

Rehablado intralingüístico y traducción automática para subtitulado en vivo: un estudio de calidad y percepción

Intralingual respeaking and machine translation for live subtitling: A study on quality and perception

Laura Ríos-Valero

Universidad de Valladolid

Valladolid, España

lauralisset.rios@estudiantes.uva.es

<https://orcid.org/0000-0002-8114-2822> 

Resumen: El presente estudio de caso analiza la calidad del subtitulado interlingüístico en vivo generado mediante un flujo de trabajo compuesto por rehablado intralingüístico y traducción automática. El enfoque fue cualitativo y el nivel, descriptivo. El corpus estuvo constituido por los subtítulos en español de una entrevista en inglés. Se aplicó el modelo NTR para medir la tasa de precisión, así como una evaluación de parámetros complementarios como el retraso, la velocidad y la coherencia audiovisual. Además, se recogió la percepción de 31 participantes a través de un cuestionario que evaluó la comprensión literal, inferencial y crítica, así como los otros factores de calidad del modelo NTR. Los resultados revelaron una tasa de precisión del 97,57 % (4/10), que no alcanzó el umbral mínimo de calidad. Los participantes también encontraron dificultades de lectura, retraso significativo y problemas en la comprensión literal. Se concluye que, en este flujo de trabajo, la traducción automática sigue representando el principal obstáculo para garantizar subtítulos de calidad, principalmente por errores de forma.

Palabras clave: calidad del subtitulado; subtitulado interlingüístico en vivo; rehablado intralingüístico; traducción automática; accesibilidad lingüística.

Abstract: This case study analyzes the quality of live interlingual subtitling produced through a workflow combining intralingual respeaking and machine translation. The study follows a qualitative, descriptive approach. The corpus consists of Spanish subtitles derived from an English-language interview. The NTR model was applied to calculate the accuracy rate and assess additional quality parameters such as delay, reading speed, and audiovisual coherence. Furthermore, the perceptions of 31 participants were collected via a questionnaire designed to measure literal, inferential, and

critical comprehension, as well as subjective evaluation of quality factors. The results revealed an accuracy rate of 97.57% (4/10), which falls below the minimum acceptable quality threshold. Participants also reported difficulties related to subtitle legibility, significant delay, and challenges in literal comprehension. The study concludes that, within this workflow, machine translation remains the main barrier to achieving high-quality subtitles, primarily due to frequent errors in form.

Keywords: subtitling quality; interlingual live subtitling; intralingual respeaking; machine translation; linguistic accessibility.

I. Introducción

Durante la pandemia de la Covid-19, las redes sociales fueron una herramienta de doble filo. Por un lado, difundieron información imprecisa y pesimista que exacerbaba los riesgos de la salud mental. Por otro lado, también fueron un canal para divulgar información verdadera que ayudó a disipar los miedos y la incertidumbre. Una de sus principales ventajas es su fácil acceso desde dispositivos electrónicos personales (Williams, 2022). Entre estas redes, destaca YouTube que, de acuerdo con Nashville Film Institute (2022), se considera una red social, ya que permite a los usuarios compartir información e intereses comunes.

En Perú, el 85 % de los habitantes usa esta plataforma a diario, principalmente para ver contenidos en video. Entre ellos, el público entre 18 y 24 años prefiere YouTube frente a la televisión nacional, y el 89 % considera que esta página les permite profundizar en temas de interés y desarrollar habilidades (Radio Programas del Perú, 2019). Según Li et al. (2020), el valor de YouTube radica en que su contenido audiovisual lo vuelve accesible a individuos con distintos perfiles demográficos.

Si se observan las realidades de otros países, como España, YouTube se volvió una fuente clave de información de medios de comunicación convencionales durante la pandemia de la Covid-19. La vista de videos aumentó ante el decreto del estado de alarma y alcanzó picos durante sus prórrogas o cuando los contagios incrementaban (Gil Ramírez et al., 2020). En el estudio de Li et al. (2020), se determinó que entre el 64 % y el 67 % de videos de YouTube en inglés contenían información veraz sobre dicha coyuntura, aunque los videos de tipo profesional o gubernamental, a pesar de tener mayor reputación, estaban poco representados. Por otro lado, Meza-Liviapoma et al. (2021) determinaron que, en América Latina y el Caribe, el 50,39 % de los videos de YouTube sobre la Covid-19 eran educativos, aunque el 25,38 % de estos infundían miedo y otro 25 % no incluía recomendaciones completas de la OMS.

El 5 de mayo de 2023, el Dr. Tedros Adhanom Gebreyesus, en representación de la OMS, declaró que la Covid-19 ya no era una emergencia sanitaria mundial. Sin embargo, recalcó que no significaba el final de la enfermedad, e instó a los países, sus poblaciones y sus sistemas de salud a mantenerse alertas (Mundasad & Roxby, 2023). A partir de este anuncio, muchos profesionales de la salud y medios de comunicación han continuado difundiendo información sobre los cuidados necesarios, pues existe el riesgo de que la población malinterprete que el peligro ha desaparecido. Además, parte de esta nueva información podría no estar disponible en español, dado que el inglés sigue siendo el idioma dominante en la comunicación científica.

Por ello, es necesario garantizar el acceso a las fuentes originales en su idioma de origen, junto con su traducción o interpretación al idioma del usuario final. Muchas de estas fuentes se

presentan en formato audiovisual, como los videos de YouTube. En cuanto a la accesibilidad en los medios de comunicación, Greco (2018) señala que esta no debe limitarse a las restricciones sensoriales, sino también lingüísticas. Romero-Fresco (2013) coincide al afirmar que, en el rodaje de películas, por ejemplo, la accesibilidad debe ser un término que incluya no solo a personas con discapacidad auditiva o visual, sino también a quienes no conocen el idioma original de la película. En ese sentido, el propio autor también defiende una visión ampliada de la accesibilidad en los medios, al destacar que los subtítulos para sordos son utilizados por audiencias más amplias, incluyendo personas que no dominan la lengua original del audiovisual, lo cual refuerza la idea de una accesibilidad lingüística de carácter universal (Villela, 2019).

En consecuencia, una posible solución en tiempo real para mejorar la accesibilidad lingüística a contenidos audiovisuales sería el subtulado interlingüístico en vivo. Esta opción es especialmente útil ante la rápida desactualización de noticias sobre la Covid-19. Muchas de las noticias de distintos países se transmiten en vivo por YouTube y se puede usar su subtulado generado automáticamente. Sin embargo, este no es preciso, no siempre es coherente, presenta omisiones y carece de puntuación. Ante estas limitaciones, se plantea la intervención humana dentro de un flujo de trabajo que combine el rehablado intralingüístico y la traducción automática. En esta modalidad, un rehablador repite oralmente el discurso original en el mismo idioma para que un software de reconocimiento de voz lo transcriba, y luego se utiliza una herramienta de traducción automática para generar subtítulos en otro idioma. Esta colaboración entre humano y máquina busca mejorar la calidad del subtulado interlingüístico en vivo y podría evaluarse aplicando el modelo NTR.

Romero-Fresco y Alonso-Bacigalupe (2022) analizaron el flujo de trabajo de hombre y máquina y obtuvieron resultados positivos en diferentes aspectos que se tratarán en este artículo. Cabe señalar que para dicho estudio se usaron dos videos de TED Talk con contenido poco técnico, baja densidad léxica y estructuras sintácticas simples. Una de las interrogantes que dejaron abiertas fue si este enfoque funcionaría con temas más especializados. Por esta razón, en el presente estudio, se optó por un tema médico de gran relevancia: el futuro de la Covid-19.

El presente trabajo de investigación tiene como objetivo analizar la calidad del subtulado interlingüístico en vivo mediante rehablado intralingüístico y traducción automática, aplicando el modelo NTR y conociendo la percepción de los usuarios. Asimismo, este estudio se justifica al aplicar el modelo NTR de Romero-Fresco y Pöchhacker (2018), que evalúa la tasa de precisión y otros factores; al seguir un adecuado método científico para el desarrollo del estudio; y al responder a una necesidad creciente de accesibilidad lingüística en transmisiones en vivo, especialmente en contextos sanitarios y educativos.

