, but with, patients from underrepresented communities. Future expansion and clinical trials may establish SOUTY as a scalable blueprint for localized, culturally adaptive AAC systems across the Global South.

This study is limited to internal testing and expert validation. Future studies involving direct testing with Arabic-speaking ALS patients in clinical environments are required to fully assess usability, emotional impact, and long-term adoption. Planned clinical trials will include ALS patients and apply validated tools such as the SUS, CETI, and quality-of-life assessments. A phased development plan includes Android porting in 2026, expanded voice bank by 2027, and multicenter clinical trials by late 2027.

Author Contributions: Conceptualization, S.A.; Methodology, S.A., S.M.A. and H.A.A.; Software development, L.A., M.A., A.A., D.A. and L.S.A.; Writing—original draft preparation, H.A.A., L.A., M.A., A.A., D.A. and L.S.A.; Writing—review and editing, H.A.A., S.M.A. and S.A.; Supervision, S.A. and H.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by King Saud University, Riyadh, Saudi Arabia, through the Ongoing Research Funding Program (ORF-2025-1206), King Saud University, Riyadh, Saudi Arabia.

Data Availability Statement: Restrictions apply to the availability of the dataset used in this study. The voice bank data were obtained under a non-disclosure agreement with a third party and cannot be shared publicly due to confidentiality obligations. Requests for access to derived data or additional details may be directed to the corresponding author, subject to permission from the data owner.

Acknowledgments: The authors are grateful to King Saud University, Riyadh, Saudi Arabia, for funding this work through the Ongoing Research Funding Program (ORF-2025-1206), King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AAC	Augmentative and Alternative Communication
AI	Artificial Intelligence
ALS	Amyotrophic Lateral Sclerosis
BCI	Brain–Computer Interface
CETI	Communication Effectiveness Index
ERD	Entity-Relationship Diagram
HCI	Human–Computer Interaction
iOS	iPhone Operating System
LLM	Large Language Model
RTL	Right-to-Left
SOUTY	Voice Identity Preserving Arabic AAC Application
SUS	System Usability Scale
TTS	Text-to-Speech
UI	User Interface
UCD	User-Centered Design

