



DISEÑO DE INTERFAZ PARA LA INTERPRETACIÓN Y SÍNTESIS DE VOZ DEL LENGUAJE DE SEÑAS MEXICANA BASADA EN DACTILOLOGÍA MEDIANTE TAP STRAP 2 Y ESP32



Conference Proceedings ICONIS – IX 2025
Tunja, Colombia, Mayo 21-23, 2025. Pag. 253-258

ISSN (Online): 2711-3310

**Paola, Veloz de la
Mora***

*Universidad Autónoma de
Ciudad Juárez. Av. del
Charro 450 Nte, Cd. Juárez,
Chih., México, CP 32580,
al204150@alumnos.uacj.mx*

**Israel Ulises, Ponce
Monárrez**

*Universidad Autónoma de
Ciudad Juárez. Av. del
Charro 450 Nte, Cd.
Juárez, Chih., México, CP
32580, israel.ulises@
uacj.mx*

**Christopher Enrique,
López Gómez**

*Universidad Autónoma de
Ciudad Juárez. Av. del
Charro 450 Nte, Cd. Juárez,
Chih., México, CP 32580,
al203761@alumnos.uacj.mx*

Resumen: Este artículo presenta un sistema que traduce la Lengua de Señas Mexicana (LSM) a voz utilizando el Tap Strap 2, el ESP32 y el módulo SYN6288. El sistema convierte el deletreo de palabras en audio, facilitando la comunicación entre personas sordas y oyentes. Aunque presenta limitaciones en la precisión de ciertos gestos, se planea mejorar el reconocimiento mediante aprendizaje

automático para incluir palabras completas, ampliando su utilidad y accesibilidad.

Palabras clave: Lengua de Señas Mexicana (LSM), Tap Strap 2, dactilología

1 INTRODUCCIÓN

La comunicación oral es esencial para el desarrollo humano en

* Citación: Veloz De La Mora, P., Ponce Monárrez, I. U. y López Gómez, C. E. (2025). Diseño de interfaz para la interpretación y síntesis de voz del lenguaje de señas mexicana basada en dactilología mediante tap strap 2 y esp32. *Conference Proceedings of the International Congress on Innovation and Sustainable*, Tunja, Colombia, Mayo 21-23, 2025, p.p. 253–258.

sociedad. Sin embargo, las personas con discapacidad auditiva enfrentan barreras significativas en ámbitos educativos, profesionales y sociales. En México, 2.3 millones de personas, equivalentes al 7.6 % de la población, tienen algún grado de discapacidad auditiva (Hernández Cruz, 2019). Aunque la Lengua de Señas Mexicana (LSM) permite la comunicación entre personas sordas, su uso no facilita una interacción efectiva con oyentes que no la dominan.

Para abordar esta problemática, se han desarrollado soluciones tecnológicas basadas en redes neuronales, sensores de movimiento y visión por computadora (Carrasco, 2024). Ejemplos incluyen dispositivos portátiles que traducen señales manuales a texto y audio (Chiguano Moreno, 2011), y guantes con sensores de flexibilidad que interpretan la LSM (López Granizo, 2020). No obstante, estas soluciones presentan limitaciones como altos costos, dependencia de internet y requerimientos de condiciones

controladas de iluminación y entorno (Universo UV, 2017).

Este artículo propone un sistema innovador que utiliza el microcontrolador ESP32, el dispositivo Tap Strap 2 y el módulo SYN6288 para traducir la LSM a voz. Este enfoque busca eliminar barreras de comunicación, permitiendo que personas con discapacidad auditiva interactúen verbalmente con oyentes y mejorando su calidad de vida.

