Aisen - Web-Based Gaze-Tracking Assistive Communication Interface with Word Cards Generated by LLMs

Wei-En Tsai*
Institute of Information Security
National Tsing Hua University
Hsinchu, Taiwan
schwarz.tsai@gmail.com

Yi-Chun Liu*
Institute of Service Science
National Tsing Hua University
Hsinchu, Taiwan
jean.liu@gapp.nthu.edu.tw

ABSTRACT

Aisen is an innovative web-based communication tool that integrates the WebGazer.js library [5] with advanced large language models (LLMs). Designed as an affordable communication solution for those with communication challenges, Aisen facilitates expression through a unique word-selection interface. Rather than using a traditional keyboard, Aisen introduces a two-tiered "word card" system. This system includes a static set of cards tailored to the patient's specific needs and a "dynamic" set where cards are intelligently generated by LLMs based on user input and preferences. Our research delineated three specific user personas, emphasizing Aisen's applicability for elderly patients. The platform integrates an eye-tracking mechanism, a gaze-responsive interface, and a word card repository enriched with LLMs. This endeavor highlights the transformative potential of web-enabled eye-tracking and LLMs in enhancing communication for individuals with impairments.

CCS CONCEPTS

• **Human-centered computing** → **Accessibility systems and tools**; **Interface design prototyping**; **Web-based interaction**; • **Applied computing** → *Health care information systems.*

KEYWORDS

gaze-controlled interface, assistive technology, communication aid, large language models(LLM)

ACM Reference Format:
Wei-En Tsai and Yi-Chun Liu. 2023. Aisen - Web-Based Gaze-Tracking Assistive Communication Interface with Word Cards Generated by LLMs. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct), October 29–November 01, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3586182.3625116

1 INTRODUCTION

Our project presents Aisen, a web-based gaze-tracking communication aid that leverages existing gaze prediction models, specifically the WebGazer.js library [5], to serve as a cost-effective alternative to dedicated eye-trackers. By utilizing the built-in front-facing camera of a laptop, Aisen enables patients with limited physical capabilities to communicate through a word-selection interface. This technology not only facilitates communication for patients but also fosters connections between patients, caregivers, and family members, enhancing the provision of care.

To accommodate the unique complexities of Chinese input methods [7], which often involve a two-stage process of character composition and selection, as opposed to phonetic-based languages like English, we have implemented a "word card" solution. This approach seeks to achieve a harmonious balance between intricate expression and convenient input methods. Furthermore, this approach can be adapted for English interfaces, benefiting those who find a standard 26-character keyboard challenging to use.

Alongside pre-defined word cards, we've designed a dynamic card section where word cards are dynamically generated by LLMs. This feature can assist patients aiming for more nuanced expression, and the LLM might even anticipate the patient's most likely needs.

Many conditions, including Parkinson's disease [4] and Amyotrophic Lateral Sclerosis (ALS) [6], as well as invasive treatments like tracheotomy and intubation [2], can impair a patient's ability to express themselves [1]. The loss of communication skills can lead to misunderstandings about a patient's needs and feelings, effectively isolating them [3]. Aisen, with its web-based interface, aims to mitigate these challenges, enhancing the accessibility and prevalence of eye-tracking communication tools.

2 IMPLEMENTATION

2.1 User Study

Through a questionnaire survey involving 13 users with communication disorders and in-depth interviews with 3 long-term nursing facilities, we identified 3 distinct personas representing different user types. These in-depth discussions provided deeper insights into the characteristics and needs of each persona.

The first persona represents elderly patients in the terminal stage of an illness. Their communication largely revolves around conveying basic needs like discomfort. Their propensity for self-expression is minimal, necessitating a language model centered on simple vocabulary tailored to essential requirements.

The second persona encompasses children with brain disorders, whose parents seek to enhance their vocabulary and cognitive skills with the assistance of speech therapists.

The third persona comprises adults with heightened awareness, desiring the ability to express complex ideas accurately and efficiently. These individuals often encounter abrupt speech loss resulting from accidents like strokes. Catering to their requirements

*Both authors contributed equally to this research.

UIST '23 Adjunct, October 29–November 01, 2023, San Francisco, CA, USA

ACM ISBN 979-8-4007-0096-5/23/10.
https://doi.org/10.1145/3586182.3625116

might necessitate a more sophisticated interface with greater precision, a level of accuracy beyond what webcam-based gaze-tracking can offer.

By incorporating LLMs, Aisen caters to the needs of the first persona, focusing on supporting dynamic communication support for the elderly.

2.2 System Design

The system is composed of four main modules: the eye-tracking mechanism, the gaze-responsive interface, the web dashboard, and the word cards repository integrated with LLMs.

2.2.1 Eye-Tracking Mechanism. The primary function of the eye-tracking mechanism is to capture the user's eye gaze on the screen accurately for input. Instead of using a dedicated eye tracker, we utilize a web gaze library called WebGazer.js in our current system version. This choice reduces costs, simplifies setup, and greatly increases the software's accessibility and distribution.

2.2.2 Gaze-Responsive Interface. This interface dictates the method by which users make selections using their gaze. It is designed to indicate the user's gaze focus, facilitate selection by maintaining a gaze on an item for over 3 seconds, and revert to a neutral state when the user diverts their gaze from the interface.