2. Consideraciones teórico-prácticas

El reconocimiento de voz o reconocimiento de voz automático (ASR, por sus siglas en inglés) es la capacidad de un software para procesar el habla humana y convertirla en texto escrito (IBM, 2021). De acuerdo con Zajechowski (2014), el proceso empieza cuando se habla al programa, que crea un archivo de ondas de sonido. Luego, se elimina el sonido de fondo y el audio se divide en fonemas. Cada fonema, como eslabón de una cadena, se analiza en secuencia. A partir de ello, el software estima la probabilidad de estadística para deducir palabras y construir oraciones. De acuerdo con Defrancq y Fantinuoli (2021), el ASR emplea algoritmos que analizan el texto y realizan

cambios para generar una transcripción del audio. Sin embargo, esta tecnología no reconoce intenciones ni significados, por lo que no puede considerarse una forma de inteligencia artificial (Romero-Fresco, 2011).

El ASR se ha empleado en diversos contextos donde la precisión es fundamental. Uno de ellos es el de los operadores virtuales, como los asistentes virtuales o sistemas automatizados de atención al cliente que interactúan con los usuarios mediante comandos de voz. Se han utilizado también en el subtítulo en vivo y en la toma de notas en entornos médicos (Rella, 2022). Además, esta tecnología se ha aplicado en el aprendizaje de idiomas para mejorar la pronunciación (Morton et al., 2012; Golonka et al., 2014; Evanini et al., 2018) y como herramienta de apoyo en la interpretación de idiomas asistida por computadora (Defrancq & Fantinuoli, 2021). Los estudios de Cucchiari et al. (2009) y Golonka et al. (2014) destacan que la precisión no es total, mientras que Morton et al. (2012) y Evanini et al. (2018) señalan tasas de error elevadas. Aun así, Defrancq y Fantinuoli (2021) afirman que la precisión es alta (96 %) y que el retraso es apropiado para los tiempos de decalaje en interpretación.

El ASR es la tecnología que se usa para generar los subtítulos automáticos (Wald & Bain, 2008), pero estos presentan tres principales limitaciones (Chan et al., 2019): la precisión (Parton, 2016), la velocidad de lectura (Romero-Fresco, 2016) y la segmentación del texto (Rajendran et al., 2013). A pesar de ello, el sistema ha evolucionado: hoy puede manejar vocabularios más amplios, es independiente del orador, reconoce discursos espontáneos y opera en entornos con ruidos (Lu et al., 2020). Estas mejoras indican que su calidad podría seguir aumentando.

Una de las técnicas más recientes utilizadas en el subtítulo en vivo es el rehablado (*respeaking*). Esta consiste en que un rehablador repita oralmente el discurso original utilizando un software de reconocimiento de voz, con el menor retraso posible e incluyendo puntuación y convenciones específicas para personas con discapacidad auditiva (Keyes, 2007; Lambourne, 2007; Romero-Fresco, 2011). Según Arma (2015), el proceso incluye un primer discurso oral que el rehablador vuelve a expresar oralmente; luego, el software lo convierte en texto escrito. Además, el rehablado permite revisar el producto generado para detectar omisiones o errores de pronunciación que hayan afectado la claridad del mensaje (Belenguer-Cortes, 2024b). Por tratarse de una tarea exigente, no se recomienda que los rehabladores trabajen más de una hora seguida (Romero-Fresco & Eugeni, 2020).

El rehablado se usa tanto en contextos médicos como en subtítulo en vivo para personas con discapacidad auditiva. En entornos médicos, se utiliza como alternativa a la transcripción manual, mediante dictado o sin conexión a internet. En el subtítulo en vivo, el rehablador debe generar subtítulos fieles al discurso original, respetando las necesidades lingüísticas y comunicativas del público, las restricciones técnicas y la falta de tiempo para editar (Arma, 2015). Esta técnica se emplea en conferencias, seminarios, eventos deportivos, televisión, transmisión por internet, entre otros. Los subtítulos generados mediante rehablado pueden ser *verbatim* (palabra por palabra), aunque en muchos casos los rehabladores deben parafrasear para adaptarse a la velocidad de lectura de la audiencia media, la velocidad de habla en el discurso original y la necesidad de incluir puntuación oral (Romero-Fresco, 2009).

Existen dos tipos de rehablado según la lengua meta: el intralingüístico (de la lengua oral A a la lengua escrita A) y el interlingüístico (de la lengua oral A a la lengua escrita B). El primero es el más común a nivel mundial. El segundo ha ganado relevancia en contextos multilingües, como

eventos internacionales o emisiones televisivas en países como Bélgica o el Reino Unido (Romero-Fresco, 2011; Romero-Fresco & Pöchhacker, 2018; Romero-Fresco & Eugeni, 2020). Esta modalidad interlingüística añade un nivel adicional de exigencia, ya que el rehablador no solo debe repetir con precisión el contenido original, sino también traducirlo en tiempo real. Como señalan Davitti y Sandrelli (2020), el rehablado interlingüístico implica coordinar simultáneamente la escucha, la oralización con puntuación y etiquetas, el control prosódico para minimizar errores del software, y la supervisión y edición del texto generado para garantizar su comprensión y legibilidad.

Según Romero-Fresco y Eugeni (2020), la implementación del rehablado y el perfil de los rehabladores varía según el contexto geográfico. En Europa, la formación tiende a ser más académica mientras que en Estados Unidos, es más técnica y está orientada a la cultura sorda, el uso de software especializado (Dragon NaturallySpeaking, SpeechCAT), la resolución de problemas técnicos y el dominio de plataformas remotas como Zoom, YouTube o Adobe Connect. Robert et al. (2020) describen que la mayoría de subtituladores en vivo son mujeres menores de 40 años, con formación en lenguas y estudios de posgrado, y trabajan de manera autónoma o para empresas especializadas. Romero-Fresco y Eugeni (2020) también detallan las herramientas comunes utilizadas: micrófonos USB, audífonos, pantallas y software para reconocimiento de voz y visualización de subtítulos. Las condiciones ideales incluyen espacios insonorizados y, en algunos países, se emplean recursos adicionales como mandos para insertar puntuación o pantallas táctiles para ingresar palabras clave.

Para evaluar la calidad del subtitulado en vivo, Romero-Fresco y Eugeni (2020) proponen considerar factores como la precisión y el retraso. La precisión depende del modelo de calidad utilizado para clasificar errores. El retraso promedio para la aparición del subtítulo suele oscilar entre 5 y 10 segundos, según el proceso de corrección. Otros parámetros importantes incluyen el formato de los subtítulos, la identificación de los interlocutores, los efectos de sonidos, los comandos y la tasa de edición.

Para medir la precisión del subtitulado intralingüístico en vivo, Romero-Fresco y Martínez Pérez (2015) desarrollaron el modelo NER. Este evalúa si los errores afectan la coherencia o alteran el contenido. Este modelo puede ser usado para analizar el subtitulado en vivo por rehablado o por subtitulado en vivo automático. Los errores se clasifican en graves (1 punto), estándares (0,5 punto) y menores (0,25 punto), reconociendo que no todos tienen el mismo impacto. El modelo se centra en la experiencia del usuario y establece que se alcanza calidad cuando la tasa de precisión es de al menos 98 %.

Figura 1: Modelo NER

$$\text{Accuracy} = \frac{N - E - R}{N} \times 100$$

CE (correct editings):
Assessment:

Fuente: Romero-Fresco y Martínez Pérez (2015)

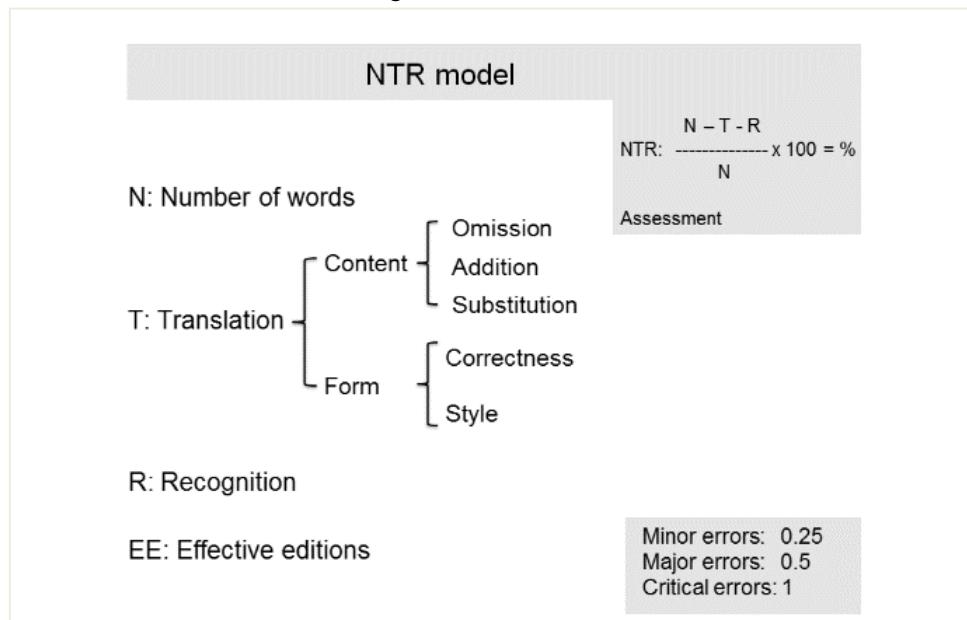
En la figura 1, se observa el modelo NER, que está compuesto por varios elementos. La letra N representa el número total de palabras rehabladas, incluyendo signos de puntuación e identificación de los interlocutores. La letra E indica los errores de edición que provienen de

