

SIGNPAL: A Human–AI Interaction Framework for Real-Time Sign Language Translation

Nadia Islam Tanha, Md Mehedi Hasan, Roksana Haque, Md Mahamudul Hasan, Jannatul Hur, and Yueqing Li

Dr. Victor A. Zaloom Department of Industrial and Systems Engineering,
Lamar University, Beaumont, TX, USA

ABSTRACT

Advances in artificial intelligence and human–computer interaction have enabled significant progress in assistive communication technologies for deaf and hard-of-hearing individuals who rely on American Sign Language (ASL). However, a persistent communication gap remains between ASL users and non-signers in everyday interactions. This study preliminarily evaluates the usability and performance of SIGNPAL, a high-fidelity AI-driven ASL interpretation system designed to provide real-time gesture-to-text translation, text-to-speech output, and customizable accessibility features to support two-way communication. Three participants with basic, moderate, and advanced ASL familiarity completed four usability tasks: real-time sign-to-text translation, accessibility customization, gesture recording and playback, and text-to-speech reply to generation. Objective measures included task completion time, translation errors, and recognition accuracy, while subjective usability was assessed using a 10-item Likert questionnaire and the System Usability Scale (SUS). The results show that SIGNPAL achieved an overall translation accuracy of 83.33%, exceeding the predefined performance threshold, and a 100% task completion rate. Response times remained below two seconds, supporting real-time interaction. Likert-scale ratings indicated high user satisfaction (overall mean = 4.5/5), and the mean SUS score of 90.83 classified the system as having excellent usability. Qualitative feedback highlighted the clarity of the interface and the usefulness of the recording-and-playback feature, with minor recommendations for improving text visibility. These findings demonstrate that integrating human factors principles with AI-driven gesture recognition can produce effective and user-centered assistive communication systems, supporting inclusive real-time interaction between ASL users and non-signers.

Keywords: Sign language recognition, Human–AI interaction, Usability evaluation, Accessibility design, Assistive communication systems

INTRODUCTION

Advances in artificial intelligence and computer vision have made automated sign language recognition an increasingly important tool for reducing communication barriers between deaf or hard-of-hearing individuals and hearing populations. Sign language translation systems have demonstrated substantial improvements in gesture recognition accuracy and continuous

sign modeling through deep learning and multimodal fusion techniques (Camgoz et al., 2018; Koller et al., 2015). However, despite these technical advances, communication accessibility remains limited in everyday educational, professional, and public contexts. Large-scale surveys report that most existing systems emphasize algorithmic performance metrics such as word error rate and BLEU (Bilingual Evaluation Understudy) scores, with limited consideration of usability, real-time interaction quality, or user satisfaction (Alaghband et al., 2023; Papastratis et al., 2021).

Human factors research emphasizes that system effectiveness depends not only on recognition fidelity but also on interface clarity, learnability, and accommodation of diverse user capabilities (ISO 9241-210 2019). Prior interdisciplinary studies further highlight that assistive communication technologies must integrate usability and accessibility measures to ensure meaningful adoption among deaf and hard-of-hearing users (Bragg et al., 2019). Although recent real-time systems report high recognition robustness, many omit structured usability testing and ergonomic performance thresholds (Alsharif et al., 2025; Yu et al., 2025).

This study investigates the usability and performance of SIGNPAL, an AI-driven sign language interpretation system designed using a human-centered framework. By combining translation accuracy, response time, task success, and standardized usability measures, this research aims to provide empirical evidence supporting inclusive human–AI communication design and address the persistent gap between technical innovation and applied human factors validation.

LITERATURE

Sign language recognition and translation systems represent a rapidly advancing domain at the intersection of computer vision, deep learning, and accessibility-oriented human–computer interaction (HCI). Significant progress has been achieved in gesture recognition accuracy, continuous sign modeling, and multimodal fusion. However, much of the literature remains predominantly focused on algorithmic performance, while systematic human-centered usability evaluation is comparatively limited. Large-scale surveys by Alaghband et al. (2023) and Papastratis et al. (2021) established the technical foundations of recognition and translation pipelines, reporting performance primarily through classification accuracy, BLEU (Bilingual Evaluation Understudy) scores, and word error rates. Although these surveys acknowledge challenges such as robustness, dataset diversity, and environmental sensitivity, they provide minimal discussion of usability, interaction quality, accessibility customization, or user satisfaction, offering limited insight into real-world communicative effectiveness.