References

1. Londral, A. Assistive Technologies for Communication Empower Patients with ALS to Generate and Self-Report Health Data. *Front. Neurol.* **2022**, *13*, 867567. [CrossRef] [PubMed]
2. Mehta, P.; Raymond, J.; Nair, T.; Han, M.; Berry, J.; Punjani, R.; Larson, T.; Mohidul, S.; Horton, D.K. Amyotrophic lateral sclerosis estimated prevalence cases from 2022 to 2030, data from the National ALS Registry. *Amyotroph. Lateral Scler. Front. Degener.* **2025**, *26*, 290–295. [CrossRef] [PubMed]
3. Arthur, K.C.; Calvo, A.; Price, T.R.; Geiger, J.T.; Chiò, A.; Traynor, B.J. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nat. Commun.* **2016**, *7*, 12408. [CrossRef] [PubMed]
4. Abuzinadah, A.R.; AlShareef, A.A.; AlKutbi, A.; Bamaga, A.K.; Alshehri, A.; Algahtani, H.; Cupler, E.; Alanazy, M.H. Amyotrophic lateral sclerosis care in Saudi Arabia: A survey of providers' perceptions. *Brain Behav.* **2020**, *10*, e01795. [CrossRef] [PubMed]
5. UNESCO. World Arabic Language Day. 2024. Available online: https://www.unesco.org/en/world-arabic-language-day (accessed on 10 July 2025).
6. World Population Review. Arabic Speaking Countries 2025. 2025. Available online: https://worldpopulationreview.com/country-rankings/arabic-speaking-countries (accessed on 10 July 2025).
7. Edughele, H.O.; Zhang, Y.; Muhammad-Sukki, F.; Vien, Q.T.; Morris-Cafiero, H.; Opoku Agyeman, M. Eye-Tracking Assistive Technologies for Individuals With Amyotrophic Lateral Sclerosis. *IEEE Access* **2022**, *10*, 41952–41972. [CrossRef]
8. Bonanno, M.; Saracino, B.; Ciancarelli, I.; Panza, G.; Manuli, A.; Morone, G.; Calabrò, R.S. Assistive Technologies for Individuals with a Disability from a Neurological Condition: A Narrative Review on the Multimodal Integration. *Healthcare* **2025**, *13*, 1580. [CrossRef] [PubMed]
9. Fischer-Janzen, A.; Wendt, T.M.; Van Laerhoven, K. A scoping review of gaze and eye tracking-based control methods for assistive robotic arms. *Front. Robot.* **2024**, *11*, 1326670. [CrossRef] [PubMed]
10. Cave, R.; Bloch, S. The use of speech recognition technology by people living with amyotrophic lateral sclerosis: A scoping review. *Disabil. Rehabil. Assist. Technol.* **2023**, *18*, 1043–1055. [CrossRef] [PubMed]
11. Neumann, M.; Roesler, O.; Liscombe, J.; Kothare, H.; Suendermann-Oeft, D.; Pautler, D.; Navar, I.; Anvar, A.; Kumm, J.; Norel, R.; et al. Investigating the Utility of Multimodal Conversational Technology and Audiovisual Analytic Measures for the Assessment and Monitoring of Amyotrophic Lateral Sclerosis at Scale. *arXiv* **2021**, arXiv:2104.07310. [CrossRef]
12. Hoyle, A.C.; Stevenson, R.; Leonhardt, M.; Gillett, T.; Martinez-Hernandez, U.; Gompertz, N.; Clarke, C.; Cazzola, D.; Metcalfe, B.W. Exploring the 'EarSwitch' Concept: A Novel Ear-Based Control Method for Assistive Technology. *J. NeuroEngineering Rehabil.* **2024**, *21*, 210. [CrossRef] [PubMed]
13. Kew, S.Y.N.; Mok, S.Y.; Goh, C.H. Machine learning and brain-computer interface approaches in prognosis and individualized care strategies for individuals with amyotrophic lateral sclerosis: A systematic review. *MethodsX* **2024**, *13*, 102765. [CrossRef] [PubMed]
14. Zhou, S.; Armstrong, M.; Barbareschi, G.; Ajioka, T.; Hu, Z.; Ando, R.; Yoshifuji, K.; Muto, M.; Minamizawa, K. Augmented Body Communicator: Enhancing daily body expression for people with upper limb limitations through LLM and a robotic arm. *arXiv* **2025**, arXiv:2505.05832. [CrossRef]
15. Hawkeye Labs, Inc. Hawkeye Labs, Inc.–Apple App Store Developer Profile. 2025. Available online: https://apps.apple.com/us/developer/hawkeye-labs-inc/id1439231626 (accessed on 10 July 2025).
16. Vocable AAC (Voiceitt Inc.). Vocable AAC (Version—); Augmentative and Alternative Communication App. 2025. Available online: https://apps.apple.com/us/app/vocable-aac/id1497040547 (accessed on 10 July 2025).
17. EyeTech Digital Systems, Inc. EyeTech Digital Systems—Eye Tracking and Speech-Generating Devices. 2025. Available online: https://eyetechds.com/ (accessed on 10 July 2025).
18. Tobii AB. Tobii—Eye Tracking and Attention Computing. 2025. Available online: https://www.tobii.com/ (accessed on 10 July 2025).
19. Benabid Najjar, A.; Al-Wabil, A.; Hosny, M.; Alrashed, W.; Alrubaian, A. Usability Evaluation of Optimized Single-Pointer Arabic Keyboards Using Eye Tracking. *Adv. Hum.-Comput. Interact.* **2021**, *2021*, 6657155. [CrossRef]
20. Ning, Y.; He, S.; Wu, Z.; Xing, C.; Zhang, L.J. A Review of Deep Learning Based Speech Synthesis. *Appl. Sci.* **2019**, *9*, 4050. [CrossRef]
21. *ISO 9241-11:2018*; Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts. International Organization for Standardization: Geneva, Switzerland, 2018.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.