decisiones del rehablador o, en el caso del subtitulado automático, errores de puntuación e identificación de los interlocutores. La letra R corresponde a los errores de reconocimiento por mala pronunciación del rehablador o mal entendimiento del software usado. Las CE son las ediciones correctas; estas son correcciones hechas por el rehablador que no afectan el contenido del discurso. Finalmente, el componente de evaluación incluye factores como la velocidad y el retraso del subtitulado, la velocidad de habla, el flujo de subtítulos en pantalla, la identificación de interlocutores, la coherencia entre la imagen/el sonido y los subtítulos, el tiempo perdido en las correcciones, entre otros. Según este modelo, la calidad de subtitulado no depende únicamente de la tasa de precisión, sino también de una evaluación integral.

Los errores se valoran según su impacto en la comprensión del mensaje. Los errores graves provocan falsos sentidos indetectables para el espectador. Los errores estándar generan sin sentidos por omisión de unidades de sentido independientes y dependientes. Por último, los errores menores incluyen errores ortotipográficos o leves adiciones u omisiones que no afectan el sentido. Este modelo es de fácil aplicación y ha sido adoptado por compañías de televisión, instituciones educativas y organismos normativos para contenidos audiovisuales en muchos países.

Romero-Fresco y Pöchhacker (2018) adaptaron este modelo al subtitulado interlingüístico en vivo, dando origen al modelo NTR. Este modelo mantiene la fórmula del modelo NER, pero reemplaza los errores de edición (E) por errores de traducción (T). Los errores graves se renombran como “críticos” y los errores estándares como “mayores”. La escala de puntuación se mantiene: 1 punto para errores críticos, 0,5 para mayores y 0,25 para menores. Además, la tasa de precisión se reajusta a una escala del 0 al 10, donde el 98 % del modelo NER corresponde a un 5 en el modelo NTR.

Figura 2: Modelo NTR



Fuente: Romero-Fresco y Pöchhacker (2018)

En la figura 2, se observa el modelo NTR. N sigue representando el número total de palabras en el subtitulado. Los errores de traducción (T) se basan en el discurso del rehablador y se dividen

en errores de contenido y errores de forma. Los errores de contenido son malas traducciones por omisión, adición o sustitución; los errores de forma se dividen en errores de corrección (terminología, gramática) y errores de estilo (naturalidad, registro, pertinencia). También se registran errores de reconocimiento (R) causados por fallos en el software. Las ediciones correctas (CE) se renombraron como “ediciones efectivas” (EE), ya que contribuyen a una comunicación efectiva.

Al igual que el modelo NER, la evaluación final abarca la tasa de precisión, las ediciones efectivas, el retraso, la velocidad y el flujo de subtítulos, la identificación de los interlocutores y la coherencia con el audio y la imagen. Si hubiera discrepancia entre la tasa de precisión y la evaluación general, esta última tiene prioridad para determinar si el subtitulado es de calidad.

Según Romero-Fresco y Eugeni (2020), las características del subtitulado que se deben evaluar incluyen la precisión, el retraso, la velocidad, la tasa de edición y la forma de presentación del subtitulado (palabra por palabra o en bloque). Estas características están respaldadas tanto por la investigación académica, como por entidades profesionales y medios de comunicación.

Por su parte, Romero-Fresco y Alonso-Bacigalupe (2022) estudiaron cinco flujos de trabajo para el subtitulado interlingüístico en vivo. Estos fueron:

- 1) rehablado interlingüístico
- 2) interpretación simultánea y rehablado intralingüístico
- 3) interpretación simultánea y reconocimiento de voz automático
- 4) rehablado intralingüístico y traducción automática
- 5) reconocimiento de voz automático y traducción automática

Los flujos 2 y 4 obtuvieron los mejores resultados con respecto a la precisión. En ambos casos, la descentralización de las tareas y la intervención humana parecen ser claves del éxito. En el flujo 2, dos personas se reparten la tarea, por lo que reduce la carga cognitiva. En el flujo 4, el proceso comienza con una intervención humana y termina con una máquina.

La traducción automática, definida por Sánchez Ramos y Rico Pérez (2020) como el uso de tecnología para traducir sin intervención humana, es una solución ampliamente adoptada por empresas que necesitan traducir grandes volúmenes de contenido (Alsan, 2022). Sin embargo, la traducción es una tarea compleja, ya que involucra creatividad, traducción de términos nuevos y mediación cultural (Arnold, 2003). Oliver et al. (2012) destacan que las computadoras solo siguen reglas de manera mecánica y literal, lo que limita la fidelidad, la inteligibilidad, la precisión y el estilo de las traducciones. La traducción automática traduce frase por frase y, por eso, no toma en cuenta la progresión temática. Esto puede afectar la coherencia del mensaje. Además, no selecciona el léxico correctamente, no capta el tono del texto y no logra evitar la lectura reiterativa o pesada. Arnold (2003) también señala que las computadoras también tienen problemas con el razonamiento del sentido común, ya que este está compuesto de conocimientos generales sobre el mundo, y enfrentan la “explosión combinatoria”, que es la gran cantidad de opciones para resolver un mismo problema.

Pese a estas limitaciones, la traducción automática también presenta ventajas: puede traducir grandes volúmenes en poco tiempo, es de bajo costo, permite traducir páginas web, facilita la comunicación en línea y esta tecnología puede ser configurada de manera más personalizada gracias a glosarios o ajustes para mantener el registro (Oliver et al., 2012). Además, es fácil de usar, se

puede integrar con memorias de traducción y, en algunos casos, ofrece resultados fiables, ya que muchos programas han sido auditados por hablantes nativos y traductores expertos (Alsan, 2022).

Según Abadou y Khadich (2019), existen dos tipos de traducción automática: la traducción automática por programa y la traducción automática en línea. La primera requiere la instalación de un programa en la computadora y funciona sin conexión a internet. La segunda, más popular, se basa en servicios web. Algunos son de pago y otros, gratuitos. Alsan (2022) menciona que los cinco mejores traductores automáticos son Google Translate, DeepL Translator, Bing Microsoft Translator, SYSTRAN Translate y Amazon Translate. Rubio (2020) coincide en destacar Google Translate como uno de los más eficaces, ya que, según Alsan (2022), este usa algoritmos sofisticados neuronales para generar traducciones de alta calidad.

3. Metodología

La presente investigación sigue un enfoque cualitativo, bajo el diseño de estudio de caso, el cual permite examinar un fenómeno específico sin pretensión de generalización (Simons, 2013). El caso analizado es la calidad del subtulado interlingüístico generado en tiempo real mediante un flujo de trabajo que combina el rehablado intralingüístico con la traducción automática.

Para generar el subtulado, se trabajó con una rehabladora profesional de Estados Unidos, quien aceptó participar, tras intentar hacer contacto con distintos rehabladores del extranjero. Esta profesional tiene una formación técnica en transcripción judicial y más de doce años de experiencia profesional. Su estilo de rehablado es *verbatim*, en consonancia con su experiencia en tribunales. El equipo que la rehabladora usó fue el habitual: computadora portátil, *stenomask*, Dragon NaturallySpeaking, CATALYST y SpeechCAT, herramienta que además permite almacenar las transcripciones del rehablado.

El video seleccionado fue una entrevista de TED titulada “Is pandemic actually over? It’s complicated”, en la que el Dr. Anthony Fauci, inmunólogo y asesor del gobierno de EE. UU. conversa con el periodista David Biello sobre la situación pospandémica de la Covid-19. El video tiene una duración de 28 minutos y 28 segundos, se encuentra en inglés, contiene terminología médica accesible y se presenta con una velocidad promedio de 160 palabras por minuto.