Algorithm-centric studies further illustrate this imbalance. Camgoz et al. (2018) introduced neural sign language translation with a landmark continuous dataset evaluated using BLEU and ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metrics, while Koller et al. (2015) advanced large-vocabulary continuous recognition through statistical modeling and curated corpora, reporting only word error rates. Neverova et al. (2016) improved

recognition via multimodal fusion and adaptive training strategies, yet relied on complex sensor configurations and benchmark datasets without end-user validation. Collectively, these technically influential works demonstrate major advances in recognition performance but provide limited evidence regarding usability, deployment feasibility, or interaction effectiveness in everyday communication contexts.

Recent real-time systems show similar trends. Alsharif et al. (2025) achieved high recognition accuracy using a YOLO–MediaPipe pipeline for alphabet recognition under varied environmental conditions, but evaluation remained model-centric without standardized usability assessment or novice-user testing. Likewise, Yu et al. (2025) developed a bidirectional avatar-based communication system with large-scale preference testing, yet emphasized realism and system comparison over task efficiency, learnability, and ergonomic performance; reported latencies near five seconds further challenge conversational naturalness. Interdisciplinary perspectives emphasize that translation technologies must integrate accessibility and usability metrics to ensure meaningful adoption (Bragg et al., 2019), yet most deep-learning-based systems, including attention-based 3D-CNN approaches (Huang et al., 2018), rarely incorporate structured human-factors evaluation.

Overall, the literature demonstrates substantial algorithmic advancement but limited empirical evidence regarding accessibility comfort, interface clarity, learnability, and user acceptance—factors essential for inclusive communication technologies. The present SIGNPAL study addresses this gap through a human-factors-oriented evaluation framework integrating objective metrics (translation accuracy, response time, error rates, and task completion) with standardized subjective measures (Likert ratings and the System Usability Scale). Real users with varying ASL familiarity performed realistic tasks in a controlled environment, enabling simultaneous assessment of technical reliability and interaction quality. Although the sample size constrains generalizability, the findings provide empirical linkage between algorithmic performance and ergonomic design, advancing sign language translation research toward a holistic framework aligned with accessibility, inclusivity, and real-world deployment.

METHODOLOGY

Participants

American Sign Language (ASL) familiarity was considered an important factor in this study, as signing proficiency can influence gesture consistency, recognition accuracy, and interaction efficiency with AI-based translation systems. To capture performance variability across different user experience levels, participants were selected to represent diverse ASL familiarity backgrounds.

Three participants (two female and one male) voluntarily participated in this usability evaluation study. Participants were recruited to represent varying levels of ASL familiarity, including basic, moderate, and advanced proficiency. All participants reported normal or corrected-to-normal vision and no physical impairments that would interfere with gesture interaction

or touchscreen use. Prior to the experiment, participants were briefed on the study objectives and procedures and provided informed consent. A short familiarization session was conducted to allow participants to become acquainted with the SIGNPAL interface before formal data collection began.

System and Equipment

The SIGNPAL prototype was developed as a modular, real-time sign language interpretation system integrating computer vision, deep learning, and an interactive user interface. The backend was implemented in Python to manage data processing, gesture recognition logic, and communication across system components. Real-time landmark extraction was performed using MediaPipe to detect hand and upper body keypoints from the laptop's built-in RGB camera, and these features were processed by a Long Short-Term Memory (LSTM) model trained with TensorFlow/Keras to classify gesture sequences into text outputs. OpenCV supported video frame capture and preprocessing to ensure stable real-time performance under controlled lighting conditions. The frontend interface was built using Streamlit, providing access to four core modules: real-time sign-to-text translation, accessibility customization, gesture recording and playback, and text-to-speech response generation. Speech output was implemented using the PyTTSX3 library to enable low-latency, offline audio feedback.