2.2.3 Web Dashboard. The design of Aisen's web dashboard takes into account the accuracy of eye-tracking, user visual sharpness, and the dimensions of the device screen. It ensures that interface elements are sufficiently large to compensate for potential tracking inaccuracies and that the text remains legible, especially for the elderly.

2.2.4 Word Cards Repository Integrated with LLMs. The content of the word cards can be tailored by either the patients or their caregivers, ensuring precise articulation of needs and enabling caregivers to provide superior care. While a fixed set of word cards can offer a more "predictive" input interface, potentially increasing efficiency, it may also restrict the range of expressions available to a patient. In light of this, we integrate the LLM with a combination of 1. preset prompts, 2. the user's past inputs, and 3. user-defined preferences. This integration will introduce a "dynamic" section to the word card system.

Aisen's vocabulary system is organized into eight categories, encompassing body parts, physical sensations, mental sensations, family members, medical care, assistance, information, and common phrases (see Figure 1). In addition to providing comprehensive care, Aisen promotes patients' mental well-being by enabling them to effectively express their thoughts and emotions. To accommodate elderly individuals with various literacy levels, Aisen's word cards include both text and illustrations.

Figure 2 depicts the input interface, segregating word cards into categories for specific content classification. Selected words appear in the upper input field, while circles and crosses serve as quick-response buttons for prompt communication with caregivers. Located in the upper right corner, the "Done" button facilitates efficient expression of gratitude and signifies service completion, aiding caregivers in monitoring patient status.

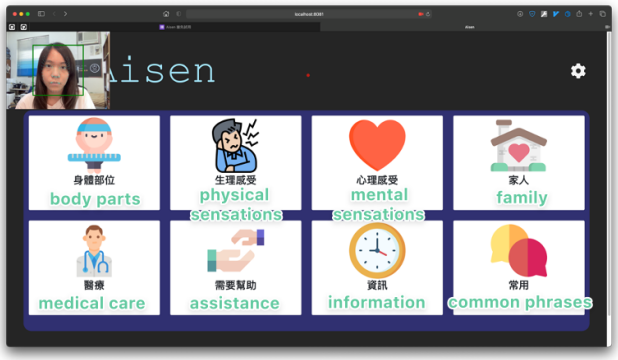

Figure 1: Aisen's Word-Card Dashboard for Comprehensive Communication and Enhanced Mental Well-being using Eye-Tracking

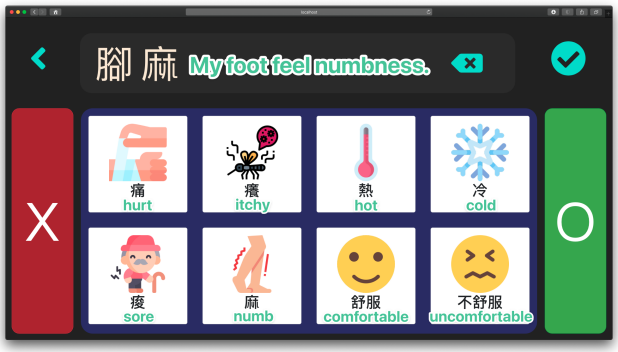

Figure 2: Aisen's Input Interface using Multi-Stage Word-Card Solution where the Word Cards are Dynamically Generated by LLMs with the User's Input History and Preference

3 CONCLUSION

This project introduces Aisen, a communication system that utilizes gaze-tracking technology for patients with communication disorders. Moreover, Aisen seamlessly integrates with Language Learning Models (LLMs). These LLMs, known for their adaptability and precision in understanding and generating human language, further amplify Aisen's effectiveness. Consequently, Aisen offers not just an efficient, affordable, and intuitive communication tool, but also a potential catalyst for improving users' quality of life. By aligning with the continually advancing and globally acknowledged LLMs, Aisen positions its users at the forefront of communication innovations.

REFERENCES

[1] Giuseppe Aceto, Valerio Persico, and Antonio Pescapé. 2018. The role of Information and Communication Technologies in healthcare: taxonomies, perspectives, and challenges. *Journal of Network and Computer Applications* 107 (2018), 125–154.

[2] John W Albarran. 1991. A review of communication with intubated patients and those with tracheostomies within an intensive care environment. *Intensive care nursing* 7, 3 (1991), 179–186.

[3] Michelle Lawton, Gillian Haddock, Paul Conroy, Laura Serrant, and Karen Sage. 2018. People with aphasia's perception of the therapeutic alliance in aphasia

rehabilitation post stroke: a thematic analysis. *Aphasiology* 32, 12 (2018), 1397–1417.

[4] Nick Miller. 2017. Communication changes in Parkinson's disease. *Practical Neurology* 17, 4 (2017), 266–274.

[5] Alexandra Papoutsaki, Patsorn Sangkloy, James Laskey, Nediyana Daskalova, Jeff Huang, and James Hays. 2016. WebGazer: Scalable Webcam Eye Tracking Using User Interactions. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*. AAAI, 3839–3845.

[6] Jayanti Ray. 2015. Real-life challenges in using augmentative and alternative communication by persons with amyotrophic lateral sclerosis. *Communication Disorders Quarterly* 36, 3 (2015), 187–192.

[7] Jingtao Wang, Shumin Zhai, and Hui Su. 2001. Chinese input with keyboard and eye-tracking: an anatomical study. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 349–356.