Una semana antes de la reunión con la rehabladora, se le informó sobre el tema para que pueda prepararse adecuadamente. Luego, se organizó una reunión por Zoom, durante la cual se proyectó el video y ella rehabló en inglés en tiempo real durante 15 minutos, siguiendo el tiempo usual para pruebas de selección laboral (Romero-Fresco, 2011). El rehablado se registró en la ventana de subtítulos de Zoom y, tan pronto como se generaban, la investigadora de este estudio los tradujo inmediatamente al español con Google Translate y los insertó manualmente en la misma ventana. La configuración predeterminada de Zoom mostró los subtítulos en bloques de hasta cuatro líneas. Esta presentación en bloque sigue el patrón de lectura descrito por Romero-Fresco (2011), quien sostiene que este formato facilita una lectura más rápida y organizada. Además, la plataforma usa la letra Lato y el tamaño de letra depende de la configuración del receptor. En este caso, se usó el tamaño pequeño predeterminado, y así se generó un video final con subtulado en español.

Este video subtulado fue proyectado a un grupo de 31 participantes hispanohablantes durante una segunda reunión de Zoom. La muestra fue seleccionada siguiendo un criterio intencional

por conveniencia, mediante contactos indirectos: alumnos y conocidos de la investigadora compartieron la invitación con familiares y amigos. Se obtuvo una muestra diversa en género y edad (hombres y mujeres entre 16 y 63 años), con distintos niveles de instrucción (entre educación básica incompleta y educación superior completa), lo que permitió observar la variación de comprensión y lectura de subtítulos entre los participantes. Desde un enfoque de accesibilidad universal, el único criterio de inclusión fue encontrarse en discapacidad comunicativa, es decir, no tener conocimientos de inglés, dado que el propósito del estudio era evaluar si el subtulado traducido facilitaba el acceso al contenido original.

Luego de ver el video, se utilizó un formulario de Google Forms, adaptado de Ríos Valero (2023), que incluía en su encabezado el consentimiento informado, donde se explicó a los participantes el propósito del estudio, la confidencialidad de los datos y la voluntariedad de su participación. El formulario constaba de tres secciones. La primera recogió datos generales (edad, sexo, grado de instrucción). La segunda evaluó la comprensión lectora mediante nueve preguntas organizadas en tres niveles según Cooper (1999): literal, inferencial y crítico. Dichas preguntas fueron:

- ¿En qué contexto se dio la entrevista? (Nivel literal, pregunta cerrada)
- ¿El Dr. Fauci cree que la pandemia terminó? (Nivel inferencial, pregunta cerrada)
- ¿Qué opinión tienes con respecto al “fin” de la pandemia que se discutió en la entrevista? (Nivel crítico, pregunta abierta)
- ¿De acuerdo con el Dr. Fauci, el número diario de muertes es grande o pequeño? (Nivel literal, pregunta cerrada)
- ¿Qué piensas sobre la situación de las personas vacunadas de EE. UU. que comentó el Dr. Fauci? (Nivel crítico, pregunta abierta)
- ¿Por qué el doctor afirma que él hubiese podido desarrollar una Covid-19 más grave si no hubiese estado vacunado? (Nivel inferencial, pregunta cerrada)
- ¿De acuerdo con el Dr. Fauci, cómo nacen las nuevas enfermedades? (Nivel literal, pregunta cerrada)
- ¿Qué cree el Dr. Fauci sobre la prevención de esta enfermedad? (Nivel inferencial, pregunta cerrada)
- ¿Qué opinas del uso de mascarillas comentado por el Dr. Fauci? (Nivel crítico, pregunta abierta)

La tercera sección indagó en la percepción de la calidad del subtulado mediante seis preguntas cerradas basadas en el modelo NTR de Romero-Fresco y Pöchhacker (2018):

- ¿El tiempo para leer los subtítulos es suficiente?
- ¿El subtítulo en el video NO tarda mucho en aparecer?
- ¿Se logra identificar a los interlocutores en los subtítulos?
- ¿El flujo de subtítulos es constante durante todo el video?
- ¿Hay coherencia entre la imagen/el sonido y el subtítulo?
- ¿Qué calificación pondrías, con total sinceridad, a los subtítulos del video del 0 al 10, donde 0 es muy malo y 10 es muy bueno?
- Si tienes algún comentario con respecto a la calidad de los subtítulos del video, indícalo. Si no, puedes dar clic en “Enviar”.

Finalmente, para analizar la calidad, se copiaron los subtítulos a un documento de Word, donde se insertaron comentarios en los errores identificados. Luego, estos errores se registraron

en una ficha de análisis adaptada de Ríos Valero (2023), que incluyó los siguientes campos: tiempo (de aparición), número de subtítulo, enunciado original, enunciado rehablado, subtítulo traducido automáticamente y descripción (del error). Se hizo el conteo de los errores y se aplicó la fórmula del modelo NTR para obtener la tasa de precisión del subtulado interlingüístico.

4. Resultados y discusión

4.1. NTR

El subtulado generado en este estudio contiene un total de 2 402 palabras (2065 palabras y 337 signos de puntuación). En la evaluación de errores, se identificaron 12 críticos (12 puntos), 49 mayores (24,5 puntos), 68 menores (17 puntos). Por tanto, la puntuación total por los errores de traducción es de 53,5. Estos son, en su mayoría, errores de forma relacionados con puntuación, morfosintaxis, selección lexical e idiomaticidad. En menor medida, se detectaron errores de contenido (omisión, adición y sustitución). Algunos ejemplos de errores de traducción se presentan a continuación en el cuadro 1:

Cuadro 1: Errores de traducción

Ejemplo	Detalle	
1	Tiempo	09:38:08 – 09:38:24
	N.º de subtítulo	24
	Discurso original	And as Dr. Tedros said, that we're seeing the light at the end of the tunnel on that. So, again, you might interpret that, well, if that's the case, is it over? Well, again, what does "over" mean? There's a lot of semantics there, David. The easiest way not to confuse people is to say we
	Discurso rehablado	As Dr. Ted rose said, we are seeing the light at the end of the tunnel on that income. Again you might interpret if that is the case, is it over? What does over mean? There are a lot of semantics there, David. The easiest way not to confuse people is to say we
	Subtítulo traducido	Como dijo el Dr. Ted Rose, estamos viendo la luz al final del túnel en cuanto a esos ingresos. Nuevamente podrías interpretar que si ese es el caso, <u>¿se</u> acabó? ¿Qué significa sobre? Hay mucha semántica ahí, David. La forma más fácil de no confundir a la gente es decir que
	Descripción	Error menor de traducción, por tipo forma, subtipo corrección. La palabra de inicio de una pregunta se escribe con mayúscula.
Ejemplo	Detalle	
2	Tiempo	09:42:49 – 09:43:07
	N.º de subtítulo	50
	Discurso original	DB: So switching gears a little bit, or taking a step back, are consecutive pandemics kind of our new reality? We obviously had the monkeypox outbreak. And if that is the case, how do we cope with consecutive pandemics?
	Discurso rehablado	>> INTERVIEWER: So switching gears a little bit or taking a step back, are consecutive pandemics kind of our new reality? We honestly had monkeypox outbreak and if that is the case, how do we cope with consecutive pandemics?
	Subtítulo traducido	>> ENTREVISTADOR: Entonces, <u>cambiar</u> un poco de tema o dar un paso atrás, <u>¿son</u> las pandemias consecutivas nuestra nueva realidad? Honestamente, tuvimos un brote de viruela simica y, si ese es el caso, <u>¿cómo</u> afrontamos pandemias consecutivas?
	Descripción	Error mayor de traducción. El tipo de error es forma, subtipo corrección. El error es por morfosintaxis y falta de idiomaticidad.

Ejemplo	Detalle	
3	Tiempo	09:46:56 – 09:47:14
	N.º de subtítulo	67
	Discurso original	Did we know how easily it was spread from human to human? No. Did we know that it was aerosol spread? No. We thought in the beginning it was like influenza, mostly droplets from a sick person. Did we know
	Discurso rehablado	Did we know how easily it was spread from human to human? No. Did we know it was aerosol spray no? We knew from the beginning it was like flu ends up mostly droplets from a sick person. Did we know
	Subtítulo traducido	¿Sabíamos con qué facilidad se transmitía de humano a humano? No. ¿Sabíamos que era aerosol, <u>no</u> ? Sabíamos desde el principio que era como si la gripe terminara en su mayoría gotitas de una persona enferma. <u>¿Sabíamos que</u>
	Descripción	Error crítico de traducción, de tipo contenido, subtipo adición. Se agrega “no” al final de una pregunta y se produce un falso sentido, ya que el discurso original indica que era una respuesta negativa, no parte de la pregunta.