The SIGNPAL prototype was deployed on a laptop computer equipped with a built-in camera and microphone to support gesture capture and speech output. The system provided four primary functional modules: real-time sign-to-text translation, accessibility customization, gesture recording and playback, and text-to-speech reply to generation. The experimental sessions were conducted in a quiet indoor laboratory environment under controlled lighting conditions to ensure stable gesture recognition and consistent testing conditions across participants. The block diagram of the system architecture is given below:

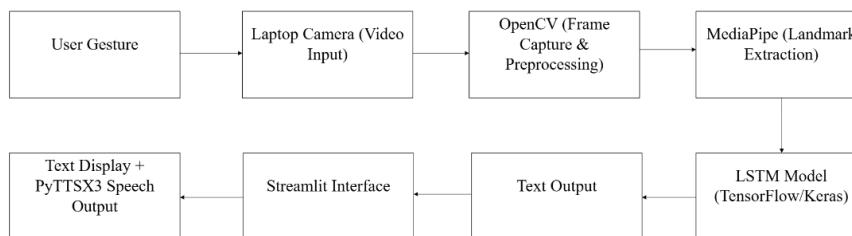


Figure 1: System architecture of the SIGNPAL prototype.

Experimental Design

A within-subject experimental design was employed. Each participant completed the same set of four usability tasks using the SIGNPAL system. The tasks were completely randomized.

Measurement

Measurement includes objective performance measure and subjective usability assessments. Objective measures consisted of task completion time, number of errors per task and translation accuracy. Subjective measures included post-test Likert-scale ratings across ten usability items and System Usability Scale (SUS) scores.

Experimental Tasks and Procedure

Each participant completed four core usability tasks: (1) real-time sign-to-text translation, (2) accessibility customization, (3) gesture recording and playback, and (4) text-to-speech reply to generation.



Figure 2: Participant performing sign language.

Participants were instructed to perform each task using natural signing behavior at their preferred pace. Task completion time and error occurrences were recorded using a standardized observation sheet. Short rest periods were provided between tasks to reduce fatigue and carryover effects. Upon completion of all tasks, participants completed the Likert-scale questionnaire followed by the SUS survey.

Measures and Data Analysis

Translation accuracy was computed using the formula $\text{Accuracy} = 1 - (\text{Errors} / \text{Total Tasks})$. Mean and standard deviation values were calculated for task completion time and Likert-scale responses. Error rates were analyzed per task to identify relative task difficulty. Overall satisfaction was calculated as

the average of item means across the Likert questionnaire. SUS scores were computed following Brooke's standard scoring procedure and interpreted using established usability benchmarks. Descriptive statistics were used to summarize system performance and user experience due to the small sample size.

RESULTS

This section presents the quantitative and qualitative outcomes of the user testing conducted to evaluate the usability, performance reliability, and user acceptance of the SIGNPAL prototype. Results are organized across subjective usability ratings, task performance metrics, and translation accuracy outcomes, supported by graphical visualizations.

Usability data were collected using pre-test questionnaires, task observation sheets, a post-test Likert survey, the System Usability Scale (SUS), and open-ended feedback forms (Appendix A–E).

Usability Ratings

Participants completed a 10-item post-test Likert questionnaire evaluating perceived usability, clarity, ease of navigation, system responsiveness, and overall satisfaction (Figure 2). Mean ratings across all items ranged from 4.00 to 4.67 on a five-point scale (5 = strongly agree, 1 = strongly disagree), indicating consistently high user agreement with positive usability statements.

Items Q1, Q2, Q4, Q8, Q9, and Q10 achieved the highest mean ratings ($M \approx 4.67$), reflecting strong satisfaction with interface clarity, interaction flow, and perceived usefulness. Slightly lower ratings were observed for Q3, Q5, and Q6 ($M \approx 4.33$), suggesting minor opportunities for refinement in navigation flow and task guidance. The lowest mean rating was recorded for Q7 ($M = 4.00$), corresponding to visual comfort and text visibility, which aligns with qualitative feedback requesting larger translation text.

The overall participant satisfaction score, computed as the average of all item means, was 4.5/5.0, exceeding the predefined usability target of 4.0 and indicating high perceived system usability.

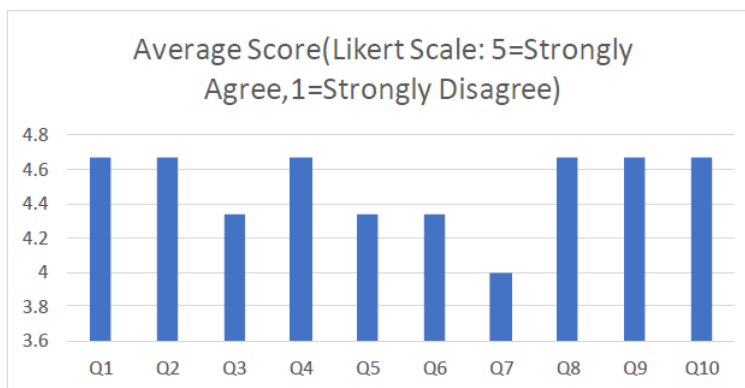


Figure 3: Usability ratings (5 = strongly agree, 1 = strongly disagree).