2 MARCO TEÓRICO Y CONCEPTUAL

La Lengua de Señas Mexicana (LSM) es el idioma utilizado por la comunidad sorda en México, compuesto por dactilología e ideogramas. La dactilología consiste en deletrear palabras utilizando el abecedario manual, mientras que los ideogramas representan palabras completas mediante movimientos y configuraciones de las manos. En este proyecto, se optó por la dactilología debido a las capacidades técnicas del Tap Strap 2, el cual está diseñado para detectar toques y combinaciones de movimientos de los dedos, pero no

tiene la capacidad avanzada para reconocer formas y desplazamientos complejos de la mano. (Secretaría de Educación Pública [SEP], 2023)

El Tap Strap 2 es un dispositivo portátil compuesto por cinco anillos interconectados. Cada anillo incluye un chip háptico, una batería, una placa de circuito y un acelerómetro de tres ejes. El anillo del pulgar incorpora una unidad IMU de seis ejes, que registra la aceleración, rotación y posición de los dedos. Los anillos están unidos por una trenza de nailon que transmite las señales de los sensores a un microcontrolador, el cual procesa los datos y los envía vía Bluetooth a otros dispositivos. Además, el Tap Strap 2 cuenta con un SDK de código abierto, lo que facilita su integración en proyectos personalizados. (Abud-Figueroa et al., 2024)

El ESP32 es un microcontrolador de alto rendimiento con capacidades de conectividad Wi-Fi y Bluetooth, ideal para aplicaciones de IoT. En este sistema, el ESP32 recibe las señales del Tap Strap 2, las procesa y

convierte los gestos en texto. Este texto se envía al módulo SYN6288, encargado de la síntesis de voz. (Pratama & Kiswantono, 2022)

El SYN6288 es un módulo que transforma texto en habla, siendo un componente clave en este proyecto. Convierte el texto generado a partir de los gestos detectados por el Tap Strap 2 y procesados por el ESP32 en una salida de voz, facilitando la traducción de la LSM a un formato accesible para personas oyentes. (ToughDev, 2017)

3 METODOLOGÍA

La metodología se basa en el diagrama de flujo de la Figura 1 y se divide en ocho etapas:

- 1 Conexión BLE entre Tap Strap 2 y ESP32: Se inicializa la conexión Bluetooth del ESP32, se escanean dispositivos cercanos y se establece la conexión con el Tap Strap 2 mediante su dirección MAC.
- 2 Recepción y decodificación de datos BLE: Se reciben datos en tiempo real del Tap Strap 2, extrayendo

información de sus sensores y analizando la estructura de transmisión para su procesamiento.

3 Mapeo de gestos y construcción de base de datos: Se registran las señales del Tap Strap 2 asociadas a cada letra dactilológica, almacenándolas y ajustando los datos para reducir variaciones y ruido.

4 Algoritmo de reconocimiento de gestos: Se comparan los datos en tiempo real con los patrones almacenados utilizando métodos de clasificación para identificar la letra ingresada.

5 Concatenación de caracteres: Se diseña una estructura que almacena letras secuenciales y detecta pausas para formar palabras correctamente.

6 Conversión de texto a voz: El texto reconocido se envía al módulo SYN6288 a través de UART, ajustando los parámetros de síntesis de voz.

7 Validación y ajuste: Se realizan pruebas con usuarios para ajustar los umbrales de clasificación y mejorar la precisión.

8 Optimización: Se perfeccionan los algoritmos para reducir el tiempo de respuesta y optimizar el consumo energético del ESP32.



FIGURA 1. METODOLOGÍA DEL PROYECTO.

4 RESULTADOS

En este trabajo, se desarrolló un sistema que traduce la Lengua de Señas Mexicana (LSM) a voz, integrando el Tap Strap 2, el ESP32 y el módulo SYN6288. La comunicación entre el Tap Strap 2 y el ESP32 se estableció mediante Bluetooth, utilizando la dirección MAC del dispositivo. El ESP32 recibió los datos, los procesó y los convirtió en texto, el cual se envió al módulo SYN6288 para generar audio.

El sistema funciona mediante deletreo, donde cada gesto captado

corresponde a una letra, y estas se concatenan para formar palabras (ver Figuras 2 y 3). Sin embargo, no reconoce palabras completas, lo que representa una limitación actual. Además, la precisión del sistema se ve afectada por la posición de la mano y la complejidad de ciertos movimientos. Por ejemplo, letras como la "J", que requiere trazar un arco con el dedo meñique, o la "Z", que implica un movimiento en el aire, presentan mayores desafíos para su reconocimiento.