Fuente: Autora (2025)

Cabe resaltar que, en los ejemplos antes expuestos, existen más errores de traducción de contenido y forma, así como errores de reconocimiento; sin embargo, cada error se analizó de manera independiente.

De acuerdo con Dawson (2020), la frecuencia y el tipo de errores presentes en el subtítulo pueden ofrecer indicios sobre las fortalezas o debilidades específicas del rehablador o del flujo de trabajo. Una menor cantidad de errores en determinada categoría puede reflejar un mayor dominio en ese aspecto (por ejemplo, la gramática o la fidelidad al contenido), mientras que una mayor concentración de errores puede evidenciar áreas problemáticas que requieren atención.

En cuanto a los errores de reconocimiento, se identificaron 3 críticos (3 puntos), 3 mayores (1,5 puntos) y 1 menor (0,25 puntos). Por tanto, la puntuación total por los errores de reconocimiento es de 4,75. Algunos ejemplos de errores de reconocimiento se presentan a continuación en el cuadro 2:

Cuadro 2: Errores de reconocimiento

Ejemplo	Detalle	
4	Tiempo	09:39:08 – 09:39:29
	N.º de subtítulo	31
	Discurso original	AF: Well, first, you've got to get your primary series. So, I mean, that's the one where I said only 67 percent of the country has gotten their primary series, which for the most part, with some exceptions, is an mRNA vaccine, either Moderna or Pfizer, given anywhere from three to four weeks apart as the primary.
	Discurso rehablado	>> DR. ANTHONY FAUCI: First you have to get your primary see years and that's why said only 67% of the country has gotten primary series. Which for the most part with some exceptions is mRNA vaccine, either Moderna or Pfizer given 3-4 make weeks apart as primary.
	Subtítulo traducido	>> DR. ANTHONY FAUCI: Primero hay que obtener los años de primaria y por eso <u>dice</u> que sólo el 67% del país ha obtenido la serie de primaria. Que en su mayor parte, con algunas excepciones, es la vacuna de ARNm, ya sea Moderna o Pfizer administrada con 3 o 4 semanas de diferencia como primaria.
	Descripción	Error menor de reconocimiento. Se omite el pronombre I y por lo tanto se genera un sin sentido. Sin embargo, este se resolvería solo corrigiendo la ortografía de “dice” por “dije”.

Ejemplo	Detalle	
5	Tiempo	09:38:08 – 09:38:24
	N.º de subtítulo	24
	Discurso original	And as Dr. Tedros said, that we're seeing the light at the end of the tunnel on that. So, again, you might interpret that, well, if that's the case, is it over? Well, again, what does "over" mean? There's a lot of semantics there, David. The easiest way not to confuse people is to say we
	Discurso rehablado	As Dr. Ted rose said, we are seeing the light at the end of the tunnel on that income. Again you might interpret if that is the case, is it over? What does over mean? There are a lot of semantics there, David. The easiest way not to confuse people is to say we
	Subtítulo traducido	Como dijo el Dr. Ted Rose, estamos viendo la luz al final del túnel en cuanto a <u>esos ingresos</u> . Nuevamente podrías interpretar que si ese es el caso, ¿se acabó? ¿Qué significa sobre? Hay mucha semántica ahí, David. La forma más fácil de no confundir a la gente es decir que
	Descripción	Error mayor de reconocimiento. Se cambia "that" (discurso original) por "that income" (discurso rehablado), lo que causa un sin sentido en español al traducirlo como "esos ingresos" cuando el contexto es sobre las noticias de salud mundial, no sobre economía u otro tema.
Ejemplo	Detalle	
6	Tiempo	09:49:34 – 09:49:53
	N.º de subtítulo	81
	Discurso original	Had we known that, I absolutely would have said right from the beginning, everybody wear a mask all the time in an indoor setting. But we didn't say that then. It was only when it became obvious. So if we had known that early on, we would have told people to wear a mask. However, I must say, David,
	Discurso rehablado	Had we known that, I Apsley would've said write from the beginning everybody wear a mask all of the time in an indoor setting. But we didn't say that then. It was only when it became obvious. If we had known that all of long, we would've told people to wear a mask. However,, I must say David,
	Subtítulo traducido	Si hubiéramos sabido eso, Apsley habría dicho que <u>escribiríramos</u> desde el principio que todos usen una máscara todo el tiempo en un ambiente interior. Pero no dijimos eso entonces. Fue sólo cuando se hizo obvio. Si hubiéramos sabido eso desde hace mucho tiempo, le habríramos dicho a la gente que usara una máscara. Sin embargo, debo decir David,
	Descripción	Error crítico de reconocimiento. Se cambia "right" (discurso original) por "write" (discurso rehablado), lo que causa un falso sentido en español al traducirlo como "escribir".

Fuente: Autora (2025)

Cabe resaltar que, en los ejemplos antes expuestos, existen más errores de reconocimiento y de traducción; sin embargo, cada error se analizó de manera independiente.

En este estudio se identificó un mayor número de errores de traducción que de reconocimiento, lo cual contrasta con los hallazgos de Dawson (2019), quien reportó más errores de reconocimiento en uno de los casos analizados. Esta diferencia podría explicarse por la calidad del rehablado, ya que durante este proceso se espera que el rehablador escuche con atención, repita el discurso en tiempo real y articule con claridad, utilizando una prosodia adecuada. Estas condiciones influyen directamente en el rendimiento del software de reconocimiento de voz automático (Belenguer-Cortés, 2024a).

En cuanto a los errores de traducción, Dawson (2019) y Sandrelli (2020, 2021) identificaron que los errores de contenido, particularmente las omisiones, fueron los más frecuentes en los procesos de rehablado interlingüístico. De forma similar, Pagano (2022) halló que, en flujos de

trabajo que combinan rehablado intralingüístico y traducción automática, los errores de omisión y sustitución también suelen tener origen en el rehablado. En este sentido, Korybski et al. (2022) destacan que las omisiones parecen ser un factor decisivo en la evaluación de la precisión del subtítulo.

Por otro lado, estudios como los de Korybski et al. (2022) y Ríos Valero (2023) atribuyen una mayor proporción de errores de forma (estilo y corrección) a la traducción automática. En ambos casos, los flujos de trabajo implicaban una etapa de traducción automática, aunque con distintas combinaciones: reconocimiento de voz automático más traducción automática en el primero, y rehablado intralingüístico más traducción automática en el segundo.

Al tener en cuenta todos los errores de este trabajo, se aplicó la fórmula para conocer la tasa de precisión:

$$\frac{2\,402 - 53,5 - 4,75}{2\,402} \times 100 = 97,57\%$$

La tasa de precisión obtenida fue de 97,57 %, lo que, de acuerdo con el modelo NTR, equivale a una puntuación de 4/10, considerada subestándar. Cabe resaltar que los resultados de otros estudios presentan una alta variabilidad: Pagano (2022) y Ríos Valero (2023) obtuvieron una puntuación de 0/10; Sandrelli (2021), una de 3/10; mientras que Korybski et al. (2022) reportaron puntuaciones entre 5/10 y 9/10, lo que confirma la influencia decisiva del flujo de trabajo y de la intervención humana en la calidad final.