Task Performance and Distribution

Task performance across the four core functionalities—real-time sign-to-text translation, accessibility customization, gesture recording and playback, and text-to-speech reply was evaluated using average completion time and mean error frequency. All participants successfully completed each assigned task, resulting in a 100% task completion rate across conditions.

The gesture recording and playback task required the longest completion time (mean = 41 s) and showed the highest error frequency (mean = 1.33), indicating greater interaction complexity compared with other system functions. In contrast, the accessibility customization and text-to-speech reply tasks demonstrated lower completion times (21.33 s and 27.33 s, respectively) and zero observed errors, suggesting more efficient and intuitive interaction. Real-time sign-to-text translation showed moderate completion time and minimal errors relative to the more complex recording task.

Overall, the results indicate that the SIGNPAL prototype supports reliable task execution across users with varying ASL familiarity levels, while identifying the gesture recording and playback module as the primary area for potential usability refinement.

Translation Accuracy

Translation accuracy was computed per participant using the formula $\text{Accuracy} = 1 - (\text{Errors} / \text{Total Tasks})$. Based on the updated error data, Participant P01 achieved an accuracy of 91.7%, Participant P02 achieved 83.3%, and Participant P03 achieved 100% (Figure 4). The overall mean translation accuracy across participants was 91.7%, exceeding the predefined performance threshold of 80%.

The observed trend shows stable accuracy across participants with varying ASL familiarity levels, with the highest performance observed for the participant with advanced ASL proficiency. The slightly lower accuracy observed for participant 02 reflects a higher number of gesture recognition errors during the recording and playback task, consistent with observational notes indicating interface confusion.

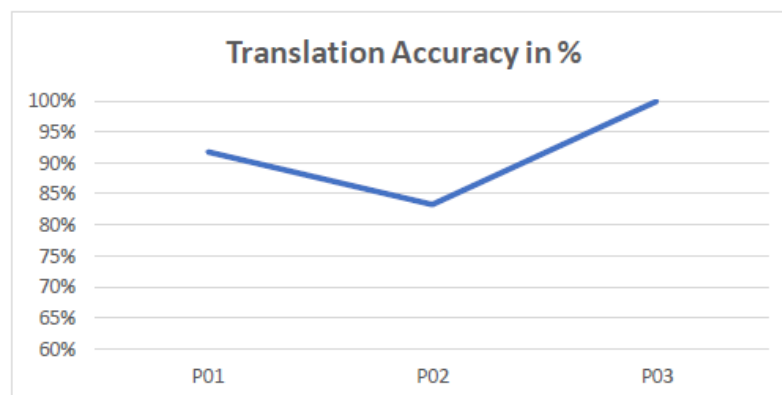


Figure 4: Translation accuracy in percentage.

Collectively, the results indicate that SIGNPAL meets or exceeds all predefined usability specifications. High Likert ratings and excellent translation accuracy demonstrate strong user acceptance and system reliability. The combination of objective performance measures and subjective usability evaluations provides converging evidence that the prototype effectively supports real-time, accessible sign language communication in realistic interaction contexts.

DISCUSSION

The results indicate that the SIGNPAL prototype demonstrates strong usability, high translation accuracy, and positive user acceptance across participants with varying levels of ASL familiarity. The convergence of objective performance measures and subjective usability ratings provides evidence that the system effectively supports real-time communication between sign language users and non-signers. Likert-scale responses showed consistently high agreement across questionnaire items (overall mean = 4.5/5.0), indicating that participants perceived the interface as clear, intuitive, and easy to navigate. In particular, items related to interface clarity, perceived usefulness, and interaction flow received strong ratings, suggesting that the core design principles successfully supported efficient human–AI interaction and rapid task comprehension. A slightly lower rating for visual comfort and text visibility aligns with qualitative feedback requesting larger translation text, highlighting a targeted interface refinement opportunity to further improve readability and accessibility.

