

Alan RAMÍREZ NORIEGA*

INTRODUCCIÓN

El Lenguaje de Señas Mexicano es una lengua natural de expresión que emplean las personas Sordas en México. Este lenguaje se basa en movimientos manuales en tres dimensiones, gestos y la precepción visual para expresar una idea, de esta manera las personas Sordas pueden comunicarse en su entorno social.

Actualmente, se está desarrollando un Sistema Web para la búsqueda de videos que muestren como traducir palabras o frases de español al Lenguaje de Señas Mexicano. El Sistema Web permite subir videos y consultarlos a público en general. Esto permite que usuarios que quieran comunicarse a través del lenguaje de señas lo puedan hacer tomando como apoyo la búsqueda de videos del repositorio. Por tal motivo, es indispensable contar con Sistema de Recuperación de Información adecuado.

Un Sistema de Recuperación de Información se define como un conjunto de procesos desarrollados con el fin de recuperar, almacenar y regresar información de acuerdo con las consultas de un usuario. Un ejemplo de este tipo de sistemas es el empleado por el buscador Google para encontrar los sitios de mayor relevancia de acuerdo con las palabras de entrada, en este caso, se buscan sitios de internet. El sitio YouTube también tiene un sistema de recuperación de información, aunque en este caso recupera videos.

Un Sistema de Recuperación de Información funcionaria muy bien si el usuario pudiera transmitir su necesidad de información de manera precisa. Sin embargo, la consulta del usuario rara vez es completa. La consulta proporcionada por el usuario es a menudo desestructurada e incompleta. Una consulta incompleta impide que un motor de búsqueda satisfaga la necesidad de información del usuario. En la práctica, es necesaria alguna representación que pueda expresar correcta y completamente la necesidad de información del usuario.

DESARROLLO

Un método para mejorar el rendimiento de la recuperación de información es complementar una consulta original con términos adicionales, es decir, reformular una consulta para tener mayor éxito en la recuperación de información, a esto se le llama expansión de la consulta. La expansión de la consulta se puede realizar de forma automática o interactiva y puede tener lugar en la

formulación inicial de la consulta, o en una etapa de reformulación de la consulta de la búsqueda en línea, o en ambas.

El Sistema Web para el Lenguaje de Señas requiere recuperar información basado en videos. Es decir, el retorno de las consultas es para ver videos, sin embargo, este es

CIENCIA
UNIVERSITARIA
PARA TODOS

Aplicando aspectos lingüísticos para la búsqueda de videos del Lenguaje de Señas Mexicano



un tipo especial de video, los cuales se caracterizan por ser cortos (dos a diez segundos), no es necesario el sonido y el video muestra señas de Lenguaje de Señas Mexicano. Por tanto, la consulta se basa solamente en la descripción y nombre que los usuarios hagan de los videos. Esto complica la búsqueda

da, por cual, para tener un mayor éxito de recuperación de videos se aplican técnicas de expansión de consultas.

Básicamente, el sistema web explora la base de información para encontrar coincidencias en las palabras o frases relacionadas al video. Un sistema de búsqueda tradicio-



nal solo buscaría coincidencias exactas de la palabra de entrada, sin embargo, es posible aplicar un procesamiento lingüístico para explotar características de la lengua y obtener mejor resultados de búsqueda, para estos casos se puede aplicar corrección de ortografía, la lematización, uso de sinónimos y antónimos para la búsqueda. Dado que se emplean estos términos extras a la consulta tradicional, se le llama expansión

Por otra parte, también es posible emplear sinónimos en la búsqueda. Dos palabras son sinónimos en un contexto cuando pueden intercambiarse, es decir son palabras con un significado en común, por ejemplo, comprar y adquirir. De esta forma, un video de una seña puede ser buscado de acuerdo con su sinónimo, de esta forma, se amplían las posibilidades de coincidencias. Así, si la búsqueda inicial es