Además, se registraron 3 ediciones efectivas, que se muestran en el cuadro 3:

Cuadro 3: Ediciones efectivas

Ejemplo	Detalle	
I	Tiempo	09:38:08 – 09:38:24
	N.º de subtítulo	24
	Discurso original	And as Dr. Tedros said, that we're seeing the light at the end of the tunnel on that. <u>So</u> , again, you might interpret that, <u>well</u> , if that's the case, is it over? Well, again, what does "over" mean? There's a lot of semantics there, David. The easiest way not to confuse people is to say we
	Discurso rehablado	As Dr. Ted rose said, we are seeing the light at the end of the tunnel on that income. Again you might interpret if that is the case, is it over? What does over mean? There are a lot of semantics there, David. The easiest way not to confuse people is to say we
	Subtítulo traducido	Como dijo el Dr. Ted Rose, estamos viendo la luz al final del túnel en cuanto a esos ingresos. Nuevamente podrías interpretar que si ese es el caso, ¿se acabó? ¿Qué significa sobre? Hay mucha semántica ahí, David. La forma más fácil de no confundir a la gente es decir que
	Descripción	Se omitieron "so" y "well", probablemente muletillas, para una transmisión más fluida del mensaje.
Ejemplo	Detalle	
2	Tiempo	09:41:36 – 09:41:50
	N.º de subtítulo	42
	Discurso original	me personally that question, <u>because</u> people have different levels of risk of severity of disease. I am a person who is relatively healthy, but I'm at an elderly age. I'm 81 years old. I'm going to be

	Discurso rehablado	me personally that question. People have different levels of risk of severity of disease. I am a person who is relatively healthy. But I'm at an elderly age. I am 81 years old and I will be
	Subtítulo traducido	Yo personalmente esa pregunta. Las personas tienen diferentes niveles de riesgo de gravedad de la enfermedad. Soy una persona relativamente sana. Pero estoy en una edad avanzada. Tengo 81 años y estaré
	Descripción	Se omitió “because” y se crearon dos oraciones. Dicha omisión no refleja una pérdida de información en el discurso.
Ejemplo		Detalle
3	Tiempo	09:45:49 – 09:45:57
	N.º de subtítulo	61
	Discurso original	AF: Well, there always is. <u>I mean</u> , it's a question of when you're involved -- Nobody is perfect, certainly not I or any of my colleagues.
	Discurso rehablado	>> DR. ANTHONY FAUCI: Well there always is. It is a question, nobody is perfect and certainly not I or any of my colleagues.
	Subtítulo traducido	>> DR. ANTHONY FAUCI: Bueno, siempre lo hay. Es una cuestión, nadie es perfecto y menos yo ni ninguno de mis compañeros.
	Descripción	Se omitió “I mean”. Dicha omisión no refleja una pérdida de información en el discurso.

Fuente: Autora (2025)

Esas ediciones efectivas pueden considerarse de tipo condensación, según la clasificación propuesta por Korybski y Davitti (2024). Este tipo de intervenciones, que implican eliminar contenido redundante o fático, no afectan la comprensión del mensaje y, por el contrario, permiten producir subtítulos más concisos y legibles. La capacidad de distinguir qué información puede ser omitida sin comprometer el significado refleja la capacidad humana de decidir en tiempo real qué elementos conservar, superando las capacidades de los sistemas automáticos actuales.

4.2. Evaluación y percepción de la calidad

En cuanto al retraso, en el estudio de Chmiel et al. (2017), el EVS promedio en rehablado interlingüístico fue de 4160 ms, mientras que, en el presente estudio, el retraso fue de 17-18 segundos, teniendo en consideración que el proceso de producción del subtitulado en vivo tuvo doble intervención: el rehablado intralingüístico y la traducción automática. De acuerdo a la percepción del retraso, 16 (52 %) participantes señalaron que estaban en desacuerdo y 11 (35 %) participantes señalaron que estaban parcialmente de acuerdo con que los subtítulos aparecían de manera rápida.

Con respecto a la velocidad del video original, esta fue de 160 palabras por minuto. Asimismo, la velocidad de los subtítulos fue de 11 cps. Sin embargo, 14 (45 %) participantes alegaron que estaban parcialmente de acuerdo y 10 (32 %) participantes señalaron que estaban en desacuerdo con que el tiempo para leer era suficiente. Por otro lado, 19 (61 %) participantes indicaron que estaban parcialmente de acuerdo con que el flujo de los subtítulos fue constante.

Aunque para la identificación de interlocutores, se emplearon las etiquetas explícitas de “Entrevistador” y “Dr. Anthony Fauci”, 12 (39 %) participantes afirmaron que están parcialmente de acuerdo y 11 (35 %) participantes afirmaron que están de acuerdo con que se logra identificar claramente a los interlocutores. Esto sugiere que el retraso influyó en la percepción.

Dado el retraso de 17-18 segundos, no hubo coherencia entre el sonido/la imagen y los subtítulos. Esto fue confirmado por 19 (61 %) participantes que indicaron que estaban en desacuerdo con que haya coherencia entre el sonido/la imagen y los subtítulos.

El promedio de la puntuación de calidad, otorgada por los participantes del estudio, es de 5,6/10. Es decir, los participantes afirman que sí hay calidad. Dentro de sus comentarios finales sobre el subtulado, los participantes señalaron el poco tiempo que tuvieron para leer, la falta de coherencia entre la imagen y el subtulado, el tamaño de la letra. Algunos de esos comentarios fueron:

- Lento y rápido por momentos, no guarda relación con la imagen, muy chiquita la letra. (Participante 8)
- Los subtítulos demoraban en aparecer en la pantalla y no iban acorde a la imagen. Pero sí hubo un tiempo prudencial para leerlos. (Participante 16)
- No se pudo comprender el video al 100 % porque los subtítulos pasaban rápido y eran muy pequeños. (Participante 21)

En cuanto a la comprensión de los participantes, hubo tres preguntas por cada nivel: literal, inferencial y crítico. En cuanto al nivel literal, menos de la mitad de los participantes respondió correctamente las preguntas. En el nivel inferencial, más de la mitad de los participantes respondió correctamente dos de tres preguntas. Estas preguntas fueron sobre si el Dr. Fauci creía que la pandemia había terminado y sobre cómo le hubiese afectado la enfermedad en caso de no haber estado vacunado. Con respecto al nivel crítico, más de la mitad de los participantes respondió congruentemente solo una de tres preguntas, la cual solicitaba la opinión del participante con respecto al fin de la pandemia según lo descrito por el Dr. Fauci en el video.

5. Conclusiones

El presente trabajo es un estudio de caso que aporta evidencia sobre las limitaciones y potencialidades del flujo de trabajo que combina el rehablado intralingüístico y la traducción automática para la generación de subtítulos interlingüísticos en vivo. Aunque los resultados muestran escasos errores críticos en total, también ponen de relieve que es complejo lograr calidad general en el subtulado.

En cuanto a los errores de reconocimiento, estos son bastante escasos. Ello implicaría que la rehabladora de este estudio no tuvo problemas de pronunciación o dicción y que el programa utilizado para el reconocimiento de voz es tecnológicamente funcional. Por ello y por los resultados de Ríos Valero (2023) y Korybski et al. (2022), la tecnología del reconocimiento de voz automático estaría bien desarrollada. En este sentido, el análisis sugiere que la traducción automática no garantiza la fidelidad del mensaje original debido al gran número de errores de traducción. Esto concuerda con lo planteado por Pagano (2022), quien advierte sobre los riesgos de permitir que la traducción automática tome la decisión final para la obtención de subtítulos traducidos. Por otro lado, dado que los errores más numerosos fueron los leves y mayores, se observa que los errores lingüísticos (selección lexical, idiomática, puntuación, etc.) son incluso más relevantes que los errores de sentido (sin sentido y falso sentido).

En el caso de las ediciones efectivas, estas son limitadas, ya que la rehabladora tiene una formación y experiencia laboral con un enfoque *verbatim*. Esto influye en la carga cognitiva del espectador. Un enfoque más centrado en la reducción podría beneficiar la comprensión y disminuir el número de caracteres en el subtulado.

Dos de los factores más comentados por los participantes del estudio fueron la permanencia del subtítulo en pantalla y el tamaño de la letra. Dado que los subtítulos estaban en bloques de cuatro líneas, la letra tenía que ser pequeña. Esta característica haría también que una lectura veloz sea una tarea compleja. Otro de los factores que generó rechazo entre los participantes fue el retraso prolongado, que hizo que no hubiera coherencia entre el sonido/la imagen. De acuerdo con Eugeni (2020), la calidad se relaciona con el menor retraso posible. La percepción de dificultad para identificar a los interlocutores, generada por el retraso, así como la falta de coherencia audiovisual reforzarían la necesidad de revisar no solo los contenidos, sino también los aspectos técnicos de visualización.

El nivel inferencial de comprensión por parte de los participantes fue el nivel que obtuvo más puntuación en promedio. Esto pone de manifiesto que los participantes entendieron de manera general y no las ideas específicas del subtulado. Esta falta de captación de ideas específicas podría deberse a que los errores leves tienen un impacto más significativo en la experiencia del usuario de lo que habitualmente se supone.