Task performance outcomes further reinforce the system’s usability and reliability. All participants successfully completed all assigned tasks, yielding a 100% task completion rate independent of prior ASL proficiency. Analysis of task time and error distribution indicated that the “Record & Playback Gestures” feature involved the greatest interaction complexity, suggesting that gesture management workflows may benefit from simplified controls or clearer visual guidance. In contrast, accessibility customization and text-to-speech reply tasks demonstrated low error rates and shorter completion times, reflecting strong interface consistency, learnability, and operational efficiency for these features. These results indicate that the prototype supports effective task execution and stable interaction across diverse user experience levels.

Translation accuracy exceeded the predefined performance threshold of 80 % for all participants, with an overall mean accuracy of 91.7%. The highest performance was achieved by the advanced ASL user, while the lowest accuracy occurred for the participant with basic signing familiarity, reflecting expected variability in gesture consistency and recognition stability. Nevertheless, consistently high accuracy across users demonstrates the robustness of the translation pipeline under realistic usage conditions and supports the system’s suitability for real-time conversational contexts. Overall, the findings indicate that SIGNPAL effectively integrates ergonomic interface design with reliable AI-driven translation, bridging the gap between algorithmic performance and applied usability validation and demonstrating strong potential for deployment in educational, service, and everyday inclusive communication environments.

CONCLUSION

This study demonstrates that the SIGNPAL prototype effectively supports real-time sign language translation through the integration of AI-driven recognition and human-centered interface design. The results indicate that high translation accuracy, low response time, and strong usability ratings can be achieved when ergonomic principles and accessibility considerations are incorporated into system development. High user satisfaction and excellent System Usability Scale scores confirm that the prototype provides a reliable and intuitive platform for inclusive communication between ASL users and non-signers. The findings encourage the continued application of human factors engineering to AI-based accessibility technologies to improve interaction quality and real-world adoption.

As a limitation, the study evaluated a small sample size under controlled laboratory conditions using a limited task set. Future studies should include larger and more diverse user populations, extended usage durations, and additional interface design alternatives to further validate performance, improve gesture recognition stability, and optimize accessibility features for deployment in everyday communication environments.

ACKNOWLEDGMENT

This study was conducted as part of the course “User Experience with AI” and the authors gratefully acknowledge the guidance and support provided by Dr. Yueqing Li throughout the design, development, and evaluation of the project. The authors also thank the participants for their time and valuable feedback, which contributed significantly to the improvement of the SIGNPAL prototype.

APPLICATION LINK

<https://ai.studio/apps/drive/1cFF43fdH1TvjrUgnifrRk19RCPZj14zO>

REFERENCES

- Alagband, M., Rahmani, M. and Sadeghi, M.T. (2023). A survey on sign language recognition and translation systems. *Engineering Applications of Artificial Intelligence*, 120, 105890.
- Alonzo, O., Elliot, L., Dingman, B., Lee, S., Al Amin, A. and Huenerfauth, M. (2022). Reading-assistance tools among deaf and hard-of-hearing computing professionals in the U.S.: Their reading experiences, interests, and perceptions of social accessibility. *ACM Transactions on Accessible Computing*, 15(2), Article 16.
- Alsharif, B., Alalwany, E., Ibrahim, A., Mahgoub, I. and Ilyas, M. (2025). Real-time American Sign Language interpretation using deep learning and keypoint tracking. *Sensors*, 25(7), 2138.
- Bragg, D., Koller, O., Ballard, M., Berke, L., Boudreault, P., Braffort, A. and Huenerfauth, M. (2019). Sign language recognition, generation, and translation: An interdisciplinary perspective. *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 16–31.

- Camgoz, N.C., Koller, O., Hadfield, S. and Bowden, R. (2018). Neural sign language translation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7784–7793.
- Huang, J., Zhou, W., Li, H. and Li, W. (2018). Attention-based 3D-CNNs for large-vocabulary sign language recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(9), 2822–2832.
- International Organization for Standardization (ISO) (2019). *ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems*. ISO.
- Koller, O., Forster, J. and Ney, H. (2015). Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers. *Computer Vision and Image Understanding*, 141, 108–125.
- Neverova, N., Wolf, C., Taylor, G. and Nebout, F. (2016). ModDrop: Adaptive multi-modal gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(8), 1692–1706.
- Papastratis, I., Chatzikonstantinou, C., Konstantinidis, D., Dimitropoulos, K. and Daras, P. (2021). Artificial intelligence technologies for sign language. *Sensors*, 21(17), 5843.
- Yu, X., Zhang, Y., Wang, H. and Li, S. (2025). AI-driven bi-directional sign language communication system with avatar-based interaction. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 1–15.