“SAWAN” Speech to Sign Conversion Avatar for Deaf People

Sanuri Himalka
Department of Computer Engineering
University of Sri Jayewardenepura
Colombo, Sri Lanka
sanurihimalka@sjp.ac.lk

Pasindu Ranasinghe
Department of Computer Engineering
University of Sri Jayewardenepura
Colombo, Sri Lanka
pasindurana111@gmail.com

Ravinath Wanniarachchi
Department of Computer Engineering
University of Sri Jayewardenepura
Colombo, Sri Lanka
ravinathwanni@gmail.com

Maleesha Hirushan
Department of Computer Engineering
University of Sri Jayewardenepura
Colombo, Sri Lanka
hirushan.KNM@gmail.com

Udaya Wijenayake
Department of Computer Engineering
University of Sri Jayewardenepura
Colombo, Sri Lanka
udayaw@sjp.ac.lk

***Keywords*—Sinhala Sign Language, HamNoSys, Deaf-Mute, 3D Avatar, Audio/Text to Sign Language**

I. Introduction

Communication is fundamental to human interaction, yet for the Deaf and Hard of Hearing (DHH) community, with more than 466 million globally and around 300,000 in Sri Lanka, sign language is the primary means of expression. In Sri Lanka, this takes the form of Sinhala Sign Language (SSL), a unique system partially influenced by British Sign Language (BSL) [1]. Despite its significance, technological solutions supporting SSL remain limited.

SAWAN is a cross-platform speech-to-sign conversion system designed to bridge this communication gap by translating Sinhala speech into SSL using a 3D avatar. It leverages HamNoSys for structured sign representation and SiGML for machine-readable animation, enabling real-time sign generation. By integrating a sign learning platform and translation tools, SAWAN provides a better solution to support both communication and education for the local Deaf community.

II. Literature Review

Several studies have been conducted on sign language generation from text or audio, mainly focused on Indian Sign Language (ISL), with only a few efforts targeting Sinhala Sign Language (SSL). These systems typically support input in the form of either text, audio, or both. Many use HamNoSys as the sign language notation, while some rely on SignWriting.

In [1], a photorealistic approach was introduced using neural machine translation, motion graphs, and GANs to produce HD video-based British Sign Language. While visually impressive, the method requires complex processing and resources. In contrast, systems like [2] and [3] provide more practical, web-based solutions that convert voice or text into ISL using 3D avatars, HamNoSys, and SiGML. Similarly, [4] supports six regional Indian languages using speech recognition and LSTM models to output ISL through an avatar.

Some systems combine speech-to-text with NLP pipelines to produce sign sequences [5][7], but most are language-specific or rely on fixed datasets that limit adaptability. Other approaches, like [6], explored SignWriting and XML-based representations to animate 3D avatars. In the case of SSL, a few studies have proposed avatar-based systems and manual video concatenation methods to represent sign gestures [8]. However, most of these are limited to static datasets or predefined sign libraries, lacking flexibility and support for dynamic input.

The SAWAN project builds upon these existing methods but specifically targets SSL with both Sinhala text and speech input support. It uses HamNoSys and SiGML to generate structured avatar animations and allows users to modify and expand the dataset, making it more adaptable and scalable compared to previous solutions.

III. Methodology

The SAWAN system features a modular and scalable architecture with four key modules.

A. Speech-to-Sign Generator

This module forms the core of real-time translation. It uses the Google Speech-to-Text API to transcribe Sinhala audio into text. The resulting text is then mapped to pre-annotated HamNoSys notations representing corresponding SSL signs. These notations are further converted into SiGML (Signing

https://doi.org/10.31705/ADScAI.2025.16

Gesture Markup Language), a machine-readable XML-based format designed to render sign animations using a 3D avatar. This process ensures the visualization of signs with accurate hand shapes, positions, and movements. To enhance the quality of the speech input, preprocessing steps such as phoneme alignment and background noise filtering are applied. The translation system references a curated dataset of frequently used Sinhala words and phrases, enabling faster and more accurate sign generation.

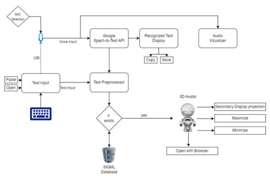

Fig. 1. Architecture of speech to sign module

B. HamNoSys Service

To support both ongoing research and dataset flexibility, the HamNoSys service provides tools for converting signs into a structured notation system. It includes:

- Automatic HamNoSys Tool: Accepts static Sinhala sign images as input and outputs a corresponding HamNoSys sequence by identifying visual sign parameters such as hand shape, position, and motion components.