La tasa de precisión del modelo NTR fue de 97,57 %, lo que en una escala del 0 al 10 equivale a 4 puntos. La puntuación de los participantes, por otro lado, fue de 5,6 sobre 10. Ambas puntuaciones dan como resultado un promedio de 4,8, indicando una calidad percibida cercana a la aprobatoria, aunque no suficiente. Esta investigación muestra que, aun con la intervención humana en el rehablado intralingüístico, la calidad del subtulado depende en gran medida de la traducción automática por los errores de traducción detectados. Para facilitar la traducción al traductor automático, Pagano (2022) indica que un rehablador intralingüístico podría simplificar su dictado. Es decir, un rehablado parafraseado, no *verbatim*, podría haber contribuido más a la calidad del subtulado (Romero-Fresco, 2009). Asimismo, la intervención de un corrector en el flujo de trabajo podría haber ayudado a reducir el rehablado intralingüístico o a corregir los errores antes de que se muestren en pantalla, aunque esto genere más retraso e imprecisión (Romero-Fresco & Eugeni, 2020).

Finalmente, el presente estudio se desarrolló antes de que las herramientas de inteligencia artificial fueran ampliamente conocidas o empleadas en entornos profesionales del subtulado en vivo. Estas tecnologías emergentes podrían representar nuevas oportunidades en el futuro. No obstante, su integración requerirá estudios rigurosos que evalúen su desempeño en condiciones reales, dado que, hasta el momento, no se ha documentado el uso profesional de estas en procesos de rehablado.

Referencias

- Abadou, F., & Khadich, S. (2019). Coherence in Machine Translation Output. *Traduction et Langues*, 18(2), 138–153. <https://doi.org/10.52919/translang.v18i2.425>
- Alsan, M. (2022, Enero 12). The best machine translation software you can try in 2022. Weglot. <https://weglot.com/blog/machine-translation-software/>

- Arma, S. (2015). Real Time Subtitling for the Deaf and Hard of Hearing: An Introduction to Conference Respeaking. In E. Perego & S. Brutti (Eds.), *Subtitling today: Shapes and their meanings* (pp. 119–134). Cambridge Scholars Publishing.
- Arnold, D. (2003). Why translation is difficult for computers. In H. Somers (Ed.), *Computers and Translation: A translator's guide* (pp. 119–142). John Benjamin Publishing Company. <https://doi.org/10.1075/btl.35.1.larn>
- Belenguer-Cortes, L. (2024a). Audiovisual Translation and Language Teaching: A Proposal on How to Use Respeaking to Improve Oral Skills. *Journal of Audiovisual Translation*, 7(1), 1–22. <https://doi.org/10.47476/jat.v7i1.2024.292>
- Belenguer-Cortes, L. (2024b). The role of accessibility in language teaching: Respeaking in the foreign language classroom. *Parallèles*, 36(1), 147–163. <https://doi.org/10.17462/para.2024.01.09>
- Chan, W. S., Kruger, J.-L., & Doherty, S. (2019). Comparing the impact of automatically generated and corrected subtitles on cognitive load and learning in a firstand second-language educational context. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 237–272. <https://doi.org/10.52034/lansts.v18i0.506>
- Chmiel, A., Szarkowska, A., Korzinek, D., Lijewska, A., Dutka, Ł., Brocki, Ł., & Marasek, K. (2017). Ear–voice span and pauses in intra- and interlingual respeaking: An exploratory study into temporal aspects of the respeaking process. *Applied Psycholinguistics*, 38(5), 1201–1227. <https://doi.org/10.1017/S0142716417000108>
- Cooper, D. (1999). *Cómo mejorar la comprensión lectora*. Editorial Visor.
- Cucchiarini, C., Neri, A., & Strik, H. (2009). Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. *Speech Communication*, 51(10), 853–863. <https://doi.org/10.1016/J.SPECOM.2009.03.003>
- Davitti, E., & Sandrelli, A. (2020). Embracing the Complexity: A Pilot Study on Interlingual Respeaking. *Journal of Audiovisual Translation*, 3(2), 103–139. <https://doi.org/10.47476/jat.v3i2.2020.135>
- Dawson, H. (2019). Feasibility, Quality and Assessment of Interlingual Live Subtitling: A Pilot Study. *Journal of Audiovisual Translation*, 2(2), 36–56. <https://doi.org/10.47476/jat.v2i2.72>
- Dawson, H. (2020). A Research-Informed Training Course for Interlingual Respeaking. *Journal of Audiovisual Translation*, 3(2), 204–225. <https://doi.org/10.47476/jat.v3i2.2020.126>
- Defrancq, B., & Fantinuoli, C. (2021). Automatic speech recognition in the booth Assessment of system performance, interpreters' performances and interactions in the context of numbers. *Target*, 33(1), 73–102. <https://doi.org/10.1075/target.19166.def>
- Eugení, C. (2020). Respeaking: aspects techniques, professionnels et linguistiques du sous-titrage en direct. *ESSACHESS-Journal for Communication Studies*, 13(1), 21–35. <https://essachess.com/3/index.php/jcs/article/view/176/176>
- Evanini, K., Timpe-Laughlin, V., Tsuprun, E., Blood, I., Lee, J., Bruno, J., Ramanarayanan, V., Lange, P., & Suendermann-Oef, D. (2018). Game-based spoken dialog language learning applications for young students. *Interspeech*, 548–549.
- Gil Ramírez, M., Gómez de Travesedo-Rojas, R., & Almansa-Martínez, A. (2020). YouTube y coronavirus: análisis del consumo de videos sobre la pandemia Covid-19. *Revista Latina de Comunicación Social*, 78, 121–153. <https://doi.org/10.4185/rcls-2020-1471>

- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Greco, G. M. (2018). The nature of accessibility studies. *Journal of Audiovisual Translation*, 1(1), 205–232. <https://doi.org/10.47476/jat.v1i1.51>
- IBM. (2021, 28 Septiembre). What is speech recognition? IBM. <https://www.ibm.com/topics/speech-recognition>
- Keyes, B. (2007, Julio). *Realtime by Voice: Just what you need to know* [Discurso]. Intersteno Congress. https://www.intersteno.it/materiale/Praga2007/praga_conferences/BettyeKeyes.htm
- Korybski, T., Davitti, E., Orăsan, C., & Braun, S. (2022). A Semi-Automated Live Interlingual Communication Workflow Featuring Intralingual Respeaking: Evaluation and Benchmarking. *Proceedings of the 13th Conference on Language Resources and Evaluation* (pp. 4405–4413). <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.468.pdf>
- Korybski, T., & Davitti, E. (2024). Human Agency in Live Subtitling Through Respeaking: Towards a Taxonomy of Effective Editing. *Journal of Audiovisual Translation*, 7(2), 1–22. <https://doi.org/10.47476/jat.v7i2.2024.302>
- Lambourne, A. (2007). Real-time Subtitling: Extreme Audiovisual Translation. *Comunicazione alla conferenza internazionale Multidimensional Translation: LSP Translation Scenarios*.
- Li, H. O.-Y., Bailey, A., Huynh, D., & Chan, J. (2020). YouTube as a source of information on Covid-19: a pandemic of misinformation? *BMJ Global Health*, 5(5), 1–6. <https://doi.org/10.1136/bmigh-2020-002604>
- Lu, X., Li, S., & Fujimoto, M. (2020). Automatic Speech Recognition. In Y. Kidiwara, E. Sumita & H. Kawai (Eds.), *Speech-to-speech translation* (pp. 21–38). SpringerBriefs in Computer Science. https://doi.org/10.1007/978-981-15-0595-9_2
- Meza-Liviapoma, J. G., Aragón-Ayala, C. J., Pérez-Chuquihuanca, K., & Quispe-Juli, C. U. (2021). Recursos de YouTube para luchar contra el Covid-19: una mirada de América Latina y el Caribe. *Revista Cubana de Medicina General Integral*, 37, 1–12.
- Morton, H., Gunson, N., & Jack, M. (2012). Interactive Language Learning through Speech-Enabled Virtual Scenarios. *Advances in Human-Computer Interaction* 1–14. <https://doi.org/10.1155/2012/389523>
- Mundasad, S., & Roxby, P. (2023, Mayo 5). Covid global health emergency is over, WHO says. BBC. <https://www.bbc.com/news/health-65499929>
- Nashville Film Institute. (2022). *Is YouTube a social media? Everything you need to know*. Nashville Film Institute. <https://www.nfi.edu/is-youtube-social-media/>
- Oliver, A., Moré, J., & Climent, S. (2012). *Traducción y tecnologías*. Editorial UOC.
- Pagano, A. (2022). *Testing quality in interlingual respeaking and other methods of interlingual live subtitling*. [Tesis de Doctorado]. Universita di Genova. <https://hdl.handle.net/11567/1091438>
- Parton, B. S. (2016). Video captions for online courses: Do YouTube's auto-generated captions meet deaf students' needs? *Journal of Open, Flexible, and Distance Learning*, 20(1), 8–18. <https://doi.org/10.61468/jofdl.v20i1.255>
- Radio Programas del Perú. (2019, Noviembre 5). 8 de cada 10 peruanos conectados a internet ingresan a YouTube todos los días, según estudio. *RPP*.