A futuro, se planea mejorar el sistema mediante aprendizaje automático, con el objetivo de reconocer palabras completas utilizando ideogramas, lo que aumentaría su precisión y utilidad en aplicaciones prácticas.



FIGURA 2. DETECCIÓN DE CADA LETRA Y CONCATENACIÓN DE LA PALABRA.



FIGURA 3. SÍNTESIS DE VOZ DE LA PALABRA.

5 CONCLUSIONES

Este sistema cumplió su objetivo al traducir la Lengua de Señas Mexicana (LSM) a voz, facilitando la comunicación entre personas sordas y oyentes. Utilizando el Tap Strap 2, el ESP32 y el módulo SYN6288, se logró convertir el deletreo de palabras en audio, eliminando barreras de comunicación. Además, el sistema sirve como herramienta de aprendizaje para oyentes, permitiéndoles visualizar y escuchar el deletreo en LSM. Aunque persisten limitaciones en la detección de ciertos gestos, la implementación de aprendizaje automático podría mejorar la precisión y permitir el

reconocimiento de palabras completas, ampliando su utilidad y alcance.

6 REFERENCIAS

- Abud-Figueroa, M. A., Juárez-Martínez, U., Rodríguez-Mazahua, L., & Muñoz-Contreras, H. (2024). *Uso de TAP STRAP 2 para el control gestual con manos*. Boletín UPIITA del IPN, (101). <https://www.boletin.upiita.ipn.mx/index.php/ciencia/1070-cyt-numero-101/2291-uso-de-tap-strap-2-para-el-control-gestual-con-manos>
- Carrasco, M. (2024). *Uso de redes neuronales para la traducción de LSM*. Instituto Tecnológico Nacional de México. https://rinacional.tecnm.mx/bitstream/TecNM/8161/1/Tesis-Martin_Carrasco.pdf
- Chiguano Moreno, D. (2011). *Desarrollo de dispositivos portátiles para la traducción de LSM*. Escuela Politécnica Nacional. https://bibdigital.epn.edu.ec/bitstream/15000/4924/1/PAPER_CHIGUANO_MORENO.pdf
- Hernández Cruz, J. C. (2019). *Datos sobre la población con discapacidad auditiva en México*. Instituto Tecnológico de Hermosillo. https://rinacional.tecnm.mx/bitstream/TecNM/1538/2/2019-149_JuanCarlos_Hernandez_Cruz.pdf
- López Granizo, K. J. (2020). *Implementación de guantes con sensores de flexibilidad para LSM*. Universidad Indoamérica. <https://dspace.ups.edu.ec/bitstream/123456789/21759/1/UPS-GT003587.pdf>
- Pratama, E. W., & Kiswanton, A. (2022). *Electrical analysis using ESP-32 module in realtime*. Journal of Electrical Engineering and Computer Sciences, 7(2), 1273-1284. https://www.researchgate.net/publication/367134207_Electrical_Analysis_Using_ESP-32_Module_In_Realtime
- Secretaría de Educación Pública (SEP). (2023). *Diccionario de la Lengua de Señas Mexicana (LSM)*. Gobierno de México. https://educacionespecial.sep.gob.mx/storage/recursos/2023/05/xzrfl019nV-4Diccionario_lengua_%20Senas.pdf
- ToughDev. (2017). *SYN6288 Chinese speech synthesis module*. <https://openlab.citytech.cuny.edu/superhero/files/2017/09/SYN6288-Chinese-Speech-Synthesis-Module--ToughDev.pdf>
- Universo UV. (2017). *Complejidad de lengua de señas dificulta su traducción automática*. Universidad Veracruzana. <https://www.uv.mx/universo/general/complejidad-de-lengua-de-senas-dificulta-su-traduccion-automatica/>