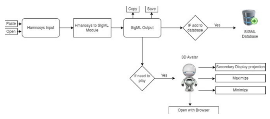

Fig. 2. Architecture of Hamnosys to SigML converter

- Manual HamNoSys Tool: Offers a user-friendly interface to manually configure HamNoSys sequences using graphical options for each component, such as handshape, palm orientation, location, and movement.
- Manual HamNoSys to SiGML Converter: Convert HamNoSys sequences into SiGML format, facilitating seamless avatar-based animation. This module ensures machine readability and is critical for visualizing newly added signs.

C. Learning Platform

To support both the Deaf community and interested learners, the system includes an interactive learning platform that categorizes signs and presents them with 3D avatar animations, structured lessons, and interactive assessments. Users can track their progress and receive feedback based on their sign recognition and reproduction abilities.

D. Sign Language Translation

This module enables text-to-sign translation by converting Sinhala text into SSL signs. It currently supports static signs and references a structured sign database. A key emphasis is placed on ensuring that the animated transitions between signs are smooth and visually natural. Techniques such as motion interpolation and real-time rendering are implemented to align animations with real-world signing standards. This component is being extended to support integration with accessibility applications such as video conferencing tools.

IV. Result And Discussion

The system was evaluated by a Sinhala Sign Language expert from the Rathmalana Deaf School, where limited (15) SSL signs were tested. An average accuracy of 80.67% was achieved.

A comparative analysis with existing ISL-based systems showed SAWAN provides improved recognition accuracy and better dataset adaptability. The structured implementation of HamNoSys contributed to consistency in generated signs. Currently, the dataset contains over 200 common Sinhala words and phrases mapped to HamNoSys. Compared to earlier SSL research attempts, SAWAN demonstrated a 20% improvement in overall translation accuracy, highlighting the value of using a modular, notation-based approach.

V. Future Work

Future work will focus on expanding the current SiGML dataset, especially to include more dynamic signs and two-handed gestures. We also plan to extend testing with students from Deaf schools to better evaluate usability in real-world learning contexts. Future versions will aim to support real-time live translation in settings like conferences.

VI. Conclusion

SAWAN is a speech-to-sign conversion system for Sinhala Sign Language (SSL), integrating HamNoSys notation and 3D avatar-based representation. The system shows promising results in translating spoken Sinhala into SSL, leveraging structured notations and real-time animations.

References

[1] S. Stoll, N. C. Camgoz, S. Hadfield, and R. Bowden, "Text2Sign: Towards Sign Language Production Using Neural Machine Translation and Generative Adversarial Networks," International Journal of Computer Vision, pp. 1–18, Jan. 2020, doi: https://doi.org/10.1007/s11263-019-01281-2.

[2] A. Yadav, R. Saxena, B. Saini, V. K. Verma, and V. Srivastava, "Audio to Sign Language Translator Web Application," IEEE Xplore, Dec. 01, 2021. https://ieeexplore.ieee.org/document/9751857

[3] B. D. Patel, H. Patel, M. A. Khanvilkar, N. Patel, and A. Thangarajah, "ES2ISL: An Advancement in Speech to Sign Language Translation using 3D Avatar Animator," 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2020, doi: https://doi.org/10.1109/CCECE47787.2020.9255783

[4] J. Peguda, V. S. S. Santosh, Y. Vijayalata, A. D. R. N, and V. Mounish, "Speech to Sign Language Translation for Indian Languages," IEEE Xplore, Mar. 01, 2022. https://ieeexplore.ieee.org/document/9784996

[5] R. San-Segundo et al.,"Speech to sign language translation system for Spanish," Speech Communication, vol. 50, no. 11–12, pp. 1009–1020, Nov. 2008, doi: https://doi.org/10.1016/j.specom.2008.02.001

[6] Y. Bouzid and M. Jemni, "An Avatar Based Approach for Automatically Interpreting a Sign Language Notation," IEEE Xplore, Jul. 01, 2013. https://ieeexplore.ieee.org/document/6601874

[7] A. Kulkarni, Arun Kariyal, V. S. Devah Dhanush, and P. B. Singh, "Speech to Indian Sign Language Translator," Jan. 2021, doi:https://doi.org/10.2991/ahis.k.210913.035

[8] M. Punchimudiyanse and R. G. N. Meegama, "3D Signing Avatar for Sinhala Sign Language," dr.lib.sjp.ac.lk, 2015, Accessed: Jun. 22, 2023.[Online].Available:http://dr.lib.sjp.ac.lk/handle/123456789/6025