- <https://rpp.pe/tecnologia/google/youtube-8-de-cada-10-peruanos-conectados-a-internet-ingresan-a-la-plataforma-todos-los-dias-segun-estudio-noticia-1228332?ref=rpp>
- Rajendran, D. J., Duchowski, A. T., Orero, P., Martínez, J., & Romero-Fresco, P. (2013). Effects of text chunking on subtitling: A quantitative and qualitative examination. *Perspectives: Studies in Translation Theory and Practice*, 21(1), 5–21. <https://doi.org/10.1080/0907676x.2012.722651>
- Rella, S. (2022, Agosto 8). Essential Guide to Automatic Speech Recognition Technology. Nvidia Developer. <https://developer.nvidia.com/blog/essential-guide-to-automatic-speech-recognition-technology/>
- Ríos Valero, L. (2023). Análisis de la calidad. El subtitulado en vivo interlingüístico de YouTube en un programa de noticias estadounidense accesible a usuarios peruanos. *Entreculturas*, 13, 143–168. <https://doi.org/10.24310/EntreculturaserTCI.v13.i15352>
- Robert, I. S., Schrijver, I., & Diels, E. (2020). Live subtitlers: Who are they? A survey study. *Linguistica Antverpiensia, New Series–Themes in Translation Studies*, 18, 101–129. <https://doi.org/10.52034/lanstts.v18i0.544>
- Romero-Fresco, P. (2009). More haste less speed: edited versus verbatim respooken subtitles. *Vigo International Journal of Applied Linguistics*, 6, 109–133.
- Romero-Fresco, P. (2011). *Subtitling through speech recognition: Respeaking*. St Jerome.
- Romero-Fresco, P. (2013). Accessible filmmaking: Joining the dots between audiovisual translation, accessibility and filmmaking. *The Journal of Specialised Translation*, 20, 201–223. <https://doi.org/10.26034/cm.jostrans.2013.410>
- Romero-Fresco, P. (2016). Accessing communication: The quality of live subtitles in the UK. *Language and Communication*, 49, 56–69. <https://doi.org/10.1016/j.langcom.2016.06.001>
- Romero-Fresco, P., & Alonso-Bacigalupe, L. (2022). An empirical analysis on the efficiency of five interlingual live subtitling workflows. *XLinguae*, 15(2), 3–16. <https://www.doi.org/10.18355/XL.2022.15.02.01>
- Romero-Fresco, P., & Eugeni, C. (2020). Live subtitling through respeaking. In L. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 269–295). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-42105-2_14
- Romero-Fresco, P., & Martínez Pérez, J. (2015). Accuracy Rate in Live Subtitling: The NER Model. In R. Baños Piñero & J. Díaz Cintas (Eds.), *Audiovisual translation in a global context: Mapping an ever-changing landscape* (pp. 28–50). Springer. https://doi.org/10.1057/9781137552891_3
- Romero-Fresco, P., & Pöchhacker, F. (2018). Quality assessment in interlingual live subtitling: The NTR model. *Linguistica Antverpiensia*, 16, 149–167. <https://doi.org/10.52034/lanstts.v16i0.438>
- Rubio, I. (2020, Mayo 29). ¿Cuál es el mejor traductor?: probamos DeepL, Google Translate y Bing. *El País*. <https://elpais.com/tecnologia/2020-05-29/cual-es-el-mejor-traductor-probamos-deepl-google-translate-y-bing.html>
- Sánchez Ramos, M. D. M., & Rico Pérez, C. (2020). *Traducción automática: conceptos clave, procesos de evaluación y técnicas de posedición*. Comares.
- Sandrelli, A. (2020). Interlingual respeaking and simultaneous interpreting in a conference setting: a comparison. *InTRALínea. Special Issue: Technology in Interpreter Education and Practice*.
- Sandrelli, A. (2021). Eventi dal vivo e accessibilità: Uno studio di caso sul respeaking interlinguistico. *Lingue e Linguaggi*, 43, 145–168. <https://doi.org/10.1285/i22390359v43p145>
- Simons, H. (2013). *El estudio de caso: teoría y práctica*. Ediciones Morata, S. L.

- Villela, L. M. (2019). Entrevista com Pablo Romero-Fresco. *Cadernos de Tradução*, 39(2), 316–325. <https://doi.org/10.5007/2175-7968.2019v39n2p316>
- Wald, M., & Bain, K. (2008). Universal access to communication and learning: The role of automatic speech recognition. *Universal Access in the Information Society*, 6(4), 435–447. <https://doi.org/10.1007/s10209-007-0093-9>
- Williams, N. (2022, Julio 27). What role has social media played in Covid-19? *News Medical Life Science*. <https://www.news-medical.net/health/What-Role-has-Social-Media-Played-in-Covid-19.aspx>
- Zajechowski, M. (2014). Automatic Speech Recognition (ASR) Software – An Introduction. *Usability Geek*. <https://usabilitygeek.com/automatic-speech-recognition-asr-software-an-introduction/>

Notas editoriales

Contribución de autoría

Concepción y elaboración del manuscrito: L. Rios-Valero

Recolección de datos: L. Rios-Valero

Ánalysis de datos: L. Rios-Valero

Discusión y resultados: L. Rios-Valero

Revisión y aprobación: L. Rios-Valero

Datos de la investigación

Los datos fueron recolectados y analizados en el marco de una tesis doctoral en curso en el Programa de Doctorado en Traductología, Traducción Profesional y Audiovisual de la Universidad de Valladolid.

Financiación

No se aplica.

Derechos de uso de imagen

No se aplica.

Aprobación del comité de ética de la investigación

En el momento de la realización y envío de este estudio (septiembre de 2024), la Universidad de Valladolid no contaba con un Comité de Ética de la Investigación con implicaciones sociales y humanas (CEISH) operativo, el cual se constituyó formalmente en diciembre de 2024. La investigación, de carácter no invasivo y de riesgo mínimo, se desarrolló siguiendo principios éticos internacionales para estudios con seres humanos, garantizando el consentimiento informado, la participación voluntaria y la confidencialidad de los datos.

Conflictos de intereses

No se aplica.

Declaración de disponibilidad de datos de investigación

Los datos de esta investigación, que no están expresados en este trabajo, podrán ser proporcionados por la autora bajo solicitud.

Licencia de uso

Autoras y autores ceden a *Cadernos de Tradução* los derechos exclusivos de primera publicación, con el trabajo simultáneamente licenciado bajo la [Licencia Creative Commons Atribución 4.0. Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). Esta licencia permite a terceros remezclar, adaptar y crear a partir del trabajo publicado, otorgando el crédito adecuado de autoría y publicación inicial en esta revista. Autoras y autores están autorizados a celebrar contratos adicionales por separado para distribuir de manera no exclusiva la versión del trabajo publicado en esta revista (por ejemplo, publicarlo en un repositorio institucional, en un sitio web personal, en redes sociales académicas, realizar una traducción o republicar el trabajo como un capítulo de libro), siempre y cuando se reconozca la autoría y la publicación inicial en esta revista.



Publisher

Cadernos de Tradução es una publicación del Programa de Posgrado en Estudios de Traducción de la Universidad Federal de Santa Catarina. La revista *Cadernos de Tradução* está alojada en el [Portal de Periódicos UFSC](#). Las ideas expresadas en este artículo son responsabilidad de su autora y no representan necesariamente la opinión del equipo editorial o de la universidad.

Edición de sección

Andréia Guerini – Willian Moura

Normalización

Alice S. Rezende – Ingrid Bignardi – João G. P. Silveira – Kamila Oliveira

Historial

Recibido el: 30-09-2024

Aprobado el: 15-05-2025

Revisado el: 15-08-2025

Publicación: 11-2025