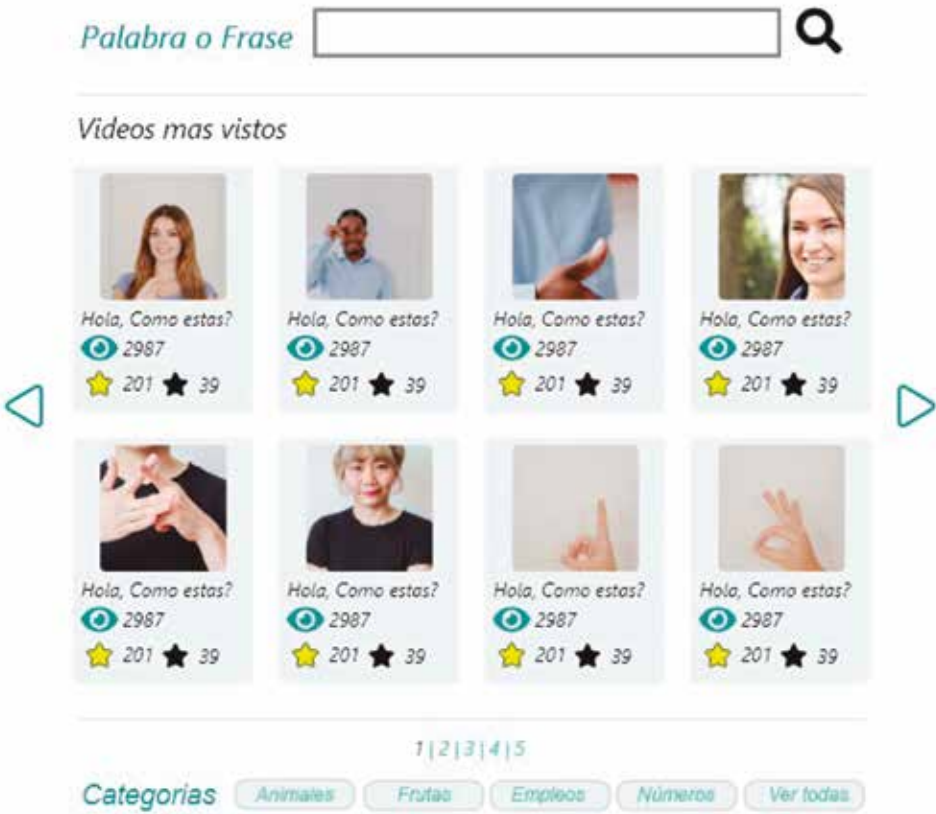
de la consulta.

De inicio, las palabras de búsqueda tienen que ser corregidas ortográficamente, ya que todo proceso fracasara si las palabras no están bien escritas. Posteriormente, el proceso de búsqueda aplica la lematización, el cual es un proceso lingüístico donde se ubica el lema o raíz de la palabra. Por ejemplo, la palabra cantante tiene como raíz el verbo cantar, por lo que el sistema buscaría todos los videos que tengan como lema la palabra cantar; de esta forma la respuesta a la consulta sería más rica. A su vez, si se busca música, algunas variaciones a música pueden aparecer como músicos o musicales.

carro, también puede buscar automóvil o coche, de esta forma se amplían las posibilidades de éxito.



*Obtuvo su maestría en Computación Aplicada en la Universidad Autónoma de Sinaloa en 2014 y su doctorado en Ciencias de la Computación en la Universidad Autónoma de Baja California en 2017, ambos temas de investigación fueron en torno a redes Bayesianas. Es Profesor e Investigador de Tiempo Completo de la Facultad de Ingeniería Mochis en la Universidad Autónoma de Sinaloa. Actualmente es miembro del Sistema Nacional de Investigadores (SNI) nivel 1 en el área IX (Interdisciplinaria), miembro adherente de la Academia Mexicana de Computación (AMEXCOMP) y Miembro de la Red Temática de Ingeniería de Software (RedMIS). Cuenta con diversas publicaciones en revistas de alto impacto (JCR) y congresos nacionales e internacionales en temas relacionados a Sistemas Tutores Inteligentes, Ingeniería de Software y Minería de texto, siendo estas últimas las áreas principales de interés. Además, ha participado en diversas direcciones y sinodalias de tesis de Licenciatura, Maestría y Doctorado.



CONCLUSIONES

Como se puede apreciar, el proceso de búsqueda de información está muy relacionada con el lenguaje natural, así el estudio del idioma abre importantes funcionalidades a los sistemas de información. Para esta propuesta, se emplean varios aspectos lingüísticos para mejorar las opciones de búsqueda, así, mostrar mejores resultados a los usuarios para expresar una frase en el Lenguaje de Señas Mexicano.

Este tipo de proyectos tiene un carácter social, ya que se trabaja con necesidades de las personas Sordas. Esto implica asesoramiento con personas que dominan el LSM y establecer una comunicación con personas Sordas para validar la aplicación. De esta forma, se apoya un aspecto educativo en el ambiente social para el aprendizaje de nuevas formas de comunicación con personas con discapacidad. 🇲🇽